

01010

UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO  
FACULTAD DE FILOSOFÍA Y LETRAS  
DIVISIÓN DE ESTUDIOS DE POSGRDO



**EXÁMENES MASIVOS:  
UN ANÁLISIS ENTRE PRUEBAS  
INTERNACIONALES Y NACIONALES**

**T E S I S**

QUE PARA OBTENER EL GRADO DE  
MAESTRA EN PEDAGOGÍA

PRESENTA  
**ROSA AURORA PADILLA MAGAÑA**

**DIRECTOR DE TESIS:  
DR. ÁNGEL DÍAZ BARRIGA**

CIUDAD UNIVERSITARIA

MAYO 2005

m 343860

**CADA PASO EN EL CAMINO, EL MEJOR MOMENTO PARA, REVALORAR, APRECIAR Y RECONOCER:  
UN AGRADECIMIENTO LLENO DE CARIÑO A:**

**QUIENES ME DIERON LA VIDA:**

*Mi Padre:*

Mi más fuerte inspirador, aunque partiste físicamente, siempre te he sentido ahí.  
Los ideales compartidos, han impulsando lo hasta hoy logrado y permanecen como una herencia invaluable  
en mi ser, sentir y actuar cotidiano.

*Mamá:* Tu presencia, cariño, sacrificio y aceptación silenciosa a mis inquietudes y "locuras" han sido  
fundamentales en mi vida.

**MIS CÓMPLICES DE VIDA Y ANDANZAS:**

*Claudia, Constanza, Fer, Belén y Pepe.*

Importantes pilares en tiempos difíciles, fundamentales para impulsar, fortalecer y acompañar...

Su presencia en la aventura de la vida, una motivación y retroalimentación constante...

Gracias por saber compartir un sin fin de experiencias y mucho más...

Su amistad sincera, uno de los más maravillosos regalos de la vida.

**MIS IMPORTANTES INTERLOCUTORES:**

*Ángel Díaz Barriga:*

El camino recorrido para la realización y construcción de este trabajo  
y la riqueza de los aprendizajes logrados; es sin duda gracias al diálogo permanente  
y a tus sabias orientaciones en materia de investigación, trabajo académico, pero sobre todo, de vida...

Tu apoyo y solidaridad en momentos de crisis y confusión, algo inigualable, imposible de olvidar...

Gracias por tu generosidad, por tu escucha, por brindarme la oportunidad...

por compartir tu tiempo, tu espacio, tu experiencia...

*Caty Inclán:*

Tu apoyo permanente, tu confianza, tus consejos, tu orientación, pero sobre todo, el hacerte presente cuando  
más lo necesite... me ayudaron a superar, aprender, continuar y llegar... Gracias mil, por todo lo compartido,  
gracias por tu solidaridad y calidez humana, por saber estar ahí...

**QUIENES DICTAMINARON ESTE TRABAJO:**

*Concepción Barrón, Frida Díaz Barriga, Catalina Inclán y Laura Mercado:*

Gracias mil, por su interés, disposición y tiempo, por sus enseñanzas, comentarios y valiosas observaciones;  
que sin duda alguna retroalimentaron, enriquecieron y ayudaron a repensar lo hoy presentado.

Pero sobre todo, gracias por su confianza y valioso apoyo.

**MIS PRINCIPALES ESPACIOS DE FORMACIÓN Y CRECIMIENTO PERSONAL Y ACADÉMICO:**

- La Universidad Nacional Autónoma de México, mi alma mater.
- Mis profesores durante la maestría, quienes me contagiaron su pasión por la investigación, la docencia y la vida académica. Un reconocimiento especial a la maestra Patricia Ducoing.
- El Centro de Estudios sobre la Universidad, de la UNAM, un importante espacio de formación, oasis y vida académica. Gracias a todos por la acogida durante mi estancia.
- El Seminario de Investigación en Educación Básica iniciado en el CESU en el 2001, un espacio fundamental para penetrar en la investigación y gestar de este trabajo.
- El estímulo económico recibido por el Sistema Nacional de Investigadores (SNI), como ayudante de investigador nacional, un apoyo, pero sobre todo un gran compromiso...

GRACIAS MIL, A TODOS AQUELLOS QUE COMPARTIERON LO HASTA HOY LOGRADO.

*A todos aquellos para quienes la docencia es su pasión  
y un espacio de creación, recreación y constante búsqueda...*

## ÍNDICE

---

	Pág.
<b>Introducción.....</b>	<b>I</b>
1. Problema y justificación.....	VI
2. Objeto de estudio.....	IX
3. Objetivos.....	IX
4. Abordaje metodológico.....	X
5. Instrumentos de investigación.....	XIV
6. Estructura y contenido.....	XVI
7. Principales hallazgos.....	XVII

### CAPÍTULO 1

<b>I. Pruebas Internacionales: Principales orientaciones.....</b>	<b>1</b>
1.1. Generación y propósito de los estudios internacionales de evaluación.....	3
Antecedentes.....	3
a. Tercer Estudio Internacional sobre Matemáticas y Ciencias: TIMSS.....	5
b. Laboratorio Latinoamericano de Evaluación de la Calidad Educativa: LLECE.....	9
c. Programa Internacional de Evaluación de Estudiantes: PISA.....	12
1.2. Contenido de las pruebas internacionales para alumnos de cuarto grado de la educación básica: (8 y 9 años) .....	17
a. Pruebas de matemáticas: TIMSS y LLECE.....	18
b. Pruebas de ciencias: TIMSS.....	28
c. Pruebas del lenguaje: LLECE.....	33
1.3. Contenido de las pruebas internacionales para alumnos que están concluyendo la educación básica.....	36
a. Pruebas de matemáticas: TIMSS y PISA.....	36
b. Pruebas de ciencias: TIMSS y PISA.....	41
c. Pruebas de lectura: PISA.....	47
1.4. Estructura de las Pruebas Internacionales: TIMSS, LLECE y PISA.....	56
a. Normas y procedimientos para la elaboración de las pruebas.....	57
Normas técnicas de las pruebas del TIMSS.....	57
Elaboración de las pruebas del LLECE.....	59
Procedimiento para elaborar las pruebas del PISA.....	60
b. Diseño de ítems en las pruebas internacionales: TIMSS, LLECE y PISA. ....	62
ítems de las pruebas del TIMSS .....	62
ítems de las pruebas del LLECE.....	65
ítems de las pruebas del PISA.....	66

c. Características de las tareas de las Pruebas Internacionales: TIMSS, LLECE y PISA .....	69
Tareas en las pruebas del TIMSS.....	69
Tareas en las pruebas del LLECE.....	69
Tareas en las pruebas del PISA.....	71
d. Sistema de calificación y puntuación de las pruebas internacionales.....	76
Procedimiento de calificación: TIMSS, LLECE y PISA.....	76
Puntuación asignada a los tipos de reactivos: TIMSS, LLECE y PISA.....	83
1.5 Evaluación de factores asociados en los estudios internacionales: TIMSS, LLECE y PISA.....	87
Evaluación de factores asociados en el TIMSS.....	87
Evaluación de factores asociados en el LLECE.....	90
Evaluación de factores asociados en el PISA.....	93
1.6 Primeras conclusiones.....	97

## CAPÍTULO 2

<b>II. El concepto de examen en las Pruebas Internacionales.....</b>	<b>101</b>
2.1. Concepto de contenido en las pruebas internacionales.....	101
a. Conceptos implícitos en las pruebas internacionales.....	102
b. El concepto de competencia en los exámenes internacionales.....	108
▪ El concepto de competencia en las pruebas del TIMSS.....	108
▪ El concepto de competencia en las pruebas del LLECE.....	108
▪ El concepto de competencia en las pruebas del PISA.....	109
c. Habilidades evaluadas en las pruebas internacionales.....	112
2.2. Tipo de ítems en los exámenes internacionales.....	114
a. Ítems del examen del TIMSS.....	115
▪ Ítems de matemáticas para cuarto grado (alumnos de 8 y 9 años)...	115
▪ Ítems de matemáticas para octavo grado (alumnos de 13-14 años que están por concluir la educación básica).....	117
▪ Ítems de ciencias para cuarto grado (alumnos de 8 y 9 años).....	120
▪ Ítems de ciencias para octavo grado (alumnos de 13-14 años que están por concluir la educación básica).....	121
b. Ítems del examen del LLECE.....	123
▪ Ítems de matemáticas para tercer y cuarto grado.....	123
▪ Ítems de lenguaje para tercer y cuarto grado.....	125
c. Ítems del examen del PISA.....	127
▪ Ítems de lectura par jóvenes de 15 años.....	128
▪ Ítems de matemáticas para jóvenes de 15 años.....	130
▪ Ítems de ciencias para jóvenes de 15 años.....	133

2.3. Contenido de los ítems de matemáticas y ciencias de los exámenes internacionales .....	136
2.4. Modelo psicométrico en los exámenes internacionales.....	139
2.5. Conclusiones preliminares: Lo que se refiere de las Pruebas Internacionales para la elaboración de exámenes.....	141

### CAPÍTULO 3

<b>III. La experiencia mexicana en pruebas nacionales.....</b>	<b>142</b>
3.1. Exámenes masivos para la educación básica: Algunas pruebas que se aplican en los salones de clase de la escuela primaria mexicana. ....	142
1) Instrumento de Diagnóstico para Alumnos de Nuevo Ingreso a Secundaria (IDANIS) .....	144
□ Origen del IDANIS.....	144
□ Propósito del IDANIS.....	144
□ Contenidos del IDANIS.....	145
□ Ítems del IDANIS.....	147
▪ Ítems de comprensión lectora para alumnos de 6° grado.....	148
▪ Ítems de habilidad verbal para alumnos de 6° grado.....	149
▪ Ítems de habilidad matemática para alumnos de 6° grado.....	151
▪ Ítems de razonamiento abstracto para alumnos de 6° grado.....	155
2) Pruebas del Sistema Nacional de Evaluación Educativa (SNEE).....	157
3) Pruebas con base en estándares, ahora Pruebas Nacionales (PN).....	163
□ Orígenes de las pruebas del SNEE.....	157
□ Propósito de las pruebas del SNEE.....	157
□ Contenidos de las pruebas del SNEE para el 3° grado de primaria.....	157
□ Ítems de las pruebas del SNEE para el 3° grado de primaria.....	159
▪ Ítems de la prueba de español.....	160
▪ Ítems de las pruebas de matemáticas.....	161
□ Orígenes de las Pruebas Nacionales.....	163
□ Propósito de las Pruebas Nacionales.....	164
□ Contenidos de las Pruebas Nacionales para la educación primaria.....	165
▪ Evaluación de la comprensión lectora en las pruebas nacionales.....	165
▪ Ítems de comprensión lectora para el 6° grado .....	171
▪ Evaluación de las matemáticas en las Pruebas Nacionales.....	174
▪ Ítems de matemáticas para el 6° grado.....	180
▪ Escalas y normas técnicas de las Pruebas Nacionales.....	183

5) Examen Nacional de Ingreso a la Educación Media Superior (EXANI-1) para alumnos que concluyen la educación básica.....	184
□ Orígenes del EXANI-1.....	184
□ Propósitos del EXANI-1.....	185
□ Contenido del EXANI-1 .....	185
□ Ítems del EXANI-1.....	186
▪ Ítems de comprensión lectora para alumnos que concluyen la educación básica.....	187
▪ Ítems de habilidad matemática para alumnos que concluyen la educación básica .....	189
▪ Ítems para evaluar conocimientos disciplinares en alumnos que concluyen la educación básica.....	190
□ Normas técnicas para la elaboración del EXANI-1.....	192
3.2. Características de los exámenes masivos nacionales.....	194
3.3. Como conclusión inicial y a manera de síntesis .....	196

## CAPÍTULO 4

<b>IV. Pruebas internacionales y pruebas nacionales: Contenidos y teorías.....</b>	<b>199</b>
4.1. Concepto de contenido en los exámenes masivos... ..	199
a. Exámenes masivos para alumnos de ocho y nueve años de edad.....	200
□ Contenido de la evaluación matemática.....	200
□ Contenido de la evaluación de la comprensión lectora.....	203
b. Exámenes masivos para alumnos que concluyen la educación básica.....	205
□ Contenido de la evaluación matemática .....	205
□ Contenido de la evaluación de la comprensión lectora .....	207
c. Conceptos estructurantes en los exámenes masivos.....	208
d. Algunas reflexiones a propósito del concepto de contenido en los exámenes masivos .....	213
4.2 Tipo de ítem en los exámenes masivos.....	216
a. Composición de los ítems de matemáticas.....	216
b. Estructura y formato de los ítems.....	218
4.3. Modelo psicométrico en los exámenes masivos.....	220
4.4. Para concluir: unas comparaciones finales .....	224
4.5. Tareas pendientes.....	226
<b>Conclusiones.....</b>	<b>228</b>
i. Una respuesta a las interrogantes planteadas .....	228
ii. Un concepto de examen diferente.....	231
iii. Lo que se puede apreciar.....	233
iv. Algunas aportaciones y sugerencias.....	235
v. Una puerta a nuevos debates y cuestionamientos.....	236

Referencias..... 240

## INTRODUCCIÓN<sup>1</sup>

---

La reflexión y discusión sobre diversas problemáticas por las que atraviesa la educación básica en América Latina en materia de política educativa, evaluación y docencia, aunadas a la inquietud intelectual con la que inicio mis estudios de maestría, misma que quedó plasmada en el ensayo que presenté como requisito de ingreso<sup>2</sup> y a mi historia personal como docente y formadora de docentes, impulsan el abordar la temática aquí presentada.

Se puede reconocer la década de los noventa, como etapa de auge en materia de evaluación y medición del rendimiento escolar en América Latina; dada la importancia que esta disciplina cobra a raíz de las políticas de evaluación en las que los sistemas educativos de México y Latinoamérica se han visto inmersos.

La evaluación, como disciplina compleja y de gran controversia, ha proliferado de una manera impresionante, entrando abruptamente a nuestra concepción para valorar lo educativo. Sin embargo, las evaluaciones que generalmente se hacen de los sistemas educativos, adolecen de una reflexión teórica y metodológica, que permita al docente identificar las actuales tendencias en materia de evaluación del aprendizaje y la evolución académica que el campo ha tenido.

Ubicar a la evaluación en su debido marco histórico, permitirá identificar los paradigmas bajo los cuales se le ha abordado y delinear, presentar y acotar las nuevas miradas que en el plano internacional se están generando.

---

<sup>1</sup> El presente trabajo, surge en el marco de un Seminario de Investigación sobre Educación Básica en América Latina, organizado y coordinado por Ángel Díaz Barriga, que inició en el 2001 con una reunión mensual, en el Centro de Estudios sobre la Universidad (CESU); las aportaciones de cada uno de los participantes, quienes también se encontraban realizando trabajos de investigación en la misma línea, ayudaron a repensar y delimitar el objeto de estudio de esta investigación.

<sup>2</sup> El ensayo con el que inicio mis estudios de maestría se denomina: "La evaluación como reflexión sobre los procesos del aprendizaje y el desempeño personal y grupal. Un enfoque cualitativo y constructivista." En este se expresa la necesidad de conceptualizar la evaluación en sus diferentes dimensiones: fines, objeto y método; delimitando el enfoque teórico desde el cual se abordará y considerándola como parte integrante del proceso educativo y didáctico y no como elemento externo. Es decir, como un espacio de reflexión en donde tanto docentes como alumnos -principales agentes del proceso educativo- identifiquen, analicen y expresen libremente los diversos factores que intervienen en el aprendizaje, para ubicarla como acción formativa, tendiente a optimizar el quehacer pedagógico.



Uno de los primeros instrumentos en materia de evaluación, que responde a muchas de las intenciones de la política educativa actual, es sin duda el *examen*, término polémico y controversial, al que se le han adjudicado situaciones de orden no sólo educativo, sino sobre todo normativas, políticas y sociales.<sup>3</sup>

Una breve revisión, de la historia que ha tenido el examen<sup>4</sup> desde su origen y primeras manifestaciones, hasta nuestros días, reflejará las perversiones sociales y educativas que este instrumento ha adquirido.

Comenio, en su *Didáctica Magna* del siglo XVII (1657), (1922/2000: 98-100) es el único en la historia del examen, que lo presentaba como una forma de revisar el método del maestro y el aprendizaje del alumno; con el fin de estimular la atención, de corregir dictados, composiciones y diversos ejercicios.

La degeneración del examen y su abandono como herramienta didáctica del maestro, comienza en el siglo XVIII, con la normalización y estandarización<sup>5</sup> de éstos; así como con la incursión de sistemas de calificación y cuantificación,<sup>6</sup> como forma de medir rendimiento.<sup>7</sup> Foucault (1977), refiere el examen, como un instrumento en el que las relaciones de saber, se invierten por las relaciones de poder, dándole un uso normalizado y presentándolo como un medio de control,<sup>8</sup> que jerárquicamente se ha utilizado para vigilar, sancionar y normar, con el fin de calificar, clasificar y castigar.

---

<sup>3</sup> Entre los textos que favorecen la reflexión sobre la evolución histórica de los exámenes, los distintos enfoques de las disciplinas que lo abordan y las discusiones que con respecto a éste se han generado se pueden citar: (Díaz Barriga 1988:65-76 y 1993) (Escudero, 2003) y (Stufflebeam y Shinkfield, 1987).

<sup>4</sup> Para esta reseña histórica del examen, se toma como base, el libro compilado por Díaz Barriga (1993), por considerarlo un material, que logra reunir textos básicos que reflejan claramente los distintos momentos, miradas y debates por los que el examen ha pasado a través de la historia.

<sup>5</sup> "La palabra "estandarizado" en el título de un test significa únicamente que todos los alumnos tienen que contestar las mismas preguntas, y un gran número de las mismas conforme a instrucciones uniformes y con límite de tiempo también uniforme, y que existe un grupo de referencia estándar o uniforme con cuya ejecución puede compararse la actuación del alumno" (Thorndike y Hagen: 272 ).

<sup>6</sup> Díaz Barriga le llama a esto una "pedagogía del examen". "Una pedagogía que se articula en función de la acreditación, descuidando notoriamente los problemas de formación, procesos cognoscitivos y aprendizaje" (1988:69)

<sup>7</sup> Nótese la vinculación de este término con los requerimientos del mundo del trabajo, lo que denota la incursión de conceptos emanados del ámbito laboral a la evaluación del aprendizaje.

<sup>8</sup> Díaz Barriga (1988:70), también habla del examen como un problema de control científico, en el siglo XX.



El siglo XIX, es la época precedente a los test<sup>9</sup> y a la medición, de donde nace un sistema de exámenes, como comprobación de una preparación específica para satisfacer las necesidades de una sociedad jerárquica y burocratizada (Escudero: 2003).

A partir del siglo XX, se institucionaliza como herramienta de medición,<sup>10</sup> con un paradigma experimental y enfoque psicométrico<sup>11</sup> que prevalece hasta nuestros días y responde a la constitución de la psicología como ciencia. Es aquí cuando emerge la Teoría de la Medición, llamada también Teoría Clásica del Test.<sup>12</sup> como medio para la elaboración y aplicación de test de todo tipo; surgen también los exámenes referidos a la norma<sup>13</sup> para comparar rendimientos de un individuo o grupo y las muy conocidas pruebas "objetivas."<sup>14</sup>

Es en la segunda mitad del siglo XX, cuando la evaluación adquiere un carácter profesionalizante, comienza a expandirse y a diversificarse, surgen nuevos objetos, finalidades, formas y enfoques; para dar paso a la evaluación de docentes, instituciones y programas.<sup>15</sup>

---

<sup>9</sup>Término con que se reemplaza el concepto de examen con miras a científizar el control ejercido por éstos. "El test fue considerado como un instrumento científico, válido y objetivo para medir factores psicológicos de un individuo: la inteligencia, las aptitudes e intereses y el aprendizaje" (Díaz Barriga, 1988: 70-76).

<sup>10</sup>"Medir es asignar numerales a objetos o acontecimientos de acuerdo con ciertas reglas" (Stevens, 1951: 17, citado por Martínez y Moreno eds. 2002: 88). Considerando que el isomorfismo, "es una regla que establece que para que la medición tenga validez debe haber una correspondencia entre las propiedades del objeto a medir y las propiedades del número, [y es] sobre lo observable y cuantificable que se pueden hacer mediciones" (Benedito, 1975: 161-166) El problema surge cuando lo que se va a medir, no es observable o cuantificable.

<sup>11</sup> "La psicometría puede definirse en términos generales como el conjunto de métodos, técnicas y teorías implicadas en la medición de variables psicológicas" (Muñiz, 1998: 17).

<sup>12</sup> "El nacimiento formal de la teoría del test puede ubicarse en los primeros trabajos de Spearman (1904, 1907, 1913), en los que establece los cimientos de la Teoría Clásica de los Test (TCT). El objetivo es encontrar un modelo estadístico que fundamentase adecuadamente las puntuaciones de los test y permitiera la estimación de los errores de medida asociados a todo proceso de medición" (Muñiz, 1998: 19).

<sup>13</sup> Las normas son marcos de referencia para la interpretación de pruebas tipificadas. Los exámenes referidos a la norma, son pruebas tipificadas, sujetas a un control experimental. También se les denomina estandarizadas, en relación al empleo de procedimientos fijos para aplicarlas, calificarlas e interpretarlas. Además son normalizadas porque se elaboran con base a normas establecidas (García, 1983: 219-221). Son pruebas estandarizadas que describen los resultados en relación a la posición que ocupa un individuo dentro de un grupo, de acuerdo con la curva normal de distribución de frecuencias. "El punto intermedio de la distribución normal o "promedio" se convierte en la base para la calificación" (Woolfok, 1996: 575).

<sup>14</sup> "El término "objetiva", en relación con la evaluación significa no abierta a muchas interpretaciones" (Woolfok, 553).

<sup>15</sup> Desde mediados de los años sesenta, se comienza a ubicar a la evaluación como medio para proveer información. Cronbach, fue de los primeros que amplió el alcance de la evaluación a la oferta de información para los responsables de decisiones, ubicándola como una actividad de servicio educativo. Como autores que han influido en la configuración de un pensamiento contemporáneo sobre una evaluación más de corte cualitativo, podemos citar a: Scriven (1967) quien promueve una evaluación formativa con miras a perfeccionar el currículo, Stufflebeam (1971) quien plantea un modelo de evaluación como servicio a directivos y administradores, Mc Donald en el Reino Unido (1974), quien se promulga por una



El siguiente cuadro permite identificar al examen en sus diferentes etapas, como instrumento de selección, control, acreditación, calificación y medición, más que de evaluación.

CUADRO1

**EL EXAMEN EN SU DESARROLLO HISTÓRICO**

Época	Características y usos
Antigua	Origen en la dinastía China como medio de selección.
Medieval	Aplicación de exámenes de corte escolástico en la universidad para determinar la idoneidad de los candidatos y poner a prueba los conocimientos adquiridos.
Comeniana (S. XVII)	Vinculados al método didáctico.
S. XVIII	Como instrumento de poder y normalización.
S. XIX	Medio de calificación y acreditación para otorgar títulos.
Primera etapa, siglo XX (1900 -1930)	Inicio y expansión de los test psicológicos, para medir capacidades y diferencias individuales.
Segunda etapa, siglo XX (1930-1945)	Vinculada a objetivos para comprobar rendimiento de estudiantes.
Tercera etapa Siglo XX (a partir 1960)	Expansión, diversificación y profesionalización, evaluación vista como servicio al cliente.

Las características y usos del examen en diferentes momentos históricos, presentan ciertas ausencias en materia de evaluación del aprendizaje, importantes a considerar :

- ⇒ La evolución y uso de los exámenes se da con fines de selección, control, eliminación, calificación, acreditación y medición, no de formación.
- ⇒ Existe una desvinculación total de los exámenes con el proceso enseñanza-aprendizaje.

evaluación democrática , como servicio de información a la comunidad sobre las características de un programa educativo y que junto con Apple (1974) y Weiss (1975), la ubican como una actividad de naturaleza política para los responsables de decisiones. Finalmente Guba y Lincon (1987), desde un enfoque naturalista, en donde la verdad depende de la naturaleza humana,(epistemología subjetivista), plantean un modelo alternativo de evaluación de corte holístico e iluminativo, como reacción contra formas tradicionales de evaluación que se muestran insuficientes para la comprensión de la complejidad de las reformas curriculares (Cfr. House, 1997).



- ⇒ Tratamiento meramente técnico, en vez de pedagógico en la elaboración de exámenes.
- ⇒ Primacía por el control, más que por la promoción y mejora.
- ⇒ Uso indistinto de la noción de examen, test, medición y/o evaluación, sin una clara delimitación conceptual.
- ⇒ Excesivo interés por la cuantificación y ausencia de fundamentos teóricos en materia de elaboración y aplicación de exámenes.

De ahí que la pedagogía y epistemología continúan siendo los grandes vacíos en materia de medición y evaluación;<sup>16</sup> lo que obedece a importantes omisiones que a continuación se enuncian:

1. La tarea sustantiva de la génesis de la escuela, como producto de un proyecto de modernidad, en donde la ilustración plantea como ideal, permitir al ser humano hacer uso de su propia razón<sup>17</sup> y demostrar sus capacidades y conocimientos, no es considerado en la época de la examinación o de la testificación (Véase, p. IV de este documento).
2. La evaluación se ha visto reducida a una visión científicista en donde el uso de ésta para verificar y mejorar los procesos de enseñanza-aprendizaje, queda totalmente ausente en la génesis y desarrollo de la disciplina; pese a que Comenio (1922/2000) en su "Didáctica Magna", presenta al examen, como la última parte del método de enseñanza, con el propósito de que si el alumno no ha aprendido, más que calificarlo, el maestro tendrá que revisar su método.
3. La reducción de la evaluación a técnicas de medición, hace que se cuantifique, en vez de que se evalúe; impidiendo el conocimiento y comprensión del proceso.
4. La primacía en la constatación del logro de los objetivos, de la acreditación de cursos, de la medición del desempeño de un sujeto o institución, a través de

---

<sup>16</sup> Díaz Barriga (1997), desde los noventa, ha venido señalado la necesidad de encontrar fundamentos epistemológicos y un replanteamiento metodológico en materia de evaluación, que supere las posturas empírico-analíticas, con que se le ha venido manejando y permita una reconstrucción teórica, que reconozca la dimensión social del campo evaluativo.

<sup>17</sup> Para una mayor información de los ideales de la ilustración se puede consultar (Kant, 1979).



exámenes, deja de lado la valoración de procesos y capacidades cognitivas de los alumnos con fines de formación.

Este es el estado que guarda la evaluación en su marco histórico, lo que da pie a la siguiente problemática.

### Problema y justificación:

La evaluación es un campo que desde mediados del siglo XX, como parte de una nueva política educativa a nivel mundial, ha tenido mayor auge y diversificación por su rápida inserción y adopción en los sistemas educativos y la multiplicidad de trabajos e investigaciones con que se cuenta.

Organismos y Agencias Internacionales de evaluación como la *Internacional Association for the Evaluation of Educational Achievement* (IEA) (Asociación Internacional para la Evaluación del Logro Educativo) que desde hace cinco décadas ha venido aplicando estudios comparados de evaluación a nivel internacional, hacen que la evaluación educativa cobre especial importancia en la última década del siglo XX.<sup>18</sup>

La difusión de los resultados obtenidos en los sistemas internacionales de evaluación, ha contribuido a la discusión sobre el tema, e impactado a la sociedad, por la evidencia en que han puesto a los sistemas educativos y en especial a México, a raíz del lugar obtenido por nuestro país en estos estudios<sup>19</sup> y difundidos con gran alarma en la prensa nacional en el

---

<sup>18</sup> Entre los estudios realizados por esta agencia se encuentran: *The Pilot Twelve-Country Study* (1959-1962), *First Internacional Study Mathematics (FIMSS, 1961-1965)*, *The First International Science Study\* (FISS)* (1966-1975), *The Study of Reading Comprehension* (1967-1973) *The Study of Literature Education*, (1966-1973) *The Study of English as a Foreign Language* (1968-1975) *The Study of French as a Foreign Language* (1968-1975) *The Study of Civic Education* (1967-1976) *The Second International Mathematics Study (SIMS)* (1976-1989) *The Classroom Environment Study* (1978-1989) *The Second International Science Study (SISS)* (1979-1991) *The Written Composition Study* (1980-1988) *The Reading Literacy Study* (1985-1994) *The Computers in Education Study*, (1985-1993) *The Languages Education Study (LES)* (1993-1996) *Progress in internacional Reading Literacy (PIRLS, 1991,2001)* *Third Internacional Mathematics and Science Study (TIMSS, 1995,1997, 1999,2003)* Para mayor información de estos estudios se puede consultar la página electrónica de la IEA <http://www.iea.nl/iea/hq/>

<sup>19</sup> México comienza su participación en 1995 con el *Tercer Estudio Internacional en Matemáticas y Ciencias (TIMSS)*, en donde obtuvo los últimos cuatro lugares (375 aciertos frente a 483 que fue el promedio) (Meigar, I. 2001, octubre, 15). En 1997, participó en el *Laboratorio Latinoamericano de Evaluación de la Calidad (LLECE)*, su promedio: 252 puntos en lenguaje y 256 en matemáticas, (por debajo de Argentina, Bolivia, Brasil, Chile, Cuba y Honduras) (Casassus, et al. 1998). En el 2000, (OCDE, 2002) el Programa Internacional para la Evaluación de los Estudiantes (PISA), nos ubica en el lugar 30 de 31 países participantes. En el 2003 "Los alumnos mexicanos obtuvieron un promedio de 400



2001;<sup>20</sup> lo que obliga como sociedad a lamentar los resultados obtenidos, pero como estudiosos interesados, especialistas en temas de educación, nos lleva a generar una reflexión profunda y sistemática que favorezca la comprensión de lo acontecido.

En esta reflexión, uno de los aspectos más importantes a considerar, reside en la situación que el docente de educación básica, como actor principal del sistema educativo, enfrenta ante esta circunstancia que resulta ambigua y contradictoria; por un lado se le ubica como uno de los principales culpables de los resultados obtenidos en los sistemas internacionales de evaluación; pero al mismo tiempo, se le deja al margen de los procesos y reflexiones que pueden surgir en torno a los sistemas evaluativos y sus resultados; Díaz Barriga, (2004:106) manifiesta al respecto: "el trabajo de los docentes en el aula es el gran ausente de la reflexión pedagógica en materia de evaluación."

Preocupados por esta situación, consideramos que lo importante en materia de evaluación en la educación básica; no es saber en qué lugar quedamos<sup>21</sup> y quedarnos con la "calificación" en el terreno de la cuantificación, desde una perspectiva determinista, que en sí no dice nada, ni a la sociedad, ni mucho menos a los docentes; saber que estamos en el lugar 30 o ahora 37, no dice en qué estamos mal y qué es lo que se debe trabajar para mejorar; es necesario pasar al terreno de la comprensión y analizar con una mirada pedagógica, el por qué de los resultados obtenidos.

---

puntos (en lectura), 22 menos que en la primera edición de PISA; en ciencias lograron 405, una disminución de 17 puntos y en matemáticas esta media bajó de 387 a 385 puntos, y en solución de problemas (una prueba nueva), alcanzaron 384 puntos." En todos los casos la media de la Organización para la Cooperación y el Desarrollo Económico (OCDE) es de 500 puntos. De nueva cuenta México ocupó el último lugar del grupo de la OCDE, pero en la lista ampliada de 40 países (que incluye a países no miembros (el país quedó en el lugar 37, apenas arriba de Indonesia, Túnez y Brasil; salvo en lectura donde se ubicó en el 38" (Herrera, 2004:42, diciembre 7).

<sup>20</sup> En el periodo de 1996-1997, la SEP, pidió formalmente a la Asociación Internacional de Evaluación en la Educación. (IEA) no incluir los resultados en la difusión masiva de las calificaciones de las naciones participantes. En 1998 el gobierno mexicano solicitó al LLECE, que quitara del informe final lo referente a los estudiantes mexicanos. Fue hasta octubre del 2001, que el periódico Reforma publicó los resultados obtenidos por México en el TIMSS de 1995, con los siguientes encabezados en primera plana: "Reprueba México calidad educativa", "Reprueba UNESCO aprendizaje" (Melgar, 2001, octubre, 15 y octubre 18:11) En el 2004, los resultados obtenidos por México en el PISA 2003, se publicaron por algunos medios de comunicación de la siguiente manera: "México es último en educación en OCDE" (Martínez, 2004:20, diciembre 7). "Reprueba la OCDE a México en educación" (Guarque, 2004, diciembre 7).

<sup>21</sup> Dice Manuel Gil, (2004:113): "Hay que salirnos del fútbol, cada vez que se presentan los resultados, los metemos como si fuese una tabla de posiciones en el fútbol".



La evaluación en México, ha entrado de manera abrupta, al formar parte importante y fundamental de una política educativa; la aparición de agencias de evaluación que desde los años cincuenta están aplicando pruebas internacionales para medir el rendimiento escolar, la incursión de México en los últimos estudios realizados por la IEA, por el *Laboratorio Latinoamericano de Evaluación Educativa (LLECE)* y por la *Organización para el Desarrollo Económico (OCDE)*, la discusión e impacto sobre los resultados obtenidos; así como el gran número de pruebas a gran escala aplicadas desde hace más de cincuenta años a nivel mundial y las diversas experiencias con las que contamos en México desde hace más de treinta años, las cuales consideramos, han estado al margen de las nuevas tendencias y teorías en materia de evaluación del aprendizaje, nos lleva a preguntarnos:

1. *¿Bajo qué orientaciones se organiza el contenido de evaluación en el plano internacional?*
2. *¿Qué concepto de contenido de evaluación está implícito en las pruebas internacionales?*
3. *¿Cuáles son los fundamentos teórico-metodológicos que sustentan a los exámenes internacionales?*
4. *¿Cuáles son los conceptos que subyacen en los exámenes internacionales?*
5. *¿Qué se puede apreciar de la comparación entre las experiencias internacionales y las nacionales, en materia de tendencias y contenidos de evaluación?*
6. *¿Cuál es el impacto que el contenido de evaluación de las pruebas internacionales tiene en materia de currículo y evaluación?*
7. *¿Qué necesidades surgen en materia de docencia?*

Para responder a estas interrogantes, se propone realizar un análisis del contenido y de la estructura de las pruebas internacionales que permita reconocer las aportaciones teórico-metodológicas, que en materia de evaluación y elaboración de exámenes se han generado; su comparación con los exámenes realizados en México, ubicará el estado que guardan las pruebas nacionales y contribuirá a una mejor comprensión de los resultados obtenidos, para favorecer el desarrollo de esta disciplina y el abordaje pedagógico de la misma al



aproximarnos al concepto de examen trabajado a nivel internacional y los nuevos enfoques, contenidos y objetos de evaluación que las Agencias Internacionales presentan.

En esta perspectiva, consideramos que la evaluación del aprendizaje, deberá adoptar nuevos contenidos, métodos y teorías, acordes con las tendencias actuales, tanto en materia de evaluación, como de aprendizaje, pues a pesar de la rápida incursión y proliferación de la evaluación en los sistemas educativos de México y Latinoamérica, la evaluación a nivel nacional, no se ha dado tiempo de reflexionar sobre los desarrollos obtenidos en esta área.

Ante esta problemática, se presenta como objeto de estudio el siguiente:

### **Objeto de estudio**

Analizar las diferencias que subyacen en el concepto de contenido, aprendizaje y evaluación en las pruebas internacionales, respecto del manejo que aparece en las pruebas nacionales y determinar el impacto que las Agencias y Organismos Internacionales de evaluación están teniendo en materia de evaluación y elaboración de exámenes.

A partir del objeto de estudio en esta investigación se plantean los siguientes objetivos:

### **Objetivos:**

La investigación tiene como objetivos:

- ✓ Dar cuenta de la concepción de contenido y evaluación que presentan las Pruebas Internacionales.
- ✓ Establecer una comparación entre el contenido de evaluación y la forma en cómo éste se organiza en el marco internacional y la situación que guardan los exámenes nacionales al respecto.
- ✓ Ofrecer elementos de análisis para comprender desde una mirada pedagógica, el porqué de los resultados obtenidos por México en los estudios internacionales.
- ✓ Presentar algunas aportaciones pedagógicas en materia de evaluación, currículum y docencia.



Los objetivos planteados, requirieron de estrategias operativas especiales, mismas que a continuación se abordan.

### Abordaje metodológico:

El estudio que se presenta, es de carácter documental y comparativo.<sup>22</sup> Atendiendo a que la comparación según Schriewer(1993), se refiere a "actos mentales dirigidos a obtener un determinado conocimiento a través del establecimiento de relaciones;" (cit., por Martínez, 2003: 18-19) se retoman elementos de la educación comparada, considerando a ésta como "un proceso ordenado que relaciona objetos y fenómenos diversos para descubrir las semejanzas, las diferencias y las relaciones entre los distintos sujetos o elementos de comparación;"(Savienses, 1973: 245-275, cit., por Martínez, 2003:19) para desarrollar una metodología propia que inicia con una descripción y análisis de los fundamentos y características de los instrumentos de medición utilizados tanto en el plano internacional como nacional, en poblaciones de edades y grados escolares similares, para a partir de aquí, dar cuenta de las semejanzas y diferencias existentes en cuanto a contenidos, elementos teóricos-metodológicos y modelos psicométricos en el campo de la evaluación y medición del aprendizaje.

En un primer momento, se recurre a las orientaciones, definiciones conceptuales, tipo de ítem y modelo psicométrico con que se construye cada una de las Pruebas Internacionales más recientes: *Tercer Estudio Internacional de Matemáticas y Ciencias. Tendencias 2003* (TIMSS), *Primer Estudio Internacional Comparativo sobre Lenguaje, Matemáticas y Factores Asociados para alumnos del tercer y cuarto grado de educación básica del Laboratorio Latinoamericano de Evaluación de la Calidad Educativa* (LLECE) y *Programa Internacional*

---

<sup>22</sup> La comparación que se realiza en esta investigación, retoma las fases y elementos que definen a una educación comparada, cuya característica principal es la descripción y análisis de prácticas educativas en una o más naciones. (Wilson, 2003:17-18)



de *Evaluación de Estudiantes (PISA)*.<sup>23</sup> Para a partir de aquí establecer una comparación con la situación que guardan, cuatro de las pruebas nacionales, aplicadas a poblaciones y edades semejantes: *Instrumento de diagnóstico para alumnos de nuevo ingreso a secundaria (IDANIS)*, Pruebas del Sistema Nacional de Evaluación Educativa (SNEE) Pruebas con base en Estándares (PE), ahora Pruebas Nacionales (PN) y Examen Nacional de Ingreso a la Educación Media Superior (EXANI-1).

Del objeto de estudio y su intencionalidad, se establecen las **categorías de análisis** y comparación; mismas que se presentan y definen a continuación.

- 1) *Concepto de contenido*. Considerado como eje organizador de la estructura de la prueba y no como un mero dato, concepto o definición operable del atributo a medir.<sup>24</sup> Se refiere la explicación que se hace del área de conocimiento y a lo que ésta implica.<sup>25</sup> Los documentos de evaluación internacional, consideran que la definición de lo que se va a evaluar debe evolucionar y cambiar conforme las transformaciones sociales, económicas y culturales, ocurridas en los últimos tiempos; con base en un consenso entre expertos de cada disciplina y junto con los grupos de asesores, especifican claramente las dimensiones del contenido; que no se limitan a reproducir los temas incluidos en los planes y programas de estudio como modelo organizador de exámenes.

---

<sup>23</sup> Estos documentos son los que se refieren en la bibliografía. Cabe destacar que la mayor parte de los documentos analizados, se encuentran disponibles en páginas Web, y aunque no son documentos dictaminados que garanticen una aportación rigurosa en materia de conocimiento, sí nos hablan de nuevas formas de acceder a la información y de la difusión con que cuentan los instrumentos y marcos teóricos de los estudios internacionales de evaluación.

<sup>24</sup> "Cuando uno revisa la prueba de la OCDE, se encuentra por ejemplo que cuando hablan de habilidades y destrezas para la lectura, dicen: la capacidad del lector es compleja, lo que significa que la valoración de estas habilidades tiene diversos elementos" (Díaz Barriga, 2004:107).

<sup>25</sup> La definición que agencias de evaluación internacional como la Organización para la Cooperación y el Desarrollo Económico (OCDE), elaboran de cada una de las áreas de evaluación, se especifica claramente en el documento *The PISA 2003. Assessment Framework: Mathematics, Reading, Science and Problem Solving Knowledge and Skills*. Una traducción en castellano de este documento la realiza el Instituto Nacional de Evaluación y Calidad del Sistema Educativo (INECSE) de España (OCDE, 2004).



- 2) *Tipo de ítems.* Considerando que la calidad y adecuación de todo test, dependen directamente de los ítems que lo conforman, es importante que éstos en una prueba, incluyan aspectos relativos al dominio y se analice el formato y contenido que lo constituyen. Existen diversos formatos para la elaboración de los ítems: de opción múltiple, de respuesta abierta, de respuesta construida, de ensayo; así como normas muy precisas para elaborarlos.<sup>26</sup> El contenido de los ítems, está delimitado tanto por el tipo de contenido que se evalúa<sup>27</sup> como por los procesos implícitos que demanda la tarea a realizar; los cuales generalmente pueden estar inscritos desde una postura tradicional, en la taxonomía de Bloom (1971) o en el análisis cognitivo de tareas (Castañeda et al., 1998 y 2002) con un fundamento más constructivista.
- 3) *Modelo psicométrico.* Se refiere a los modelos psicométricos de la teoría de la medición o del test, en ausencia de una teoría de la evaluación.<sup>28</sup> A pesar de la proliferación de exámenes masivos, tanto en el plano internacional como nacional, éstos continúan haciendo uso de los elementos y conceptos básicos de la teoría de la medición,<sup>29</sup> aunque con diferentes modelos psicométricos para su análisis. La teoría de la medida o teoría de test, se constituye como la metodología empleada para la validación estadística de los ítems de una prueba, en donde se considera los elementos y cualidades del proceso de

---

<sup>26</sup> Entre los documentos que podemos citar como referentes para elaboración de ítems se encuentran: Thorndike, y Hagen (1978); Haladyna, Haladyna, Merino, C., (s.f); Hernández, (1991) y Woolfolk, (1996).

<sup>27</sup> Ahumada (2003), aborda la evaluación desde una concepción de aprendizaje significativo y presenta algunas sugerencias y procedimientos para evaluar los distintos tipos de contenidos.

<sup>28</sup> Díaz Barriga, (1997:155-157) expresa: "El discurso actual de la evaluación se fundamenta en la teoría de la medición; en la medición se encuentra la propuesta actual para realizar evaluación (...) la teoría de la medición se toma como referente único y sostén del discurso relacionado con la evaluación".

<sup>29</sup> Dice Thorndike y Hagen (1978: 39) "El término evaluación (...) está estrechamente ligado con el concepto de medición." Inclusive en los documentos del PISA, se especifica claramente: "El propósito del proyecto PISA, es analizar las características de los estudiantes de cada país de modo colectivo, no examinar a los alumnos individualmente. En consecuencia se utiliza el término "evaluación", para describir el proyecto PISA, aunque las características de las pruebas a las que contestan los alumnos sean similares a las de un examen normal" (OCDE, 2000:32).



medición<sup>30</sup> y los modelos psicométricos que permitan la estimación de los errores asociados a éste, mediante el análisis correspondiente; que permite obtener las características de los ítems y determinar si éstos cumplen con la regla de medir adecuadamente lo que se pretende medir. Para esto existen dos modelos: Teoría Clásica del Test (TCT)<sup>31</sup> o Teoría de Respuesta al Ítem (TRI), enfoque psicométrico alternativo, con un tratamiento matemático diferente;<sup>32</sup> que aunque sus propósitos son similares a los de la Teoría Clásica, su nombre responde a que se centra más en lo que mide cada ítem y en su probabilidad de respuesta que en las propiedades globales del test. De acuerdo con lo que el mismo Lord, (1952, 1953, cit. por Muñiz, 1997:48) plantea en sus trabajos pioneros en donde sistematiza los principales conceptos de esta Teoría, este modelo se puede considerar más bien como un enfoque que complementa lo trabajado por la Teoría Clásica.

La intencionalidad de estas categorías, es realizar un análisis comparativo para identificar semejanzas y diferencias en el concepto de contenido, composición de los ítems y tratamientos estadísticos que sustentan a los exámenes masivos y comprender el estado que guarda la evaluación del aprendizaje a nivel nacional con referencia al ámbito internacional, a manera de acercamiento cualitativo a la fase en la que se encuentra el desarrollo de esta disciplina en un momento dado.

---

<sup>30</sup>Validez y confiabilidad son dos cualidades básicas de todo sistema de medición. "Por validez se entiende el grado en que un test mide lo que queremos medir realmente. La confiabilidad es la exactitud y precisión de un procedimiento de medición" (Thorndike y Hagen, 1978: 177).

<sup>31</sup> La Teoría Clásica de los Test, es un modelo estadístico de tipo lineal que tiene como objetivo, fundamentar adecuadamente las puntuaciones de los test y permitir la estimación de los errores de medida asociados a todo proceso de medición... (Muñiz, 1998: 19) y asume que la puntuación empírica de un sujeto en un test, son afectadas por errores aleatorios de origen, cuyas causas podrían estar en el propio sujeto, en el ambiente externo, en el instrumento de medida, o en el proceso mismo de medición (Muñiz, 1997).

<sup>32</sup> La Teoría de Respuesta al Ítem, es un modelo teórico de la medida, de incursión reciente en la psicometría, que representa un nuevo enfoque dentro de la Teoría de los Test. Surge con Lord en 1952 con la idea de subsanar los errores de la teoría clásica, comenzó en los años sesenta, se desarrolla en los 70 y se aplica masivamente en los 80. EL término proviene de un enfoque basado en propiedades de los ítem, más que en las del test global para reflejar el funcionamiento real de estos (Muñiz, 1997:15).



Lo anterior ubica a esta investigación como un estudio documental, de carácter conceptual, y de tipo comparativo, con enfoque interpretativo. Con base en los documentos existentes sobre pruebas internacionales y nacionales, se recuperan premisas del método comparado aplicado a la educación, para dar cuenta del contenido y estructura de estos exámenes masivos. Considerando que la descripción es una de las metodologías base de la educación comparada, (García: 1986:115), se procede a dar cuenta de las orientaciones, conceptos, modelos y teorías que sustentan a los instrumentos de evaluación de corte internacional en un primer momento y posteriormente en el ámbito nacional; para a partir de aquí realizar un análisis<sup>33</sup> del contenido, constitución y enfoque psicométrico, de las diferentes experiencias de evaluación, que permita establecer las relaciones y comparaciones correspondientes en materia de evaluación del aprendizaje en ambas esferas: la internacional y la nacional. Con el fin de encontrar significados (Erickson, 1989) y penetrar en la complejidad de un acontecimiento (Pérez, 1998) que en este caso refiere a los estudios internacionales y nacionales de evaluación.

### **Instrumentos de investigación:**

Dado el carácter documental de esta investigación, se recurre a las definiciones conceptuales con las que se construye cada examen y a los ítems muestra de los más recientes exámenes masivos internacionales y nacionales:

En el plano internacional se trabajó con los siguientes documentos: <sup>34</sup>

---

<sup>33</sup> El análisis, como "operación mental mediante la cual un todo es dividido en sus partes para mejorar su comprensión" (García, 1986:124) constituye uno de los primeros elementos del proceso metodológico de corte comparativo.

<sup>34</sup> Dada la abundancia de material existente al respecto, fue necesaria hacer una selección de los mismos y trabajar sólo con los documentos base. Es decir aquellos que ofrecían marcos teóricos y muestra de ítems.



- Marcos teóricos e ítems del *Tercer Estudio Internacional de Matemáticas y Ciencias (TIMSS) Tendencias 2003*, elaborado por la IEA. (véase Mullis, I.V.S., Martín, M. O., González, E. J. Gregory, K.D. Garden, R.A., O'Connor, K.M., Chrostowski, S.J., Smith, T.A. 2002).
- Informes y documentos del *Primer Estudio Internacional Comparativo sobre Lenguaje, Matemáticas y Factores Asociados* para alumnos del tercer y cuarto grado de educación básica, elaborado por el LLECE de la UNESCO. (véase LLECE 1997, LLECE, 2002).
- Marcos teóricos y reactivos muestra del Programa Internacional para la Evaluación de Estudiantes PISA, elaborado por la OCDE. (véase OCDE 2000, 2001a 2002a y 2002b).

En el ámbito nacional se recurrió a los documentos existentes:<sup>35</sup>

- Guía de estudio del Instrumento de Diagnóstico para Alumnos de nuevo Ingreso a Secundaria IDANIS.
- Documentos e Informes de la Dirección General de Evaluación Educativa (DGEE), del Sistema Nacional de Evaluación Educativa (SNEE) de la Secretaría de Educación Pública (SEP) (véase SEP, 1999; 2000a; 2000b; 2002).
- Pruebas en red del Sistema Nacional de Evaluación Educativa. (véase <http://snee.sep.gob.mx/>).
- Informes de las Pruebas Nacionales realizadas por el Instituto Nacional de Evaluación Educativa.

---

<sup>35</sup> Cabe destacar lo difícil que fue el localizar estos documentos y sobre todo acceder a reactivos muestra, dada la escasez de material que fundamente y explique las acciones evaluativas a nivel nacional y el carácter confidencial y secreto con que en México se ha venido manejando la evaluación hasta el momento.



- Guías de estudio para el Examen Nacional de Ingreso a la Educación Media Superior (EXANI-1) elaborada por el Centro Nacional para la Evaluación de la Educación Superior (véase CENEVAL, 2004).

### **Estructura y contenido:**

El contenido del trabajo presentado se organizó en cuatro capítulos:

En el capítulo uno, denominado: *Pruebas internacionales: Principales orientaciones*, se describe el origen y propósitos de éstas; se analiza y compara el contenido de evaluación y su estructura en cuanto al diseño de ítems, así como las normas y procedimientos para su elaboración y sistema de calificación; todo, conforme a las dos poblaciones referidas en los estudios internacionales: alumnos de 8 y 9 años que están cursando la escuela primaria y alumnos de 14 y 15 años que están por concluir la educación básica. También se aborda lo referente a la evaluación de factores asociados y se establecen las primeras conclusiones.

En el capítulo dos, cuyo título es: *Concepto de examen de las pruebas internacionales*, se refiere el concepto de examen de estas pruebas, mismo que se extrae de las categorías de análisis establecidas a partir de la descripción realizada en el primer capítulo: Concepto de contenido, tipo de ítem y modelo psicométrico, mismas que se abordan también con base en la edad de las dos poblaciones de estudiantes.

En el capítulo tres, al que se le denomina: *La experiencia mexicana en pruebas nacionales*, se aborda la descripción y análisis de los exámenes masivos a nivel nacional, en cuanto a orígenes, propósitos y diseño de ítems de cuatro de las experiencias mexicanas en materia de pruebas a gran escala. Las experiencias seleccionadas fueron: el *Instrumento de Diagnóstico para Alumnos de Nuevo Ingreso a Secundaria*. (IDANIS), las pruebas del *Sistema Nacional de Evaluación Educativa* (SNEE), las pruebas con base en estándares, ahora *Pruebas Nacionales*, y el *Examen Nacional de Ingreso a la Educación Media Superior*, elaborado y aplicado por el Centro Nacional de Evaluación para la Educación Superior



(CENEVAL); lo que permitió obtener una caracterización de las pruebas realizadas en México.

Finalmente en el capítulo cuatro cuyo nombre es: *Pruebas Internacionales y Pruebas Nacionales, una comparación en cuanto a su contenido y estructura*, se establece una comparación de estas experiencias en los distintos ámbitos que conforman el estudio: el internacional y nacional, de acuerdo a las categorías creadas, se establecen las comparaciones finales y se formulan algunas tareas pendientes.

El desarrollo de este trabajo, permitió dar cuenta de algunos hallazgos:

### **Principales hallazgos:**

Los estudios de evaluación internacional retoman funciones, objetivos y planteamientos, de la evaluación y medición en sus diferentes paradigmas y enfoques; así como algunos planteamientos de campos afines, como el currículo, en donde las percepciones curriculares de John Dewey referentes a contenidos inmersos en la experiencia o la práctica y de Franklin Bobbit, quien desde 1918, planteaba la necesidad de articular el mundo laboral a contenidos curriculares, son concepciones inmersas en las Pruebas Internacionales.

Las Agencias Internacionales de evaluación, ofrecen un concepto de contenido de corte más funcional y procedimental. Presentan pruebas que descartan el manejo enciclopédico del contenido curricular, para dar paso a ítems contextualizados que incluyen tareas cotidianas, que demandan ubicadas en situaciones reales, que demandan la activación de procesos mentales de orden superior, como el razonamiento, la reflexión, el pensamiento lógico, la interpretación y la transferencia para la aplicación del conocimiento en la resolución de problemas y situaciones de la vida cotidiana. En la estructura y conformación de las pruebas, se incorporan enfoques psicométricos alternativos para el análisis de sus ítems, en donde la validación de los mismos se realiza conforme a las propiedades y probabilidad de respuesta de cada ítems y no en relación a la totalidad de la prueba y al porcentaje de alumnos que contestaron adecuadamente el reactivo. Se trata de pruebas criterioles que



retoman paradigmas actuales de la evaluación para su sistema de calificación, como es la evaluación por niveles de desempeño.

Los estudios para evaluar el rendimiento escolar, realizados por Organismos Internacionales, presentan a la evaluación, desde otra mirada conceptual; y además consideran importante evaluar, no sólo los resultados que están teniendo los sistemas educativos, sino los factores asociados al rendimiento y desempeño escolar.

En México, los programas de evaluación generados por institutos dedicados a la realización y aplicación de exámenes masivos a nivel nacional, desde hace tres décadas: Dirección General de Evaluación Educativa. (DGEE), Centro Nacional de Evaluación para la Educación Superior (CENEVAL) y el Instituto Nacional para la Evaluación de la Educación (INEE),<sup>36</sup> continúan elaborando pruebas centradas en la reproducción de contenidos y verificación de objetivos curriculares, que mantienen como fin primordial, determinar posibilidades de éxito en los estudios, calificar, acreditar y seleccionar. Su creación y aplicación, ha estado vinculada a constatar y conocer el logro obtenido por los estudiantes de acuerdo con lo establecido en planes y programas de estudio, así como a comprobar la efectividad de los programas compensatorios y de estímulos. Esto refleja la función enciclopédica, sumativa y fiscalizadora que prevalece en los exámenes de corte nacional y la enorme necesidad de trabajar con los docentes, una mirada pedagógica de la evaluación con enfoque más conceptual, formativo y cognitivo que ayude a mejorar los procesos de enseñanza-aprendizaje y por ende, los resultados obtenidos en estudios de evaluación comparada a nivel internacional.

---

<sup>36</sup> Entre las experiencias más importantes de evaluación y aplicación de exámenes masivos en educación básica a nivel nacional en México se encuentran: *Instrumentos de Diagnóstico para alumnos de nuevo ingreso a secundaria*. IDANIS, cuyos antecedentes datan de 1970, inicia en 1986 y se edita por primera vez en 1989, *Evaluación de alumnos de escuelas apoyadas por programas compensatorios, entre los que se encuentra el Programa para abatir el rezago educativo PARE*, (1992) *Evaluación del aprovechamiento escolar del programa Carrera Magisterial*, (1994) *Evaluación de la Educación Primaria*. EVEP (1996-2000)0. *Examen para alumnos de nuevo ingreso a secundaria EXANI 1 elaborado y aplicado por CENEVAL a partir de 1996 y las más recientes pruebas nacionales, antes llamadas, pruebas de estándares, aplicadas anualmente desde el ciclo 1997-1998 por el INNE* (cfr. SEP, 2002 y s.f.)

## CAPÍTULO UNO

### PRUEBAS INTERNACIONALES: PRINCIPALES ORIENTACIONES

---

El objetivo del capítulo es presentar una aproximación a las orientaciones teórico-metodológicas que se están impulsando a nivel mundial en materia de medición y evaluación del aprendizaje escolar.

Mediante una descripción y análisis de las pruebas internacionales<sup>1</sup> elaboradas por las agencias encargadas de realizar y publicar estudios comparados sobre rendimiento escolar; se pretende identificar y caracterizar elementos de las experiencias internacionales en materia evaluación y medición que contribuyan a mejorar los procesos y sistemas de evaluación en la educación básica y por ende, de aprendizaje y enseñanza.

En un primer momento, hemos situado el origen de las pruebas internacionales, auspiciadas por Organismos Internacionales como la *Organización para la Cooperación y el Desarrollo Económico* (OCDE) y la *Organización para la Educación la Ciencia y la Cultura* (UNESCO);<sup>2</sup> para posteriormente proceder a la descripción, análisis y comparación del contenido y estructura de los exámenes realizados por estas Agencias principalmente; aunque también se refieren los estudios de carácter contextual que contemplan estos sistemas de evaluación. En un segundo momento, a partir de las categorías establecidas, se exploran los conceptos

---

<sup>1</sup> Se conoce como Pruebas Internacionales a los estudios de evaluación comparada realizados por Agencias y Organismos Internacionales. Se trata de "instrumentos de medición estandarizados que se aplican en varios países simultáneamente, en grados o grupos seleccionados y que recogen información sobre logros de aprendizaje conceptual y procedimental en diferentes disciplinas y sobre los factores contextuales que se presume influyen en el logro académico de los estudiantes" (Ferrer y Arregui, 2003: 8) El empleo de instrumentos estandarizados para comparar los sistemas educativos, habla de unificación de finalidades, modalidades, métodos y normas para valorar lo educativo.

<sup>2</sup> La relación de los Organismos Internacionales con la educación, es un tema reciente; a partir de 1990 cobra cada vez más importancia, por el papel que juegan en materia de política educativa en las acciones y proyectos que últimamente se han realizado en materia de financiamiento, investigación, e informes sobre temas específicos; de ahí la necesidad de considerarlos como "actores" importantes de los actuales procesos educativos. De acuerdo con el último Estado de Conocimiento organizado por el *Consejo Mexicano de Investigación Educativa* (COMIE), sobre Actores y Organizaciones, en donde por primera vez es abordado el tema, la UNESCO, el Banco Mundial (BM) y la OCDE son los Organismos que a partir de 1990, notoriamente han tenido una relación más directa con el campo educativo (Ducoing, coord., 2003).



estructurantes de examen que están proliferando a nivel internacional. Con lo trabajado, se retoma la experiencia nacional, para llegar a una comparación entre los factores que caracterizan los exámenes masivos<sup>3</sup> en el plano internacional y la situación que México guarda con respecto a este tipo de exámenes.

Se seleccionan como ejes básicos para la descripción y análisis de las Pruebas Internacionales, cuatro elementos:

- 1) *Generación y propósito*: Pretende dar cuenta de la gestación e intención que guardan los estudios de evaluación a nivel internacional. Incluye la explicación de su origen, concepto organizador, preguntas a las que responde, finalidad, población objetivo (edades que atiende) y áreas de evaluación.
- 2) *Contenido*: Como contenido de evaluación se ha considerando como base, la conceptualización del área, los campos temáticos, y los procesos cognitivos contemplados como objeto de evaluación. La descripción y análisis de este contenido, se hace, conforme con la edad y grado de escolaridad de los alumnos a quienes se aplica la prueba y de acuerdo al área de evaluación que comprenden cada uno de los estudios. Esto permitirá realizar las comparaciones correspondientes.
- 3) *Estructura*: Se refiere a las normas y procedimientos de tipo técnico, que contemplan los estudios internacionales en materia de elaboración de pruebas, diseño de ítems, modelo estadístico de medición y sistema de calificación.
- 4) *Evaluación de factores asociados*: Contempla los cuestionarios aplicados a la par de los exámenes en estos estudios internacionales, con el fin de conocer los

---

<sup>3</sup> Se conoce como exámenes masivos a las pruebas aplicadas en gran escala, que cumplen con las cualidades psicométricas básicas y las características técnicas que permitan comparaciones transversales y la aplicación a muestras numerosas en forma casi simultánea. La aplicación de exámenes masivos en México, comienza a partir de 1970. Una de las experiencias pioneras al respecto en México, fue la realizada por la Facultad de Medicina de la UNAM, que a mediados de los setenta, desarrollo un banco de reactivos de opción múltiple para el examen de titulación de la carrera de Médico Cirujano. Los exámenes de selección para ingreso a la UNAM, se consideran como las más importantes al respecto. En materia de educación básica la SEP, a finales de los setenta comienza la aplicación de pruebas a muestras nacionales de alumnos para valorar el aprendizaje. (Cfr. Martínez Rizo, 2001:80-83)



componentes de corte sociocultural y escolar que se considera, pueden afectar los resultados obtenidos.

## 1.1. Generación y propósito de los estudios internacionales de evaluación:

### Antecedentes.

Los primeros estudios internacionales para medir el rendimiento educativo surgen con las Agencias Internacionales de Evaluación, que tienen por finalidad, establecer comparaciones sobre el rendimiento escolar en varios países y comprobar de esta manera, tanto la efectividad de los sistemas, como el aprendizaje de los alumnos.

Como pionera tenemos a la *Internacional Association for the Evaluation of Educational Achievement* (Asociación Internacional para la Evaluación del Logro Educativo) (IEA)<sup>4</sup> con sede en Ámsterdam, cuyos orígenes datan de los años cincuenta, constituye un consorcio cooperativo independiente de institutos y agencias de investigación en más de 45 países, que desde hace 40 años; se ha ocupado de realizar mediciones y estudios 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>5</sup> Su misión, al realizar estudios comparados sobre las políticas y prácticas educativas de los países participantes, es influir en el aprendizaje dentro y a través de los sistemas de evaluación y no sólo medir el rendimiento escolar, en cuanto a resultados de aprendizaje, sino examinar los efectos de los programas, la organización de las escuelas y las aulas, la relación entre rendimiento y actitudes, los efectos de las prácticas en algunas asignaturas, el tiempo pasado en clase, la

---

<sup>4</sup> Su sitio en internet es: <http://www.iea.nl/>

<sup>5</sup> La 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 varió hacia las técnicas aplicadas de evaluación educacional, que permitiera hacer comparaciones válidas entre naciones. La IEA, se formalizó como asociación en 1954 y ha continuado funcionando hasta el presente" (Frank, 1997:117).



enseñanza de la redacción en el idioma materno, así como diferentes prácticas educativas y los resultados obtenidos por grupos especiales. Entre sus objetivos evaluativos, destaca un nuevo concepto de rendimiento, que incluye la valoración no sólo de los resultados, sino de los factores a los que obedecen dichos resultados.<sup>6</sup> “El rendimiento educativo es valorado tomando en consideración diversas variables de contexto y de proceso, con el objetivo de buscar explicaciones e interpretaciones a las diferencias de rendimiento y no limitarse solamente a la construcción de indicadores o a la difusión de tablas de clasificación de los países participantes”(Tiana, 2000:14).

El primer estudio internacional de la IEA fue:<sup>7</sup> *La Enseñanza, Elemental y Secundaria de las Matemáticas en los Países*, efectuado entre 1959 y 1966. En 1964, se realizó el *Primer Estudio Internacional sobre Matemáticas (FIMS)* Las ciencias se evaluaron por primera vez en los años 1970-1971 con el *First International Science Study, (Primer Estudio Internacional sobre Ciencias)*. A este le siguieron el *Study of Civic Education (Estudio sobre Educación Cívica)* llevado a cabo en 1971.<sup>8</sup> Posteriormente en 1980-1982, se efectuó el *Second International Mathematics Study (SIMS)*, aplicado a 224 países, para continuar con el *Second International Science Study (SISS)* realizado en 1984. Entre 1989 y 1992 se realizó el *Computers in Education* (estudio realizado sobre Tecnologías de la Información y la

---

<sup>6</sup> Cabe señalar que este objetivo, todavía no es algo consolidado en virtud de que a pesar de los esfuerzos de las Agencias Internacionales de evaluación por incluir estudios de contexto para conocer los factores asociados al rendimiento escolar, la forma en como hasta ahora se han presentado los resultados del rendimiento escolar y de los cuestionarios de contexto, no contribuye a una comprensión del porqué de los mismos. Creemos que para que este objetivo se pueda cumplir satisfactoriamente, se hace necesario pasar de los puntajes, porcentajes, índices y gráficas, a una significación pedagógica de lo ocurrido en dichas evaluaciones.

<sup>7</sup> La IEA, fue la primera organización internacional que utilizó las mismas pruebas objetivas en más de un país.

<sup>8</sup> La descripción del estudio realizado en 28 países se encuentra en: IEA s/f *Estudio de Educación Cívica*, consultado en junio 2003, en [http://www2.rz.hu-berlin.de/empir\\_bf/iea\\_s.html](http://www2.rz.hu-berlin.de/empir_bf/iea_s.html). Un resumen de los resultados obtenidos están en el documento de: Torney-Purta, J. y Orozco, L. s/f *Ciudadanía y educación cívica en 28 Países: Resumen de los resultados del estudio comparado de la IEA*. E.U.A. Universidad de Maryland del College Park, Documento electrónico, consultado en octubre 2003, en: [http://www.wam.umd.edu/~iea/exec\\_summ/ExecutiveSummSpanish.htm](http://www.wam.umd.edu/~iea/exec_summ/ExecutiveSummSpanish.htm)



comunicación). En 1990-1991, se aplicó el *Reading Literacy Study (RLS)* (Primer Estudio sobre Lectura).<sup>9</sup>

#### a. Tercer Estudio Internacional sobre Matemáticas y Ciencias: TIMSS

Entre 1994 y 1994 la Asamblea General de la IEA decidió evaluar conjuntamente las matemáticas y las ciencias de manera regular cada cuatro años. Esta decisión supuso el primero de una serie de estudios internacionales a gran escala para medir tendencias en el rendimiento del alumnado. Fue así como en 1995 surge el *Tercer Estudio Internacional sobre Matemática y Ciencias (TIMSS)*, con aplicaciones subsecuentes: TIMSS repetido en 1999 y TIMSS 2003.<sup>10</sup>

Los estudios de la IEA, han sido realizados con la intención de servir como documentos de consulta tanto a responsables en materia de educación, como a investigadores que deseen estudiar las vinculaciones existentes entre rendimiento escolar, crecimiento económico y productividad laboral; y a especialistas en programas de estudio y docentes que deseen mejorar sus conocimientos sobre las implicaciones existentes entre los programas de estudio y los resultados conseguidos por los estudiantes.

Su función es presentar tanto a los tomadores de decisiones, como a la comunidad en general, lo que sucede en el sistema educativo de cada país y el estado que éste guarda dentro de un contexto internacional; de tal manera que las autoridades y responsables de la educación puedan conocer lo que sucede en otros sistemas educativos, en cuanto a materiales utilizados, enfoques pedagógicos-didácticos, forma de trabajo y preparación de los profesores y al relacionarlos con los resultados obtenidos, puedan efectuar los ajustes pertinentes. Queremos subrayar, que éste sería el propósito más importante de los estudios de evaluación comparada; sin embargo, dado la ausencia de un tratamiento de corte más

---

<sup>9</sup> Otros estudios de evaluación internacional: El *International Assessment of Educational Progress*, (IAEP-1). Estudio sobre matemáticas y ciencias realizado por Educational Testing Service en 1988, el cual se repitió con el IAEP-2 en 1991.

<sup>10</sup> La aplicación de los tres estudios del TIMSS, en más de 40 países, ha estado a cargo de la Escuela de Educación de Lynch del Centro Internacional de Estudios de la Universidad de Boston.



pedagógico que estadístico y formativo que informativo, a la forma de presentar los resultados, aún no se alcanza del todo.<sup>11</sup>

Desde su origen, los estudios de la IEA, han cambiando en cuanto a forma y contenido, pero en general siempre, de acuerdo con los programas de estudio; han elaborado y aplicado instrumentos de evaluación (pruebas y cuestionarios) para tres poblaciones de alumnos conformadas según su edad: estudiantes cuyas edades oscilan entre nueve y diez años (población uno), trece y catorce años (población dos) y diecisiete y diecinueve años (población tres).<sup>12</sup>

Los estudios de la IEA y las pruebas que ésta realiza, cuentan con un marco conceptual centrado en el currículo, en tres ejes: currículo deseado, currículo enseñado y currículo asimilado, los cuales se trabajan en tres niveles, que van desde el más global, que aborda lo social como el local, que es el aula y el personal centrado en el alumno.

El primero de estos ejes: *Currículo deseado*, se ubica en un nivel macro y se refiere a lo que se considera que deben enseñar las escuelas y aprender los alumnos. Este se realiza mediante el análisis de programas oficiales, programas de curso y libros de texto.

En el segundo aspecto: *Currículo enseñado*, se sitúa a nivel local y se refiere al contenido y forma de enseñanza de las escuelas y los profesores. Es decir, a lo que se enseña y sucede

---

<sup>11</sup> Un país que ha utilizado los estudios comparados para estudiar su situación, es Suiza, al analizar hasta qué punto sus prácticas, influyen en los resultados alcanzados en cuanto aprendizaje escolar. Suiza revisó las prácticas pedagógicas de Japón, un país que destaca en resultados de aprendizaje y llegó a la conclusión de que sus prácticas pedagógicas podían mejorar. Existen otros estudios realizados por personal del Departamento de Desarrollo Sostenible del Banco Interamericano de Desarrollo (BID) como el de "Imágenes de la enseñanza: El proyecto de grabaciones del video TIMSS," realizado por Laurence Wolf, en donde a raíz de ver que los puntajes obtenidos por los estudiantes estadounidenses, se encontraban muy por debajo de los de los estudiantes japoneses y alemanes, se procedió a revisar los métodos pedagógicos utilizados por dichos países y su impacto en el desempeño estudiantil. Para esto se realizaron las videograbaciones de las clases impartidas en escuelas muestra de estas naciones y se llegó a la conclusión de que los docentes estadounidenses tienden a impartir clases más fragmentadas y deshilvanadas, al mismo tiempo los alumnos pasan menos tiempo realizando actividades prácticas que requieran la aplicación de los conceptos teóricos aprendidos (Wolf, 2002).

<sup>12</sup> En el Sistema Educativo Mexicano, las poblaciones uno y dos comprenden el nivel de Educación Básica. La población uno equivale al tercero y cuarto grado de la escuela primaria, la población dos al segundo grado de la escuela secundaria, dentro del mismo nivel básico, o Educación Primaria, de acuerdo con la clasificación hecha por la UNESCO. Mientras que la población tres se ubica en el nivel medio superior, que en otros países correspondería al último curso de la Educación Secundaria, según la UNESCO.



realmente en las aulas en cuanto a: contenidos, tiempo asignado y estrategias didácticas. Este se realiza mediante cuestionarios y observaciones.

El tercer aspecto: *Currículo asimilado* ubicado en un nivel micro, se refiere a lo que en verdad aprenden los alumnos en cuanto a contenidos y habilidades cognitivas.<sup>13</sup>

Este marco conceptual delimita a la vez, el tipo de variables a utilizar en el estudio. Por ejemplo, el "*Segundo Estudio Internacional de Matemáticas*," comprendía diferentes elementos, como: características y condiciones, entorno institucional, proceso de la escuela y del aula y comportamiento de los alumnos.<sup>14</sup>

TIMSS, es uno de los más importantes estudios de la IEA por su cobertura y magnitud aplicado por vez primera en 1995;<sup>15</sup> ha tenido aplicaciones subsecuentes: TIMSS repetido 1999, enfocado a medir "*Las tendencias en el logro de las matemáticas y las ciencias alrededor del mundo*"<sup>16</sup> y TIMSS Tendencias 2003.<sup>17</sup> Estos estudios aplicados en las tres ocasiones a las poblaciones uno y dos, constituyen la continuación e integración de los realizados previamente y son de gran cobertura por la magnitud de su aplicación.<sup>18</sup>

---

<sup>13</sup> cfr. Plomp. T. (1998) "El potencial de los estudios internacionales comparados para el control de la calidad de la educación" *Perspectivas. Revista trimestral de educación*. Vol. XXVIII, No. 1, Santiago de Chile: OREALC UNESCO pp. 50-67.

<sup>14</sup> Plomp (1992:300), elabora un diagrama de la manera en que la IEA examinó los programas de estudio, tanto a nivel sistema, como escuela, clase y alumno.

<sup>15</sup> Para mayor información sobre este estudio y los documentos elaborados con ítems muestra derivados del mismo, se puede consultar su sitio en internet: <http://timss.bc.edu/timss1995.html> , <http://timss.bc.edu/timss1995i/TIMSSPublications.html> y [http://timss.bc.edu/timss1995i/t95\\_study.html](http://timss.bc.edu/timss1995i/t95_study.html)

<sup>16</sup> TIMSS 1999 es el segundo estudio de la IEA cuyo objetivo es medir tendencias en las matemáticas de los estudiantes y el logro de la ciencia. Este proyecto fue aplicado a estudiantes de 38 países. Es un estudio en donde también se recogió información de estudiantes, de profesores, y de directores de escuela sobre: planes de estudios de las matemáticas y de la ciencia, instrucción, contextos caseros, características y políticas de la escuela. De los 38 países que participaron, 26 de ellos también estuvieron en TIMSS 1995.

<sup>17</sup> El texto completo de este estudio se localiza en línea en la siguiente dirección electrónica: [http://timss.bc.edu/timss2003i/t03pdf/t03\\_af\\_book.pdf](http://timss.bc.edu/timss2003i/t03pdf/t03_af_book.pdf) . Una versión en español de este documento se puede encontrar en la página del Instituto Nacional de Calidad de la Educación (INCE) del Ministerio de Educación y Cultura de España, en la siguiente dirección electrónica <http://www.ince.mec.es>

<sup>18</sup> Los estudios realizados por la IEA, alcanzan una cobertura actual de 60 países y más medio millón de estudiantes examinados en más de 30 lenguajes diferentes; de ahí que TIMSS se haya ubicado como el más ambicioso de los estudios educativos con fines comparativos.



El TIMSS, cuenta con documentos que integran toda una base de datos<sup>19</sup> que presentan las variables educativas recogidas en más de 40 países, de acuerdo a la información proporcionada por profesores y directores de escuela; así como con los resultados arrojados en las pruebas de matemáticas y ciencias.<sup>20</sup> Lo que asegura que la intención principal del estudio no es sólo presentar datos, sino proporcionar información sobre el porqué de los resultados obtenidos.<sup>21</sup>

El objetivo de estos tres últimos estudios, es comprobar el rendimiento<sup>22</sup> de los estudiantes, en dos de las áreas disciplinarias: matemáticas y ciencias, que a nivel internacional, se han considerado como fundamentales para el desarrollo del razonamiento y el pensamiento crítico y al mismo tiempo constituyen la base para la integración de los jóvenes en un mundo cada vez más tecnificado.

La intención, es responder a preguntas tales como:

- ¿Qué matemática y ciencia pretende la sociedad que aprendan los estudiantes?
- ¿Qué es lo que realmente se imparte?

---

<sup>19</sup> La base de datos del TIMSS, está compuesta por documentos que contienen suplementos sobre: "Versiones internacionales de los cuestionarios" y "Variables de los profesores y alumnos que de ellos se derivaron.". Así como "Guías para el usuario" tanto para los primeros años escolares, como para los medios y finales: *User Guide for the TIMSS International Database Primary and Middle School Years (Population 1 and Population 2) Data Collected in 1995*. Este es un texto electrónico que se puede obtener en: [http://timss.bc.edu/timss1995i/database/UG\\_1and2.pdf](http://timss.bc.edu/timss1995i/database/UG_1and2.pdf). Los suplementos y demás publicaciones se encuentran también disponibles en: <http://timss.bc.edu/timss1995i/TIMSSPublications.html>

<sup>20</sup> Los resultados de TIMSS 1995 fueron publicados en 1996 y 1997, en informes que contienen la descripción del logro de los estudiantes en cuanto a matemáticas y ciencias y de los factores que en este influyen. Los informes técnicos de TIMSS 1995 comprenden tres volúmenes que documentan el desarrollo, diseño, análisis, y actividades que garantizan la calidad de TIMSS. Estos se pueden encontrar en las siguientes direcciones electrónicas <http://timss.bc.edu/timss1995i/TIMSSPDF/TR2book.pdf> <http://timss.bc.edu/timss1995i/TIMSSPDF/TR3book.pdf> <http://timss.bc.edu/timss1995i/TIMSSPDF/TRall.pdf>

<sup>21</sup> Estos materiales se pueden encontrar en el sitio: <http://timss.bc.edu/timss1999i/database.html>

<sup>22</sup> El término rendimiento, emanado del mundo del trabajo, lo podemos encontrar en algunos de los textos clásicos de Didáctica, en donde es considerado como aspecto importante de la verificación y evaluación que se realiza a través de los exámenes y es un término que algunos autores retoman y conceptualizan como: "la suma de transformaciones que se operan en: a) el pensamiento, b) en el lenguaje técnico, c) en la manera de obrar y d) en las bases actitudinales del comportamiento de los alumnos en relación con las situaciones y problemas de la materia que enseñamos" (Mattos, 1963). He aquí una situación que nos muestra la vinculación existente desde tiempos remotos entre empresa y educación, al incorporar conceptos laborales para tratar lo educativo.



- ¿Quién y cómo se imparte?
- ¿Cuáles son los conceptos, procesos y aptitudes que realmente han logrado desarrollar los alumnos?

También, en su carácter de evaluación comparativa pretende averiguar o tener puntos para definir ¿cómo varían los objetivos de aprendizaje del currículo oficial de matemáticas y ciencias de un país a otro? y ¿cuáles son las características del sistema educativo de cada país que influyen en el desarrollo de esos objetivos? Todo esto, para establecer las relaciones correspondientes entre currículo, contexto social y educativo.

#### **b. El Laboratorio Latinoamericano de Evaluación de la Calidad Educativa: LLECE**

Una experiencia de evaluación en el contexto Latinoamericano, es la que desarrolló el *Laboratorio Latinoamericano de Evaluación de la Calidad Educativa (LLECE)*.<sup>23</sup> Sistema de evaluación creado en 1994 especialmente para los países de la Región Latinoamericana y del Caribe por la Oficina Regional de Educación para América Latina y el Caribe (OREALC)<sup>24</sup> de la UNESCO; concebido más que como un estudio propiamente

---

<sup>23</sup> Los orígenes de la prueba LLECE se le atribuyen a la UNESCO cuando en el año de 1989, este organismo lleva a cabo a través de pruebas de rendimiento y encuestas un estudio comparativo exploratorio en siete países de la Región, cuyos resultados dejaron de manifiesto las deficiencias existentes en los sistemas educativos. En la presentación de estos resultados, los representantes regionales plantearon a la UNESCO-OREALC, la necesidad de contar con un mecanismo de evaluación que permitiera la obtención de información sistemática a través de una organización estable. Fue así como surge la idea de crear el Laboratorio Latinoamericano para la Evaluación de la Calidad de la Educación, después más comúnmente conocido como "Laboratorio," a partir del cual comienza a gestarse la idea de realizar una prueba comparativa propia para la región latinoamericana. Otro de los acontecimientos que por su impacto originó la creación de este "Laboratorio" fue el de, Proyecto Principal de Educación en América Latina y el Caribe (PROMEDLAC) iniciado en 1980 para terminar en el año 2000, creado como resultado del consenso de los representantes gubernamentales de los países de la región sobre la necesidad de un esfuerzo renovado, intenso y sostenido para atender, teniendo como horizonte el año 2000, carencias y necesidades educativas fundamentales no satisfechas e indispensables de enfrentar para el desarrollo de los países, respondiendo a objetivos compartidos por todos y mediante acciones en que el esfuerzo propio de cada país se beneficie de la cooperación horizontal, subregional, regional e internacional. (UNESCO/OREALC 1981-2000). De ahí que en 1993 los ministros reunidos en Chile discutieran la necesidad de que UNESCO impulsara el desarrollo de la evaluación y la creación de estándares en la Región (Ferrer, 2002).

<sup>24</sup> Los programas promovidos por la OREALC/UNESCO, obedecen a los principios del Informe Delors para el siglo XXI (Véase Delors, 1996).



comparativo, como una Red de los Sistemas Nacionales de Evaluación surgida a partir de la década de los noventa en el marco de los objetivos y recomendaciones del Proyecto Principal de Educación en América Latina y el Caribe en el periodo 1993-1996;<sup>25</sup> representa la parte operativa y concreta de uno de los programas estratégicos de este organismo internacional denominado: "Evaluación de la Calidad de los Aprendizaje."<sup>26</sup>

Este *Primer Estudio Internacional Comparativo sobre Lenguaje, Matemáticas y Factores Asociados para alumnos de Tercer y Cuarto grado* de la Educación Básica realizado por el Laboratorio Latinoamericano de Evaluación de la Calidad, se aplicó a un número aproximado de 48, 688 niños, y a 41, 077 familiares, 3,675 docentes y 1,387 directores (Tiana, 2000: 7-8) de trece países de la Región,<sup>27</sup> cuyos resultados fueron publicados en el año de 1998.<sup>28</sup>

Entre los objetivos del LLECE, está el generar estándares<sup>29</sup> regionales de aprendizajes, medir el estado de los niveles de aprendizaje logrados, tener información respecto a la situación del aprendizaje escolar y las variables que en éste inciden, obtener información sobre la calidad y estado de las variables, y experimentar el enfoque innovador de algunos métodos pedagógicos; para elaborar una comparación del rendimiento académico y analizar los factores asociados a ese rendimiento. Así como, fortalecer los sistemas nacionales de medición y evaluación por vías prácticas y generar una vía de investigación a

---

<sup>25</sup> Para tener una idea más clara de los alcances de este proyecto se puede consultar (UNESCO 2001).

<sup>26</sup> La puesta en marcha de sistemas de evaluación en la Región está estrechamente vinculada con la prioridad política de mejorar la calidad a través de la evaluación, lo que comenzó a darse primero en países más desarrollados y luego se hizo extensiva a América Latina con los llamados Sistemas Nacionales de Evaluación surgidos en la última década del siglo XX. "La evaluación es concebida como un poderoso instrumento al servicio de la mejora cualitativa de la educación, gracias al conocimiento valorativo que proporciona sobre la realidad" (Tiana 2000:4).

<sup>27</sup> Los países que participaron en este primer estudio fueron: Argentina, Bolivia, Brasil, Colombia, Cuba, Chile, Costa Rica, Honduras, México, Paraguay, Perú, República Dominicana y Venezuela.

<sup>28</sup> Los resultados de este primer estudio que presenta los logros obtenidos por once de los trece países participantes (El Informe no incluye los resultados de Perú y Costa Rica por circunstancias técnicas y de autorización) se publicaron en el año de 1998 (Casassus, J., Froemel, J. E. y Palafox, J. K. Cusato, S., 1998) en donde se presenta una visión comparada de los logros obtenidos en materia de aprendizaje en países con características culturales semejantes.

<sup>29</sup> El Laboratorio Latinoamericano de Evaluación de la Calidad de la Educación, define a los estándares como "Construcciones (constructos teóricos) de referencia...útiles para llevar adelante acciones en algún ámbito determinado" (Casassus, 1997: 3).



partir de los resultados obtenidos<sup>30</sup>. Lo que se busca es establecer niveles de calidad y eficiencia en los sistemas educativos en aras de determinar los conocimientos y destrezas requeridas para promover la participación activa de jóvenes en el mercado de trabajo y apreciar los efectos de la enseñanza, sobre el aprendizaje.

La intención del LLECE es incorporar a los Planes y Programas de Estudio los niveles de logro e indicadores de calidad del aprendizaje, mediante el establecimiento de estándares, que intentan servir como contrucciones teóricas de referencia para la acción y para la rendición de cuentas.<sup>31</sup> Éstos comprenden cuatro dimensiones: lo prescrito, (objetivos educativos que constituyen el elemento básico, lo deseable, que alude a elementos de excelencia en cuanto a lo que se espera de la educación, lo observable, es decir, lo que efectivamente se logra y lo factible referente a las condiciones para su realización. (Cassasus, 1997: 6).

En el año 1997, surge la necesidad de aplicar pruebas de Lenguaje y Matemáticas para realizar el *Primer Estudio Internacional Comparativo sobre Lenguaje, Matemáticas y Factores Asociados en Tercero y Cuarto grado* (LLECE, 1998).

Las preguntas básicas a las que responde este estudio son:

---

<sup>30</sup> Los Sistemas Nacionales de Evaluación de la Calidad surgen a partir de la década de los noventa, como parte importante de una política educativa que cobra vigor ante los niveles de expansión experimentados por los Sistemas Educativos en la década de los ochenta en América Latina; pues una vez alcanzada la cobertura deseada en cuanto a oportunidades de acceso; lo que interesa ahora son los contenidos y procesos de la experiencia escolar, para así mejorar la calidad de la educación en cuanto a procesos y variables que inciden en los niveles y resultados de aprendizaje obtenidos por los estudiantes. Existen varios estudios y documentos que se han dedicado al análisis y difusión de estos sistemas de evaluación. (Horn, Wolf y Velez, 1992; Tiana, 1996; Rojas y Esquivel, 1998; Schwartzman, 2001 y Arregui, 2003). Actualmente los Sistemas Nacionales de Evaluación se encuentran en una etapa de revisión y mejora en cuanto a su función, calidad técnica, uso y difusión de los resultados. El Programa de Promoción de la Reforma Educativa en América Latina y el Caribe (PREAL), en conjunto con Grupo GRADE (Grupo de Análisis para el Desarrollo) tienen varios documentos al respecto, (véase Esquivel, 2001 y Ravela, 2001). que se pueden consultar en: <http://www.preal.cl/GTEE/> o en la biblioteca digital de la OEI <http://www.campus-oei.org/oeivirt/evaleduc.htm>

<sup>31</sup> *Rendición de cuentas* es concepto emanado del mundo de trabajo que destaca una vez más, la inserción de conceptos laborales al ámbito educativo y el enfoque pragmático, que encierra para la educación y la evaluación el establecimiento de estándares. "Si no existiera esta relación con la acción, no tendría mucho sentido elaborar estándares en la educación, pues la observación de la calidad en la educación, conlleva a la necesidad de formular acciones para mejorarla... al no existir estándares nacionales, la sociedad no tiene los elementos de información necesarios para exigir *accountability* al sistema" (Cassasus, 1997: 6).



- ¿Qué aprenden los alumnos?
- ¿Cuál es el nivel en el que ocurren estos aprendizajes?
- ¿Qué competencias han desarrollado?

Se generaron instrumentos de medición y evaluación, consistentes en pruebas de matemáticas y lengua y cuestionarios de contexto, con la intención de conocer las aptitudes, capacidades, competencias y destrezas<sup>32</sup> de los alumnos; conceptos que se constituyen como estándares básicos, relacionados con el aprendizaje, que refieren lo que se espera de un sistema educativo.<sup>33</sup>

### c. Programa Internacional de Evaluación de los Estudiantes: PISA

El más reciente estudio evaluativo de corte internacional es el *Programa de Evaluación Internacional de estudiantes*, (*Programme for International Student Assessment*) PISA por sus siglas en inglés, realizado básicamente con todos los países miembros de la OCDE y con algunos países no miembros<sup>34</sup> que aceptaron unirse.<sup>35</sup>

El estudio se concibió en el año de 1997 y se puso en marcha al año siguiente. Fue aplicado por primera vez en el 2000 a un número aproximado de 250 000 estudiantes de quince años, aquellos que se encuentran concluyendo la educación básica a nivel mundial. Una

---

<sup>32</sup> Los documentos del Laboratorio, refieren la aptitud como disposición innata relacionada con el rendimiento que permite diferenciar a los individuos con un mismo nivel mental. La capacidad se refiere a poder realizar un acto físico o mental, puede ser innata o desarrollada mediante el aprendizaje. Y la competencia alude a la formación o preparación de un individuo para intervenir de un modo eficaz en un proceso de contexto. (Cassasus, 1997: 9).

<sup>33</sup> Según documentos del Laboratorio: "cuando se quiere generar estándares, es inevitable referirse a lo que se espera que los alumnos aprendan y que los profesores enseñen. Es necesario establecer de forma clara y pública, cuáles son las competencias conceptuales y prácticas que se espera que alcancen los alumnos" (Cassasus, 1997: 6).

<sup>34</sup> Entre los primeros países no miembros de la OCDE, que decidieron incorporarse están: Brasil, Latvia Liechtenestein y Federación Rusa.

<sup>35</sup> Queremos destacar que este es el estudio cuyos resultados dados a conocer en nuestro país por primera vez, en el mes de diciembre del 2001 causaron gran controversia. Algunas de los encabezados de la prensa con respecto a los resultados de México en la prueba de la OCDE publicadas tanto en diarios nacionales como extranjeros fueron: "Truena OCDE a México en Educación". (Delgado, 2001) "Mala nota en examen de calidad"(Cevallos, 2001) .



muestra pequeña de los 17 millones de jóvenes de los 32 países participantes.(OCDE, 2002: 16) <sup>36</sup> Su objetivo, obtener indicadores sobre el rendimiento en las áreas de lectura,<sup>37</sup> matemáticas y ciencias de los estudiantes de 15 años; para por una parte, contar con elementos que permitan influir en la definición de las políticas educativas y por otra, identificar y difundir los factores que en cada sistema educativo, favorecen el desarrollo de competencias y aptitudes.

Todo esto con el fin de apoyar a los países a mejorar su sistema escolar, en miras de una mejor preparación de los jóvenes para ingresar al mundo cambiante y globalizado.

La finalidad y diseño de estas pruebas, pretende responder a las siguientes cuestiones (OCDE, 2002: 3):

- ¿Tienen los estudiantes la preparación adecuada para enfrentar los retos del futuro?
- ¿Son capaces de analizar, razonar y comunicar sus ideas eficazmente ?
- ¿Tienen la capacidad de continuar aprendiendo durante toda su vida? <sup>38</sup>

---

<sup>36</sup> Posteriormente en el año 2001, 11 países decidieron incorporarse al proyecto PISA. Estos fueron Argentina, Bulgaria, Chile, Hong Kong-China, Indonesia, Israel, FYR Macedonia , Perú, Rumania y Tailandia. Los resultados de este estudio que implica una cobertura más amplia se publicaron en (OECD- UNESCO, 2003).

<sup>37</sup> PISA 2000, estuvo principalmente orientado a la evaluación de la lectura.

<sup>38</sup> Cuestionamiento base que la OCDE establece en sus documentos y que responde a un modelo de aprendizaje dinámico entendido como "aprendizaje a lo largo de la vida en el que se van añadiendo nuevos conocimientos y destrezas que son necesarios para la adaptación con éxito a las circunstancias de cada momento" (OCDE, 2000: 20).

El propósito que la OCDE se plantea, es realizar una evaluación de tipo prospectiva, enfocada a comprobar la capacidad de los jóvenes para emplear sus conocimientos y competencias al enfrentar los retos que presenta la vida real y la sociedad del conocimiento, más que comprobar el dominio de un plan de estudios (OCDE, 2002: 14). La intención, según se especifica en sus documentos, va más en un sentido de evaluar "competencias transversales" (OCDE, 2000:17). Llama la atención el que en los marcos teóricos que sustentan los estudios de evaluación de la OCDE, se recurra a este término para referir destrezas como comunicación, capacidad de adaptación, flexibilidad, solución de problemas y empleo de las tecnologías de la información, a manera de competencias esenciales para el desarrollo de los alumnos que requieren una evaluación con enfoque transversal.(OCDE, 2000:20). Un autor que ha abordado con mayor profundidad el problema de las competencias transversales en la educación es Bernard Rey (Véase Rey, 2001).

Sin embargo, asumimos que indudablemente tales cuestionamientos, responden a una intencionalidad de comprobar la efectividad de los sistemas educativos en términos de su funcionalidad para insertar a los jóvenes al mundo del trabajo y a la vida productiva. Este es el transfondo sociopolítico y económico de los estudios de evaluación de la OCDE, aspecto que no abordamos con profundidad en este trabajo, por considerar que escapa a las intención central del mismo, de realizar un análisis de los estudios de evaluación de la OCDE, más desde lo pedagógico que de lo social; pero que no podemos dejar de reconocer



El estudio es una evaluación comparada de tipo estandarizada de corte prospectivo, cuya intención es conocer el grado de preparación de los estudiantes para la vida adulta, en cuanto al manejo de ciertos conocimientos básicos, los procesos y principios que se necesitan para la comprensión de los conceptos implicados en éstos y la capacidad de los alumnos para responder a situaciones diferentes en contextos que demandan la aplicación de conocimientos y destrezas desarrolladas hasta el momento( OCDE, 2000).

En su carácter de estudio de evaluación prospectiva, el PISA, pretende dar continuidad a las evaluaciones realizadas, aplicándose consecutivamente cada tres años. En el 2003( véase OECD, 2003) se llevó a cabo la segunda evaluación del proyecto PISA;<sup>39</sup>participaron 42 países miembros de la OCDE<sup>40</sup> y tres países Latinoamericanos además de México;<sup>41</sup> el énfasis en este año se centró en el área de matemáticas.<sup>42</sup>Dado que en las aplicaciones subsecuentes, el PISA pretende incorporar la evaluación de distintas áreas del Plan de estudios, las pruebas del PISA 2003, incluyen la evaluación de aptitudes para la resolución de problemas, considerando a éstas como la capacidad para usar los procesos cognitivos en la solución de problemas cotidianos; en donde la solución no es tan obvia y se requiere la aplicación de los dominios de las áreas de matemáticas, ciencias y lectura. (OCDE, 2004:147).

El siguiente cuadro sintetiza el origen, orientaciones y finalidades de los estudios de evaluación de estas tres agencias internacionales.

---

y señalar que ha sido motivo de polémica y debate para algunos sociólogos, sobre todo por el significado sociocultural de los mismos, en cuanto a que si los alumnos que no obtienen resultados satisfactorios en estos estudios, significa que "no están preparados para la vida" (véase Gil, 2005).

<sup>39</sup> En México se aplicó del 2 al 6 de junio con la participación de 35 000 jóvenes pertenecientes a 1130 escuelas de toda la República. El Instituto Nacional para la Evaluación de la Educación (INEE), fue el encargado de coordinar dicha aplicación; así como de hacer la traducción de los instrumentos. Los resultados de esta aplicación se dieron a conocer a finales del 2004 y México contará con resultados por entidad federativa (con excepción de Michoacán).

<sup>40</sup> Cabe recordar que en PISA 2000, 32 fueron los países participantes y después en PISA 2001 se incorporaron 11 más.

<sup>41</sup> Los países latinoamericanos que participaron además de México son: Brasil, Perú y Uruguay.

<sup>42</sup> Conviene señalar que en PISA 2000 el énfasis de la evaluación estuvo orientada hacia la lectura.



CUADRO 1.1

**GENERACIÓN Y PROPÓSITOS DE LOS ESTUDIOS INTERNACIONALES DE EVALUACIÓN**

AGENCIA	ESTUDIO	ORIGEN	CONCEPTO ORGANIZADOR	PREGUNTAS A LAS QUE RESPONDE	FINALIDAD	EDAD DE LA POBLACIÓN QUE ATIENDE	ÁREAS Y CAMPOS A EVALUAR
IEA	TIMSS	1994	Curriculo	¿Qué es lo que se pretende que la sociedad aprenda? ¿Qué es lo que realmente se imparte? ¿Quién y cómo se imparte? ¿Qué es lo que realmente aprenden los alumnos?	Comprobar rendimiento Establecer relaciones entre: Currículo, Contexto social Y educativo	Tres poblaciones: Población 8-9 años (población uno) 13-14 años (población dos) 17-19 años (población tres)	Matemáticas Ciencias Contexto social y educativo
UNESCO	LLECE	1994	Calidad	¿Qué es lo que aprenden los alumnos? ¿Cuál es el nivel en el que ocurren dichos aprendizajes? ¿Qué competencias se han desarrollado?	Establecer estándares Comparar rendimiento y factores asociados	8- 9 años	Lenguaje Matemáticas Factores asociados
OCDE	PISA	1997	Conocimientos y aptitudes para la vida	¿ Hasta qué punto están preparados los jóvenes para enfrentar los retos del futuro? ¿ Son capaces de analizar, razonar y comunicar sus ideas correctamente? ¿ Disponen de la capacidad para seguir aprendiendo a lo largo de sus vidas?	Conocer el grado de preparación de los estudiantes para la vida y los factores que influyen	15 años	Lectura Matemáticas Ciencias Factores que favorecen desarrollo de la competencia

Hasta aquí, queremos destacar dos características fundamentales encontradas en el análisis y comparación de las orientaciones y finalidades de estos tres estudios de evaluación de corte internacional efectuados, durante la última década del siglo XX.



- 1) Se puede decir que los tres ejes básicos de su organización: Currículo, calidad y aptitudes para la vida, marcan un rumbo importante para los sistemas educativos y procesos de enseñanza-aprendizaje de fin de siglo; ahora lo importante ya no sólo es enseñar, sino enseñar con calidad de acuerdo a ciertos estándares establecidos y sobre todo con miras a desarrollar las capacidades necesarias para una aplicación de lo que se enseña y aprende en la escuela, a la vida cotidiana.
- 2) El fin con que se realizan dichas evaluaciones, tiene la intención no sólo de medir en términos de rendimiento, sino de conocer los factores asociados al rendimiento obtenido; lo que nos habla de un concepto más amplio de evaluación. Sin embargo, cabe señalar que la forma en como se presentan los resultados, no contribuye a la realización de tal objetivo;<sup>43</sup> los informes oficiales, incluyen lo referente a perfil de estudiantes y situaciones contextuales de los alumnos y de la escuela, pero se quedan en un nivel meramente descriptivo, los resultados obtenidos, continúan mostrándose en términos de puntajes, porcentajes, gráficas y escalas.<sup>44</sup> Pese a que en el informe sobre los resultados de PISA 2000, se incluye un capítulo que presenta las relaciones existentes entre desempeño de los estudiantes y entorno familiar y escolar del estudiante, (OCDE, 2002: 202-234) aún queda pendiente lo referente al significado pedagógico de los resultados obtenidos y a particularidades de cada país en materia de sistemas de enseñanza y organización curricular. Lo que invita a retomar la evaluación del aprendizaje y sus resultados, en un sentido más

---

<sup>43</sup> Algunos países como España, Uruguay y México, se han dado a la tarea de traducir los documentos oficiales (véase OCDE 2000, 2001 a y b, 2004 y Mullis, Martin, Smith, Garden, Gregory, Gonzales, Chrostowski, y O Conoor, 2002, OCDE/UNESCO, 2003) y de elaborar pequeños informes nacionales que permitan dar a conocer a docentes, alumnos y público en general, además de los resultados obtenidos por los alumnos de ese país en relación con los del resto del mundo, qué es lo que las pruebas evalúan y que implican los reactivos muestra utilizados. (véase López y Moreno, 1997a y b; INCE, 2004; Pajares, Pajares, R., Sanz, A., y Rico, L, 2004; ANEP, 2004 a, b, c y s/f a, b, c y d; Martínez, Vidal, R., Díaz, M. A. y Loyola del Rio, J. 2003; Vidal y Díaz, 2004) Sin embargo estos informes no contemplan la difusión de los resultados de contexto, ni la posible vinculación que pudiera existir, lo que nos remite a pensar en el predominio de un paradigma cuantitativo más que cualitativo en materia de presentación de resultados, lo que a su vez limita e impide que se lleve a cabo la función formativa y de retroalimentación de la evaluación.

<sup>44</sup> El trabajo de Ravela (2002), puede ayudar a tener más claridad al respecto.



significativo y formativo;<sup>45</sup> que permita a partir de un análisis pedagógico de los resultados, comprender y dar a conocer el porqué de los mismos y sus implicaciones en términos de contenidos, formas de enseñanza y sistema escolar.<sup>46</sup>

La intención de los estudios realizados por tres de las agencias internacionales de evaluación más importantes; nos lleva ahora a analizar el contenido, diseño y sistema de calificación de los instrumentos que se han utilizado para evaluar y comparar el rendimiento escolar en la educación básica.

Consideramos que lo más importante ante los resultados obtenidos, es dar cuenta de las características conceptuales y metodológicas de estos estudios y su relación con la situación de nuestro sistema educativo y de evaluación.

## 1.2. Contenido de las pruebas internacionales para alumnos de cuarto grado de la educación básica (8 y 9 años).

La revisión y análisis de los documentos existentes sobre Pruebas Internacionales, (véase LLECE, 2001; Mullis, Martin, Smith, Garden, Gregory, Gonzales, Chrostowski y O Conoor, 2002; OCDE, 2000 y 2002) permite dar cuenta de que referir el contenido de una prueba, es abordar no sólo los temas y áreas incluidos en el examen; sino sobre todo: la conceptualización, organización e implicaciones cognitivas de lo que se evalúa.

En primer término presentamos el contenido de los estudios realizados con poblaciones de estudiantes de 8 y 9 años, ubicados en el tercero y cuarto grado de la educación básica; como son el TIMSS 1995 y los del LLECE, para en un segundo momento, trabajar los que se ocupan de estudiantes que están por concluir la educación básica:<sup>47</sup> TIMSS y PISA .

---

<sup>45</sup> El término evaluación formativa se le atribuye a Scriven (1967). Scriven distinguió la función formativa de la evaluación, en contraposición a la sumativa, con la idea de contribuir a una curricula alternativa. "La función formativa de la evaluación, se proyecta no sobre los resultados, sino sobre el proceso didáctico. Sirve para determinar la naturaleza del desarrollo del mismo y constituye el punto de partida para decisiones de perfeccionamiento" (Rosales, 1990:34).

<sup>46</sup> El trabajo realizado por Bertoni, A., Poggi, M. y Teobaldo, M. (1995) en Argentina, constituye una referencia importante para pensar en un tratamiento más pedagógico de la evaluación.

<sup>47</sup> Los estudiantes que están por concluir la educación básica, son estudiantes de 13,14 y 15 años.



### a. Pruebas de matemáticas: TIMSS y LLECE

En el caso del TIMSS, la selección de los contenidos en cada una de sus áreas de evaluación, se realiza con base en el currículo de la mayoría de los países participantes y en común acuerdo con el Consejo Internacional de Expertos en Educación Matemática y en Ciencia que vertieron su opinión en cuanto a la importancia de estos contenidos para el futuro y la adaptación de los mismos a las características de los estudiantes con quienes se trabajó y su factibilidad para ser evaluados.

Las matemáticas, son concebidas como la capacidad de pensamiento lógico y sistemático que implica la transferencia de conocimientos y destrezas a nuevas situaciones.

La finalidad de la evaluación matemática en TIMSS, es medir la capacidad de los estudiantes para utilizar procedimientos rutinarios en la resolución de problemas y comunicar su conocimiento y razonamiento matemático.

El contenido de las pruebas se centra en dos dimensiones: *dominios de contenidos*<sup>48</sup> y *dominios cognitivos* que es a lo que comúnmente llaman comportamientos.<sup>49</sup>

Cinco son los temas en que se organizaron los contenidos de TIMSS 2003:<sup>50</sup>

1. Números
2. Álgebra (identificada para el cuarto grado como patrones, ecuaciones y relaciones)<sup>51</sup>

---

<sup>48</sup> Para algunos autores (Buer, 1993), el término dominio de contenido va relacionado con la interpretación de datos que fundamenta una teoría o diseño.

<sup>49</sup> Aunque todavía no está resuelto el debate en cuanto a la conceptualización y diferencia entre capacidad, habilidad y competencias, algunos autores, utilizan el término comportamientos para definir a una competencia (Rey, 1996). Perrenoud (1999), establece que una competencia exige la activación tanto de conocimientos como de operaciones mentales complejas y esquemas.

<sup>50</sup> En el TIMSS 1999 los contenidos incluidos fueron: Fracciones y sentido numérico, medición, representación de datos, análisis y probabilidad, geometría y álgebra. En el TIMSS 1995 consistieron en: números naturales, fracciones, proporcionalidad, medición, estimación y sentido numérico, representación de datos, análisis y probabilidad, geometría y patrones, relaciones y funciones. (Mullis, I V S, Martin, M O, Smith, T A, Garden, R A, Gregory, K D, Gonzales, E J, Chrostowski, S J. y O Conoor, K, 2002).

<sup>51</sup> Esto obedece a que el álgebra no es un área que se trabaje en la escuela primaria (en México, solamente en el Plan de Estudios de Educación Secundaria, se incluyen temas álgebra).



3. Medición
4. Geometría
5. Datos

Cada uno de estos temas incluyen objetivos y áreas apropiados a las diferentes poblaciones, expresados en términos de destrezas implícitas a deducir en los ítems, de acuerdo a las tareas que hay que realizar.

El tema de *números* demanda una comprensión del conteo, de la forma de representarlos, y relacionarlos, de los sistemas numéricos. Se subdivide en cuatro áreas:

- Números naturales
- Fracciones y decimales
- Números enteros
- Razón, proporción y porcentaje

En *álgebra* las áreas incluidas para la población uno, son tres:

- Patrones
- Ecuaciones y fórmulas
- Relaciones

*Medición*, es un tema enfocado a la comprensión de los atributos mensurables y al conocimiento de las unidades y los procesos de medida empleados en diversos atributos, que se subdivide en dos áreas:

- Atributos y unidades
- Herramientas técnicas y fórmulas

El tema de *geometría*, rebasa la simple identificación de formas y demanda saber analizar las propiedades y características de una variedad de figuras geométricas y explicar las relaciones entre éstas. Las áreas que incluye son cinco:

- Líneas y ángulos
- Formas bidimensionales y tridimensionales
- Congruencia y similitud
- Localización y relaciones



- Simetría y transformaciones

Los *datos* están enfocados a comprender cómo se recopilan, organizan, representan e interpretan éstos en situaciones concretas. Está dividido en cuatro áreas:

- Recopilación y organización de datos
- Representación de datos
- Interpretación de datos
- Incertidumbre y probabilidad

Estas áreas, en que se han agrupado los contenidos es semejante para los alumnos de ocho y nueve años de edad, como para los que están por concluir la educación básica; lo que varía es el nivel matemático y cognitivo de cada ítem; sin embargo las áreas referentes a *números naturales, expresiones algebraicas, incertidumbre y probabilidad*, no están contempladas para los niños del cuarto curso.

Cuatro son los dominios cognitivos, del área de matemáticas, incluidos en los contenidos y expresados en términos de destrezas y habilidades:

1. Conocimiento de hechos y de procedimientos
2. Utilización de conceptos
3. Resolución de problemas
4. Razonamiento

Cada uno de estos dominios está también claramente definido y comprende comportamientos específicos, aplicables a ambas poblaciones (población uno y dos) pero con diferente grado de complejidad en cada una.<sup>52</sup>

Los *hechos*, referidos a un conocimiento factual, son considerados como un lenguaje que forma parte del pensamiento matemático y los *procedimientos*,<sup>53</sup> constituyen el puente entre el conocimiento básico y la aplicación de las matemáticas para la solución de problemas de

---

<sup>52</sup> Cabe aclarar que los diferentes niveles de complejidad de los dominios cognitivos en cada una de las poblaciones se refieren a una complejidad de corte cognitivo y no al nivel de dificultad de los ítems.

<sup>53</sup> Conviene recordar aquí que Cesar Coll, junto con otros colaboradores, dentro del movimiento de Reforma español de la educación básica, han manejando la enseñanza y aprendizaje de hechos, conceptos, procedimientos y aptitudes y la clasificación o denominación de contenidos curriculares en estas tres categorías (Véase Coll, Pozo, Sarabia, y Valls, 1992).



la vida diaria. Los comportamientos incluidos en el conocimiento de hechos y procedimientos son:

- Recordar
- Reconocer
- Calcular
- Usar herramientas

Los *conceptos*, son representaciones o ideas que constituyen el centro de la comunicación matemática y la capacidad para crear representaciones y constituyen la base para llegar a la aplicación y comprensión de las matemáticas. Los comportamientos incluidos en este dominio son:

- Saber
- Clasificar
- Representar
- Formular
- Distinguir

La *resolución de problemas*, considerado como propósito central y muchas veces como un medio en la enseñanza de las matemáticas, comprende problemas tanto de la vida diaria como de preguntas netamente matemáticas que demanda el empleo de expresiones numéricas o algebraicas; por lo tanto es un dominio incluido aquí de manera específica y como parte del dominio del razonamiento. Las destrezas requeridas son:

- Seleccionar
- Representar
- Interpretar
- Aplicar
- Verificar o comprobar

El *razonamiento matemático*, comprende la capacidad de pensamiento lógico y sistemático, e implica tanto el razonamiento intuitivo como el inductivo para lograr la resolución de problemas no familiares para el estudiante, ya sea meramente matemáticos o inmersos en



situaciones reales, que requieren de exigencias cognitivas superiores a las utilizadas en problemas de la vida diaria; lo que demanda la transferencia del conocimiento y destrezas a nuevas situaciones. En este dominio los comportamientos requeridos son:

- Formular hipótesis, conjeturar o predecir
- Analizar
- Evaluar
- Generalizar
- Conectar
- Sintetizar o integrar
- Resolver problemas no habituales
- Justificar o demostrar

El contenido de evaluación de las pruebas del TIMSS, sugiere algunas reflexiones:

El contenido aún contempla temas comunes del currículo escolar; sin embargo una aportación importante de las pruebas del TIMSS, es que no sólo se queda en contenidos, sino que la evaluación de esos contenidos demanda el ejercicio de procesos cognitivos. Lo que su vez da pie para pensar en la relación existente entre contenido de evaluación y sistemas de enseñanza, es decir, ¿qué tanto en las escuelas se enseña a utilizar procedimientos y conceptos, resolver problemas y emplear el razonamiento?

El contenido de evaluación de las pruebas del TIMSS, presenta dos retos importantes: El primero gira en torno a pugnar por la evaluación de procesos cognitivos y de aprendizajes no sólo de contenidos curriculares; pero cuidando que los sistemas de enseñanza incluyan el ejercicio de estos procesos para la aplicación del conocimiento a problemas de la vida cotidiana.

Dado que las matemáticas es un área de evaluación común tanto en el TIMSS, como en el LLECE, presentamos ahora el contenido de evaluación matemática de las pruebas del LLECE.

En el caso de las matemáticas, lo que se pretende es identificar la capacidad de los estudiantes para *matematizar* el mundo a través de la construcción de modelos, para una



evaluación del desarrollo de la *competencia matemática*<sup>54</sup> para el uso de procedimientos matemáticos. Esta competencia es evaluada a través de cinco áreas temáticas:

1. Numeración
2. Operaciones con números naturales
3. Fracciones comunes
4. Geometría
5. Medición

Cada uno de estos, incorpora además, algunas habilidades aritméticas de nivel superior como:

- Interpretación de gráficos
- Reconocimientos de patrones
- Manejo de probabilidades
- Establecimiento de relaciones entre datos

La prueba está organizada en tres rangos denominados *niveles de desempeño*:<sup>55</sup>

- *Reconocimiento y utilización de hechos y relaciones matemáticas*: Implica saber leer y escribir números y establecer relaciones de orden en el Sistema Decimal, así como reconocer figuras planas y utilizar algoritmos rutinarios usuales. Se trata de contenidos y habilidades que constituyen la base para la comprensión matemática.
- *Reconocimiento y utilización de estructuras matemáticas simples*. Implica la capacidad para reconocer patrones, establecer regularidades y aplicar operaciones en situaciones no convencionales. Se trata del manejo de estructuras matemáticas simples que subyacen a situaciones matemáticas cotidianas; lo que implica el trabajo tanto de operaciones algorítmicas y ejercicios usuales como de situaciones problemas a las que les subyacen estructuras aditivas.

---

<sup>54</sup> En el siguiente capítulo presentamos lo que el LLECE entiende por competencia matemática.

<sup>55</sup> De acuerdo con lo estipulado por el Laboratorio Latinoamericano de Evaluación de la Calidad, "cuando se habla de nivel de desempeño, se hace referencia a un espacio caracterizado por el reagrupamiento de preguntas que cumplen con ciertos rasgos particulares en razón de la dificultad y la habilidad del estudiante que las responde" (LLECE, 1998).



- *Reconocimiento y utilización de estructuras matemáticas complejas:* Implica reconocer estructuras matemáticas complejas que subyacen a situaciones matemáticas o a situaciones cotidianas matematizables. Se trata de abordar problemas, el reconocimiento de la estructura del sistema de numeración y el manejo del valor posicional para el establecimiento de equivalencias. Los problemas de este nivel, requieren además del uso de los algoritmos, estrategias y procedimientos comunes, la creación y aplicación de estrategias, razonamientos y planes no rutinarios que requieren usar su conocimiento matemático y su competencia para darle solución a los problemas presentados (LLECE, 1998).

En las dimensiones del contenido se incluyen habilidades y destrezas adquiridas por los educandos en cuanto a pensamiento<sup>56</sup> lógico-matemático.<sup>57</sup>

En el siguiente cuadro se presenta una síntesis de los elementos que tanto el TIMSS, como el LLECE, consideran en la organización del contenido de las pruebas de matemáticas para alumnos de ocho y nueve años.

---

<sup>56</sup> La conceptualización de pensamiento en las pruebas del Laboratorio está fundamentadas en la definición del psicólogo alemán Friedhart Klix: "el pensamiento es la creación de información mediante procesos cognoscitivos" (Klix, 1983, cit. en LLECE, 2001:20).

<sup>57</sup> El pensamiento lógico matemático para el Laboratorio, implica "desarrollar la capacidad para obtener información desconocida utilizando procedimientos matemáticos de procesamiento de la información que se conoce.(LLECE, 2001:20).



CUADRO 1.2

**PRUEBAS DE MATEMÁTICAS PARA ALUMNOS DE 8 Y 9 AÑOS:  
TIMSS Y LLECE**

PRUEBA	CONCEPTO	PROPÓSITO DE EVALUACIÓN	DIMENSIONES	CONTENIDO DE LA PRUEBA	TEMAS O ÁREAS	PROCESOS COGNITIVOS	DESTREZAS IMPLÍCITAS
TIMSS	Capacidad de pensamiento lógico y sistemático que implica la transferencia de conocimientos y destrezas a nuevas situaciones	Medir la capacidad de los estudiantes para utilizar procedimientos rutinarios en la resolución de problemas y comunicar su conocimiento y razonamiento matemático.	Dominios de contenidos  Dominios cognitivos	Cinco campos:		Comportamientos	
				Números	Naturales Fracciones y decimales Razón, proporción y porcentaje	Conocimiento de hechos y procedimientos	Recordar Reconocer Calcular Usar herramientas
				Álgebra	Patrones Ecuaciones y fórmulas Relaciones	Utilización de conceptos	Saber Clasificar Representar Formular Distinguir
				Medición	Atributos y unidades Herramientas, técnicas y formas	Resolución de problemas	Seleccionar Representar Interpretar Aplicar Verificar o comprobar
				Geometría	Líneas y ángulos Formas bidimensionales y tridimensionales Congruencia y similitud Localizaciones y relaciones espaciales Simetría y transformación	Razonamiento	Formular hipótesis Analizar Evaluar Generalizar Conectar Sintetizar Resolver problemas Justificar
				Datos	Recopilación y organización Representación Interpretación		



LLECE	Uso de procedimientos matemáticos que se le ha dado en llamar: competencia matemática	Identificar la capacidad de los estudiantes para matematizar el mundo	Habilidades y destrezas para un pensamiento lógico-matemático	<p><b>3 RANGOS O NIVELES DE DESEMPEÑO :</b></p> <p>(1) Reconocimiento y utilización de hechos y relaciones matemáticas</p> <p>(2) Reconocimiento y utilización de estructuras matemáticas simples</p> <p>(3) Reconocimiento utilización de estructuras matemáticas complejas</p>	<p>Numeración</p> <p>Operación con números naturales</p> <p>Fracciones comunes</p> <p>Geometría</p> <p>Medición</p>	<p><b>HABILIDADES MATEMÁTICAS DE NIVEL SUPERIOR</b></p> <p>Interpretación de gráficos</p> <p>Reconocimientos de patrones</p> <p>Manejo de probabilidades</p> <p>Establecimiento de relaciones entre datos</p>	<p><b>Nivel (1):</b> (contenidos y habilidades básicos para la comprensión matemática) Leer y escribir no. naturales Establecer relaciones de orden en el sistema decimal Reconocer figuras planas Utilizar algoritmos rutinarios</p> <p><b>Nivel (2):</b> (Manejo de estructuras matemáticas simples) Reconocer patrones Establecer regularidades Aplicar operaciones en situaciones no convencionales</p> <p><b>Nivel 3:</b> (Reconocer estructuras matemáticas complejas que subyacen a situaciones matemáticas cotidianas) Creación y aplicación de estrategias, razonamientos y planes no rutinarios para usar su conocimiento y resolver problemas</p>
-------	---	---	---	--	---	---	--



Hasta aquí, podemos decir que las pruebas de matemáticas del TIMSS y LLECE para alumnos de ocho y nueve años, manejan un concepto de matemáticas aplicadas a situaciones cotidianas, su propósito es verificar la capacidad de los estudiantes para demostrar y comunicar su conocimiento y pensamiento matemático. El contenido de la evaluación se da en dos dimensiones: conocimientos temáticos y habilidades implícitas en los conocimientos a las que comúnmente llaman destrezas. En cuanto a los campos temáticos, se puede decir que los números naturales, las fracciones, la medición y la geometría son áreas comunes en ambas pruebas y en el currículo escolar. Aunque la organización de la prueba el TIMSS, está conforme a campos temáticos y el LLECE, con base en niveles de desempeño; existe cierta jerarquización en los temas abordados. Sin embargo, se puede distinguir una diferencia en cuanto al nivel de conocimientos que se contemplan en cada una: Las pruebas del TIMSS, incluyen cuestiones de álgebra y manejo de datos, mientras que el LLECE, se queda en el manejo y aplicación de conocimientos y estructuras matemáticas básicas. Los procesos cognitivos, implícitos en los contenidos también varían un tanto: el TIMSS, los centra en los requeridos para la solución de problemas y el LLECE se enfoca más al reconocimiento, interpretación y establecimiento de relaciones. Lo que nos lleva a plantear que en virtud de que las matemáticas, en ambas pruebas son conceptualizadas como el empleo de conocimientos y procedimientos matemáticos al mundo real; éstas son evaluadas a través de la aplicación de conceptos, relaciones y estructuras matemáticas básicas a situaciones que requieren no sólo del conocimiento o reconocimiento de ciertos conceptos; sino de la aplicación de destrezas de orden superior como la relación, la interpretación y el razonamiento. Dado que tanto el TIMSS, como el LLECE, tienen otra área de evaluación, a continuación presentamos el contenido a evaluar en cada una de estas.



## b. Pruebas de ciencias: TIMSS

Las *ciencias*, en el TIMSS 2003 se conceptualizan como “proceso utilizado para aprender sobre el mundo físico, que implica observar, investigar y explicar fenómenos naturales, lo que incluye tanto la demostración del conocimiento de los contenidos, como la habilidad de aplicar y comunicar la comprensión conceptual en la resolución de problemas, elaboración de explicaciones y desarrollo de investigaciones”. (Mullis, Martin, Smith, Garden, Gregory, Gonzales, Chrostowski y O Conoor, 2002: 73).

Su evaluación comprende dominios de contenidos agrupados en cinco campos temáticos;<sup>58</sup>

1. Ciencias de la vida
2. Química
3. Física
4. Ciencias de la tierra
5. Ciencias medioambientales

Al igual que en las matemáticas, cada dominio comprende áreas temáticas definidas en términos de objetivos específicos para cada una de las poblaciones.<sup>59</sup>

Para el caso de la población uno, (4° grado, 8 y 9 años) las áreas incluidas en cada uno de los dominios son:

*Ciencias de la vida:*

- Tipos, características y clasificación de los seres vivos
- Estructura, función y procesos vitales de los organismos
- Desarrollo y ciclos de vida de los organismos

---

<sup>58</sup> La organización de los temas en general es la misma que se utilizó en el TIMSS 1995 y TIMSS 1999; sólo hubo algunas variaciones en el contenido de las áreas de ciencias medioambientales y de investigación científica. En TIMSS 1995 existió un área de evaluación que combinaba las cuestiones medioambientales con las de la Naturaleza para la Ciencia y en 1999 esta categoría fue reemplazada por dos más: Cuestiones medioambientales y de recursos, Investigación científica y Naturaleza de la Ciencia.

<sup>59</sup> Para el caso de 4° grado (8 y 9 años), los objetivos propuestos, encaminados al logro de una comprensión conceptual son de tipo observable, mientras que los de 8° grado (13-14 años), son de corte más abstracto.



- Reproducción y herencia
- Diversidad, adaptación y selección natural
- Ecosistemas
- La salud humana

*Química:*

- Clasificación y composición de la materia
- Propiedades y usos del agua
- El cambio químico

*Física:*

- Estados físicos y cambios en la materia
- Tipos, fuentes y conversiones de energía
- Calor y temperatura
- Luz
- Electricidad y magnetismo
- Fuerzas y movimiento

*Ciencias de la tierra:*

- Estructura y rasgos físicos de la tierra
- Procesos, ciclos e historia de la Tierra
- La tierra en el sistema solar y en el universo

*Ciencias medioambientales:*

- Utilización y conservación de los recursos naturales
- Cambios en el medio ambiente

Cada uno de estos dominios requiere de destrezas y habilidades cognitivas implícitas en los ítem para su evaluación.



Los dominios cognitivos<sup>60</sup> explícitos en los documentos que refieren el marco teórico para la evaluación la evaluación de las ciencias, son básicamente tres: (Mullis, Martin, Smith, Garden, Gregory, Gonzales, Chrostowski y O Conoor, 2002: 49).

1. Conocimiento factual
2. Comprensión conceptual
3. Razonamiento y análisis

Estos al igual que en las matemáticas, son aplicables a ambas poblaciones pero con habilidades cognitivas acordes para cada nivel.

El *conocimiento factual*, comprende los conocimientos básicos que deben tener los estudiantes en cuanto a hechos, datos, herramientas y procedimientos relevantes. Sin embargo éste, va más allá del simple hecho de memorizar datos aislados, supone el empleo y comprensión de conceptos en términos científicos.<sup>61</sup> Los comportamientos requeridos en este dominio son:

- Recordar o reconocer
- Definir
- Describir
- Usar herramientas y procedimientos

La *comprensión conceptual* se refiere a "captar las relaciones que explican el comportamiento del mundo físico y relacionar lo observable con conceptos científicos más abstractos o más generales." (Mullis, Martin, Smith, Garden, Gregory, Gonzales, Chrostowsk y O Conoor, 2002:74). Los comportamientos de este dominio implican:<sup>62</sup>

---

<sup>60</sup> Cabe señalar que el Marco Teórico de las pruebas del TIMSS, presenta como *dominios cognitivos* para la evaluación de las Ciencias, lo que Coll, trabaja como contenidos procedimentales, al plantear la necesidad de incluir en el currículo, no sólo contenidos referentes a hechos y conceptos (Véase Coll, Pozo, Sarabia, y Valls, 1992).

<sup>61</sup> Si bien, de acuerdo con los planteamientos de Coll, (1992), el conocimiento de hechos y conceptos, más que un dominio cognitivo se refiere a contenidos de hechos y principios, las pruebas del TIMSS, lo presentan como dominio cognitivo por la inclusión de procesos y procedimientos mentales que se requiere para el manejo de hechos y conceptos en términos científicos.

<sup>62</sup> En términos de Coll, éstos vendrían a ser más bien contenidos de tipo procedimental, aunque de distinto tipo y nivel (Véase Coll, Pozo, , Sarabia, y Valls, 1992:94).



- Ilustrar con ejemplos
- Comparar, contrastar o clasificar
- Representar o modificar
- Relacionar
- Obtener o aplicar información
- Hallar soluciones
- Explicar

El *razonamiento* y *análisis* son dos acciones implícitas tanto en la resolución de problemas científicos, como en el desarrollo de explicaciones y elaboración de conclusiones para la toma de decisiones; lo que requerirá en muchas ocasiones del análisis del problema para encontrar los principios que contempla y llegar a la solución mediante la aplicación de un concepto o relación científica. Los comportamientos incluidos en este dominio se refieren a:

- Analizar, interpretar o resolver problemas
- Integrar o sintetizar
- Formular hipótesis o predecir
- Diseñar o planificar
- Recopilar, analizar e interpretar datos
- Sacar conclusiones
- Generalizar
- Evaluar
- Justificar

Las pruebas del TIMSS, también comprenden la *investigación científica*, como un área de evaluación separada aplicada a las dos poblaciones (cuarto y octavo curso de la educación básica). Se trata de una dimensión más que incluye conocimientos destrezas y habilidades , evaluados en los ítems o tareas de los diferentes contenidos. De ahí que las tareas o ítems designados a evaluar la comprensión y las habilidades relacionadas con la investigación científica, se vinculan con ciertos dominios de contenido y cognitivos.



En el siguiente cuadro, podemos ver cómo la organización de las pruebas de ciencias del TIMSS, presenta cierta relación entre conceptualización, dominios de contenido y dominios cognitivos:

CUADRO 1.3

**LA PRUEBA DE CIENCIAS EN EL TIMSS PARA ALUMNOS DE 8 Y 9 AÑOS**

CONCEPTO DE CIENCIA	IMPLICACIONES COGNITIVAS	REQUERIMIENTOS	CAMPOS TEMÁTICOS	ÁREAS	DOMINIOS COGNITIVOS	DEFINICIÓN	COMPORTAMIENTOS
Proceso utilizado para aprender sobre el mundo físico	Observar Investigar Explicar fenómenos naturales	Demostración de conocimientos  Habilidad de aplicar y comunicar la comprensión conceptual	Ciencias de la vida	Tipos, características y clasificación de los seres vivos Estructura, función y procesos vitales, Desarrollo y ciclos de vida de los organismos Reproducción y herencia Diversidad, adaptación y selección natural Ecosistemas La salud humana	Conocimiento factual	Hechos Datos Herramientas Procedimientos	Recordar Reconocer Definir Describir
			Química	Clasificación y composición de la materia Propiedades y usos del agua El cambio químico	Comprensión conceptual	Captar relaciones que explican el comportamiento del mundo físico y relacionarlo observable con conceptos científicos	Ilustrar con ejemplos Comparar, contrastar o clasificar Representar o modificar Relacionar Obtener o aplicar información Hallar soluciones Explicar
			Física	Estados físicos y cambios en la materia Tipos, fuentes y conversiones de energía Calor y temperatura Luz Electricidad y magnetismo Fuerzas y movimiento	Razonamiento y análisis	Acciones implícitas tanto en la resolución de problemas científicos como en el desarrollo de explicaciones y elaboración de conclusiones	Analizar interpretar Resolver problemas Integrar o sintetizar Formular hipótesis Diseñar o planificar Recopilar, analizar e



			Ciencias de la tierra	Estructura y rasgos físicos Procesos cíclicos e historia  La tierra en el sistema solar y el universo		para tomar decisiones	interpretar datos Sacar conclusiones Generalizar Evaluar Justificar
			Ciencias Medioambientales	Utilización y conservación de los recursos naturales Cambios en el medio ambiente			

Hasta aquí, queda claro que el hecho de que las ciencias sean conceptualizadas como una comprensión del mundo físico; los dominios cognitivos y comportamientos señalados, consisten en procesos y procedimientos que favorecen dicha comprensión, con el propósito de evaluar qué tanto se saben usar los hechos, datos, herramientas y procedimientos científicos para explicar los acontecimientos de la vida.

### c. Pruebas del lenguaje: LLECE

En el LLECE, la evaluación del lenguaje está enfocada principalmente al desarrollo de la competencia comunicativa.<sup>63</sup> Con base en esto, las pruebas demandan en los estudiantes procesos concretos de comunicación que requieren el ejercicio de ciertas habilidades, porque lo que importa no sólo es demostrar qué tanto sabe del lenguaje, sino cómo se utiliza.

Los aspectos a observar en la construcción de las pruebas fueron la lectura y la escritura como formas de hacer, en donde el estudiante aplica su saber y saber hacer con el lenguaje.

<sup>63</sup> En el capítulo siguiente se presenta qué es lo que entiende el LLECE, por competencia.



La evaluación está centrada en el proceso de comprensión lectora<sup>64</sup> y de producción escrita.

La organización de la prueba se realizó en tres niveles de desempeño:

- *Lectura literal primaria*: Se refiere a la reconstrucción semiótica del texto en su manifestación superficial. Las preguntas que se ubican en este nivel demandan el reconocimiento de estructuras explícitas a nivel local. Se plantean preguntas que proponen al lector, el reconocimiento de los actores de un relato, la identificación de los fragmentos claves en la argumentación y las relaciones explícitamente planteadas entre proposiciones.
- *Lectura de carácter literal a modo de paráfrasis*: Representa un grado mayor de complejidad, porque lo que se requiere es una traducción semántica regulada por el sentido literal de éste. Se trata de "decir" el texto con otras palabras.
- *Lectura de carácter inferencial*: En este nivel el lector tendrá que llenar los espacios vacíos del texto explicando los supuestos sobre los que está estructurado, lo que implica relacionar micro y macro proposiciones textuales identificando las distintas formas de implicación causación, temporalidad, espacialidad, inclusión y exclusión. Se trata de preguntas que exigen al estudiante relacionar porciones del texto alrededor de un tópico parcial y del reconocimiento de algunas siluetas textuales. (LLECE, 2002:24).

Las pruebas del LLECE están sustentadas teóricamente en un enfoque constructivista, fundamentalmente se destacaron los aportes de Vigotsky, Bruner y Ausubel, en relación a la construcción del conocimiento y los planteamientos para un aprendizaje significativo<sup>65</sup>

El siguiente cuadro muestra la organización de la prueba del lenguaje del LLECE y las relaciones existentes entre definición y contenido.

---

<sup>64</sup> Según los planteamientos teóricos en que se sustentan las pruebas del LLECE; desde una perspectiva cognitiva el proceso de comprensión requiere poner en juego tanto conocimientos previos como representaciones y saberes para construir y constatar hipótesis de lectura.

<sup>65</sup> Para profundizar en la teoría constructivista del aprendizaje y en los aportes de estos autores al aprendizaje significativo (Véase Díaz Barriga, Arceo y Hernández, 2002).



CUADRO 1.4

**LA PRUEBA DEL LENGUAJE DEL LLECE**

CONCEPTO	REQUERIMIENTOS	CAMPOS	CONTENIDOS	PROCESOS	ORGANIZACIÓN	DEFINICIÓN	IMPLICACIONES
Lenguaje entendido como competencia comunicativa.	Procesos concretos de comunicación que requieren el ejercicio de ciertas habilidades. Lo que importa no es sólo demostrar qué tanto se sabe del lenguaje, sino cómo se utiliza.	Comprensión lectora	Estrategias de pensamiento	Posibilidades interpretativas para construir o develar el significado de un texto, formulando hipótesis globales de lectura	Tres niveles de desempeño:		
					Lectura literal primaria.	Reconstrucción semiótica del texto en su manifestación superficial	Reconocimiento de estructuras explícitas Reconocimiento de los actores de un relato Identificación de fragmentos claves en la argumentación Relaciones explícitamente planteadas entre proposiciones.
		Producción escrita	Capacidad de los alumnos para integrar un texto complejo.	Producción de elementos interpretantes	Lectura a manera de paráfrasis	Decir el texto con otras palabras.	Traducción semántica regulada por el sentido literal
					Lectura de carácter literal		
Lectura de carácter inferencial							



La forma en como está organizada la prueba del lenguaje del LLECE, nos permite afirmar que son pruebas centradas en una interpretación, más que en la decodificación literal de un texto, lo que demanda al estudiante, capacidades para reconocer, traducir y relacionar estructuras de una producción escrita; esto con el fin de evaluar la comprensión y organización de los significados lingüísticos, definidos como competencia comunicativa.

### 1.3. Contenido de las Pruebas Internacionales para alumnos que están concluyendo la educación básica.

#### a. Pruebas de matemáticas: TIMSS Y PISA

La selección y conceptualización de contenidos de las pruebas de matemáticas del TIMSS para alumnos que están concluyendo la educación básica, son los mismos que para los de la población uno, (8 y 9 años, cuarto grado), sólo varían algunas áreas y el nivel de los objetivos, los cuales están adecuados a cada población.

Las áreas incluidas para la población dos (13 y 14 años, ubicada en el 8° grado)<sup>66</sup> en cada uno de los temas son:

- En *números: números enteros*
- En *álgebra: expresiones algebraicas*
- En *datos: incertidumbre y probabilidad*

Los dominios cognitivos de las pruebas del TIMSS, también son los mismos para ambas poblaciones.

En las pruebas de matemáticas del PISA, lo que se mide es la *aptitud matemática*, entendida como "aptitud para plantear, formular, resolver e interpretar problemas a través de las matemáticas en diferentes situaciones y contextos (OCDE, 2004:27).

Dos son los campos de evaluación:

---

<sup>66</sup> En México, equivaldría al 2° grado de secundaria.



1. "Grandes ideas"<sup>67</sup>

2. Competencias matemáticas

Estos campos, incluyen tres áreas:

- Contenidos
- Procesos o procedimientos
- Situaciones o tareas en las que se emplean las matemáticas

Los *contenidos*, definidos como: "conjuntos de conceptos matemáticos relevantes y conectados que se presentan en situaciones y contextos reales"(OCDE 2000:24). Son temas generales presentados como *grandes ideas* que incluyen:

- Cantidades
- Espacios
- Formas
- Cambios
- Relaciones e incertidumbre.

Los *procesos* o *procedimientos*, se refieren a aptitudes agrupadas en tres clases de "conjuntos de competencias:"<sup>68</sup>

- Reproducción.
- Conexiones
- Reflexión

La *reproducción*, incluye cálculos simples o definiciones comunes. Las *conexiones* demandan la reunión de ideas y procedimientos matemáticos para resolver problemas. La *reflexión* consiste en el razonamiento matemático y la generalización e introspección. Es decir son destrezas generales que incluyen la argumentación, diseño, planteamiento y resolución de problemas, representación, simbolización, formalismo, comunicación y uso de herramientas, todas ellas, enfocadas a la resolución de problemas, el uso del pensamiento

---

<sup>67</sup> En el proyecto OCDE-PISA, "las grandes ideas matemáticas representan conjunto de conceptos matemáticos fundamentales y conectados que aparecen en situaciones y contextos matemáticos reales" (OCDE 2000: 74).

<sup>68</sup> Recordemos que el término competencia, también es utilizado en las pruebas del LLECE.



matemático y el diseño de éste; las cuales no se evalúan por separado, sino de manera conjunta.

Cada una de estas competencias se evalúan a través de tareas que a su vez demandan un tipo de destreza específico.

Para el primer nivel de competencia, que es el más elemental, las tareas se refieren al: recuerdo de objetos y propiedades matemáticas, desarrollo de procedimientos de rutina y la manipulación de expresiones que contengan símbolos, es decir destrezas de corte simbólico, formal y técnico.

En el segundo nivel de competencia, el intermedio, las tareas que se solicitan son: manejo de diferentes métodos de representación de acuerdo a ciertas situaciones y objetos; así como distinguir y relacionar diferentes afirmaciones; lo que requiere de destrezas de argumentación diseño, planteamiento, decodificación e interpretación del lenguaje simbólico y formal y comprensión de sus relaciones con el lenguaje natural.

En el tercer nivel de competencia que es el más complejo, lo que se pretende es la *matematización* implícita en todo pensamiento matemático.<sup>69</sup> Las tareas que se incluyen son: análisis, interpretación y desarrollo de los propios modelos y estrategias para la resolución de problemas, así como presentación de argumentos matemáticos que incluyen demostraciones y generalizaciones. Para lo cual se requiere de la capacidad para expresar y comunicar ideas matemáticas en distintas maneras, así como de la comprensión y razonamiento de las expresiones matemáticas en diversidad de formas. (OCDE, 2000: 75-78). Las *situaciones*, se refieren a "tareas auténticas" que representan los tipos de problemas en donde se emplean las matemáticas, que se encuentran en la vida diaria.

---

<sup>69</sup> Nótese que el término *matematizar*, también es utilizado en las pruebas del LLECE. La *matematización* dentro del proyecto PISA, se concibe como la Organización de la realidad percibida a través de la utilización de ideas y conceptos matemáticos para lo cual se requiere de diversas competencias. Esta *matematización* puede acontecer en dos direcciones: *Matematización horizontal*, que implica trasladar al mundo real el matemático, mediante la organización empleada para descubrir regularidades, relaciones y estructuras con base en los conocimientos y destrezas matemáticas adquiridas. La *matematización vertical*, que requiere a trabajar sobre un problema y emplear herramientas para resolverlo, mediante la representación de una relación, la comprobación de regularidades, el refinado combinación e integración de ajustes y modelos (OCDE, 2000: 79).



El siguiente cuadro muestra de manera conjunta la organización del contenido en las pruebas de matemáticas del TIMSS y PISA para alumnos que están por concluir la educación básica.

CUADRO 1.5

**PRUEBAS DE MATEMÁTICAS DEL TIMSS Y DEL PISA PARA ALUMNOS QUE ESTÁN POR CONCLUIR LA EDUCACIÓN BÁSICA**

PRUEBA	DEFINICIÓN	DIMENSIONES	CAMPOS	TEMAS	PROCESOS COGNITIVOS	DESTREZAS IMPLÍCITAS
TIMSS	Capacidad de pensamiento lógico y sistemático que implica la transferencia de conocimientos y destrezas a nuevas situaciones	Dominios de contenidos  Dominios cognitivos	Números	Naturales Fracciones y decimales Razón, proporción y porcentaje Números enteros	Conocimiento de hechos y procedimientos  Utilización de conceptos	Recordar Reconocer Calcular Usar herramientas
			Álgebra	Patrones Ecuaciones y fórmulas Relaciones Expresiones algebraicas	Resolución de problemas	Saber Clasificar Representar Formular Distinguir
			Medición	Atributos y unidades Herramientas, técnicas y formas	Razonamiento	Seleccionar Representar Interpretar Aplicar Verificar o comprobar
			Geometría	Lineas y ángulos Formas bidimensionales y tridimensionales Congruencia y similitud Localizaciones y relaciones espaciales Simetría y transformación		Formular hipótesis Analizar Evaluar Generalizar Conectar Sintetizar Resolver problemas Justificar
			Datos	Recopilación y organización Representación Interpretación Incertidumbre y Probabilidad		



PRUEBA	DEFINICIÓN	DIMENSIONES	CAMPOS	TEMAS	PROCESOS COGNITIVOS	DESTREZAS IMPLÍCITAS
PISA	Capacidad de identificar, comprender y practicar las matemáticas	Grandes ideas  Competencias	Contenidos  Procesos y procedimientos (Competencias)  Situaciones o contextos	Cantidades Espacios y formas Cambios Relaciones e incertidumbre  Reproducción  Conexiones  Reflexión  Tareas auténticas	Cálculos simples  Reunión de ideas y procedimientos  Razonamiento matemático Generalización  Problemas donde se empleen las matemáticas	Recuerdo Desarrollo de procedimientos Manipulación de expresiones  Manejo de métodos Distinguir y relacionar Interpretar Argumentar Diseñar Plantear Decodificar  Análisis Interpretación Desarrollo de modelos Demostración generalización

Al comparar la organización de las pruebas del TIMSS y del PISA, encontramos que ambas incluyen en su evaluación dos dimensiones: Contenidos y procesos cognitivos. Si hablamos de contenidos, notamos que las pruebas del TIMSS, tienen una mayor demanda de éstos, mientras que las pruebas del PISA, están centradas más en la aplicación de conocimientos básicos a situaciones concretas; sin embargo tenemos que las cantidades, el espacio y forma y el manejo de la incertidumbre son temas comunes. En cuanto a procesos cognitivos



se refiere, el ejercicio de procedimientos y el razonamiento, en la resolución de problemas, son las destrezas intelectuales más importantes que se demandan en los estudiantes.

### **b. Pruebas de ciencias: TIMSS Y PISA**

Al igual que en las matemáticas, los contenidos de las pruebas de ciencias del TIMSS para los alumnos que están por concluir la educación básica, son los mismos que para los de la población uno (cuarto grado, 8 y 9 años); sólo se hace una adecuación en las áreas y el nivel de los objetivos.

Las áreas que se incluyen para los alumnos que están por concluir la educación básica son:

- En *ciencias de la vida* : Las células y sus funciones
- En *química*: La estructura de partículas de la materia y ácidos y bases
- En *física*: sonido y vibración
- En *ciencias medioambientales* : Cambios de la población

Los dominios cognitivos, son los mismos que para cuarto grado.

En PISA el marco de evaluación de la aptitud científica, comprende tanto la formación científica, como la conceptual y procedimental y considera a la formación científica como la combinación del conocimiento científico con los procesos con los que se desarrolla y la capacidad o aptitud científica se concibe como "la capacidad para emplear el conocimiento científico, para identificar preguntas, y extraer conclusiones basadas en evidencias con el fin de comprender y apoyar la toma de decisiones acerca de la naturaleza y los cambios que se realizan a través de la actividad humana" (OCDE:2002:25).

Aborda tres campos temáticos con sus respectivas áreas de aplicación:

- *Ciencias de la vida y la salud*. Incluye: Salud, enfermedad y nutrición mantenimiento y uso sostenible de las especies, interdependencia de los sistemas físicos biológicos.
- *Ciencias de la tierra y del medio ambiente*. Se refiere a : efusión, producción y pérdida de suelo, tiempo y clima.



*Ciencias de la tecnología.* Comprende: biotecnología, empleo de materiales y recogida de los residuos, utilización de la energía y el transporte (OCDE, 2000:99-104).

La organización de la prueba incluye tres dominios:<sup>70</sup>

1. *Conceptos científicos.* Se trata de ideas generales integradoras, que dan sentido a los aspectos del mundo natural y tecnológico. Algunos de los temas seleccionados para evaluar estos conceptos son:

- Estructura y propiedades de la materia
- Cambio atmosférico
- Cambios físicos y químicos
- Transformaciones en la energía
- Fuerza y movimiento
- Función y forma
- Biología humana
- Cambio fisiológico
- Biodiversidad
- Cuadro genético
- Ecosistemas
- La tierra y su lugar en el universo
- Cambio geológico

Estos fueron seleccionados conforme a la aplicación que pudieran tener a situaciones cotidianas, la relevancia de su manejo para toda la vida, la aplicabilidad de las situaciones, y la factibilidad de poder combinar conceptos con procesos científicos.

2. *Procesos científicos:* Comprenden acciones mentales empleadas en la concepción, obtención, interpretación y uso de las pruebas o datos para obtener conocimiento o comprensión. Los procesos que se evalúan son los siguientes:

---

<sup>70</sup> El término dominio no es utilizado precisamente en el PISA. Dada la similitud de lo que comprende, aquí se utiliza con fines de comparación.



- *Reconocimiento de cuestiones científicamente investigables* Implica: identificación de los tipos de preguntas que la ciencia puede intentar responder y la distinción de los tipos de preguntas que pueden responderse mediante la investigación científica. La forma de evaluarse reside en la presentación de una situación con cuestiones que sólo pueden responderse científicamente, las cuales deberán ser identificadas por el estudiante.
- *Identificación de las evidencias en una investigación.* Implica la ubicación de evidencias para responder a preguntas planteadas, así como de la información necesaria para probar de manera válida una determinada idea y de los elementos que se van a comparar, de las variables que deberán cambiar y de la información que se necesita. Para esto se presenta una investigación en donde los estudiantes deberán identificar la evidencia o el procedimiento que hace falta para obtener la información deseada.
- *Extracción de evaluación de conclusiones obtenidas en una investigación.* requiere: relacionar las conclusiones con la evidencia que las fundamenta, elaborar una conclusión a partir de unas pruebas o datos científicos, seleccionar una conclusión entre varias, elaborar argumentos a favor o en contra de una conclusión, de acuerdo con los datos que se obtengan. Por lo que se proporciona a los estudiantes el informe de una investigación con sus conclusiones para que emitan una valoración de éstas y a partir de aquí deduzcan una o varias conclusiones a manera de alternativa.
- *Comunicar conclusiones válidas.* Implica valorar las evidencias que se dan, elaborar argumentos de acuerdo con las situaciones y datos que aportan lo relevante de la información. Para esto se les presenta a los estudiantes una situación en la que se requiere reunir información y datos de fuentes diferentes para apoyar un determinado proceder o conclusión.
- *Demostrar la comprensión de conocimientos científicos.* Requiere la comprensión y aplicación de éstos a situaciones distintas, para lo cual se



necesita no sólo recordar el conocimiento, sino saber la importancia del mismo para hacer predicciones o dar explicaciones, explicar las relaciones y las causas de los posibles cambios, hacer predicciones como resultados de los posibles cambios, al identificar los factores que influyen en determinados resultados. Y la forma de evaluarse es a través de las explicaciones o predicciones emitidas por los estudiantes.

3. *Situaciones científicas.* Referidas como contexto o entorno de las tareas.

En el siguiente cuadro conjuntamos con fines de comparación, la organización de las pruebas de ciencias del TIMSS y PISA, para alumnos que están por concluir la educación básica.

CUADRO 1.6

**PRUEBAS DE CIENCIAS  
PARA ALUMNOS QUE ESTÁN POR CONCLUIR LA EDUCACIÓN BÁSICA:  
TIMSS Y PISA**

PRUEBA	CONCEPTO	CAMPOS TEMÁTICOS	ÁREAS	DOMINIOS	CONCEPTUALIZACIÓN	DEFINICIÓN
TIMSS	Proceso utilizado para aprender sobre el mundo físico	Dominios de contenidos	Tipos, características y clasificación de los seres vivos Estructura, función y procesos vitales, Desarrollo y ciclos de vida de los organismos Reproducción y herencia Diversidad, adaptación y selección natural Ecosistemas La salud humana Las células y sus Funciones	Cognitivos	Hechos Datos Herramientas Procedimientos	Comportamientos
		Ciencias de la vida		Conocimiento factual		Comportamientos
				Comprensión conceptual	Captar relaciones que explican el comportamiento del mundo físico y relacionar lo observable con conceptos científicos	Ilustrar con ejemplos Comparar, contrastar o clasificar Representar o modificar Relacionar Obtener o aplicar información Hallar soluciones Explicar



*Exámenes masivos*  
*Un análisis entre pruebas internacionales y nacionales*

---

		Química	Clasificación y composición de la materia Propiedades y usos del agua El cambio químico Estructura de las partículas de las materias Ácidos bases			
		Física	Estados físicos y cambios en la materia Tipos, fuentes y conversiones de energía Calor y temperatura Luz Electricidad y magnetismo Fuerzas y movimiento Sonido y vibración	Razonamiento y análisis	Acciones implícitas tanto en la resolución de problemas científicos como en el desarrollo de explicaciones y elaboración de conclusiones para tomar decisiones	Analizar interpretar Resolver problemas Integrar o sintetizar Formular hipótesis Diseñar o planificar Recopilar, analizar e interpretar datos Sacar conclusiones Generalizar Evaluar Justificar
		Cambios medioambientales	Utilización y conservación de los recursos medioambientales Cambios en el medio ambiente Cambios en la población			



*Exámenes masivos*  
*Un análisis entre pruebas internacionales y nacionales*

PISA	Capacidad para emplear el conocimiento científico, identificar preguntas, y extraer conclusiones basadas en evidencias con el fin de comprender y apoyar la toma de decisiones acerca de la naturaleza y los cambios que se realizan a través de la actividad humana	Ciencias de la vida	Salud Enfermedad Nutrición Mantenimiento Usos sostenibles De las especies Interdependencia de los sistemas físicos y biológicos	Conceptos científicos	Ideas generales integradoras que dan sentido a los aspectos del mundo natural y tecnológico	Temas: Materia Cambio atmosférico Cambios físicos y químicos Transformaciones en la energía Fuerza y movimiento Función y forma Biología humana Cambio fisiológico Biodiversidad Cuadro genético Ecosistemas La tierra Cambio geológico
		Ciencias de la tierra	Efusión Producción Pérdida de suelo Tiempo y clima	Procesos científicos		
		Ciencias de la tecnología	Biotecnología Empelo de Materiales Recogida de residuos Utilización de la energía y el transporte.			
				Situaciones científicas	Acciones mentales empleadas en la concepción, obtención, interpretación y uso de las pruebas	Reconocimiento de cuestiones científicamente investigables Identificación de las evidencias en una investigación Extracción de evaluación de conclusiones obtenidas en una investigación Comunicar conclusiones válidas Demostrar la comprensión de conocimientos científicos
					Contexto o entorno de las tareas	

Hasta aquí, se puede decir que en cuanto al concepto de ciencia, las pruebas del TIMSS, manejan más un concepto de ciencia como explicación, con cierto corte experimental; mientras que PISA, la concibe sí como una comprensión del mundo natural, pero para la aplicación e intervención; lo que nos habla de un enfoque más funcional.



En síntesis, las pruebas de ciencias para alumnos que están por concluir la educación básica manejan a la ciencia como aclaración y aplicación. En cuanto a los campos temáticos que incluyen, tenemos que *Ciencias de la vida*, y *factores ambientales* es un área común en ambas pruebas; sin embargo las pruebas del TIMSS, quizá por su corte experimental incluyen contenidos de química y física que PISA no contempla, porque los temas que incluye, son con un corte más funcional y práctico. Vemos en las dos pruebas el manejo tanto de conocimientos factuales (conceptos científicos), pero además del uso de procesos cognitivos, que por la orientación de cada prueba en TIMSS, van más hacia la comprensión y el razonamiento y en PISA, se centran en empleo e interpretación. Algo importante, en las pruebas PISA, se incluyen además como dominio las situaciones científicas, que TIMSS, no contempla. Finalmente queremos destacar, que dado lo hasta aquí presentado, queda claro que la forma de organizar cada prueba, mantiene una estrecha relación con el concepto que se maneja del área evaluada.

### c. Pruebas de lectura: PISA

La aptitud para la lectura o capacidad de lectura en PISA, va más allá de la simple decodificación y transcripción literal, ésta se conceptualiza como "la capacidad de comprender, emplear y reflexionar sobre textos escritos con el fin de lograr las metas individuales, desarrollar el conocimiento y potencial y participar eficazmente en la sociedad". (OCDE, 2002:22). Es decir, se concibe a la capacidad lectora, como el uso personal y público que se le da a la lectura. Lo que implica demostrar capacidad para emplear información escrita con fines personales y utilizar esta información para desempeñarse eficazmente.

Esta capacidad o aptitud para la lectura, se evalúa a través de tres dimensiones:

- La forma y estructura del material de lectura o del texto
- El tipo de reactivo de la lectura
- El propósito y uso para el cual se elaboró el texto



La forma y estructura del material de lectura del texto, se refiere a la inclusión de diferentes tipos de textos: narrativos, expositivos, literarios, científicos, publicitarios e informativos, por ejemplo.

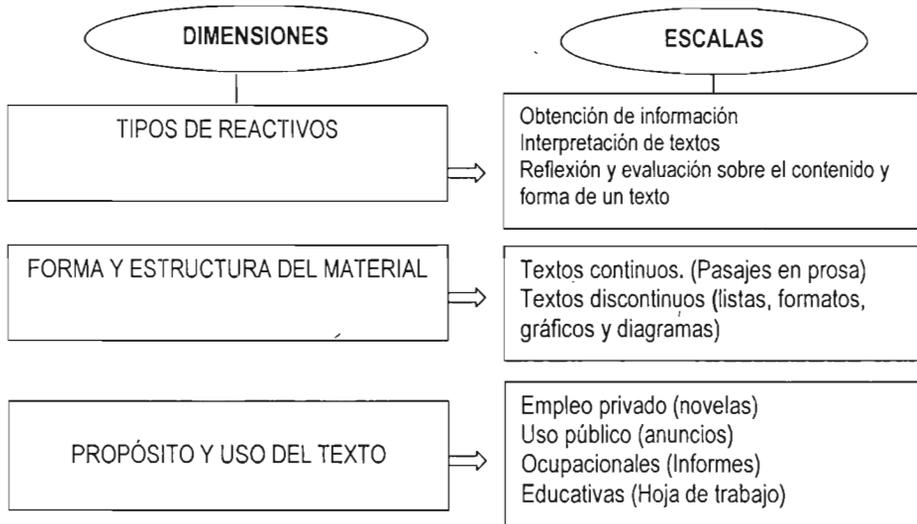
El tipo de reactivo de la lectura, comprende aptitudes cognoscitivas que se requieren para saber utilizar adecuadamente la lectura.<sup>71</sup>

El propósito y uso de la elaboración del texto, versa en torno al contexto o situación de los textos.

Cada una de estas dimensiones se mide con base en escalas o niveles de dominio.<sup>72</sup>

El siguiente esquema presenta las dimensiones y escalas<sup>73</sup> que se utilizan para la evaluación de la capacidad lectora en las pruebas del PISA (OCDE, 2002: 36-38).

### LAS TRES DIMENSIONES DE LECTURA Y SUS ESCALAS DE EVALUACIÓN EN PISA



<sup>71</sup> El lema para la evaluación de la lectura en PISA es "leer para aprender", más que "aprender a leer". (OCDE, 2000: 23).

<sup>72</sup> En la parte de sistema de calificación, presentamos los niveles de dominio establecidos por las pruebas del PISA.

<sup>73</sup> Nótese que aquí el término escala tiene una connotación diferente a la empleada en la estadística y es utilizado como un técnica semiformal de evaluación. Según Díaz Barriga Arceo y Hernández: "Podemos definir las escalas como instrumentos que permiten establecer estimaciones cualitativas dentro de un continuo sobre ejercicios o productos realizados por los alumnos" ( 2002: 394).



Destacamos que el marco para la evaluación de las pruebas PISA, contempla la definición de cada una de estas escalas:

"La *obtención de información* se define como la ubicación de uno o más fragmentos de información en el texto. La *interpretación de textos* se define como la construcción de significados y la generación de inferencias a partir de una o más secciones de un texto. La *reflexión y evaluación* se define como la capacidad de relacionar un texto con la experiencia, los conocimientos y las ideas propias" (OCDE, 2002:38).

Los tipos de texto, base de la organización de las pruebas de lectura de PISA, comprende tanto *textos continuos* como *textos discontinuos*. Se llama *textos continuos* a los textos compuestos por oraciones, que a su vez están organizados en párrafos. Algunos textos considerados como continuos son: Descripciones, narraciones, exposiciones, argumentaciones, instrucciones y textos unidos a apoyos visuales, a modo que las unidades puedan ser leídas en distintos órdenes y no necesitan del uso de estrategias lineales por parte del lector, que es a los que comúnmente se les llama "hipertextos".<sup>74</sup> Los *textos discontinuos* son los que se presentan como textos compuestos por listas, que pueden ir desde las listas sencillas, hasta las combinadas, las de intersección las intercaladas y las de combinación. Considerando que los textos discontinuos pueden ser clasificados con base en su formato, los formatos de textos discontinuos que se contemplan en PISA son: *formularios, hojas informativas, vales y cupones, certificados, avisos y anuncios, cuadros y gráficos, diagramas, tablas y matrices, listas y mapas*.

El *empleo privado de la lectura* se refiere a cuando la lectura se realiza para satisfacer un interés propio tanto práctico como intelectual, además de la lectura que se realiza para mantener o desarrollar las conexiones personales. Algunos ejemplos de textos considerados en este rubro son: cartas personales, narraciones de ficción, biografías y textos de carácter informativo, leídos por curiosidad o como parte de una actividad de ocio. La *lectura para uso público* es la que se efectúa con fines sociales, comprende tanto documentos oficiales como

---

<sup>74</sup> Se conoce como hipertextos a los textos que pueden ser objeto de varias lecturas.



de información. La lectura *para el trabajo u ocupacional* es la que cumple con el objetivo de "leer para hacer". Y las *lecturas educativas*, son las que se realizan para cumplir una tarea de aprendizaje (OCDE, 2000: 44-51).<sup>75</sup>

La evaluación de dichas escalas se realiza por medio de tareas, incluidas cinco aspectos de la lectura que cubren distintos niveles de conocimiento. Estos son: (OCDE, 2000:52-58).

1. Lograr una comprensión global
2. Obtener y recuperar información
3. Elaborar una interpretación
4. Reflexionar sobre el contenido del texto
5. Reflexionar sobre la forma del texto

*Lograr una comprensión global*, consiste en considerar una perspectiva amplia del texto en su conjunto. Algunas de las tareas requeridas en este aspecto consisten en identificar el tema o mensaje principal, algunos ejemplos de este tipo de tareas: Seleccionar o elaborar un título o una reflexión escrita para el texto, explicar el orden de instrucciones simples, identificar las principales dimensiones de un gráfico, describir el personaje principal, escenario o ambiente de una historia, identificar el tema o mensaje de un texto literario, o bien explicar el objetivo de un mapa, cuadro, mensaje de correo electrónico, o página de Internet.

La *obtención de información*, se trata de la obtención de fragmentos aislados de información, para lo cual se requiere repasar y buscar en el texto, para localizar y seleccionar la información relevante, mediante una comprensión inmediata. Las tareas incluidas en este aspecto son: Detectar o identificar los elementos esenciales de un mensaje, seleccionar información explícita, como una referencia de tiempo o lugar en una tabla.

La *elaboración de una interpretación* requiere enlazar información entre las distintas partes y centrarse en detalles específicos como partes del conjunto. Entre las tareas comprendidas para evaluar este aspecto se encuentran: comparar y contrastar información, deducir la

---

<sup>75</sup> Los reactivos muestra para estos tipos de lectura en sus diferentes niveles de dificultad, se pueden encontrar en (OECD, 2000 a, b y d; OCDE 2001a y 2002 b).



intención del autor, identificar las pruebas que sustenten esta intención y la deducción del significado a partir del contexto.

La *reflexión sobre el contenido de un texto* se realiza conectando información del texto con conocimientos precedentes de otras fuentes y evaluar las afirmaciones del texto, contrastándolas con su propio conocimiento: Las tareas que se encuentran dentro de este aspecto residen en: la aportación de pruebas o argumentos externos al texto y la evaluación de la importancia de fragmentos de información, para lo cual se le pide al alumno que identifique fragmentos alternativos de información que sirvan para reforzar la postura del autor, o que evalúe la validez de las pruebas que aporta el texto.

La *reflexión sobre la forma de un texto*, consiste en tomar distancia del texto para realizar una consideración objetiva y una evaluación de su calidad y conveniencia. Aquí se requiere una evaluación crítica y una apreciación de la pertinencia y alcance de ciertas características textuales, como la ironía, el humor y la organización lógica. Las tareas implícitas en este texto se refieren a: determinar la utilidad de un texto para un fin específico y evaluar el empleo por parte del autor de rasgos textuales, identifiquen el estilo del texto y las actitudes o propósitos del autor.

El siguiente cuadro ayuda vincular de manera gráfica los aspectos evaluados con las tareas requeridas.



CUADRO 1.7

**ASPECTOS Y TAREAS PARA LA EVALUACIÓN DE LA LECTURA EN PISA**

ASPECTOS	TAREAS
Comprensión global	<ul style="list-style-type: none"><li>• Seleccionar o elaborar un título o una reflexión escrita para el texto.</li><li>• Explicar el orden de instrucciones simples.</li><li>• Identificar las principales dimensiones de un gráfico. Describir el personaje principal, escenario o ambiente de una historia.</li><li>• Identificar el tema o mensaje de un texto literario.</li><li>• Explicar el objetivo de un mapa, cuadro, mensaje de correo electrónico, o del Internet.</li></ul>
Obtención de información	<ul style="list-style-type: none"><li>• Identificar los elementos esenciales de un mensaje.</li><li>• Seleccionar información explícita, como una referencia de tiempo o lugar en una tabla.</li></ul>
Elaboración de una interpretación	<ul style="list-style-type: none"><li>• Comparar y contrastar información.</li><li>• Deducir la intención del autor.</li><li>• Identificar las pruebas que sustenten una intención.</li><li>• Deducir el significado a partir del contexto.</li></ul>
Reflexión sobre el contenido de un texto	<ul style="list-style-type: none"><li>• Aportación de pruebas o argumentos externos al texto.</li><li>• Evaluación de la importancia de fragmentos de información.</li><li>• Selección de fragmentos alternativos de información que sirvan para reforzar la postura del autor.</li><li>• Evaluación de la validez de las pruebas que aporta el texto.</li></ul>
Reflexión sobre la forma de un texto.	<ul style="list-style-type: none"><li>• Determinar la utilidad de un texto para un fin específico.</li><li>• Evaluar el empleo por parte del autor de rasgos textuales.</li><li>• Identificar el estilo del texto y las actitudes o propósitos del autor.</li></ul>

La desagregación de los distintos aspectos en tareas, permite conocer a qué se refiere y qué es lo que se evalúa en cada aspecto y la distribución jerárquica de éstos conforme a la dificultad de los procesos mentales requeridos para su ejecución.

Las tareas, incluidas en cada uno de los ítems, se determinan conforme a diferentes niveles de dificultad y desempeño.<sup>76</sup>

El siguiente cuadro presenta las tareas jerarquizadas por niveles de dificultad en las escalas para evaluar la lectura de la prueba del PISA.

<sup>76</sup> "La división de escalas en niveles de dificultad y desempeño hace posible no sólo catalogar el desempeño de los estudiantes, sino también describir lo que son capaces de hacer"(OCDE, 2002:37 ).



CUADRO 1.8

<b>NIVEL DE DIFICULTAD DE LAS ESCALAS DE LA PRUEBA DE LECTURA EN PISA, CONFORME AL TIPO DE REACTIVO. Fuente: (OCDE, 2002:38)</b>			
<b>N</b>	<b>OBTENCIÓN DE INFORMACIÓN</b>	<b>INTERPRETACIÓN DE TEXTOS</b>	<b>REFLEXIÓN Y EVALUACIÓN</b>
5	Ubicar, ordenar y combinar fragmentos anidados profundamente y que se encuentran fuera del texto. Tratar con información extensa.	Interpretar	Emitir juicios e hipótesis, apoyándonos en un conocimiento especializado y en el entendimiento de textos largos
4	Ubicar y ordenar fragmentos de acuerdo a varios criterios y con contextos y formas no familiares	significado del lenguaje y demostrar una comprensión total y detallada	Emplear conocimientos formales o públicos en la comprensión precisa de un texto
3	Ubicar y reconocer la relación entre fragmentos	Interpretar significados de una sección en relación a todo el texto.	Realizar conexiones, dar explicaciones con base a características
2	Ubicar fragmentos de información		Realizar comparaciones en relación al conocimiento exterior
1	Tomar en cuenta un solo criterio para ubicar fragmentos independientes expresados explícitamente	Identificar , una idea central y significado de frases o palabras	Realizar conexiones simples entre información y conocimiento común

En este cuadro vemos cómo las mismas tareas van creciendo en grado de complejidad, en cada una de las tres escalas establecidas para cada tipo de reactivo.

Cabe mencionar que el conocer los niveles de desempeño establecidos para cada una de las escalas de evaluación, permite identificar que es lo que puede hacer y qué es lo que no puede hacer un alumno de acuerdo a los rangos establecidos.

A manera de síntesis se elabora un cuadro que recoge la manera en que el PISA organiza su prueba de lectura; lo que permite establecer las relaciones correspondientes y tener una perspectiva general de qué es lo que el PISA demanda en materia de lectura de los jóvenes que están concluyendo la educación básica.



CUADRO 1.9

**LA PRUEBA DE LECTURA DEL PISA**

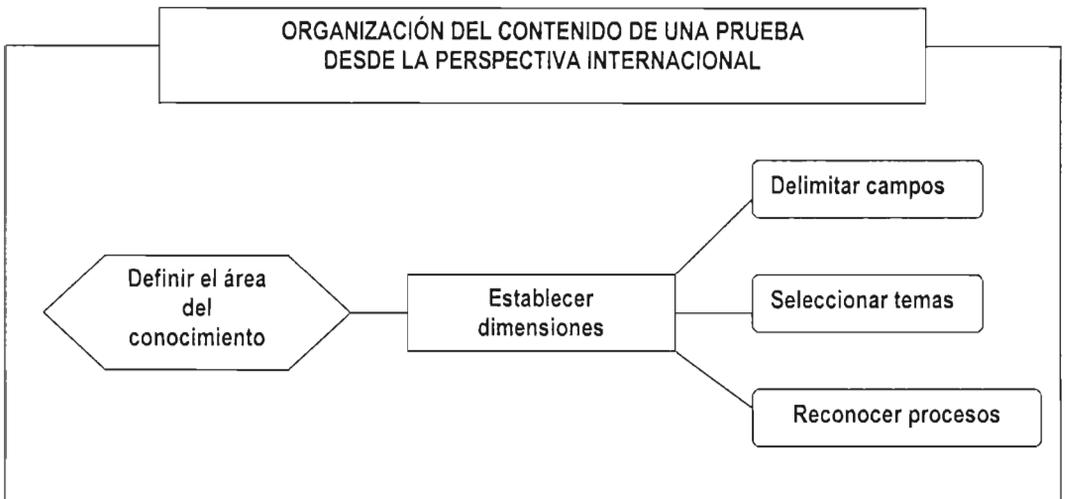
CONCEPTO	DIMENSIONES	ESCALAS	ASPECTOS	TAREAS
Capacidad de comprender, emplear y reflexionar sobre textos escritos con el fin de lograr las metas individuales, desarrollar el conocimiento y potencial y participar eficazmente en la sociedad	La forma y estructura del material de lectura o del texto  El tipo de reactivo de la lectura  El propósito y uso para el cual se elaboró el texto	Textos continuos Textos discontinuos  Obtención de información Interpretación de textos Reflexión y evaluación sobre el contenido y forma de un texto  Empleo privado Uso público Ocupacionales Educativas	Lograr una comprensión global  Obtener información  Elaborar una Interpretación  Reflexionar sobre el contenido del texto  Reflexionar sobre la forma del texto	Seleccionar Elaborar Explicar Identificar Describir Comparar Deducir Aportar Evaluar Determinar

La forma en como PISA organiza las pruebas de lectura presenta una muy particular forma de concebir y estructurar su evaluación. Los elementos considerados para definir y medir aptitud para la lectura, dan cuenta de que leer para el proyecto PISA no es acumular información, sino que implica comprender y usar la información de una lectura, mediante la activación de procesos mentales de orden superior como la interpretación y reflexión. Nótese que los términos comprensión y reflexión están incluidos tanto en el concepto como en los aspectos y las tareas incluidas en éstas tienen una estrecha relación con dichas operaciones intelectuales. Además las escalas y dimensiones, tienen una relación estrecha entre sí: Por ejemplo, para evaluar la forma y estructura, hay que utilizar distintos tipos de texto; los ítems deben ser de distinto tipo conforme a las tres escalas establecidas; el propósito y uso del texto considera diferentes ámbitos; además de que tanto la forma, el tipo y usos de los textos utilizados; responden al concepto de lectura establecido por el proyecto PISA.



- Hasta aquí, se pueden distinguir algunos de los elementos<sup>77</sup> que las agencias internacionales consideran para la organización del contenido de sus pruebas:
- Definir el área
- Establecer dimensiones
- Delimitar campos
- Seleccionar temas
- Reconocer procesos (llámense destrezas o comportamientos) implícitos:

El siguiente esquema, representa la organización de dichos elementos:



Si comparamos estos elementos con los señalados en la teoría de la medida para la elaboración de una prueba psicológica, tenemos que aunque no hay evidencias de un seguimiento directo; lo que hemos encontrado en el análisis de las Pruebas Internacionales como factores a considerar para la organización de su contenido, correspondería en cierta forma a uno de los pasos necesarios que Thorndike y Hagen (1978:47) indican para la planeación de una prueba y al primero de los elementos a considerar en el proceso de

<sup>77</sup> Cabe aclarar que a pesar de que en cada prueba se incluye el empleo de términos diversos; todas comprenden aspectos comunes que se pueden tomar como base.



medición;<sup>78</sup> lo que en sí, ya representa una aportación importante, no tanto en materia de medición y elaboración de una prueba, porque en sí retoma uno de los primeros pasos, sino más bien de la definición de éste. Cabe destacar además, que la especificación que se hace del atributo u objeto, no se refiere a lo estipulado en los planes y programas de estudio, pues a pesar de que retoma y selecciona temas generales del mismo, las Pruebas Internacionales, cuentan con una conceptualización propia del área de conocimiento, lo que en sí de acuerdo con Díaz Barriga (2004), induce a pensar en el desarrollo de una Concepto de contenido en los exámenes internacionales.

Hasta aquí, tres aportaciones importantes en materia de contenido, en las Pruebas Internacionales:

- ✓ La primera consiste en una definición propia del contenido a evaluar.
- ✓ La segunda se refiere a la inclusión en su contenido de evaluación, no sólo de temas propios del currículo escolar, sino de los procesos cognitivos implícitos en esos contenidos.
- ✓ La tercera reside en una desagregación de los elementos a considerar para definir el área de conocimiento.

Una vez analizado el contenido de una prueba desde la perspectiva internacional, cabe ahora reparar en la estructura que éstas guardan.

#### 1.4 Estructura de las Pruebas Internacionales: TIMSS, LLECE y PISA

Para presentar la estructura que las pruebas internacionales contienen, consideramos las normas y procedimientos para la elaboración de las pruebas, el diseño de los ítems en cuanto a composición, constitución del formato y el modelo estadístico de medición que los

---

<sup>78</sup> Desde la postura clásica, el proceso de medición contempla tres pasos a seguir: Definición de la cualidad o atributo que habrá de medirse, determinar las *operaciones* mediante las cuales el atributo pueda manifestarse y *cuantificación* del atributo (Thorndike y Hagen, 1978:17).



sustenta; así como las particularidades de las tareas<sup>79</sup> requeridas en las áreas de evaluación que comprenden.

**a. Normas y procedimientos para la elaboración de las pruebas.**

**Normas técnicas de las pruebas del TIMSS:**

La elaboración de las pruebas del TIMSS responde a ciertas normas técnicas, (véase IEA, 1999) entre las que se encuentran:

- Contar con un armazón que incorpore las características centrales de los estudios de la IEA y que se desarrolle un juego de normas de tipo criterial. El armazón o modelo conceptual de los instrumentos debe definir el dominio del área que se evaluará en el estudio, las estructuras subyacentes y los procedimientos, e incluir las habilidades que los estudiantes deben manejar para conectar conocimiento y procesos de cada disciplina; especificando las dimensiones, áreas, temas y subtópicos dentro del dominio, es decir, qué es lo que los estudiantes deben saber y poder hacer en cada una de las áreas sujetas a evaluación; así como el propósito para el que se deben utilizar los resultados de las pruebas y encuestas.
- Tener una estructura diferente para cada país.
- Los instrumentos son traducidos en los idiomas correspondientes a cada país; pero es importante mencionar que estas traducciones son verificadas para asegurar que el significado y nivel de dificultad no cambien la idea a nivel internacional.
- En las normas técnicas se deberá incluir información sobre las tareas que los participantes deben realizar en las pruebas.
- Describir los equipos especiales que serán incluidos tales como: calculadoras, computadoras, equipos de ciencias, libros y folletos.

---

<sup>79</sup> En materia de evaluación, llamamos tarea a las actividades que se les pide a los alumnos realicen y que incluye aspectos cognitivos de regulación.



- Descripción de las características psicométricas deseadas en las pruebas y encuestas, lo que incluye el nivel de dificultad de la prueba.
- Todas las pruebas deben incluir la definición del dominio a evaluar.
- Los artículos individuales y tareas deben ser apropiados al propósito de la prueba y a la población que la responderá.
- Los niveles de dificultad deben estar acordes con los propósitos de las pruebas.
- Los ítems deben estar formulados de manera clara y apropiada para cada población y cuidar el nivel gramatical y sensibilidad cultural.
- Las traducciones de las pruebas se deberán realizar bajo normas muy precisas; de tal manera que las versiones internacionales, no cambien el significado o dificultad, pues la meta es obtener calidad en la traducción de instrumentos para proporcionar datos comparables en países y culturas. Por lo que se debe cuidar que la traducción de los ítems de la prueba, permanezca con la intención de la versión internacional en cuanto a significado de la pregunta, nivel de lectura del texto, dificultad del artículo y probabilidad de respuestas correctas. Es así como la traducción se realiza mediante una identificación de las diferencias culturales, una búsqueda de las palabras equivalentes y frases, un aseguramiento de que el nivel de lectura es el mismo que el internacional y una certeza de que el significado esencial de los artículos es el mismo.
- Tener un procesamiento, análisis e información de los datos obtenidos para lograr los objetivos del estudio, ya que una de las normas de la IEA radica en que sus estudios son valiosos sólo en la medida en que los resultados son ampliamente conocidos y entendidos, pues el propósito es asegurar que los lectores de informes de la IEA, cuentan con todos los requisitos para tomar las decisiones pertinentes. Por lo que se deberá contar con un plan de publicación que satisfaga las necesidades de los lectores.



## Elaboración de las pruebas del LLECE

En el caso de las pruebas del LLECE, para garantizar que los ítems de las pruebas fueran significativos y tuvieran una fuerte vinculación con los currículos de los países participantes se invirtieron aproximadamente dos años y medio en la elaboración de pruebas en donde se realizaron reuniones para seleccionarlos, de acuerdo al estado de avance de la disciplina, la relevancia de ciertas competencias en relación a su utilidad para la vida y los estándares internacionalmente reconocidos; de ahí que se pidieran "Coordinadores generales" de cada uno de los países para que cada miembro representante de éste aportara al menos 10 ítems de cada una de las áreas a evaluar y se formara así la matriz curricular. Los contenidos incluidos en las pruebas debían representar al menos al menos un 80% de los currículos de lenguaje y matemáticas de cada uno de los países participantes.

Una vez que las pruebas fueron aceptadas por todos los Coordinadores Nacionales, éstas fueron certificadas por evaluadores externos del Educational Testing Service

La elaboración de las pruebas del Lenguaje y Matemáticas aplicadas por el Laboratorio se realizó de la siguiente manera (véase LLECE, 1998:16-17):

1. Análisis de los documentos que permitieran conocer lo que se enseña en los cuatro primeros años del primer ciclo de educación básica
2. Construcción de una matriz de objetivos curriculares
3. Aprobación de dicha matriz por parte de los países participantes
4. Validación de los instrumentos a partir de una aplicación piloto
5. Presentación de resultados



## Procedimiento para elaborar las pruebas del PISA

Las pruebas del PISA, se elaboran de acuerdo al siguiente procedimiento (Véase OCDE, 2002a: 264):

1. Elaboración de los marcos conceptuales de referencia y especificaciones de evaluación de cada área.
2. Desarrollo de una definición funcional del ámbito y la descripción de los supuestos subyacentes a esa definición.
3. Operacionalización de la definición elaborada, con base en tres dimensiones: Contenidos, procesos y situaciones.
4. Elaboración por parte de comités específicos del contenido de las pruebas.
5. Aportación por parte de los países participantes de preguntas revisadas por expertos en el área.
6. Organización del conjunto de ítems.
7. Identificación del conjunto de características clave para la construcción de los ítems.
8. Aplicaciones piloto.
9. Validación de las variables y evaluación por parte de cada una de las partes de la dificultad del reactivo y de la aceptabilidad cultural y relevancia curricular.
10. Traducción de los instrumentos al idioma de cada país.
11. Preparación de un esquema de interpretación de resultados.

Considerando que la especificación del contenido que ha de abarcarse y la preparación del esquema de ejecución de la prueba, constituyen dos de los pasos que Thorndike y Hagen (1978) delimita como parte de la planeación de una prueba,<sup>80</sup> se puede decir que los exámenes de PISA, consideran estos pasos en su elaboración, pero le dan un tratamiento

---

<sup>80</sup> "El contenido y el enunciado de los objetivos del proceso representan las dos dimensiones en las que debe acomodarse el plano de una prueba" (Thorndike y Hagen, 1978:47 ).



diferente al no limitarse sólo a señalar el contenido conforme a los programas establecidos o preparar el esquema de ejecución de la prueba, enunciando de los objetivos curriculares el proceso que representan; sino que especifican y desarrollan una definición propia del área y delimitan el contenido y procesos cognoscitivos a evaluar no sólo conforme a lo estipulado en el currículo, sino con base en el consenso entre especialistas y en la definición que hacen del contenido.

El siguiente cuadro sintetiza los principales elementos a considerar en la elaboración de una prueba desde la mirada internacional.

CUADRO 1. 10

**PROCESO DE ELABORACIÓN DE LAS PRUEBAS INTERNACIONALES**

	TIMSS	LLECE	PISA
Análisis de los planes y programas de estudio de cada país participante		x	
Construcción de una matriz de objetivos curriculares	x	x	
Definición del dominio especificando las dimensiones, áreas, temas y subtópicos dentro del dominio	x	x	x
Operacionalización de la definición	x	x	x
Información sobre las tareas que los participantes deben realizar	x	x	x
Inclusión y definición de conocimientos y habilidades	x		x
Descripción de características psicométricas	x	x	x
Aportación de ítems por los países participantes		x	x
Traducción y estructura propia según el país	x		x
Aplicación de una prueba piloto		x	x
Preparación de un esquema para la interpretación de resultados	x	x	x

De acuerdo con lo presentado en el cuadro, tenemos que aunque cada estudio difiere en los elementos que privilegia en su proceso de elaboración: el TIMSS está centrado más en el aspecto curricular, técnico y conceptual, el LLECE en la unificación de contenidos a través de estándares y el PISA en la elaboración de ítems contextuales.<sup>81</sup> Encontramos

<sup>81</sup> Se trata de una evaluación auténtica, caracterizada según Herman, Aschbacher y Winters (1992), en "Demandar que los aprendices resuelvan activamente tareas complejas y auténticas, mientras usan sus conocimientos previos, el aprendizaje reciente y las actividades relevantes para la solución de problemas reales" (cit. en: Díaz Barriga Arceo y Hernández, 2002:365 ).



como elementos comunes en la elaboración de las pruebas: la definición del dominio y las tareas, la operacionalización de esta definición, la descripción de las características psicométricas y la presentación de resultados, conforme a un esquema y procedimientos específicos; mismos que coinciden con lo especificado en la teoría clásica de la medida<sup>82</sup> pero las aportaciones de las Agencias Internacionales en cuanto a elaboración de pruebas, reside, como hemos venido mencionando, en no partir de objetivos netamente curriculares, sobre todo en las pruebas del PISA y del LLECE y en no sólo retomar lo que los planes y programas de estudio establecen; sino en presentar una definición propia del contenido a evaluar, hacer una selección de los temas básicos de los planes de estudio conforme a un consenso internacional y a su relación con lo que según especialistas son los más importantes del área de acuerdo con la definición elaborada, operacionalizar sus ítems no sólo con base en contenidos, sino de los procesos cognitivos que se requiere para el manejo de esos contenidos, e incluir en las tareas, situaciones auténticas.

#### **b. Diseño de ítems en las pruebas internacionales: TIMSS, LLECE y PISA**

##### **Ítems de las pruebas del TIMSS**

En las pruebas del TIMSS, los ítems están compuestos por contenidos concretos que a su vez requieren de la aplicación de conocimientos y destrezas cognitivas de diferentes áreas a las que se les llama comportamientos.

La constitución de sus ítems, depende del área, grado y tipo de rendimiento que deben mostrar los estudiantes, pero básicamente los ítems están constituidos con un formato de

---

<sup>82</sup> Thomdike y Hagen (1978: 43-60), establecen como pasos para la elaboración de una prueba definición de objetivos, especificación del contenido que ha de abarcarse, preparación del esquema de ejecución de la prueba, determinación del tipo o tipo de ítems que ha de usarse y determinación del nivel adecuado de dificultad de los ítems.



opción múltiple,<sup>83</sup> aunque también incluyen preguntas de respuesta corta<sup>84</sup> y abierta en donde el alumno deberá explicar el proceso que siguió para llegar a la respuesta; el fin es evaluar los conocimientos y destrezas para explicar fenómenos e interpretar datos. En este tipo de preguntas, se pide a los estudiantes, la redacción por escrito de su respuesta con el propósito de que expliquen fenómenos o interpreten datos sobre la base de su experiencia y conocimiento.<sup>85</sup>

Las pruebas del TIMSS, recurren al modelo estadístico de Teoría de Respuesta al ítem (TRI)<sup>86</sup> para reportar los resultados de los estudiantes de manera individual y ofrecer

---

<sup>83</sup> Cabe mencionar que las dos terceras partes del número total de puntos en las pruebas del TIMSS, están representados por preguntas de elección u opción múltiple. En este tipo de preguntas se ofrece a los estudiantes cuatro o cinco opciones de respuesta, de las cuales sólo una es la correcta y pueden medir, cualquiera de los rendimientos en los dominios cognitivos; sin embargo está claro que éstas no son muy apropiadas para evaluar la capacidad de los estudiantes para realizar interpretaciones.

Para la elaboración de reactivos de opción múltiple, existe una amplia variedad de documentos, que determinan ciertos lineamientos. (Véase Haladyna, Haladyna, T. M., Haladyna, R. Merino, C. (s/f).

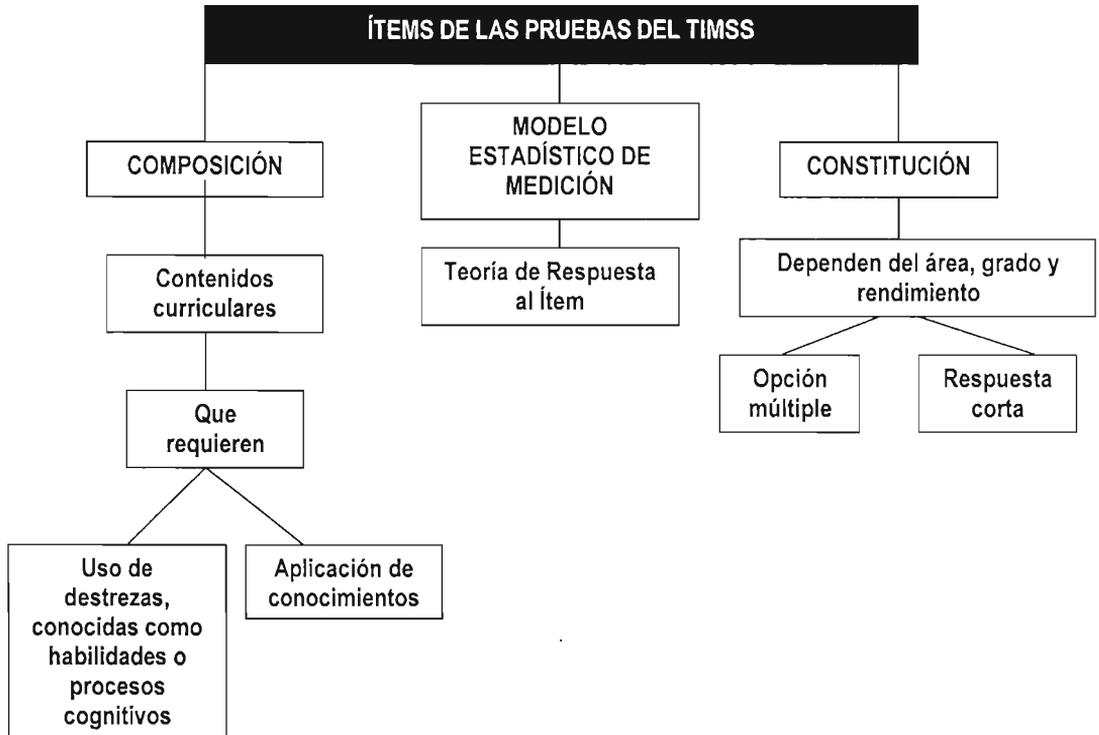
<sup>84</sup> Las preguntas de opción múltiple, los ejercicios de relación de conceptos, los enunciados verdaderos/falsos y las preguntas de respuesta corta o de ensayo, son tipos de ítems, comúnmente conocidos como "prueba objetiva". Esta supuesta objetividad en la calificación de los exámenes es cuestionada por Díaz Barriga (1988, 1993 y 1997) por varias razones, entre las que destacan: la dimensión de las relaciones que en ésta se establecen; la anulación de una concepción de evaluación y por ende de aprendizaje y docencia; el reduccionismo cuantitativo en el que se ve inmersa y el desconocimiento de la complementariedad necesaria entre subjetividad y objetividad en todo proceso de conocimiento.

<sup>85</sup> Como se sabe estas preguntas ofrecen a los estudiantes cuatro o cinco opciones de respuesta, de las cuales sólo una es la correcta, y pueden medir, cualquiera de los rendimientos en los dominios cognitivos; sin embargo está claro que estas no son muy apropiadas para evaluar la capacidad de los estudiantes para realizar interpretaciones.

<sup>86</sup> Hablar de Teoría de Respuesta al ítem, (TRI) es hablar de un tratamiento matemático alternativo dentro de la psicometría que plantea ajustes estadísticos conforme a la dificultad de los ítems, más que a las de la prueba. (cfr. Hambleton, y Van der Linden, 1996). "El propio nombre de Teoría de Respuesta al Ítem alude a que bajo la óptica de la TRI, la unidad básica de análisis es el ítem y no el test. En la Teoría Clásica, el índice de dificultad o su índice de discriminación, índices con los que se describe a los ítems, depende de la muestra de personas utilizadas para estimarlo. En cambio, mediante la TRI, se pueden realizar mediciones invariantes, con respecto al instrumento y a las personas seleccionadas y es más resistente a las variaciones en cuanto al número de sujetos y reactivos, lo que le da mayor coeficiente de consistencia interna." "La denominación *Teoría de Respuesta a los ítems*" hace referencia a un conjunto de modelos matemáticos que asumen la existencia de una relación funcional, en principio entre la variable que miden los ítems de un test y las probabilidades de respuesta a éstos, es decir entre los distintos niveles de constructo, que se está midiendo y la probabilidad de responder adecuadamente a un ítem concreto. Esta relación se representa gráficamente en las Curvas características de los ítems (CCI)". (Martínez, y Moreno, 2002:103) En la TRI, "La dificultad y la discriminación...modifican la forma de la curva variando la relación entre el nivel de aptitud y la probabilidad de acertarlo. El índice de dificultad, es el valor de la aptitud. La dificultad determina la posición



información sobre el rendimiento en cada uno de los dominios cognitivos , en términos del porcentaje medio de los estudiantes que responden correctamente a los ítems de cada dominio (Mullis, Martin, Smith, Garden, Gregory, Gonzales, Chrostowsk y O Conoor, 2002).



de la CCI con respecto a la escala de aptitud, (eje X). Otro parámetro es la discriminación, que representa la pendiente o inclinación de la CCI." ( Arce, 2005) .



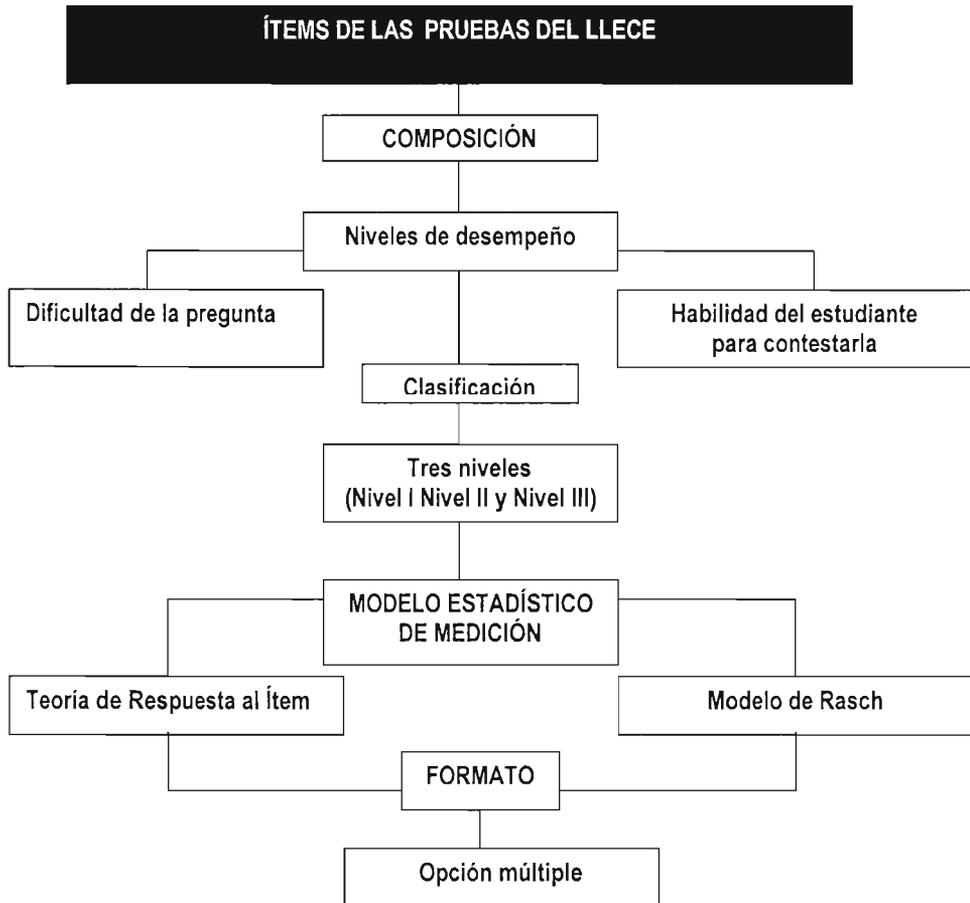
## Ítems de las pruebas del LLECE

En el LLECE, los ítems de las pruebas están compuestas por ítems establecidos conforme a diferentes niveles de desempeño que contemplan tanto la dificultad de la pregunta, como la habilidad de los alumnos para poder contestarla. Se encuentran clasificados en tres grupos que miden diferentes escalas, establecidos en orden ascendente (Nivel I, Nivel II y Nivel III) de manera inclusiva de acuerdo con los postulados de la teoría constructivista del aprendizaje; es decir, que si un alumno llega al nivel tres, es porque ha superado los otros dos.

La constitución de los ítems de la prueba, conserva un formato de opción múltiple y está fundamentado de manera general en la TRI y desde una visión más específica en el modelo de Rasch.<sup>87</sup>

---

<sup>87</sup> "El modelo de Rasch es un modelo probabilístico que permite estimar las probabilidades de respuesta de un sujeto ante un estímulo o reactivo dado" (Tristán, 1998:126). En un capítulo posterior hablaremos de manera más detallada sobre este modelo y su empleo en la metodología de la evaluación y medición.



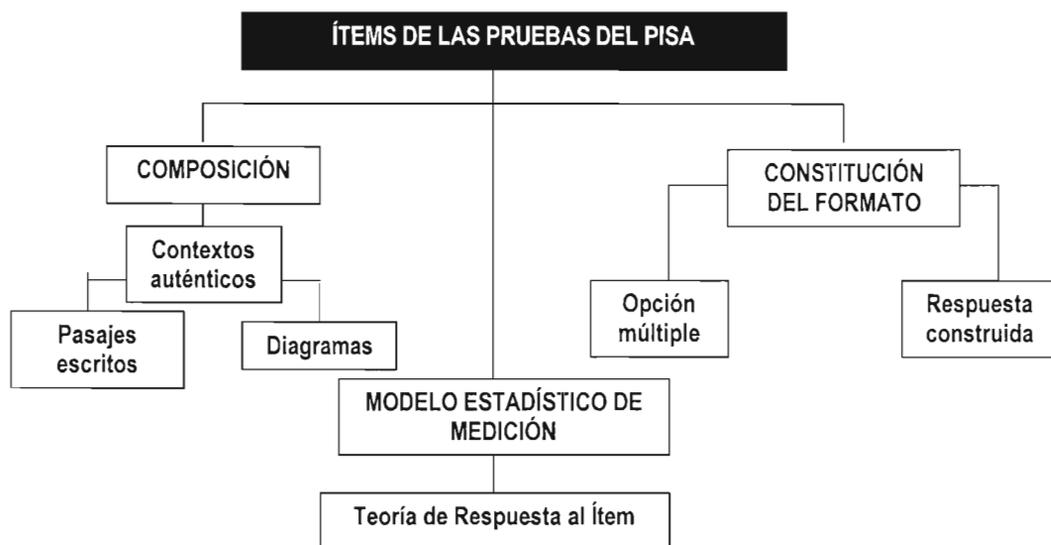
### Ítems de las pruebas del PISA

Los ítem del proyecto PISA, están compuestos por contextos auténticos. Es decir, por pasajes escritos que describen situaciones de la vida real. Los ítems están constituidos por varios rubros, consisten en preguntas que incluyen pasajes escritos y diagramas, a partir de los cuales los estudiantes tienen que contestar ciertas preguntas; todo esto con la finalidad



de conocer qué tan capaces son de reflexionar y pensar activamente sobre el tema presentado. De ahí que el formato de los ítems es tanto de opción múltiple como preguntas en donde se pide a los estudiantes construir sus propias respuestas. Cada una de las preguntas indica: proceso requerido, concepto científico y área de aplicación.

La dificultad de los ítems se definió de acuerdo a la complejidad de los conceptos empleados, la cantidad de datos proporcionados, el tipo y secuencia del razonamiento, el contexto de la información y la presentación de las preguntas; pero el modelo utilizado para la selección de los ítems también corresponde al de la Teoría de Respuesta al Ítem. (TRI)



En un esfuerzo por sintetizar y comparar el diseño de las pruebas internacionales se presentan los siguientes cuadros y puntualizaciones:



CUADRO 1. 11

**DISEÑO DE ÍTEMS EN LAS PRUEBAS INTERNACIONALES**

PRUEBA	COMPOSICIÓN DE LOS ÍTEMS	CONSTITUCIÓN DEL FORMATO	MODELO TEÓRICO
TIMSS	Contenidos concretos que demandan destrezas y conocimientos	Opción múltiple Respuesta abierta	Teoría de la Respuesta al Ítem
LLECE	Niveles de desempeño con base en dificultad de la pregunta y habilidades del estudiante	Opción múltiple	Modelo de Rasch Teoría de Respuesta al Ítem
PISA	Contextos auténticos Pasajes y diagramas	Opción múltiple Respuesta corta Respuesta construida	Teoría de respuesta al ítem

Hasta aquí podemos decir las pruebas internacionales presentan en su diseño elementos importantes a considerar:

- En cuanto a su composición, los ítems de las pruebas del TIMSS y LLECE, incluyen tanto el conocimiento de hechos, principios y procedimientos, como el ejercicio de procesos y habilidades cognitivas; las pruebas PISA, destacan por ítems formados por situaciones auténticas.
- La estructura de los ítems mantiene el formato de opción múltiple, pero dos de éstas pruebas (TIMSS y PISA), incluyen el de respuesta construida.
- El modelo de medición que fundamenta la selección de ítems de las tres pruebas es el de la Teoría de Respuesta al ítem.



**c. Características de las tareas en las Pruebas Internacionales: TIMSS, LLECE, PISA.**

**Tareas en las pruebas del TIMSS**

Las tareas de las pruebas del TIMSS, son específicas de las dos áreas de conocimiento que este estudio abarca: matemáticas y ciencias. En el caso de matemáticas, las tareas, residen en problemas de la vida real y problemas matemáticos que requieren de la aplicación de todo un conjunto de acciones, mediante el uso del razonamiento.

Los ítems de ciencias, elaborados para utilizar los dominios de contenidos y cognitivos propios de esta área, están encaminados a conducir a los estudiantes hacia el proceso de investigación científica.

Es importante mencionar que la IEA, después de cada periodo de evaluación, publica algunos de los ítems utilizados, lo que ocasiona que éstos sean sustituidos por nuevas preguntas en ciclos subsecuentes. Se procura que los nuevos ítems, demanden de los estudiantes, síntesis de conocimientos y destrezas de diferentes áreas. Las capacidades analíticas, de resolución de problemas y de investigación, son destrezas que se incluirán de manera especial en las pruebas del TIMSS a partir del 2003. Además, con el propósito de mantener actualizados los instrumentos, las pruebas del TIMSS pretenden incorporar tareas relacionadas con la investigación y la producción.

**Tareas en las pruebas del LLECE**

En las pruebas del LLECE la evaluación de las matemáticas, comprende 32 ítems de opción múltiple enfocadas a conocer la capacidad en el uso de procedimientos matemáticos. Las tareas implícitas en los ítems de las pruebas pretenden promover:



“Una praxis en la que se tenga la posibilidad de matematizar el mundo a partir de la modelación de situaciones con las que se entre en diálogo [Esto] implicará: interpretar datos, establecer relaciones y conexiones sobre la información encontrada en la situación, poner en juego conocimiento matemático, analizar regularidades, establecer patrones de cambio, buscar argumentos y justificaciones, proponer leyes de transformación o modelos para matematizar o predecir acontecimientos asociados a la situación estudiada”(LLECE, 2001:20).

Las pruebas del lenguaje en el LLECE, enfocadas a la comprensión textual; constan de un ejercicio de producción de texto, como tema a desarrollar y 19 ítems de opción múltiple, en donde se les pregunta a los alumnos sobre un texto complejo, en el que deberán identificar y reconocer información a nivel macro, meso y micro. Las preguntas se encuentran agrupadas por niveles de complejidad; con fines de una interpretación semántica que lleve a una interpretación semántico-crítica, de acuerdo a los niveles establecidos para tal efecto.

Un elemento central en el Diseño de la Prueba, es el *Diario Mural* (semejante a los que existen en los salones de clase), con esto se pretende evitar medir sólo los saberes gramaticales de los niños como verbos, adjetivos o sustantivos; lo que importa es conocer la capacidad de los alumnos para integrar un texto complejo en donde se encuentran tipos de textos, que requieren de diversos modos de leer. De ahí que el proceso de lectura<sup>88</sup> demande una serie de posibilidades interpretativas en donde el estudiante tendrá que poner en juego estrategias de pensamiento para construir o develar el significado de un texto. En

---

<sup>88</sup> La lectura en el LLECE, es definida como. “Proceso semiótico en el que entran en diálogo los saberes del lector y los del texto”. (LLECE, 2001:18).



los procesos de lectura que tiene que realizar el estudiante en las pruebas del lenguaje, se le demanda al lector trabajar sobre el discurso del texto en función de hipótesis globales de lectura para responder a preguntas tales como: ¿Qué dice el texto? y ¿Qué pienso yo sobre lo que dice el texto?

### **Tareas en las pruebas del PISA**

En las pruebas del PISA, las características de las tareas depende de la naturaleza y definición del área de conocimiento a evaluar: Lectura, Matemáticas y Ciencias.

En el caso de la lectura existe una distribución de tareas conforme a las situaciones, formatos, textos, aspectos y tipo de preguntas. (OCDE, 2000: 62-64)

En las situaciones de lectura, la distribución de las tareas se realizó conforme al siguiente criterio:

- 28% para la lectura personal
- 28% para la lectura educativa
- 16% para la lectura ocupacional
- 28% para la lectura pública.

En textos continuos la distribución se hace por tipo de texto:

- 20% narrativo
- 33% expositivo
- 20% descriptivo
- 20% argumentativo- descriptivo
- 7% mandatorio

En los textos discontinuos se siguió la siguiente distribución:

- Diagramas 33%
- Tablas 33%
- Esquemas 10%
- Mapas 10%



- Formularios 8%
- Anuncios 6%

En los aspectos de lectura que se evalúan, la distribución quedó de la siguiente manera:

- Recuperación de información 20%
- Comprensión global 20%
- Desarrollo de una interpretación 30%
- Reflexión sobre el contenido 15%
- Reflexión sobre la forma 15%

En la distribución de ítems por formato, destaca el alto porcentaje de ítems en donde se le pide al alumno que elabore su respuesta, empleado para la evaluación de la lectura; estamos ante un 45% de ítems respuesta construida y el 55 % restante es de opción múltiple. Nótese además que el formato de opción múltiple es utilizado con un mayor porcentaje para la recuperación de la información, la comprensión global y el desarrollo de una interpretación.

En las pruebas de matemáticas, las tareas de evaluación, están situadas en contextos<sup>89</sup> matemáticos de distinto tipo, obtenidos de fuentes diversas como: otras materias, áreas de la práctica profesional o vocacional, vida cotidiana, vida en comunidad y sociedad, y contextos de tiempo libre (deportes, juegos): Todo esto con el fin de demostrar los distintos usos de las matemáticas en la vida cotidiana. De tal manera que las tareas de evaluación, están situadas en *contextos auténticos*.<sup>90</sup>

Considerando el impacto que puede tener el formato de las tareas en el rendimiento del alumno, el proyecto OCDE/PISA, considera importante, considerar la definición del

---

<sup>89</sup> Para el proyecto OCDE-PISA, un contexto matemático es el entorno extra-matemático o intra-matemático dentro del cual se interpretan los elementos de un *complejo matemático* (por ejemplo, un problema, una tarea o una colección de objetos matemáticos, relaciones, fenómenos, etc.). Un contexto es tanto el entorno en el cual un complejo matemático dado ya se ha establecido (entorno intra-matemático), como un entorno que se presta a la activación de tal complejo y que entonces se establece en ese contexto (entorno extra-matemático) (OCDE, 2000: 85).

<sup>90</sup> "Se considera un contexto auténtico, aquel que reside en experiencias y prácticas reales en el entorno del mundo real por parte de los participantes(...) Preguntas acerca del rendimiento de unos ahorros guardados en un banco con un interés real, pueden ser auténticas" (OCDE, 2000: 86).



constructo para la selección del formato de los ítems. En el caso de las matemáticas, éstas son evaluadas con una combinación de preguntas de opción múltiple,<sup>91</sup> de tipo cerrado y de tipo abierto. Para objetivos de mayor nivel lo más recomendable es utilizar preguntas de tipo cerrado. Las preguntas de tipo abierto, requieren procesos de orden superior, y son preguntas en donde no sólo se les pide a los alumnos que den una respuesta; sino que más bien expliquen el cómo llegó a esa respuesta. “La característica importante de las preguntas de tipo abierto es que permiten que los alumnos demuestren sus habilidades mediante la aportación de soluciones en una serie de niveles de complejidad matemática.” (OCDE, 2000:89) En las pruebas de matemáticas del PISA, un buen porcentaje de las preguntas (del 25% al 35%) son de tipo abierto.

Para operacionalizar la evaluación de las competencias matemáticas mediante la construcción de preguntas, el proyecto OCDE/PISA, organiza las destrezas implícitas en las competencias matemáticas en tres grandes grupos:

1. Reproducción de definiciones y cálculos
2. Conexiones e integración para la resolución de problemas
3. Pensamiento matemático, generalización y comprensión súbita.

Con base en esta clasificación distribuye el número de preguntas y puntos designados para evaluar las *grandes ideas*; asignado el mayor número de preguntas a la clase uno y dos, es decir a la reproducción de definiciones y cálculos y a la conexión e integración para la resolución de problemas, en temas de cambio y crecimiento y espacio y forma.

En las pruebas de ciencias, las tareas incluidas en las preguntas, evalúan más de un proceso y concepto científico. Las tareas se configuran lo más realísticamente posible con el fin de reflejar en cierta forma la complejidad de vida diaria. Se trata de tareas extensas en donde se incluyen aproximadamente ocho preguntas que requieren de la comprensión y el conocimiento de los conceptos en cuestión. En las pruebas del PISA, presentan problemas

---

<sup>91</sup> Para la conformación de las preguntas de opción múltiple, el proyecto OCDE/PISA, tomando en cuenta recomendaciones internacionales, considera que: “Cabe emplear el formato de elección múltiple, pero sólo en casos contados y únicamente para los objetivos y conductas más bajos y los resultados del aprendizaje” (OCDE, 2000: 87).



concretos, que demandan la ejecución de procesos y del conocimiento científico para la lectura de algún texto, cuadro o diagrama que muestre la aplicación de conceptos científicos a situaciones de la vida real. (Véase reactivos muestra (OCDE; 2001a, y 2002b, OECD, 2000a, 2002d)

Para la evaluación de las situaciones científicas, se indican fenómenos del mundo real en el que puedan aplicarse las ciencias; lo que implica la realización de tareas en donde se encuadren situaciones de vida que incluyan problemas que puedan afectarnos. Las preguntas se refieren a cuestiones relacionadas con uno mismo, con la familia y la comunidad; así como con la vida en el mundo y su importancia histórica. La distribución de preguntas, se realiza de manera equitativa y las tareas implícitas, se refieren a las cuatro situaciones contempladas: personales, comunitarias, globales e históricas; las cuales están compuestas por fragmentos escritos de textos breves, acompañados de una tabla, esquema o gráfico de un diagrama. Se trata de preguntas independientes que requieren de una comprensión de los conceptos científicos, mediante la aplicación de estos en las distintas áreas de evaluación.

Una de las características más importantes de las pruebas PISA, es que son tareas extensas que evalúan uno o más de los procesos científicos, algunas de las preguntas incluyen contenidos de lectura y matemáticas, contribuyendo así a la evaluación de las áreas. Algo significativo es que no existen preguntas que requieran sólo del recuerdo o repetición puntual de información aislada.

De acuerdo a la distribución de preguntas realizadas para cada uno de los procesos requeridos, las pruebas del PISA, consideran que tanto la comprensión conceptual como la combinación del conocimiento científico, con la capacidad de sacar conclusiones basadas en evidencias; es lo que demuestra un aprendizaje científico.

Las pruebas del proyecto OCDE/PISA, fueron distribuidas a los estudiantes a través de nueve folletos que incluían preguntas de matemáticas, lectura y ciencias, pero no todos los folletos evaluaban los mismos dominios. El folleto 1, 3 y 5 contenía cuestionamientos de matemáticas, los folletos 2,4 y 6 contenían cuestionamientos de ciencia, el folleto 7 de



lectura y el 8 y 9 de matemáticas y ciencias. La cantidad de ítems de cada uno de los cuadernillos o folletos variaba de 55 a 67, dependiendo de la carga de lectura.<sup>92</sup>

A manera de síntesis, presentamos en el siguiente cuadro las características más sobresalientes de las tareas implícitas en las pruebas internacionales.

CUADRO 1. 12

**CARACTERÍSTICAS DE LAS TAREAS DE LAS PRUEBAS INTERNACIONALES**

<b>P</b>	<b>MATEMÁTICAS</b>	<b>CIENCIAS</b>	<b>LENGUAJE Y LECTURA</b>
<b>TIMSS</b>	Problemas de la vida real Problemas matemáticos	Uso de dominios de contenidos y cognitivos para conducir al estudiante al proceso de investigación científica.	Área no evaluada
<b>LLECE</b>	Con fines de mate matización	Área no evaluada	Comprensión e interpretación semántica
<b>PISA</b>	Contextos auténticos	Fenómenos reales en los que puedan aplicarse procesos y conceptos científicos	Conforme a las situaciones, formatos, textos, aspectos y tipo de preguntas

De acuerdo con lo presentado en el cuadro, tenemos que tanto las pruebas del TIMSS, como del PISA y del LLECE, incluyen en sus diferentes áreas de evaluación tareas propias de la disciplina, que demandan la aplicación del conocimiento a situaciones reales.

Una vez definida el contenido y estructura de una prueba, es importante dejar claro, el procedimiento de calificación y los puntajes que se asignarán a los diferentes ítems; esto permite contar con los elementos necesarios para un análisis e interpretación de resultados.

<sup>92</sup> "En total en la evaluación PISA 2000 se utilizaron 141 reactivos en aptitud para la lectura, 32 reactivos en aptitud para las matemáticas, y 32 reactivos en aptitud para las ciencias". Cada reactivo fue distribuido en los diferentes cuadernillos, con el fin de garantizar que una muestra representativa de estudiantes contestara cada uno de los reactivos (OCDE 2002b: 20).



#### d. Sistema de calificación y puntuación de las pruebas internacionales.

Para determinar el rendimiento de los estudiantes, las pruebas internacionales establecen dentro de su marco teórico los criterios y procedimientos de calificación y fijan los puntajes para los ítems; éstos se determinan de acuerdo a la naturaleza de la tarea, área y nivel establecido.

#### Procedimiento de calificación: TIMSS, LLECE y PISA

En el TIMSS la calificación de los resultados se realiza conforme al formato de los ítems. Para las preguntas de respuesta abierta o construida, se establecen las características esenciales de las respuestas apropiadas, para lo cual existe una *guía codificada* que señala el tipo de comportamiento que la pregunta evalúa y describe el tipo de respuesta parcial o totalmente correcta, lo que permite que cada ítem cuente con la identificación de diversas aproximaciones de éxito total, parcial o nulo. Estas guías contienen además una muestra de respuestas en cada nivel de comprensión que sirven como pauta para quienes califican las respuestas de los estudiantes.

Al considerar que uno de los objetivos importantes del estudio es diagnosticar las dificultades de aprendizaje más comunes en matemáticas y ciencias, las cuales quedan en evidencia a través de los errores y conceptos equivocados de los estudiantes, las preguntas de respuesta construida, constituyen una parte importante de la evaluación y forman parte de la medición de tendencias, por lo cual es muy importante que se apliquen consistentemente las guías de puntuación en todos los países y en cada ciclo de recopilación de datos.<sup>93</sup>

---

<sup>93</sup> La IEA ha archivado muestras de respuestas dadas por estudiantes de cada país, con el fin de garantizar la aplicación consistente de estas guías. (Mullis, Martin, Smith, Garden, Gregory, Gonzales, Chrostowski, y O'Conor, 2002:94).



En las pruebas del LLECE, la corrección de las pruebas, sobre todo para los ítems referentes a la "Producción de texto", se hizo mediante la entrega de *pautas guía* a profesores contratados especialmente para tal efecto.

La calificación e interpretación de los resultados se realiza conforme a un criterio,<sup>94</sup> previamente determinado, que representa las competencias alcanzadas por los alumnos de acuerdo a los niveles de desempeño<sup>95</sup> y estándares preestablecidos.

El procedimiento de calificación de Pruebas Internacionales como las del LLECE, corresponde a una nueva tendencia en evaluación surgida a partir de los años sesenta, que se conoce como pruebas referidas a un criterio<sup>96</sup> y en el análisis de las tareas<sup>97</sup> que requiere la conducta en cuestión; lo que las ubica en un paradigma cognoscitivo, en oposición al conductual; pues permite conocer lo que un estudiante puede o no puede realizar en relación a un punto de referencia y a los procesos cognoscitivos que se requieren para realizar una tarea o subtarea.

---

<sup>94</sup> La evaluación referida a criterio, surge con los artículos publicados por Glaser, (Véase Glaser 1963, cit. en Castañeda, S., Lugo, E. Pineda, L. y Romero, N., 1998:50). Se le llama pruebas criteriosales a las pruebas en donde se establecen con anterioridad, los resultados que deben alcanzar los estudiantes. El establecimiento de criterios para evaluar el aprendizaje, surge como una necesidad de responder a las nuevas orientaciones en materia de evaluación: la evaluación por objetivos de aprendizaje y como manera de contrarrestar la evaluación referida a la norma de principios de siglo, cuya calificación se asignaba de acuerdo a la posición que ocupaba un sujeto en relación a las puntuaciones obtenidas por el grupo de referencia (García, 1993:275).

<sup>95</sup> El vincular la educación con el desempeño, es una tendencia que se manifiesta fuertemente en los países de América Latina a partir de la década de los ochenta como descripción de los requisitos de calidad. Dice Eisner: Las pruebas de desempeño intentan trasladarse desde las prácticas evaluativas que exigen a los estudiantes seleccionar la única respuesta correcta entre una serie de cuatro o cinco distractores hacia una práctica que requiere que los alumnos generen evidencia, mediante su propio desempeño, que permita a los examinadores realizar juicios válidos sobre "lo que ellos saben y pueden hacer" en situaciones que importan (Eisner, 1999:3).

<sup>96</sup> "El modelo de evaluación referida a un criterio propone evaluar más bien el desempeño individual en relación a criterios específicos, que en relación al desempeño de otros individuos" (García, 1983:273).

<sup>97</sup> "El análisis cognitivo de tareas es un análisis recursivo que descompone una tarea determinada en aquellos componentes considerados como esenciales para llevar a cabo cada una de las subtareas que la constituyen(...)Le interesa conocer con mayor precisión los factores que componen una tarea global y estar en capacidad de identificar en qué paso específico de la tarea el estudiante puede tener problemas (...)Este tipo de análisis añade el elemento cognitivo, es decir, incluye la identificación de los procesos, las estructuras y las estrategias psicológicas que subyacen a cada subtarea componente." (Castañeda, Lugo, Pineda, y Romero, 1998:30).



La calificación con base en criterios y/o desempeños, denota la adopción de un paradigma evaluativo diferente a la norma, para calificar el aprendizaje y valorar lo educativo.<sup>98</sup> Se trata de exámenes que se califican con un enfoque interpretativo, en relación a logros alcanzados de acuerdo con un dominio y que actúan como forma de contrarrestar los sistemas de calificación unificados de acuerdo a normas y medidas estadísticas. Dice L. Mabry (1999), “las prácticas de las pruebas de desempeño que ofrecen oportunidades para respuestas abiertas y que permiten a los jóvenes demostrar sus fortalezas, chocan contra los supuestos sobre la uniformidad” (cit. en Eisner, 1999: 5).

Del empleo de pruebas estandarizadas que se calificaban con base en los puntajes alcanzados por el grupo de referencia; las pruebas internacionales recurren al uso de estándares relacionados con el ejercicio de habilidades cognitivas y con el dominio de ciertas tareas relacionadas con el aprendizaje.<sup>99</sup>

Aunque en páginas anteriores presentamos la descripción de los niveles de desempeño de las Pruebas del LLECE en las dos áreas que evalúa: Matemáticas y Lenguaje, mostramos ahora los desempeños incluidos en cada nivel, mismos que han sido empleados para la calificación y presentación de resultados.

---

<sup>98</sup> Una metodología para elaborar exámenes criteriosales, pero con referencia a un currículo, se puede encontrar en: Contreras, y Backhoff, (2004). Actualmente también existen sistemas y *software* para calificar pruebas referidas a criterios, conocido como *Kalt-Criterial*, desarrollado desde 1997, (véase Tristán, 2004).

<sup>99</sup> Un análisis de los componentes y competencias para la selección de pruebas de habilidades complejas, lo podemos encontrar en: Ferrer y Díaz (2004).



CUADRO 1.13

**NIVELES DE DESEMPEÑO EN MATEMÁTICAS Y LENGUAJE ESTABLECIDOS EN LAS PRUEBAS DEL LLECE PARA LA CALIFICACIÓN DE LOS RESULTADOS**

Nivel	Matemáticas	Lenguaje
III	<p>Reconocimiento y uso de estructuras matemáticas complejas que implica: Abordar situaciones que requieran tanto procedimientos usuales como más elaborados que implica:</p> <ul style="list-style-type: none"><li>▪ Abordar problemas que impliquen reconocer la estructura del sistema de numeración decimal y del manejo del valor posicional para establecer equivalencias.</li></ul>	<p>Lectura de carácter inferencial que implica:</p> <ul style="list-style-type: none"><li>▪ Explicitación de supuestos sobre los que está estructurado el texto.</li><li>▪ Vinculación de proposiciones a nivel micro y macrotextual.</li><li>▪ Identificación de distintas formas de relaciones implícitas en el texto.</li></ul>
II	<p>Reconocimiento y uso de estructuras matemáticas simples que implica:</p> <ul style="list-style-type: none"><li>▪ Realización tanto de ejercicios rutinarios del aula como situaciones matemáticas simples que requieran realizarlas cuatro operaciones básicas</li></ul>	<p>Lectura de carácter literal que implica:</p> <ul style="list-style-type: none"><li>▪ Traducción de las palabras reguladas por el sentido literal del texto</li></ul>
I	<p>Reconocimiento y empleo de hechos y relaciones matemáticas básicas, que implica:</p> <ul style="list-style-type: none"><li>▪ Reconocimiento de estructuras explícitas de nivel local como: Identificar los actores de un relato, los fragmentos claves en la argumentación, y las relaciones explícitamente planteadas.</li></ul>	<p>Lectura literal primaria que implica:</p> <ul style="list-style-type: none"><li>▪ Reconocimiento de estructuras explícitas de nivel local, mediante la identificación de los actores de un relato, los fragmentos claves en la argumentación y las relaciones explícitamente planteadas.</li></ul>

Adaptado de: LLECE, 2001:28

En las pruebas del PISA, además de desempeños, los criterios que se utilizan para evaluar las preguntas residen en: correcto o incorrecto y en el caso de las preguntas abiertas: totalmente correcta, parcialmente correcta o incorrecta; para esto se establecen muestras de



respuestas y palabras claves que representan aspectos importantes a considerar en las respuestas de los estudiantes.

Las calificaciones se determinan conforme a una escala de desempeño<sup>100</sup> que representan niveles de dominio en cada una de las áreas de evaluación. Una característica importante de la calificación de las pruebas del PISA consiste en que:

“La división de las escalas en niveles de dificultad y desempeño hace posible no sólo catalogar el desempeño de los estudiantes, sino también describir lo que son capaces de hacer... Por ejemplo, una calificación baja en la escala de interpretación indica que el alumno tiene capacidad limitada en la comprensión de relaciones, en la construcción de significados o en la obtención de inferencias de una o más partes de un texto. En contraste, una calificación alta en la escala de interpretación indica que el estudiante tiene habilidades avanzadas en este aspecto” (OCDE:2002:37).

Los niveles de desempeños establecidos para la evaluación de la lectura son: <sup>101</sup>

---

<b>Nivel 5</b>	Capacidad para responder ítems de lectura sofisticada mediante: <ul style="list-style-type: none"><li>▪ Manejo de información difícil de encontrar en textos con los que no están familiarizados.</li><li>▪ Comprensión detallada de dichos textos.</li><li>▪ Inferencia de información relevante en el texto.</li><li>▪ Evaluación crítica y establecimiento de hipótesis.</li><li>▪ Empleo de conocimiento especializado e inclusión de conceptos que pueden ser contrarios a las expectativas.</li></ul>
<b>Nivel 4</b>	Capacidad para responder reactivos de lectura difíciles mediante: <ul style="list-style-type: none"><li>▪ Ubicación de información anidada.</li><li>▪ Interpretación de significados a partir de sutilezas del lenguaje.</li><li>▪ Evaluación crítica de un texto.</li></ul>

---

---

<sup>100</sup> De acuerdo con Díaz Barriga Arceo y Hernández, las escalas, “son instrumentos que permite establecer estimaciones cualitativas dentro de un continuo sobre ejecuciones o productos realizados por los alumnos”. Estas pueden ser cualitativas, basadas en criterios o estándares amplios o numéricas, que consiste en la asignación de valores numéricos a la escala establecida. (2002:394-95). Ante esto, cabe destacar que los sistemas de calificación del PISA, contemplan los dos tipos de escalas.

<sup>101</sup> Fuente: (OCDE, 2002:41-51)



---

<b>Nivel 3</b>	Capacidad para manipular reactivos de complejidad moderada mediante: <ul style="list-style-type: none"><li>▪ Ubicación de fragmentos múltiples de información.</li><li>▪ Vinculación de distintas partes de un texto, relacionándolo con conocimientos familiares cotidianos.</li></ul>
<b>Nivel 2</b>	Capacidad para responder reactivos básicos de lectura mediante: <ul style="list-style-type: none"><li>▪ Ubicación de información directa.</li><li>▪ Realización de inferencias sencillas de diverso tipo.</li><li>▪ Determinación de lo que significa una parte definida en el texto.</li><li>▪ Empleo de cierto nivel de conocimientos externos para comprenderla</li></ul>
<b>Nivel 1</b>	Incapacidad para realizar el tipo de lectura básico que el PISA pretende medir por: <ul style="list-style-type: none"><li>▪ Presentar dificultades para emplear la aptitud para la lectura, como una herramienta eficaz que les permita ampliar y aumentar sus conocimientos y destrezas en otras áreas.</li></ul>

---

Para el caso de las matemáticas existen tres niveles de desempeño, establecidos de acuerdo a un puntaje promedio de 500 puntos:<sup>102</sup>

---

<b>Nivel más alto</b>	Postura activa y creativa en su acercamiento a los problemas matemáticos que consiste en: <ul style="list-style-type: none"><li>▪ Interpretar y formular problemas en términos matemáticos.</li><li>▪ Capacidad para manejar información compleja y negociar una serie de pasos de procesamiento.</li><li>▪ Identificar y aplicar conocimientos y herramientas relevantes, en problemas con los que no están familiarizados.</li><li>▪ Empleo de la perspicacia para identificar maneras adecuadas de encontrar una solución.</li><li>▪ Demostración de procesos cognoscitivos de alto nivel como: la generalización, el razonamiento y la argumentación para explicar y comunicar resultados.</li></ul>
<b>Nivel medio</b>	Capacidad para interpretar, vincular e integrar distintas representaciones de un problema o diferentes fragmentos de información al: <ul style="list-style-type: none"><li>▪ Manipular o emplear un modelo dado que requiere el uso del álgebra u otras representaciones simbólicas.</li><li>▪ Verificar y revisar proposiciones o modelos dados.</li><li>▪ Seleccionar y aplicar el conocimiento matemático necesario para resolver un problema que puede requerir de un número pequeño de pasos de procesamiento.</li></ul>
<b>Nivel bajo</b>	Capacidad para completar sólo un paso de procesamiento, mediante: <ul style="list-style-type: none"><li>▪ La reproducción de elementos matemáticos básicos.</li></ul>

---

<sup>102</sup> Fuente: (OCDE, 2002:82)



- Aplicación de habilidades simples de cálculo.
  - Reconocimiento de información a partir de un material diagramático o de texto que es familiar y directo y en el cual se proporciona o aparece de manera clara la fórmula matemática.
  - Reconocimiento de un solo elemento familiar en un problema cuya solución requiere de la aplicación de procedimientos rutinarios en un solo paso de procesamiento.
- 

En las pruebas de ciencias el desempeño se califica con base en una sola escala promedio de 500 puntos, a partir de la cual se establecen tres escalas:<sup>103</sup>

---

<b>Escala superior</b>	<ul style="list-style-type: none"><li>▪ Creación o empleo de modelos conceptuales para hacer predicciones o dar explicaciones</li><li>▪ Análisis de investigaciones científicas con el objeto de comprender por ejemplo el diseño de un experimento o identificar una idea que está puesta a prueba</li><li>▪ Comparación de datos con el fin de evaluar puntos de vista alternativos o perspectivas distintas.</li><li>▪ Comunicación de argumentos científicos o descripciones detalladas y precisas.</li></ul>
<b>Escala media</b>	<ul style="list-style-type: none"><li>▪ Empleo de conceptos científicos para hacer predicciones o dar explicaciones</li><li>▪ Reconocimiento de preguntas que pueden ser respondidas mediante la investigación científica.</li><li>▪ Identificación de los detalles involucrados en una investigación científica</li><li>▪ Selección de información relevante entre datos que compiten o cadenas de razonamiento para obtener o evaluar conclusiones.</li></ul>
<b>Escala baja.</b>	<ul style="list-style-type: none"><li>▪ Recordar conocimientos científicos simples.</li><li>▪ Emplear el conocimiento científico común para obtener o evaluar conclusiones.</li></ul>

---

Una consideración importante, el conocer los desempeños establecidos en cada una de las niveles o escalas con los que se califican las pruebas PISA, nos ayuda a comprender que los resultados obtenidos por nuestro país, obedece en gran parte, a que lo que se enseña en

---

<sup>103</sup> Fuente: (OCDE, 2002:21)



nuestras escuelas, está centrada más en la reproducción y el recuerdo, más que en el desarrollo de procesos cognitivos superiores.

La descripción de los niveles de desempeño con base en criterios, que incorporan las Pruebas Internacionales, mediante los cuales se evalúan y califican los resultados obtenidos por los estudiantes, nos hace pensar en un tipo de evaluación auténtica<sup>104</sup> que junto con las rúbricas (véase Díaz Barriga Arceo y Hernández 2002 y Díaz Barriga Arceo, 2004) representa una nueva tendencia dentro de la evaluación contemporánea (Castañeda, 1998: 51), consistente en que a los alumnos se les pide que demuestren las habilidades y conocimientos adquiridos en un contexto real, como es el caso de las pruebas PISA.

### **Puntuación asignada a los tipos de ítems: TIMSS, LLECE y PISA**

En el TIMSS, se han creado bloques de ítems que establezcan una media de 15 puntos para la población de octavo grado (14 y 15 años) y 12 para la población de cuarto grado (8 y 9 años).

A los ítems de elección múltiple se les asigna un punto. Las preguntas de respuesta construida generalmente valen, 1, 2 o 3 puntos, dependiendo de la naturaleza de la tarea y de las destrezas necesarias para llevarla a cabo. La resolución de problemas extensos y los ítems de investigación pueden valer hasta 5 puntos, pues requieren que los estudiantes trabajen y manipulen distintos materiales.

El LLECE, se utilizan puntajes estandarizados, cuyo promedio es de 250 puntos y en donde los más bajos se sitúan en 65 puntos para lenguaje y 53 para matemáticas y los más altos se ubican en 397 para lenguaje y 492 para matemáticas.

---

<sup>104</sup> Sin embargo, dentro de las actuales tendencias en materia de evaluación educativa y de acuerdo con Díaz Barriga Arceo, y Hernández (2002:388), es importante hacer la distinción entre evaluación auténtica y evaluación por desempeño: "En la evaluación de desempeño se pide a los alumnos que demuestren ciertas habilidades, destrezas o conductas en una situación de prueba. Mientras que la evaluación auténtica demanda que los aprendices demuestren dichas habilidades, destrezas o conductas en situaciones de la vida real" El interés por la evaluación auténtica condujo al desarrollo de varios planteamientos nuevos basándose en la meta del desempeño en el contexto" (Woolfolk, 1996: 565).



El PISA establece puntajes para cada una de las áreas de evaluación, que representan diversos niveles de desempeño. Por ejemplo, para evaluar la aptitud para la lectura, existen tres escalas, que representan cinco niveles de conocimiento y destreza. El nivel 5 comprende una puntuación de más de 625 puntos, el nivel 4 contempla un puntaje entre 553 y 625, el nivel 3 se ubica entre 481 y 552, el nivel 2 se asigna con 408 y 480 y el nivel 1 se refiere con 335 y 407 puntos. En matemáticas se establece una sola escala con un puntaje promedio de 500 puntos. Las puntuaciones del nivel más alto llegan hasta 750 puntos, las del nivel medio se ubican en 570 puntos y las de nivel bajo contemplan 380 puntos. Para ciencias, la escala más alta corresponde a 690 puntos, la media a 550 puntos, la baja a 400 puntos.

En un esfuerzo por comparar los criterios de calificación y puntuación utilizados en las pruebas internacionales, presentamos el siguiente cuadro.

CUADRO 1. 14

**SISTEMA DE CALIFICACIÓN Y PUNTUACIÓN DE LAS PRUEBAS INTERNACIONALES**

	Calificación	Puntuación
<b>TIMSS</b>	Por <i>guía codificada</i> : señala el tipo de comportamiento que la pregunta evalúa y describe el tipo de respuesta parcial o totalmente correcta	15 puntos como media para 8° curso 12 puntos como media para 4° curso Dependiendo de la naturaleza de la tarea, se asigna uno, dos o tres puntos a las respuestas de respuesta construida.
<b>LLECE</b>	<i>Pautas guía</i> para la calificación del reactivo de producción de texto. Conforme a un <i>criterio</i> . Establecimiento y descripción de tres <i>niveles de desempeño</i> .	Puntajes estandarizados con un promedio de 250 puntos
<b>PISA</b>	Se establecen criterios de: <i>Correcto o incorrecto</i> para preguntas cerradas <i>Totalmente correcta, parcialmente correcta o incorrecta</i> , en el caso de las preguntas abiertas: <i>Muestras de respuestas y palabras claves</i> que	<i>Puntajes por niveles de desempeño</i> : Para lectura, existen tres escalas, que representan cinco niveles de conocimiento y destreza: Nivel 5 (625)



representan aspectos importantes a considerar en las respuestas de los estudiantes.	Nivel 4 (553 y 625)
Las calificaciones se determinan conforme a escalas de desempeño que representan niveles de dominio en cada una de las áreas de evaluación.	Nivel 3 (481 y 552)
	Nivel 2 (408 y 480)
	Nivel 1 (335 y 407).
	En matemáticas se establece una sola escala con un puntaje promedio de 500 puntos.
	Nivel más alto (750 puntos)
	Nivel medio (570 puntos)
	Nivel bajo (380 puntos)
	Para ciencias:
	Escala más alta (690 puntos)
	La media (550 puntos)
	La baja a (400 puntos)

---

Considerando que los rangos son medidas de dispersión que se obtienen con base en una distribución de puntajes que van desde el más alto al más bajo, es importante destacar todavía, el uso de medidas normalizadas para medir, comparar y reportar lo educativo.

El que se recurra todavía a escalas numéricas de intervalo<sup>105</sup> -según la clasificación realizada por Stevens,<sup>106</sup> ubica a estas pruebas, no sólo como instrumentos de evaluación, sino también de medición y lleva a reflexionar que aún en sistemas actuales de evaluación, la medición es un paradigma que sigue prevaleciendo hasta nuestros días. para asignar puntuaciones y presentar resultados. Sin embargo, queremos destacar que la presentación

---

<sup>105</sup> Se habla de escalas de intervalo por el uso de puntuaciones medias, en donde las distancias entre los pares de categorías o grados presentan una relación de igualdad (cfr. Camilloni, 1998:136).

<sup>106</sup> El uso de escalas en materia de medición, es abordado por Stevens en 1946, como una forma de asignar números a un conjunto de modalidades distintas y así cumplir con la regla del isomorfismo relacionando las propiedades de los objetos o sucesos a medir y las propiedades del sistema numérico. Algunas de estas propiedades y sus usos en la medición son: Identidad: Se refiere al uso de los números para identificar clases o artículos. Orden: Sirve para identificar un rango ordenado. Intervalos. Permite ubicar diferencias entre los artículos. Proporciones: Indica proporciones entre los artículos. Tales propiedades permiten asignar números para representar los aspectos del estado que guardan las cosas de acuerdo con ciertas operaciones, que pueden ser empíricas, pero que determinan el tipo de escala a utilizar. Con base en esto, Stevens plantea cuatro tipos de escalas; Nominal, Ordinal, Intervalos y Proporciones. La *nominal* implica el uso de operaciones para determinar la igualdad, mediante números de casos, modo, medidas de información y correlación de contingencias. En la *ordinal*, se utilizan operaciones para una determinación de mayor o menor, haciendo uso de la mediana, el tanto por ciento, la correlación de orden. En la de *intervalos*, las operaciones que se realizan determinan la igualdad de los intervalos o de las diferencias, a través de desviaciones estándar y correlaciones, un ejemplo típico de esta son los resultados estandarizados en las pruebas de inteligencia y en la de *proporción*, se llevan a cabo operaciones para determinar la igualdad de proporciones a través de medias geométricas, medias armónicas y variaciones de tanto por ciento (Stevens, 1959).



de los criterios de calificación y puntuación de las pruebas internacionales revela elementos importantes a considerar como los siguientes:

- La calificación de las preguntas de respuesta abierta o construida mediante guías, presenta una alternativa para inducir la aplicación y calificación de exámenes masivos no meramente objetivos.
- El que en las pruebas del LLECE y PISA se establezca una calificación por niveles de desempeño, da cuenta de que estamos ante un cambio de parámetros en materia de calificación de exámenes, en donde se sustituye al número que en realidad no dice nada, por el desempeño que permite darle significado al resultado obtenido y conocer en qué están bien o mal los estudiantes y cuál es el nivel alcanzado por éstos.
- En materia de puntuación se suprime el sistema decimal y se establecen rangos promedio, los cuales también se establecen de acuerdo al nivel de dificultad de las preguntas.

Sin embargo no podemos dejar de señalar que el que en Pruebas Internacionales recurran a rangos y escalas como sistemas de calificación y puntuación, habla del predominio de un paradigma cuantitativo en materia de evaluación y de la gran necesidad de encontrar un paradigma diferente al de la medición en materia de calificación y presentación de resultados, además de los aquí presentados.

Con el propósito de conocer los factores que influyen en los resultados obtenidos, las agencias internacionales de evaluación internacional realizan además, una evaluación contextual, de los llamados factores asociados al rendimiento.



## **1.5 Evaluación de factores asociados en los estudios internacionales: TIMSS, LLECE y PISA**

Una característica importante de los estudios de evaluación realizados por estas agencias internacionales, es la aplicación de cuestionarios de contexto para explicar los factores asociados al rendimiento y aprendizaje de los alumnos, los cuales contemplan diversos aspectos del entorno sociocultural y escolar de los estudiantes.

### **a. La evaluación de factores asociados en el TIMSS:**

Con el propósito de obtener una visión más general de lo que significan los resultados obtenidos en las pruebas y que elementos se deben considerar para mejorar el rendimiento de los estudiantes, el TIMSS 2003 pretende conocer los contextos en los que aprenden los estudiantes, a través de un marco contextual que comprende cinco áreas:

1. El currículum
2. Los centros educativos
3. Los profesores y su preparación
4. Las actividades del aula
5. Los estudiantes

En cuanto al currículum, lo que se pretende es conocer la organización, seguimiento y evolución del currículum aplicado, así como los materiales utilizados para facilitar la aplicación del mismo.

En la organización de los centros educativos, lo que se espera es establecer ciertos indicadores que permitan identificar la forma de aplicación de los objetivos curriculares en: tiempo asignado a la docencia y en específico a la instrucción en matemáticas y ciencias, objetivos de los cursos, funciones del director, recursos para apoyar el aprendizaje de las matemáticas y las ciencias, y participación de los padres de familia.



Por lo que respecta a los profesores y su preparación, la evaluación se refiere a los métodos y programas para la formación y actualización del profesorado, lo que incluye la participación de las universidades y centros educativos en dicha formación y el establecimiento de estándares de competencia, además de otros aspectos tales como: preparación académica, autorizaciones legales para ejercer la docencia, número de alumnos por cada profesor, forma de iniciar a los profesores recién egresados, así como la experiencia con la que cuentan los que ya llevan varios años de servicio, estilos didácticos y desarrollo profesional de los mismos que incluye observación de y por otros profesores e incursión de éstos en actividades de investigación.

Llama la atención la importancia que se le da a las condiciones y preparación de los profesores ya que la diversidad y profundidad de aspectos que incluye su evaluación, hace pensar en el peso y la significatividad que las agencias internacionales les atribuyen en los resultados obtenidos por los estudiantes.

Para las actividades del aula se incluyen tanto actividades del salón de clases, como las de las llamadas extracurriculares, tales como: encargos de biblioteca o trabajo de campo, enfoque pedagógico empleado, materiales y equipos disponibles, condiciones en que se da el aprendizaje que incluye tamaño y composición de la clase y cantidad de tiempo lectivo dedicado a la enseñanza y el clima en que se realiza la misma; así como el número de clases dedicado a cada tema y el tipo de evaluaciones que los profesores realizan con sus estudiantes.

En relación a los estudiantes, lo que se aborda es su ambiente y condiciones familiares tales como: número de libros espacios y otros materiales con los que cuentan para estudiar, así como la actitud de los padres y la participación de estos en los estudios de sus hijos. También se consideran la oportunidad que se les proporciona a los estudiantes de participar en actividades recreativas y el tipo de experiencias anteriores con las que cuentan, que incluye: conocimientos previos, relación con los profesores, facilidad o dificultad para aprender ciertos contenidos o asignaturas y actitudes de éstos hacia el estudio.



La evaluación contextual del TIMSS pretende identificar la estructura, contenido y alcance del currículo de cada uno de los países participantes, la preparación, experiencia y actitudes de los profesores, los contenidos de matemáticas y ciencias que realmente se imparten y el tiempo y forma en que se trabajan, la organización y recursos del aula y las experiencias, actitudes y medios de los estudiantes y padres dentro de los centros; lo que le da un sentido más amplio tanto a la evaluación del aprendizaje y desarrollo de la docencia en un sistema escolar.

Se dice que en los estudios realizados por la IEA interesa conocer, además del nivel de logro obtenido por los estudiantes, los factores que influyeron en éste a fin de realizar las intervenciones y ajustes necesarios para mejorar la eficacia en el aprendizaje; los informes del TIMSS;<sup>107</sup> sin embargo, no ofrecen una correlación entre desempeño y factores asociados. Éstos, sólo contemplan la exposición de resultados, en cuanto al logro obtenido por los estudiantes en cada una de las áreas de evaluación y brindan información sobre el contexto en el que se realiza el aprendizaje de las matemáticas, pero no la relación existente. En algunos países como Chile,<sup>108</sup> este análisis se efectúa a manera de informe nacional.

La evaluación contextual del TIMSS se lleva a cabo, pidiendo a directores de escuela, docentes y estudiantes que contesten cuestionarios sobre el contexto en el que se realiza el aprendizaje de la matemáticas y las ciencias para los cuales se les da a los estudiantes de 15 a 30 minutos extras.

---

<sup>107</sup> Los informes internacionales del TIMSS 2002/2003, están disponibles en: <http://timss2003.iaea.gov/timss2003.html> (véase IEA. 2002; Mullis, Martin, Smith, Garden, Gregory, Gonzáles, J, Chrostowski, y O Conoor, 2004 y IVEI s.f.).

<sup>108</sup> Véase artículos de la revista del Ministerio de Educación y Cultura de Chile: (2001) ". en donde se realiza una análisis del caso Chileno en cuanto a resultados obtenidos e indicadores de contexto socioeconómicos en relación con otros países, además de que se compara los resultados obtenidos en el TIMSS 1999 que se realizaron previos a la Reforma Curricular y su relación con los más recientes resultados, declarando que la participación de Chile en el estudio ratificó la decisión de reformar los currículos nacionales en las asignaturas evaluadas por TIMSS.



### b. Evaluación de factores asociados en el LLECE:

Los estudios de evaluación contextual realizados por el Laboratorio, también incluyen cuestionarios aplicados a alumnos, padres, tutores, profesores y directores con los cuales se pretende conocer los factores asociados al aprendizaje y con el fin aportar elementos que ayuden a mejorar la actual situación educativa de los países. responder a las siguientes preguntas: (Casassus, 2000: 5)

- ¿Cuál es el rendimiento de los alumnos en la región?
- ¿Cómo mejorar la calidad de la educación?
- ¿Cómo hacer que ella sea más pertinente a las necesidades de la población?
- ¿Cómo mejorar mediante la educación las posibilidades de millones de estudiantes?

Cuatro fueron los factores estudiados los cuales presentan una relación estrecha con el desarrollo económico y social de cada país.

1. El alumno y su contexto familiar: Incluye rasgos de compromiso, hábitos y destrezas del alumno y su núcleo familiar.
2. El profesor y el ámbito educativo: Incluye variable de currículum y gestión pedagógica llevada a cabo por el docente en el aula.
3. El director y el microcosmos escolar: Incluye variables que sustentan los procesos de gestión institucional de la escuela.
4. Las autoridades públicas y el macrocosmo: Incluye aspectos relacionados con la gestión pública de las instituciones de nivel nacional (Casassus, 2000:12).

Los datos obtenidos en las pruebas y cuestionarios de contexto, del *Primer Estudio Internacional Comparativo sobre Lenguaje, Matemáticas y Factores Asociados para Alumnos de Tercer y Cuarto Grado de Educación Básica* permitieron identificar escuelas a las que el



Laboratorio denominó como *escuelas destacables*.<sup>109</sup> A partir de esta información y como complemento a los informes del Primer estudio, en el año 2002 el Laboratorio publicó un “Estudio cualitativo de escuelas con resultados destacables. (véase LLECE, 2002) en siete países latinoamericanos,” que considera a las escuelas en donde padres de familia presentan bajo nivel de escolaridad, pero cuyos alumnos obtuvieron un resultado en matemáticas por arriba de lo esperado.

Con base en esto, el interrogante principal del estudio giró en torno a: “¿Qué factores provocan que estas escuelas se orienten positivamente en el nivel de los desempeños de sus alumnos a pesar de las condiciones de mayor adversidad en que se desenvuelven?” (LLECE, 2002:6).

Otros cuestionamientos que guiaron el estudio fueron:

- ¿Cuáles son los factores permanentes, propios y contextuales de las escuelas?
- ¿Qué tipo de transformaciones se están llevando a cabo en las escuelas?
- ¿Cuáles son las características especiales de los actores (profesores y alumnos), que permiten la obtención de “resultados destacables”?

Dos fueron los objetivos del estudio:

- Indagar a profundidad los factores que inciden en los resultados de los alumnos.
- Rescatar aspectos cualitativos de los procesos e interacción de los actores.

La metodología del estudio consistió en técnicas cualitativas de observación y entrevista a grupos focales y la muestra estuvo conformada por países con mayor producción en materia de evaluación e innovación.<sup>110</sup>

El análisis de resultados consistió en el establecimiento de categorías significativas que reflejaran la experiencia de los actores y permitiera entender su comportamiento, estado psíquico y modelo motivacional.

---

<sup>109</sup> El Laboratorio entiende como escuelas destacables aquellas cuyos alumnos obtuvieron un rendimiento en matemática por encima del que era posible esperar según el nivel educacional de sus padres (LLECE, 2002).

<sup>110</sup> Los países participantes de este estudio fueron: Argentina, Bolivia, Cuba, Chile, Colombia, Costa Rica y Venezuela.



Estas categorías giraron en torno a dos aspectos básicamente: *Gestión escolar y práctica pedagógica*.

En cuanto a Gestión Escolar los hallazgos del estudio fueron:

- Proceso de gestión colectivo e inclusivo en donde la palabra y acción de los actores es muy importante. Interrelación positiva de los actores que los llama a la reflexión.
- Nueva forma de relaciones caracterizada por clima armónico y emociones positivas.
- Las normas constituyen un medio de construcción no de destrucción.
- Modelo integrativo de decisión que involucra a los actores en la estructura de la gestión.
- Empleo de recursos con visión pedagógica.
- Papel protagónico del alumno en los ambientes de aprendizaje. Importancia de los niños como individuos.
- Trabajo en equipo entre padres de familia, directores y docentes.

En materia de Práctica pedagógica los hallazgos del estudio consistieron en:

- Uso eficiente de espacios pedagógicos y físicos.
- Importancia de actividades extracurriculares.
- Vinculación de escuela, familia y hogar.
- La clase como proceso de construcción conjunta.
- Docentes preocupados por la elaboración de materiales.
- Administración eficiente de recursos.
- Adecuación de programas conforme a necesidades.
- Lectura, elemento central de estrategias pedagógicas.
- La tarea como continuación del proceso escolar.
- Métodos heurísticos de corte constructivo.
- Forma de evaluación no clasificatoria.
- Alto compromiso y vocación por parte de los docentes. Los docentes se sienten gratificados por su trabajo.



- Funcionamiento autónomo y descentralizado de escuelas.

Los resultados del estudio fueron redactados por temas en forma de declaraciones; son situaciones ejemplificadas de los diversos países de la muestra. Sobresale el caso de Cuba, Chile y Argentina.

Estamos ante un estudio pionero de evaluación de corte cualitativo, que no sólo reporta el rendimiento de los alumnos y la situación sociocultural y escolar de los alumnos evaluados; sino que hace uso de la información obtenida tanto en materia de rendimiento como en situación contextual, para penetrar al interior de las escuelas que obtuvieron resultados destacables y peculiares en estos rubros, para conocer el porqué de los resultados obtenidos a pesar de presentar situaciones desfavorables en cuestión sociocultural. Los hallazgos emanados del estudio se constituyen como verdaderas aportaciones pedagógicas que pueden contribuir a mejorar la situación educativa de escuelas en condiciones socioculturales similares.

### c. Evaluación de factores asociados en el PISA

El objetivo de la evaluación contextual de PISA, es recabar datos sobre el estudiante, su familia y los factores institucionales que pueden contribuir a explicar las diferencias en dicho desempeño y orientar de esta manera las políticas educativas.

Todo esto con la intención de contar con elementos para saber: ¿Cómo puede la escuela fomentar la disposición para aprender en los estudiantes?<sup>111</sup> Y si los jóvenes al dejar la escuela ¿tienen la capacidad para seguir aprendiendo durante toda la vida?

El proyecto OCDE/PISA, como parte de su evaluación contextual aplica cuestionarios diseñados para recabar información sobre los siguientes factores:

---

<sup>111</sup> Otro estudio de la OCDE que revela un interés especial por estas cuestiones, lo podemos encontrar en el Proyecto realizado por el Centro para la Investigación e Innovación Educativa (CERI) de la OCDE, cuyo objetivo es resaltar la necesidad de vincular las ciencias del aprendizaje con la investigación cerebral, en un afán de lograr una visión interdisciplinaria que permita conjuntar avances e intereses en la comprensión del aprendizaje humano que permita establecer líneas de acción en materia de política educativa y práctica docente (Véase Moreno, 2003:79).



- Familia
- Medio ambiente
- Escuela
- Actividades diarias del estudiante, en cuanto a su motivación hacia el estudio y sus prácticas de autoaprendizaje.<sup>112</sup>

La evaluación contextual realizada por el proyecto OCDE/PISA, tiene el propósito de analizar las actitudes de los educandos hacia la lectura y las matemáticas y su dedicación general al aprendizaje y la instrucción, así como la confianza de estos en sus propias capacidades. Adicionalmente el programa incluía dos cuestionarios para recabar información sobre las prácticas de estudio los alumnos y su vinculación con el uso de las computadoras. El primero de estos cuestionarios incluían preguntas sobre estrategias de aprendizaje, motivaciones, preferencias y autoconcepto.<sup>113</sup> El segundo cuestionario contenía preguntas sobre el uso de las computadoras por parte de los estudiantes y sus habilidades al respecto. También había cuestionarios para la escuela, especialmente dirigidos a los directores de éstas. Las sesiones para la aplicación de estos instrumentos estuvieron distribuidos de la siguiente manera: dos horas para evaluar la lectura, las matemáticas y la ciencia, veinte a treinta minutos para los cuestionarios a estudiantes, diez minutos para las preguntas de opinión acerca de sí mismo y sus competencias y cinco minutos para la opinión sobre la familiaridad con el uso de computadoras. Los cuestionarios para los directores de escuela fueron contestados en un tiempo de 20 a 30 minutos.<sup>114</sup>

---

<sup>112</sup> Los resultados obtenidos en los cuestionarios de contextos en cuanto a preferencias de los estudiantes, motivación, estrategias, entorno familiar y ambiente de aprendizaje se pueden consultar en: OCDE (2002). También existe otro documento en donde se presenta resultados acerca de la manera en que los estudiantes de los diferentes países y sexos, se acercan al aprendizaje, (OECD, 2003b). Este es un informe dedicado a los tomadores de decisiones y responsables de la política educativa; pues da algunas aportaciones interesantes sobre la manera en que los sistemas educativos pueden favorecer el aprendizaje eficaz entre sus estudiantes.

<sup>113</sup> Un documento que presenta el compromiso de los estudiantes con la escuela y su nivel de participación en ésta y en sus estudios a través de los resultados obtenidos en los cuestionarios es OECD (2003c). El cual, es un informe que examina varios aspectos relacionados con el nivel de participación y compromiso que los estudiantes tienen con su escuela.

<sup>114</sup> Algunas preguntas muestra de los cuestionarios a estudiantes se pueden encontrar en: OECD (2002c) Publicación que proporciona información para entender la base de datos 2000 de PISA y para realizar



Cabe señalar que en el informe realizado por la OCDE (2002) sobre los resultados de PISA 2000, se presenta, la relación existente entre desempeño de los estudiantes en lectura y matemáticas y entorno familiar y socioeconómico; y se concluye que la situación laboral de los padres, entorno familiar (2002:170), prácticas escolares relacionadas con el entusiasmo, dedicación de los maestros, autonomía escolar, relaciones entre maestros y alumnos, ambiente de disciplina en el salón de clases y presión para obtener logros, según opiniones de directores y maestros, así como la situación socioeconómica media de la escuela (2002:225), contribuye a explicar las diferencias en el desempeño estudiantil. Sin embargo en el informe del PISA también se concluye que "no existe un solo factor que explique porqué algunas escuelas o algunos países obtienen mejores resultados que otros. Más bien el resultado exitoso se puede atribuir a una constelación de factores que incluye a los recursos escolares, las políticas y prácticas de la escuela y las prácticas en el salón de clases" (OCDE: 2002: 228).

Finalmente, con fines de una posible comparación, se presentan de manera conjunta los propósitos, cuestionamientos y factores asociados contemplados en la evaluación contextual de los tres estudios internacionales de evaluación.

---

análisis de acuerdo con las metodologías utilizadas para recoger y para procesar los datos. También incluye las respuestas de los estudiantes a los cuestionarios y a las preguntas de la prueba.



CUADRO 1.15

**EVALUACIÓN DE FACTORES ASOCIADOS  
EN LOS ESTUDIOS INTERNACIONALES**

PRUEBA	PROPÓSITO	CUESTIONAMIENTOS A LOS QUE RESPONDE	FACTORES QUE EXPLORA
TIMSS	Obtener una visión general de lo que significan los resultados de las pruebas	¿Cómo son los contextos en los que aprenden los estudiantes?	El currículum Los centros educativos Los profesores y su preparación Las actividades del aula Los estudiantes
LLECE	Aportar información útil para la formulación y ejecución de políticas educacionales.	¿Cuál es el rendimiento de los alumnos en la región? ¿Cómo mejorar la calidad de la educación? ¿Cómo hacer que ella sea más pertinente a las necesidades de la población? ¿Cómo mejorar mediante la educación las posibilidades de millones de estudiantes?	El alumno y su contexto familiar. El profesor y el ámbito educativo. El director y el microcosmo escolar. Las autoridades públicas y el macrocosmo
PISA	Orientar de esta manera las políticas educativas.	¿Cómo puede la escuela fomentar la disposición para aprender en los estudiantes? ¿Tienen la capacidad para seguir aprendiendo durante toda la vida?	Familia Medio ambiente Escuela Actividades diarias del estudiante, en cuanto a su motivación hacia el estudio y sus prácticas de autoaprendizaje

Los tres estudios de evaluación mantienen un propósito de evaluación formativa, es decir, de utilizar los resultados obtenidos para mejorar las condiciones de los sistemas educativos, sea en materia de rendimiento o de política educativa, nótese que tanto los estudios del LLECE, como del PISA, tienen una orientación política. Los cuestionamientos que los generan, están encaminados también a este fin: conocer la situación de las escuelas para mejorarla y propiciar mejores condiciones. El alumno, como sujeto de aprendizaje y las condiciones



personales a nivel de estrategias, educativas a nivel de escuela y sociales a nivel de familia y entorno es el eje central de estos estudios.

## **1.6 Primeras conclusiones**

A manera de conclusiones preliminares, el análisis de las principales orientaciones de las pruebas internacionales, da cuenta de que:

1. En cuanto a generación y propósitos, de acuerdo con lo que se muestra en el cuadro “Generación y propósitos de los estudios internacionales de evaluación” (véase cuadro 1.1), tenemos indudablemente que estos estudios obedecen a una política de calidad y eficiencia de los sistemas educativos impulsada a partir de los años noventa; su propósito está enfocado a tres aspectos fundamentales:
  - a) Comprobación del currículo, mejora de la calidad educativa y efectividad de la escuela, en cuanto a aprendizaje y preparación de los estudiantes para la vida; ya que los conceptos organizadores en los que hacen énfasis cada uno de los estudios de las distintas agencias son: currículo, eje rector de las pruebas del TIMSS, calidad, como concepto base de los estudios del Laboratorio Latinoamericano y habilidades para la vida, planteamiento principal de los estudios de la OCDE.
  - b) Elaboración de estándares, como factor de calidad educativa, idea central de los estudios del LLECE. (véase en este mismo capítulo, lo referente a Generación y propósito de los estudios internacionales de evaluación, inciso b, p. 9).
  - c) Vinculación entre rendimiento de los estudiantes y situación sociocultural y escolar de los sistemas educativos. Componente fundamental trabajado en los tres a través de la evaluación de factores asociados, (véase en este mismo capítulo “Evaluación de los factores asociados en los estudios



internacionales," punto 1.5, p. 85,) que además resulta una aportación importante a la concepción de evaluación que revelan estos estudios.

2. El contenido y objeto de evaluación, está claramente conceptualizado, dimensionado y delimitado e incluye tanto temas, como procesos cognitivos a los que les llaman comportamientos o destrezas. Es decir, como se muestra en los cuadros correspondientes (véase cuadros 1.2, 1.3, 1.4, 1.5 y 1.6) en cada estudio se define claramente lo que se evalúa y los elementos que incluyen en su evaluación.
3. La estructura de las pruebas contempla tanto su procedimiento de elaboración, referente a la revisión del dominio, área, tema y tópicos, como las características psicométricas y la forma de presentar los resultados. El proceso de elaboración de pruebas, de acuerdo con lo que se presenta en el cuadro: "Proceso de elaboración de las pruebas internacionales"(véase cuadro No.1.10), coincide con la teoría del test, pues contempla una clara definición del dominio, tareas, características psicométricas y operacionalización de la definición; tal y como se muestra en el cuadro: "Diseño de ítems de las pruebas internacionales" (véase cuadro 1.11). El diseño de los ítems presenta elementos de vanguardia: Los ítems de las pruebas internacionales, sobre todo en las del PISA, son ítems elaborados con base en tareas auténticas que incluyen la aplicación de conocimientos y habilidades a la vez; conservan el formato de opción múltiple, pero además incorporan el de respuesta construida y el modelo estadístico mediante el que se analizan es el de TRI. El sistema de calificación que se muestra en el cuadro: "Sistema de calificación y puntuación de las pruebas internacionales" (véase cuadro 1.14), incorpora ciertas aportaciones importantes a considerar: son pruebas criteriosales, que responden a recientes paradigmas en materia de evaluación y en donde la calificación con base a la norma ha quedado atrás; el establecimiento de niveles de desempeño, mediante el análisis de tareas, le da otro sentido y significado a la evaluación pues permite conocer logros y dificultades de los estudiantes y no sólo el número obtenido; el establecimiento de puntajes medios ofrece una alternativa más de calificación,



aunque sigue prevaleciendo el paradigma cuantitativo en la presentación de resultados.

4. La evaluación de factores asociados, condensada en el cuadro: "Evaluación de los factores asociados en los estudios internacionales" (véase cuadro 1.15), aparece como una preocupación por no sólo medir desde una perspectiva meramente cuantitativa, sino sobre todo evaluar, ofreciendo información que permita comprender el porqué de los resultados obtenidos y detecte elementos que permitan mejorar la situación y efectividad de los sistemas educativos.

Sin embargo, cabe señalar que la presentación y publicación que se hace de los resultados obtenidos en estos estudios de evaluación comparada de corte internacional, aún no contribuye con la función de retroalimentación, propósito fundamental en toda acción evaluativa y a la reorientación de políticas y prácticas educativas, fin con el que fueron creados.<sup>115</sup> A excepción de Chile, (véase MEC, 2001) quien impulsó fuertemente sus reformas curriculares a raíz de los resultados obtenidos en TIMSS, no se tienen evidencia hasta la fecha, de cambios significativos en los sistemas educativos a raíz de los resultados obtenidos en estos estudios. La "cultura de la evaluación" a la que tanto se ha hecho alusión en últimas fechas, queda reducida a una mera rendición de cuentas y comparación de los sistemas educativos. Pese a la proliferación de los estudios de evaluación internacional y a la inserción de muchos de los países en éstos, el uso que se le da a los mismos adolece todavía de un tratamiento pedagógico serio, en materia de contenido y difusión de resultados de las evaluaciones internacionales.

---

<sup>115</sup> Sólo en algunos países de entre los que destacan España, (véase INCE, 2004; Pajares, R., Sanz, A., y Rico, L, 2004); Chile (MEC 2001) y Costa Rica, (Véase ANEP 2004 a, b y c) se cuenta con documentos que a partir de las traducciones que se hacen de los informes oficiales, no sólo se informa cuál es la posición que ocupa el país en relación a otros; sino mediante un análisis de los aciertos y errores de los estudiantes se pretende dar a conocer a la población en general el porqué de los resultados obtenidos. Costa Rica presenta además una experiencia interesante con docentes a raíz de los resultados obtenidos en PISA 2003 (Véase ANEP, s.f a, b, c y d).



El análisis realizado, permitirá en el siguiente capítulo presentar el concepto de examen de las pruebas internacionales.

## CAPÍTULO DOS

### EL CONCEPTO DE EXAMEN EN LAS PRUEBAS INTERNACIONALES

---

Este capítulo ofrece un análisis de los conceptos que sustentan a los exámenes internacionales en materia de contenido y estructura conceptual proponen las Pruebas Internacionales.

En el capítulo anterior, se reconocen tres categorías que sustentan el concepto de examen a nivel internacional:

- Concepto de contenido<sup>116</sup>
- Tipo de ítems
- Modelo psicométrico

#### 2.1. Concepto de contenido en las Pruebas Internacionales

Se entiende como Concepto de contenido, la definición que las Agencias y Organismos Internacionales de evaluación hacen del objeto a evaluar y su relación con el tipo de ítems que aparece en las pruebas internacionales.

De las tres pruebas analizadas en este trabajo, cabe destacar que son en las del PISA(2000), donde se puede observar con mayor claridad esta Concepto de contenido (véase OCDE, 2000 y 2004b).

Esta primer categoría referente a la Concepto de contenido, se identifica con base en los conceptos más recurrentes del Marco Teórico que sustenta a las Pruebas Internacionales y que deja ver una muy particular forma de concebir el examen.

Los conceptos encontrados fueron los siguientes:

---

<sup>116</sup> Se refiere el tratamiento que las Pruebas Internacionales le dan a los contenidos de evaluación a partir de una previa definición de los mismos (Díaz Barriga, 2004).



### a. Conceptos implícitos en las Pruebas Internacionales

Los conceptos que se pueden identificar de la revisión y análisis que se hizo del contenido y estructura de las Pruebas Internacionales, fueron los siguientes:

- Dominio de contenido
- Aptitud
- Conocimiento
- Rendimiento
- Formación
- Capacidad
- Interpretación
- Razonamiento

El cuadro que a continuación se presenta, reúne y sintetiza, con fines de comparación los conceptos localizados en los documentos de evaluación internacional, conforme a los cuáles se organiza el contenido de evaluación.

CUADRO 2.1

#### CONCEPTOS DE LAS PRUEBAS INTERNACIONALES

CONCEPTO	TIMSS	LLECE	PISA
<b>DOMINIO DE CONTENIDO</b>	Temas o aspectos expresados en términos de rendimiento y resultados que demandan uso de estrategias para la realización de tareas Demanda conocimiento factual y comprensión (cfr Mullis, Martin, Smith, Garden, Gregory, Gonzales, Chrostowski, y O Conoor, 2002: 25).	Se refiere a estándares de contenido y tiene que ver con el dominio sustantivo de objetivos educacionales (cfr. LLECE, 1997,11).	Temas, expresados en matemáticas como "grandes ideas" (OCDE, 2001:81).
<b>DOMINIO COGNITIVO</b>	Comprensión y destreza (cfr . Mullis, Martin, Smith, Garden, Gregory, Gonzales, Chrostowski, y O Conoor 2002: 25).	No se emplea	No se emplea



*Exámenes masivos*  
*Un análisis entre pruebas internacionales y nacionales*

<b>APTITUD</b>	No se emplea	Se retoman las concepciones propias de la Psicología, para definirlo como un concepto que implica vinculación de la disposición innata con el rendimiento (cfr. Casassus, 1997: 9)	Patrones continuos . Gama de competencias con las que se cuenta. (OCDE , 2002: 19)
<b>CAPACIDAD</b>	No se emplea	No se emplea	Como formación o competencia, que se mide en continuo, no como algo que alguien posee o no (cfr. OCDE, 2001a:15)
<b>CONOCIMIENTO</b>	Se refiere a hechos y conceptos y constituyen parte de un lenguaje. (cfr. Mullis, Martin, Smith, Garden, Gregory, Gonzales, Chrostowski, y O Conoor 2002:39).	Proceso de transformación y desarrollo y no como una bolsa de contenidos (LLECE, 2002:17)	Comprensión de conceptos clave. (cfr. OCDE 2001:16,19)
<b>RENDIMIENTO</b>	Se refiere al conocimiento y comprensión de contenidos temáticos y a las actitudes y logros educativos de estudiantes (cfr. Mullis, Martin, Smith, Garden, Gregory, Gonzales, Chrostowski, y O Conoor 2002:49).	Se maneja como resultados (cfr. LLECE, 2002:49, 268)	No se emplea
<b>FORMACIÓN</b>	No se emplea	No se emplea	Grado en el que los alumnos han adquirido conocimientos y destrezas.
<b>PROCEDIMIENTO</b>	Puente entre el conocimiento básico y su uso para la resolución de problemas (cfr. Mullis, Martin, Smith, Garden, Gregory, Gonzales, Chrostowski, y O Conoor 2002:39)	No se emplea	"Consisten en la capacidad de los estudiantes para analizar, razonar y comunicar ideas de manera efectiva mediante el planteamiento, la formulación y la resolución de problemas matemáticos" (OCDE, 2001a :77)
<b>CONCEPTOS</b>	Forma parte de un conocimiento factual. (hechos, datos, herramientas y procedimientos) (cfr. Mullis, Martin,	No se emplea	Se refiere a conceptos clave amplios, factibles de aplicación. Como parte de los atributos que permitan



	Smith, Garden, Gregory, Gonzales, Chrostowski, y O Conoor 2002: 75).		una plena participación en la vida como adultos. (cfr. OCDE, 2001: 16,24)
<b>RAZONAMIENTO</b>	Capacidad de pensamiento lógico y sistemático que implica la transferencia de conocimientos y destrezas a nuevas situaciones para resolver problemas. (cfr. Mullis, Martin, Smith, Garden, Gregory, Gonzales, Chrostowski, y O Conoor 2002: 44).	No se emplea	Se maneja como destreza necesaria para un tipo de competencia (cfr. OCDE, 2000: 28)
<b>INTERPRETACIÓN</b>	No se emplea	Se refiere al proceso de interpretación textual como una activación de operaciones sobre una manifestación, lineal para la producción de elementos interpretantes. (cfr. LLECE, 2002:21)	Se maneja como la reflexión sobre el contenido de un texto (cfr. OCDE: 200:25)
<b>NIVEL DE DESEMPEÑO</b>	No se emplea	"Espacio caracterizado por el reagrupamiento de preguntas que cumplen con ciertos rasgos particulares en razón de la dificultad y la habilidad del estudiante que las responde". (LLECE, 2002:25)	No se emplea
<b>PENSAMIENTO LÓGICO</b>	"El razonamiento matemático, implica la capacidad de pensamiento lógico y sistemático. Incluye el razonamiento intuitivo e inductivo, basado en patrones y regularidades que se pueden utilizar para llegar a soluciones para problemas no habituales" (Mullis, Martin, Smith, Garden, Gregory, Gonzales, Chrostowski, y O Conoor, 2002:44)	"Capacidad para obtener información desconocida utilizando procedimientos matemáticos de procesamiento de la información que se conoce" (LLECE, 2002:24)	No se emplea
<b>SOLUCIÓN DE PROBLEMAS</b>	Fin y medio de la enseñanza matemática y pueden ser de la vida real o preguntas meramente matemáticas en donde es necesario saber aplicar todo un conjunto de acciones (cfr. Mullis, Martin, Smith, Garden, Gregory, Gonzales, Chrostowski, y O Conoor, 2002:108).	"Implica obtener información desconocida a partir de información conocida aplicando reglas de procesamiento matemático" (LLECE, 2002, 24) y requiere la puesta en práctica de habilidades matemáticas de complejidad superior. (LLECE, 2002:4)	Se maneja como competencia matemática y destreza.(OCDE,;2000:74)



Los exámenes en el plano internacional, conciben los *dominios de contenidos*, como temas básicos del currículo escolar expresados en el PISA como "grandes ideas", en términos de rendimiento o desempeños; lo que elimina el uso enciclopédico de contenidos, a los que se les considera, sólo como elementos fundamentales que permitirán realizar una tarea.

El *dominio cognitivo* como tal, se encuentra explícitamente sólo en el TIMSS, pero las tres pruebas incluyen el uso de destrezas y procesos cognitivos.

El concepto de *aptitud*, al que algunos autores definen como capacidad potencial para aprender o desempeñar una tarea (Agut y Grau, 2001); coincide con lo que el LLECE maneja y está en estrecha relación con la habilidad o destreza para realizar algo; en PISA se presenta como sinónimo de competencia.<sup>117</sup>

La *capacidad*, es trabajada no como algo estático que alguien posee y en donde lo que interesa es ver si se tienen o no, sino como algo que la educación puede y tiene que desarrollar.

El *conocimiento*, concepto implícito en los tres estudios, se maneja en términos de funcionalidad más que de acumulación, se considera como un lenguaje que permite hacer; lo que refuerza el hecho de que lo que interesa no es ver cuanto se sabe, sino que es lo que se sabe hacer con lo que se tiene; que en sí es lo que se maneja como *rendimiento* en el TIMSS, o *formación* en el caso del PISA.

El *rendimiento*, de acuerdo con el campo del cual fue incorporado, se refiere básicamente a resultados y en PISA se maneja como *formación*.

Los *procedimientos* se relacionan con el uso de operaciones mentales para la aplicación de *conceptos*, los cuales aluden a conocimientos básicos de tipo factual.

El *razonamiento y la interpretación* son dos operaciones mentales que actúan aunadas y vinculadas al pensamiento lógico, lo que implica la posibilidad de activar procesos mentales para transferir, *resolver problemas* y producir elementos interpretantes.

---

<sup>117</sup> El manejo indistinto de los términos habilidad y aptitud y su cercanía con el de competencia es frecuente encontrarlo, no sólo en el marco teórico de los exámenes internacionales, sino en la reciente literatura sobre aprendizaje y formación profesional (véase San Martín ( s.f.).



El *nivel de desempeño*, es el término empleado sobre todo por el LLECE, para la calificación y presentación de resultados.

El *pensamiento lógico y la solución de problemas*, están relacionadas con el uso de procedimientos matemáticos y en PISA se manejan como competencias.

Un análisis de lo que implica cada uno de estos términos a los que recurren las Pruebas Internacionales, permite identificar tras de sí dos conceptos fundamentales:

- Competencia
- Habilidad

Estos conceptos, son la base de su organización y se encuentran de manera implícita en cada una de las pruebas, abordados de diferente manera:

CUADRO 2.2

**CONCEPTOS IMPLÍCITOS EN LAS PRUEBAS INTERNACIONALES**

	<b>TIMSS</b>	<b>LLECE</b>	<b>PISA</b>
<b>Competencia</b>	Destrezas implícitas	Aplicación de un saber	Conocimientos y destrezas para la vida adulta, aplicadas a contenidos y situaciones de la vida diaria.
<b>Habilidad</b>	Comportamientos	Capacidad y aptitud para realizar una tarea con maestría y eficacia.	Destreza para aplicar procedimientos, considerados también como competencias

El concepto de competencia<sup>118</sup> es la noción principal, que sobresale tanto implícita como explícitamente en la estructura y marco teórico de las Pruebas Internacionales, aunque se

<sup>118</sup> Concepto emanado del mundo laboral, fruto de los procesos de globalización de las economías y de la transformación de las reglas de competitividad (véase Barrón, 2000) que al igual que la evaluación ha entrado de manera proliferante y rápida al ámbito educativo, pero sin una clara definición y análisis del mismo y sobre todo, de sus implicaciones y articulaciones con las prácticas educativas. En la Unión Europea este concepto surge con la idea de mejorar la calidad de la formación profesional para una mejor adaptación al mercado laboral y sistema de empleo. En el Reino Unido el debate se inicia vinculado a la evaluación y el



maneja más bien como destreza. En PISA se habla de competencias para la vida adulta y aunque el término competencia ha sido incluido desde la educación preescolar<sup>119</sup>, PISA se refiere a las competencias que el estudiante que está por concluir la educación básica necesita para desarrollarse en su vida como adulto (OCDE, 91: 23, 22, 24).

La habilidad, con terminaciones distintas en cada una de las pruebas, se refiere en sí a los comportamientos, aptitudes, destrezas o competencias, para la aplicación de un conocimiento a una situación o problema.

El modo en cómo se aplican y conciben estos conceptos en cada uno de los exámenes internacionales se aborda como siguiente punto.

---

rendimiento. En Alemania las competencias están ligadas al ámbito profesional. En Francia el debate emerge como una crítica a la pedagogía tradicional basada únicamente en conocimientos teóricos. (CEDEFOP, 1994). En México, el último Estado de Conocimiento sobre Investigación curricular, en el campo del currículo (véase Díaz Barriga, coord. 2003), refiere que este concepto comienza a incursionar en el ámbito del desarrollo curricular desde el año de 1993. "Se le ubica como un enfoque que combina la formación teórica con la aplicación del conocimiento y su pretensión es lograr las capacidades para realizar una determinada labor, tarea o actividad con un alto nivel de calidad y eficacia" (Díaz Barriga Arceo, y Lugo, 2003: 70). Sin embargo al considerar que de acuerdo con Romainville (2000) "evaluar una competencia significa fundamentalmente determinar la capacidad de deducción del alumno" (cit. en EURYDICE, 2002: 20); la evaluación del aprendizaje por competencias en que se sustenta la idea de examen de las Pruebas Internacionales, es una tendencia poco trabajada, sobre todo a nivel de educación básica; la mayor parte de los trabajos que abordan el tema de la evaluación por competencias, se enfoca a la formación profesional y al desempeño laboral. (véase Huerta, Pérez, Castellanos, 2000, San Martín, s.f., Posada 2004, Roe, 2003) En algunos países como Colombia, país donde el concepto, ha tenido mayor abordaje en términos de discusión y análisis, apenas se pueden detectar algunos trabajos en relación con la evaluación de competencias como parte del proceso de formación integral, (véase Fernández y Quiroz S/f). En la Unión Europea, existe un interesante debate en torno a la identificación de competencias clave que han de trabajarse y evaluarse durante la educación obligatoria; sin embargo, en cuanto a la identificación y certificación de éstas, también se afirma que: "A menudo las pruebas están demasiado orientadas a la memorización de la información y a la producción de respuestas correctas, más que a la aplicación de conocimientos mediante la reflexión crítica y creativa(...)El resultado es que los conocimientos de hechos se aprenden y se memorizan con el fin de aprobar un examen y no para que sea de utilidad en la vida adulta(...)" (EURYDICE, 2002:20).

<sup>119</sup> Un ejemplo de la inclusión del término competencia en la educación infantil lo podemos encontrar en (Garduño, Guerra, Silva y Rodríguez, 2001).



## **b. El concepto de competencia en los exámenes internacionales.**

### **□ El concepto de competencia en las pruebas del TIMSS**

El concepto de competencia en el TIMSS, no es un concepto que sea trabajado ni definido propiamente dentro del marco teórico de este estudio<sup>120</sup>; sin embargo en éste se utiliza el término competencias para referir las "competencias básicas" a considerar dentro de los dominios cognitivos de las matemáticas, que aunadas a la comprensión, no deben ser reemplazadas por el uso de calculadoras. (cfr. Mullis, Martin, Smith, Garden, Gregory, Gonzales, Chrostowski y O Conoor, 2002: 20). Y Se emplea el término de "estándares de competencia" (cfr. Mullis, Martin, Smith, Garden, Gregory, Gonzales, Chrostowski y O Conoor, 2002: 86), como medio para mejorar la preparación académica de los profesores. Esto habla de que el término competencia es utilizado en el TIMSS para referirse a procesos y metas alcanzables en cuanto a formación de profesores.

### **□ El concepto de competencia en las pruebas del LLECE**

En el LLECE, como referimos en el capítulo anterior, el concepto de competencia se emplea como objeto y centro de la evaluación del Lenguaje y Matemáticas.

Las pruebas del lenguaje en el LLECE, están centradas en la evaluación de la *competencia comunicativa* definida como:

- "El conjunto de posibilidades que tiene un estudiante para comprender, organizar, negociar y producir actos significativos a través de distintos sistemas de signos lingüísticos y no lingüísticos "(LLECE, 2001:16).

---

<sup>120</sup> El que el término competencia se incluya pero no se defina en los marcos teóricos de las pruebas internacionales, sobre todo en las del TIMSS, que fue el primer estudio evaluativo de corte internacional, no es algo exclusivo de esta experiencia, cabe señalar que sobre todo en el ámbito del diseño curricular, se adopta el término competencia pero sin una clara revisión del concepto. (véase p.e. Corvalan, 1993; INSAFORP, 2001) Lo que quizá obedece a la tendencia de incorporar nuevos términos como sinónimo de innovación sin una reflexión teórica del significado. Algunos autores que sí han abordado la discusión conceptual del concepto son Perrenoud (1999) y Rey (2001).



Las pruebas de matemáticas, tienen como objeto de evaluación la *competencia matemática* entendida como:

- "La capacidad que tiene el estudiante de utilizar procedimientos matemáticos para comprender e interpretar el mundo real" (Casassus, 2001:19).

Esto promueve un examen centrado en el uso de procedimientos matemáticos.

La competencia en los exámenes del LLECE, es entendida como "posibilidades intelectuales" y como "uso de procedimientos."

#### □ El concepto de competencia en las pruebas del PISA

En los exámenes del PISA, la competencia se equipara con el concepto de formación o capacidad y se trabaja como un campo de evaluación en las áreas de comprensión lectora y matemáticas y como procesos científicos en el área de ciencias.

- *"La competencia lectora se define como la comprensión, 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>a</sup>:32).
- *Las competencias matemáticas se conciben como destrezas generales para la resolución de problemas, el uso del lenguaje matemático y el diseño matemático, (OCDE, 2001a:73) enfocadas a la construcción de modelos y uso de procedimientos.*

Una característica importante del proyecto PISA es que las competencias no se evalúan en términos de currículo escolar, sino en términos de preparación de los jóvenes para la vida.

En un intento por comparar el concepto de competencia que se maneja en las pruebas internacionales, el siguiente cuadro concentra los términos utilizados en las pruebas internacionales para referirse al concepto de competencia.



CUADRO 2.3

**EL CONCEPTO DE COMPETENCIA EN LOS EXÁMENES INTERNACIONALES**

PRUEBA	LENGUAJE	MATEMÁTICAS
TIMSS	Procesos o metas alcanzables	
LLECE	Posibilidades	Procedimientos
PISA	Comprensión, empleo y reflexión de textos	Destrezas para la solución de problemas

Hasta aquí, se puede decir que el término competencia incursiona en las Pruebas Internacionales como activación de procesos que permiten la aplicación de un conocimiento. La inclusión del concepto competencia en los exámenes internacionales se refiere al uso de procesos, procedimientos y destrezas, en términos de habilidades para la resolución de problemas; lo que refleja cierta confusión y falta de una conceptualización clara del término. Autores como Perrenoud (1999), definen a las *competencias* como "capacidad de actuar de manera eficaz en un tipo definido de situación que se apoya en conocimientos pero no se reduce a ellos" (1999:7). Cuando hablamos de *capacidades*, de acuerdo con Monereo (1998) se trata de "un conjunto de disposiciones de tipo genético"(1998:18); Coll y Martín (2003), las identifican como intenciones educativas que se pueden construir a lo largo de la escolaridad; incluyen contenidos y se expresan a través de ejecuciones o actividades que realizan los alumnos. Éstas se pueden considerar en cinco grandes ámbitos: cognitivas, intelectuales, motrices, emocionales o de equilibrio personal. En cambio la *habilidad*, (*skill*) considerada también como destreza por su traducción al español, se refiere a la *pericia* para realizar algo (Augut y Grau, 2001). Dice Perrenoud (1999), "a partir del momento en el que él haga lo que debe hacer, sin siquiera pensar, ya no se hablará de competencias sino de habilidades" (1999:33). Otro de los conceptos que frecuentemente se confunde con el de competencia, habilidad y en algunos casos con el de capacidad, en su carácter de innato, es el de *aptitud*, (véase OCDE, 2002a). Augut y Grau, (2001) proponen tres criterios para



diferenciarlas: Las *aptitudes* tienen un carácter permanente, mientras que las habilidades son modificables, las aptitudes tienen un carácter innato, no aprendido, mientras que las habilidades pueden ser aprendidas y por último las habilidades se infieren a partir de la conducta, mientras que las habilidades son observables (2001:3).

De acuerdo con la conceptualización que se hace de estos términos, se puede decir que la competencia es un concepto amplio que incluye tanto capacidades, como destrezas, pero que no significan lo mismo, ni se reduce sólo al empleo de éstas. La competencia se refiere al uso de capacidades, ya sea innatas o construidas a través de la escolarización que incluye la activación, o como diría Perrenoud (1999) la *movilización de conocimientos* para poder realizar un acto físico o mental y resolver situaciones; se pueden manifestar a través de aptitudes como base, o capacidades en términos de disposiciones, pero no son lo mismo ni se restringe a éstas. Dice Roe (2003) "las competencias deben distinguirse también de las aptitudes" (2003:5).

Lo que lleva a pensar que hablar de competencia es hablar de la conjunción de capacidades en términos de potencia y de conocimientos es decir de saberes, que se ponen en juego en situaciones originales. Nótese que estamos hablando de conjunción más no de igualdad de conceptos, o reducción.

A diferencia de lo que se refiere en las Pruebas Internacionales como competencias, éstas no son ni habilidades, ni destrezas, ni aptitudes solamente; sino que incluye todos estos elementos aunados a la activación y puesta en práctica de conocimientos. Dice Roe (2003) los conocimientos, las habilidades y las actitudes, son los ingredientes a partir de los cuales se construyen las competencias; de ahí que se puede identificar cierta confusión en la definición del término dentro de los documentos internacionales de evaluación; aunque, como se verá más adelante en la organización del contenido y tipo de reactivos, los exámenes internacionales sí contemplan la conjunción de conocimientos y capacidades en contextos diferentes .

Dado el énfasis encontrado en las Pruebas Internacionales en cuanto a evaluación de competencias, consideradas en este caso de alguna forma como habilidades, aunque como



ya se aclaró, no son lo mismo, hemos considerado conveniente, analizar el tipo de habilidades evaluadas por estos exámenes.

### c. Habilidades evaluadas en las pruebas internacionales.

Una característica importante en el concepto de examen de las pruebas internacionales, consiste en la inclusión de habilidades como dimensión dentro de su contenido.

Un análisis de tareas incluidas en los ítems de las pruebas internacionales, permite identificar las siguientes habilidades.

CUADRO 2.4

## HABILIDADES A EVALUADAS EN LAS PRUEBAS INTERNACIONALES

PRUEBA	ÁREAS		
	LENGUAJE	MATEMÁTICAS	CIENCIAS
TIMSS	No aplica	Recordar, reconocer, identificar, calcular, usar herramientas, clasificar, representar, formular, distinguir, seleccionar, representar, interpretar, aplicar, verificar, comprobar, observar y hacer conjeturas y deducciones lógicas para justificar resultados	Explicar, comparar, contrastar, enunciar, clasificar, relacionar, localizar, describir, asociar, manifestar, utilizar, demostrar o aplicar el conocimiento, interpretar, reconocer, dar ejemplos, diferenciar, seleccionar, ordenar, demostrar, ilustrar, describir transformaciones, resolver problemas prácticos, usar o interpretar diagramas, datos y mapas, dibujar, identificar rasgos esenciales, predecir cambios, representar, calcular, analizar tendencias y comentar factores y/o ventajas.
LLECE	Establecimiento de hipótesis, interpretación, Reconocimiento, identificación, establecimiento de relaciones, traducción semántica	Interpretación de gráficos, reconocimientos de patrones, manejo de probabilidades y establecimiento de relaciones entre datos	No se evalúa



PISA	Comprensión global Obtención de información Elaboración de una interpretación Reflexión sobre el contenido de un texto Reflexión sobre la forma de un texto	Argumentación, diseño, planteamiento y resolución de problemas, representación, simbolización y formalismo, comunicación, uso de herramientas y pensamiento matemático Competencias: reproducción de definiciones y cálculos, la conexión e integración para la resolución de problemas y la generalización y comprensión de situaciones matemáticas	Reconocimiento de cuestiones científicamente investigables Identificación. Evaluación de conclusiones relacionar las conclusiones con la evidencia que las fundamenta. Comunicar conclusiones válidas. Comprensión y aplicación de sus conclusiones a situaciones distintas.
------	---	---	--

De la gama de procesos implícitos en las diferentes tareas de la pruebas internacionales, las habilidades intelectuales más recurrentes, según el análisis realizado son:

- reconocimiento
- identificación
- clasificación
- interpretación
- representación
- establecimiento de relaciones
- aplicación
- comprensión<sup>121</sup>

Todos, son conceptos vinculados a la activación del pensamiento para la aplicación y uso del conocimiento; lo que implica una concepción cognitiva de aprendizaje. Para los principales

<sup>121</sup> Todos estos, son procesos que actúan como operadores intelectuales sobre los conocimientos para transformarlos y generar nuevas estructuras cognitivas. Por ejemplo la identificación y reconocimiento, están vinculados a la observación que es “un proceso de identificación permanente en la interacción del sujeto con su medio ambiente” Sánchez 1999: 44). La clasificación es un proceso de análisis y síntesis que ayuda a poner orden a la existencia. La interpretación viene a ser el proceso mediante el cual se pueden extraer significados. (Cfr Raths, 1993). La representación actúa a manera de código, como una propiedad funcional y estructural del pensamiento, el establecimiento de relaciones y la aplicación, están vinculados con la comprensión, que puede considerarse como “un proceso cognitivo complejo de alto nivel que requiere la intervención de los sistemas de memoria y atencionales, de los procesos de codificación y percepción y en fin de operaciones inferenciales basadas en los conocimientos previos y en sutiles factores contextuales” (De Vega, 1984: 367).



representantes de la teoría cognitiva, el aprendizaje es visto como una activación e interrelación de elementos conceptuales para la resolución de problemas.<sup>122</sup>

Estamos ante un cambio de parámetros en la concepción de examen que presentan las pruebas internacionales en sus ítems. Lo relevante ahora, no es medir lo que se sabe o lo que se puede retener para el examen; el examen desde un enfoque internacional consiste en demostrar qué es lo que se puede y se sabe hacer<sup>123</sup> con el conocimiento adquirido y sobre todo como se utiliza, representa, organiza y relaciona, interpreta y transforman los contenidos del currículum escolar para resolver problemas de la vida cotidiana.

El análisis de algunos ítems muestra, ha permitido identificar el contenido y los principales conceptos estructurantes de los exámenes en el plano internacional.

## 2.2 Tipo de Ítems en los exámenes internacionales

Los siguientes ejemplos, publicados por las agencias encargados de elaborarlos,<sup>124</sup> justifican la forma en como en cada reactivo, se manifiestan los conceptos que hemos identificado

---

<sup>122</sup> "La resolución de problemas, se entiende como aquellas tareas que exigen procesos de razonamiento relativamente complejos y no una mera actividad asociativa y rutinaria...Normalmente existen tres fases en la solución de un problema: Preparación, producción y enjuiciamiento. La preparación supone un análisis e interpretación de los datos disponibles, la producción comprende operaciones diversas como recuperación de la información de la (memoria a largo plazo) MLP, y el enjuiciamiento evalúa la solución generada contrastándola con el criterio de solución"(cfr. De Vega, 1984: 494 y 495)

<sup>123</sup> Conviene Recordar que ya desde 1996 en el informe de Delors de la UNESCO se proponen como pilares de la educación : aprender a conocer, a ser, a hacer y a convivir. (Véase Delors, 1996)

<sup>124</sup> Conviene destacar que el hecho de que las Agencias de Evaluación internacional publiquen una muestra de los ítems utilizados en su exámenes, presenta una visión distinta de evaluación; sobre todo si se contempla que la mayoría de los exámenes nacionales, son manejados con absoluta confidencialidad, lo que les da un "carácter de misterioso secreto" imposible de revelar.



**a. Ítems de un examen del TIMSS.**

Para el análisis de los ítems muestra de un examen del TIMSS, se seleccionaron de los ítems publicados por esta agencia,<sup>125</sup> ítems de las dos áreas y poblaciones a las que va dirigido este estudio. Cada reactivo especifica, campo temático y dominio cognitivo o comportamiento al que pertenece.

**□ Ítem de matemáticas para cuarto grado (alumnos de 8-9 años)**

**Ejemplo 1\***

Juan quiso usar su calculadora para sumar 1463 y 319. El insertó  $1263 + 319$  por error. ¿Qué puede él hacer para corregir su error?

- A. Sumar 200
- B. Sumar 2
- C. Restar 2
- D. Restar 200

Contenido	Comportamiento esperado
Números	Resolución de problemas

\*Traducido de: IEA's 1995b, 45

En este reactivo el alumno tendrá recurrir a sus conocimientos de aritmética, específicamente los relacionados con el significado y propiedades de la suma y la resta, además de establecer las relaciones correspondientes. Se trata de un ejemplo que requiere tanto de conocimiento numérico, como de las habilidades o capacidades para resolver un problema.

<sup>125</sup> Ítems traducidos para este capítulo. La traducción fue realizada para poder comprender el contenido y estructura de los ítems, pero no constituye la versión tipificada al español (véase IEA's 1995b)



### Ejemplo 2\*

Aquí está un clip.



Aproximadamente ¿qué tan largo es el clip en comparación con el largo de esta línea?

\_\_\_\_\_

Responde: \_\_\_\_\_

Contenido	Comportamiento
Medición	Uso de procedimientos

\* Traducido de IEA'S 1995b, 57

Este reactivo, no sólo implica conocer el sistema de medida y lo que este significa, sino que además el alumno tendrá que observar, calcular y representar numéricamente una medida. Ahora se presentan, ítems muestra de un examen de matemáticas del TIMSS para alumnos que están por concluir la educación básica:



- Ítem de matemáticas para octavo grado,  
(alumnos de 13-14 años, que están por concluir la educación básica).

### Ejemplo 3\*

La tabla muestra una relación entre  $x$  y  $y$ . ¿Cuál es el número faltante en la tabla?

X	Y
1	1
2	?
4	7
7	13

A. 2  
B. 3  
C. 4  
D. 5  
E. 6

Contenido	Comportamiento
Álgebra	Procedimientos de rutina

\*Tomado de: Backhoff y Solano, 2003, p.139

Este es un reactivo que requiere de la lectura de patrones y de la aplicación de los procedimientos necesarios para establecer las relaciones correspondientes entre éstos.



### Ejemplo 4 \*

Estas figuras están colocadas en orden.



¿Cuál de los siguientes conjuntos de figuras tiene el mismo orden?

- A. \*□\*□\*□\*□\*□\*□□
- B. □\*□□\*□□□\*□□□□
- C. \*□\*□\*□□\*□\*□□□
- E. □□\*□\*□□\*□\*□\*

Contenido	Comportamiento esperado
Álgebra	Conocimiento

\*Tomado de: Backhoff y Solano, 2003, p. 152

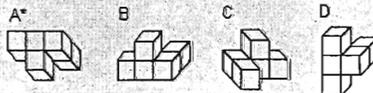
Ítem que demanda no sólo tener conocimiento de relaciones algebraicas, sino de encontrar la relación entre representaciones gráficas.

### Ejemplo 5\*

Vamos a girar esta figura a otra posición.



¿Cuál de las siguientes figuras podría ser la figura anterior después de ser girada?



Contenido	Comportamiento
Geometría	Procedimiento complejo

\*Tomado de: López Varona y Moreno Martínez, 1997.<sup>126</sup>

<sup>126</sup> Conviene destacar que estos reactivos, fueron extraídos de documentos elaborado por técnicos del INCE de España y publicados por el Ministerio Español a partir de los resultados obtenidos por este país en el TIMSS 1995. La difusión de este trabajo es una muestra de la preocupación de los especialistas en



En este reactivo el alumno tendrá que aplicar procesos complejos de representación y razonamiento para visualizar una figura en tres dimensiones.

Los ítems muestra de un examen de matemáticas del TIMSS tanto para alumnos de 8 y 9 años que se encuentran cursando el 4º grado de la educación básica seleccionados, como para los que están por concluir la educación básica, presentan características importantes para la evaluación matemática:

- ❖ El contenido y proceso cognitivo a evaluar, está inmerso en problemas cotidianos de numeración y medición, que requieren de la aplicación y transferencia de conocimientos y procedimientos matemáticos básicos y de uso común.
- ❖ Las capacidades requeridas para la solución de los ítems, están vinculadas al pensamiento lógico-matemático.
- ❖ Existe una clara vinculación entre el tipo de ítems de las pruebas de matemáticas del TIMSS y la definición que TIMSS da a las matemáticas, concebidas por esta agencia como capacidad de pensamiento lógico que implica la transferencia de conocimientos y destrezas a nuevas situaciones.

La siguiente área de evaluación de los exámenes del TIMSS es la ciencia:

---

evaluación de España, no sólo por presentar el lugar ocupado por su país en las evaluaciones internacionales; sino por dar a conocer, con fines de mejora las fallas del currículum de la Educación Básica en España.



□ Ítem de ciencias para 4° grado (alumnos de 8 y 9 años)

**Ejemplo 6\***

El siguiente cuadro muestra información acerca de la temperatura y la precipitación (lluvia o nieve) en cuatro diferentes ciudades el mismo día.

	Ciudad A	Ciudad B	Ciudad C	Ciudad D
Temperatura más baja	13° C	-9° C	22° C	-12° C
Temperatura más alta	25° C	-1° C	30° C	-4° C
Precipitación (lluvia o nieve)	0 cm	5 cm	2.5 cm	0 cm

¿Dónde nevó?

- A. Ciudad A
- B. Ciudad B
- C. Ciudad C
- D. Ciudad D

Contenido	Comportamiento esperado
Ciencias de la tierra	Uso de herramientas, procedimiento de rutinas y procesos científicos

\* Tomado de Backhoff y Solano, 2003., p. 86

En este ítem el alumno tendrá que localizar, contrastar, relacionar y aplicar su conocimiento para un reconocimiento de las diferentes temperaturas y lo que éstas significan ambiental y climáticamente.

Este es un reactivo que permite observar cómo los exámenes de ciencias del TIMSS para alumnos que cursan el cuarto grado de la educación básica, presentan información de uso común que el alumno tendrá que explicar, recurriendo a ciertos datos y utilizando herramientas y procedimientos cognitivos que le permitan demostrar lo que sabe con respecto a la naturaleza y el mundo físico.

A continuación, una muestra de los ítems del examen de ciencias para el octavo:



- Ítem de ciencias para octavo grado: (alumnos de 13 y 14 años que están por concluir la educación básica)

**Ejemplo 7\***

¿Cuál es la mejor razón para incluir frutas y vegetales en una dieta saludable?	
A. Tienen un alto contenido de agua	
B. Son la mejor fuente de proteínas	
C. Son ricos en minerales y vitaminas	
D. Son la mejor fuente de carbohidratos	
<b>Contenido</b>	<b>Comportamiento esperado</b>
Ciencias de la vida	Comprensión de información simple

\*Tomado de Backhoff y Solano, 2003., p. 204

Reactivo relativamente sencillo en el que el alumno tendrá que demostrar su comprensión de conceptos sobre nutrición para explicar el valor alimenticio de las frutas y verduras.

**Ejemplo 8\***

¿ Por qué existe el día y la noche? Indica qué opción es la correcta	
A. La tierra gira sobre su propio eje	
B. El sol gira sobre su propio eje	
C. El eje de la tierra está inclinado	
D. La tierra se mueve alrededor del sol	
<b>Contenido</b>	<b>Comportamiento esperado</b>
Ciencias de la tierra	Comprensión de información compleja

\* Tomado de Backhoff y Solano, 2003., p. 251

Para responder acertadamente a esta pregunta, el alumno tendrá que demostrar su comprensión de procesos de rotación de la tierra para explicar un fenómeno natural. Nótese que no se pregunta en qué consiste dichos movimientos; más bien se pide que aplique la información obtenida al respecto y la relacione con un hecho físico.



El análisis de la constitución de los ítems para los exámenes de ciencias del TIMSS, tanto en alumnos que cursan el 4° grado, como en los que están por concluir la educación básica, da cuenta de que:

- ❖ Los exámenes de ciencias de TIMSS para los grados de cuarto y octavo, incluyen, en algunos de sus ítems, hechos y fenómenos cotidianos, relacionadas con la vida del estudiante y el medio ambiente, que requieren del uso de procesos cognitivos que permitan la representación, abstracción y aplicación de un conocimiento científico para explicar lo que ocurre a su alrededor.
- ❖ Las características de los ítems de un examen del TIMSS, muestran una estrecha relación con el concepto establecido por TIMSS para la evaluación de las ciencias, conceptualizadas como procesos utilizados para aprender sobre el mundo físico.

El análisis de la constitución de ítems en las pruebas internacionales, permite identificar tres nociones básicas en el concepto de examen: conocer, rendimiento y razonamiento; las cuales, ofrecen un enfoque distinto para la evaluación del aprendizaje.

Los ítems de los exámenes del TIMSS dan cuenta de que:

- Conocer es utilizar tanto el pensamiento como el conocimiento.
- Rendimiento es igual a comprensión de contenidos temáticos.
- El razonamiento debe ser utilizado para trasladar el conocimiento a situaciones problemas, medio principal de la enseñanza y la evaluación.

Los exámenes del TIMSS tanto de matemáticas como de ciencias, presentan otra manera de concebir los contenidos de aprendizaje y su evaluación, pues no se trata sólo de medir la información, sino de las capacidades implícitas en el proceso de conocimiento

Remitimos ahora la experiencia Latinoamericana.



## b. Ítems de los exámenes del LLECE

Para el análisis de los ítems de las pruebas de matemáticas y lenguaje para 3° y 4° grado del LLECE, se seleccionaron ítems de cada uno de los niveles de desempeño establecidos por esta agencia.

### □ Ítem de matemáticas para tercero y cuarto grado:

**Ejemplo 9\***

Nivel de desempeño I:

**Reconocimiento y empleo de hechos y relaciones matemáticas:**

Cuatro amigos escribieron números en la pizarra

Lucía escribió	820
María escribió	802
Ana escribió	811
Raúl escribió	801

¿Quién escribió el número mayor?

- B) Lucía
- C) María
- D) Ana
- E) Raúl

**Contenido:** Numeración

\*Tomado de LLECE, 2001, p. 204

Este ítem, implica el reconocimiento y establecimiento de relaciones entre cantidades y requiere del dominio de conocimientos sobre sistema de numeración decimal y el valor absoluto y relativo de las cifras en cuanto a su valor posicional, como de la aplicación de procedimientos para establecer series numéricas.



### Ejemplo 10\*

Nivel de desempeño II: <b>Reconocimiento y utilización de estructuras matemáticas simples.</b>
La diferencia de dos números es 300 el número menor es igual a 500. El otro número es igual a
a) 200
b) 300
c) 500
d) 800
<b>Contenido:</b> Operaciones con números naturales

\*Tomado de LLECE, 2001, p. 199

En este ítem el alumno tendrá que encontrar relaciones entre datos y aplicar procedimientos para establecer estructuras matemáticas simples. Es un reactivo con una tarea semejante a la anterior, sólo que incluye el empleo de dos procedimientos aritméticos (suma y resta) para identificar las diferencias entre cantidades.

### Ejemplo 11\*

Nivel de desempeño III: <b>Reconocimiento y utilización de estructuras matemáticas complejas.</b>
Consuelo leyó la mitad de un libro de 40 páginas. Yolanda leyó la cuarta parte de un libro de 20 páginas. Eugenio leyó lo mismo que Consuelo y Yolanda juntas.
¿ Cuántas páginas leyó Eugenio?
a) 60
b) 50
c) 25
d) 20
<b>Contenido:</b> Operación con números naturales

\*Tomado de LLECE, 2001, p. 201

Este es un ítem que permite reconocer que el grado de complejidad y nivel de desempeño al que pertenece, obedece a que requiere de la activación de procesos de razonamiento y del



uso de procedimientos para establecer relaciones entre datos, que implica mayor grado de complejidad.

El análisis de los ítems de un examen de matemáticas del LLECE da cuenta de que :

- ❖ El nivel de desempeño al que pertenecen cada uno de estos tres ejemplos, está en gran parte determinado por el tipo y nivel de proceso y procedimiento cognitivo que requiere.
- ❖ Los procesos cognitivos implícitos en cada uno de los ítems, presenta una relación importante entre lo que el LLECE concibe como competencia matemática y las tareas establecidas para su evaluación; en donde lo que se requiere es del uso de procedimientos y de la activación del pensamiento lógico matemático, lo que coincide con el concepto, propósito y dimensiones que el LLECE establece como contenido de la evaluación matemática.

#### □ Ítem de Lenguaje para tercer y cuarto grado

Dos son los campos de evaluación que abordan de las pruebas del Lenguaje del LLECE: comprensión lectora y producción escrita.

El análisis de los exámenes de Lenguaje del LLECE, se basa en ítems de lectura de cada uno de los tres niveles determinados por le LLECE :

<b>Ejemplo 12*</b>	
<b>Nivel de desempeño 1</b>	
<b>Lectura literal primaria</b>	
Si el sol _____ al colibrí, se ven sus colores.	
a. Ilumina	
b. Pinta	
c. Lleva	
d. mira	

\*Tomado de LLECE, 2001, p. 195



Este es un ítem que requiere una identificación semántica de la información presentada en el texto, para lo cual, el alumno tendrá que reconocer estructuras gramaticales y las relaciones que entre estas existen.

**Ejemplo 13\***

<b>Nivel de desempeño II</b> <b>Lectura a modo de paráfrasis</b>
¿Cuál es el mensaje del poema "El patio de mi escuela?" a. Nunca hay basura en el patio b. Hay que cuidar la limpieza del patio c. No hay que correr en el patio d. Hay que tirar la basura en el patio

\*Tomado de LLECE, 2001, p. 196

En este ítem, el alumno tendrá que realizar una apropiación semántica, de acuerdo a lo que se dice en un texto.

**Ejemplo 14\***

<b>Nivel de desempeño III</b> <b>Lectura de carácter inferencial</b>
El perrito de Timoteo está: a. Alegre b. Contento c. Enojado d. Triste

\*Tomado de LLECE, 2001, p. 194

Con este ítem, lo que se pretende es la realización de supuestos por parte del alumno para inferir e interpretar mensajes no explícitos en un texto.

Los ítems analizados, dan cuenta de que para la evaluación de la lectura, los exámenes del LLECE:



- ❖ Presentan una estructura cuyos niveles de desempeño están determinados por los requerimientos semánticos del tipo de lectura que se pretende realice el alumno. En el primer nivel de desempeño se habla de una traducción semántica, en el segundo de una apropiación y en el tercero de una interpretación.
- ❖ Los ítems que el LLECE emplea para la evaluación de la lectura, demandan la ejecución de procesos de comunicación en distintos niveles, que le implica al alumno demostrar sus posibilidades para encontrar el significado a signos lingüísticos y no lingüísticos; que en sí es lo que el LLECE define como competencia comunicativa.

Los conceptos encontrados en el contenido y estructura de los ítems de las pruebas del LLECE dan cuenta de que:

- Conocer consiste en un proceso de transformación y desarrollo que se evalúa mediante el uso de procedimientos y posibilidades intelectuales, a las que llaman competencia.
- El rendimiento está asociado a la aptitud, definida en términos de potencia como "posibilidad."
- El razonamiento está asociado al pensamiento lógico que requiere de la activación de procesos o "posibilidades cognitivas."

### c. Ítems del examen del PISA

De los exámenes del PISA, un examen para jóvenes de 15 años que están por concluir la educación básica, se presentan ítems muestra de cada una de las tres áreas que aborda el estudio; lo que permitirá dar cuenta tipo de ítems que emplea esta agencia para la evaluación de la lectura, las matemáticas y las ciencias y de los conceptos implícitos en su estructura y contenido.

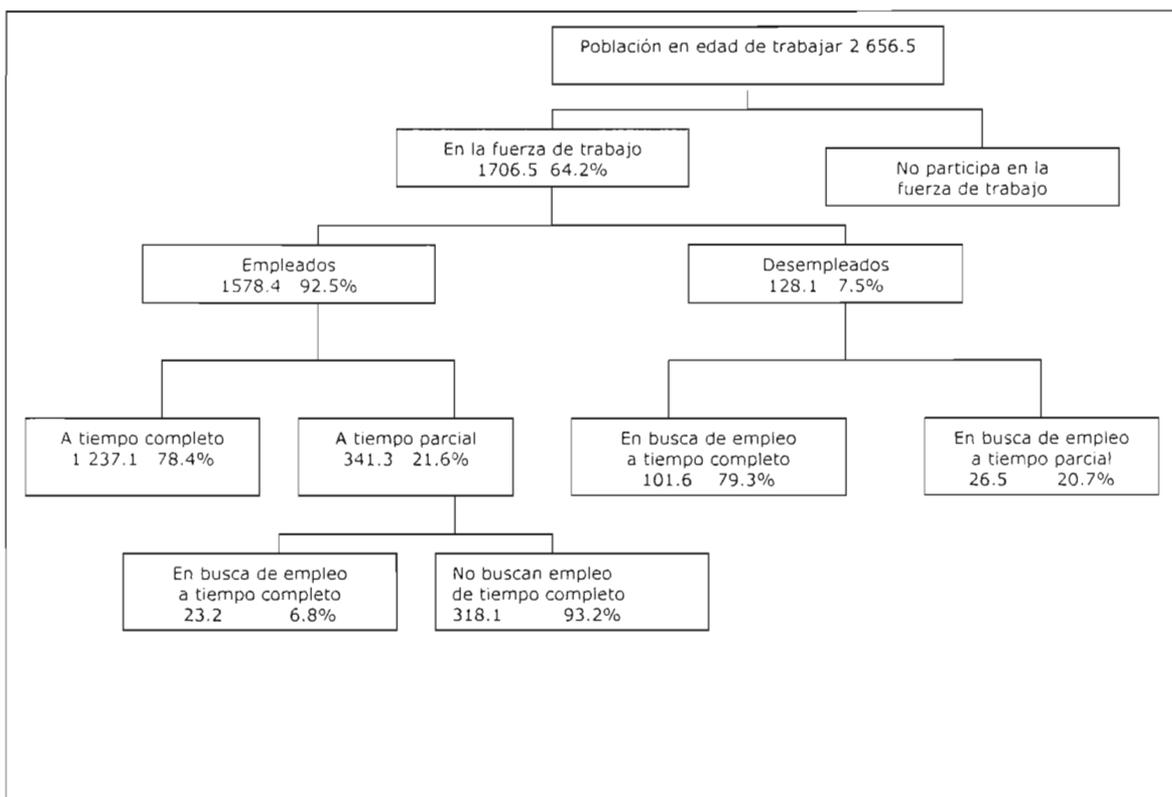
Los exámenes de lectura del Proyecto OCDE/PISA para jóvenes de 15 años que están por concluir la educación básica están compuestos por ítems del siguiente tipo:



□ Ítem de lectura para jóvenes de 15 años:

**Ejemplo 15\***

El diagrama que ves muestra la estructura de la fuerza de trabajo de un país o la "población en edad de trabajar". La población total del país en 1995 era aproximadamente de 3.4 millones de personas





1. ¿Cuántas personas en edad de trabajar no se encontraban en la fuerza de trabajo?

**Aspecto que evalúa: Obtención de información**

2. ¿Cuáles son los dos principales grupos entre los que se divide la población en edad de trabajar?

- A. Empleados y desempleados.
- B. En edad de trabajar y no en edad de trabajar.
- C. Empleados a tiempo completo y a tiempo parcial.
- D. En la fuerza de trabajo y fuera de la fuerza de trabajo.

**Aspecto que evalúa: Interpretación de textos:**

3. La información sobre la estructura de la fuerza de trabajo se presenta como un diagrama jerárquico, pero podría haberse presentado de varias otras maneras, tales como una descripción escrita, una gráfica de pastel, una gráfica simple o un cuadro. El diagrama jerárquico se eligió probablemente porque es particularmente útil para mostrar:

- A. Los cambios en el tiempo
- B. El tamaño de la población total de un país
- C. Categorías dentro de cada grupo
- D. El tamaño de cada grupo

**Aspecto que evalúa: Reflexión y evaluación**

En este reactivo, las preguntas que lo componen demandan no sólo el reconocimiento e información, sino el establecimiento de relaciones que permita demostrar la comprensión e interpretación del contenido del texto y la reflexión sobre su estructura.

Este reactivo muestra da cuenta de que:

- ❖ La evaluación de la lectura en las pruebas del PISA, tiene un enfoque funcional.
- ❖ Los jóvenes de 15 años deben darle un uso personal y eficaz a la lectura, mediante una comprensión, empleo y reflexión de diversos tipos de textos; lo que en sí corresponde al concepto de lectura elaborado por PISA.

Algunos ítems de los exámenes de matemáticas en el Proyecto OCDE/PISA se presentan a continuación:



□ Ítem de matemáticas para jóvenes de 15 años:

Los ítems seleccionados de los exámenes del PISA para su análisis, pertenecen a temas similares a los de las pruebas del TIMSS, con miras a una posible comparación entre los mismos. El primero de éstos:

**Ejemplo 16\***  
**Manzanas**

Un agricultor planta manzanos en un patrón cuadrado. Para proteger su árboles contra el viento, planta coníferas alrededor de la huerta.

A continuación puedes ver un diagrama de esta situación donde puedes observar el patrón de manzanos y coníferas para cualquier número (n) de filas de manzanas:

x = coníferas

• = manzano

n = 1    x x x  
          x • x  
          x x x

n = 2    x x x x x  
          x • • x  
          x        x  
          x • • x  
          x x x x x

n = 3    x x x x x x x  
          x • • • x  
          x        x  
          x • • • x  
          x        x  
          x • • • x  
          x x x x x x x

n = 4    x x x x x x x x  
          x • • • • x  
          x        x  
          x • • • • x  
          x        x  
          x • • • • x  
          x        x  
          x • • • • x  
          x x x x x x x x

\*Tomado de OCDE, 2002, p. 105



## Pregunta 1

Completa la siguiente tabla

n	Número de manzanos	Número de coníferas
1	1	8
2	4	16
3	9	24
4	16	32
5	25	40

## Pregunta 2

Existen dos fórmulas que puedes emplear para calcular el número de manzanos y el número de coníferas para el patrón descrito arriba:

$$\text{Número de manzanos} = n^2$$

$$\text{Número de coníferas} = 8n$$

Donde  $n$  es el número de filas de manzanos.

Existe un valor de  $n$  para el cual el número de manzanos es igual al número de coníferas

Encuentra el valor de  $n$  y muestra el método para calcularlo.

## Pregunta 3

Supón que el agricultor desea agrandar mucho más su huerta y sembrar varias filas de manzanos. Conforme el agricultor agranda su huerta, ¿cuál crecerá más rápido el número de manzanos o el no de coníferas? Explica cómo encontraste tu respuesta.

Contenido (Idea):	Procesos (Comportamientos):
Cambio y relaciones	Preguntas 1 y 2: Conexiones e integración para la solución de problemas Pregunta 3: Construcción matemática, pensamiento matemático, generalización y planteamiento de ideas.

Este es un multireactivo,<sup>127</sup> en donde las preguntas que lo conforman, contienen diferente grado de complejidad: En la pregunta 1, los alumnos deberán a partir de un patrón dado, interpretar la información presentada y encontrar las relaciones correspondientes. En la pregunta 2, los alumnos necesitarán, interpretar expresiones que contengan palabras y símbolos y vincular diferentes representaciones (gráfica, verbal y algebraica) en dos relaciones (cuadrática y lineal), además de usar su razonamiento para explicar su respuesta.

<sup>127</sup> Se conoce como multireactivo, a los ítems que dependen de un texto fuente, del cual se derivan varias preguntas.



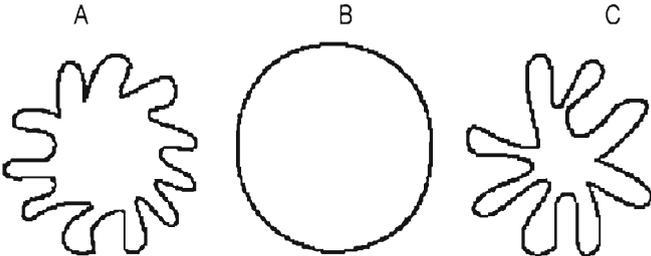
En la pregunta 3, se requiere demostrar la comprensión de las funciones matemáticas mediante la comparación de una función lineal con una cuadrática. Se trata de una pregunta que implica un pensamiento matemático, en donde el estudiante deberá reconocer y extraer las matemáticas incluidas en una situación dada, para el planteamiento y resolución de problemas, en situaciones que requieren del análisis, la interpretación y el desarrollo de los propios modelos matemáticos, que incluyan la demostración y generalización.

Otro reactivo con un contenido diferente:

**Ejemplo 16\***  
**Las figuras**

¿Cuál de las figuras tiene mayor área?  
Muestra tu razonamiento

A B C



Contenido (Idea)	Procesos (comportamiento)
Espacio y forma	Reproducción y cálculo

\*Tomado de OCDE, 2001, p. 95

Este reactivo requiere de procesos de reproducción y cálculo y del recuerdo de ciertas expresiones y procedimientos de medida relacionados con el espacio y la forma. (geometría).

Los ítems de un examen de matemáticas del PISA para jóvenes de 15 años, dan cuenta de que:

- ❖ Las matemáticas en los exámenes del PISA, son evaluadas con ítems que demandan capacidades de análisis, interpretación y representación de situaciones matemáticas que se puedan encontrar en la vida diaria.



Los exámenes de ciencias del proyecto OCDE/PISA contemplan el siguiente tipo de ítems:

□ **Ítem de ciencias para jóvenes de 15 años**

**Ejemplo 17\***

Lee el siguiente resumen de un artículo del periódico Daily Mail del 30 de marzo de 1998 y responde a las preguntas que siguen:

Un artículo de periódico contaba la historia de una estudiante de 22 años, llamada Jessica, que siguió «una dieta basada en el chocolate». Pretendía mantenerse saludable, con un peso estable de 50 kilos, mientras comía 90 barras de chocolate a la semana y prescindía del resto de la comida, con la excepción de una «comida normal» cada cinco días. Una experta en nutrición comentó: «Estoy sorprendida de que alguien pueda vivir con una dieta como ésta. Las grasas le proporcionan la energía necesaria para vivir, pero no sigue una dieta equilibrada. En el chocolate existen algunos minerales y nutrientes, pero no obtiene las vitaminas suficientes. Más adelante, podría sufrir serios problemas de salud.»

En un libro en el que se habla de valores nutricionales se mencionan los siguientes datos acerca del chocolate. Supón que todos estos datos son aplicables al tipo de chocolate que come, frecuentemente, Jessica. También, considera que cada barra de chocolate que come tiene un peso de 100 gramos.

**Tabla I**  
**Contenido nutritivo de 100 g de chocolate**

Proteínas	Grasas	Hidratos de carbono	Minerales		Vitaminas			Energía total
			Calcio	Hierro	A	B	C	
5 g	32 g	51 g	50 mg	4 mg	-	0,20 mg	-	2142 kJ

Pregunta No. 1

Según los datos de la tabla, 100 gramos de chocolate contienen 32 gramos de grasas y proporcionan 2142 kJ de energía. La nutricionista afirmó: «Las grasas le proporcionan la energía para vivir...». Si alguien come 100 gramos de chocolate, ¿toda su energía (2142 kJ) procede de los 32 gramos de grasas? Explica tu respuesta utilizando los datos de la tabla.

Pregunta No. 2

Los expertos en nutrición afirman que Jessica «... no obtiene las vitaminas suficientes». Una de esas vitaminas que no contiene el chocolate es la vitamina C. Quizás podría compensar esta carencia de vitamina C incluyendo algún alimento que contenga un alto porcentaje de vitamina C en «la comida normal que hace cada cinco días».



Aquí tienes una lista de tipos de alimentos.

1. Pescado.
2. Fruta.
3. Arroz.
4. Vegetales.

¿Qué dos tipos de alimentos, de los que aparecen en esta lista, recomendarías a Jessica para que pudiera compensar la carencia de vitamina C?

- A. 1 y 2
- B. 1 y 3
- C. 1 y 4
- D. 2 y 3
- E. 2 y 4
- F. 3 y 4

Contenido (Idea)	Procesos (comportamientos)
Ciencias de la vida y de la salud	Mostrar comprensión de conceptos científicos

\* Tomado de OCDE, 2001,p.137

Se trata de otro multireactivo, cuyas preguntas que lo componen requieren no sólo demostrar comprensión de conceptos sobre nutrición, sino aplicarlos y darles sentido a una situación real apoyándose en la consulta de otras fuentes.

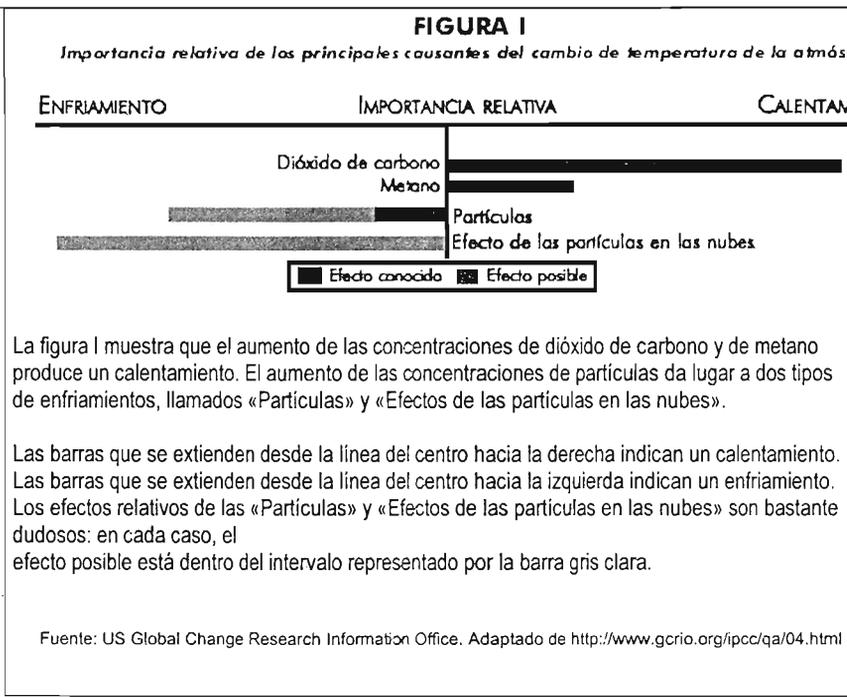
### Ejemplo 18 \*

Lee la siguiente información y contesta las preguntas que aparecen a continuación.

¿Qué actividades humanas contribuyen al cambio climático?

La combustión del carbón, la gasolina y el gas natural, así como la deforestación y diversas prácticas agrícolas e industriales, están alterando la composición de la atmósfera y contribuyendo al cambio climático. Estas actividades humanas han llevado a un aumento de la concentración de partículas y gases del efecto invernadero, en la atmósfera.

La importancia relativa de los principales causantes del cambio de temperatura se presenta en la figura 1.



Utiliza la información de la figura No. 1 para desarrollar un argumento que apoye la reducción de la emisión de dióxido de carbono por las actividades has mencionadas.

Contenido (idea)	Procesos (comportamiento)
Ciencias de la tierra y del medio ambiente (tiempo y clima) Cambio atmosférico	Comunicar conclusiones validas

\*Tomado de OCDE, 2001,p.133

Reactivo que requiere de una comprensión del concepto de cambio atmosférico, para extraer y comunicar conclusiones validas.

Estos ítems dan cuenta de que en los exámenes de ciencias del PISA:

- ❖ Los ítems analizados para la evaluación de las ciencias, presentan una estrecha vinculación entre la concepción de ciencia del Proyecto PISA y las tareas y procesos requeridos para su evaluación.



- ❖ La estructura y contenido de los ítems evalúan no la reproducción de un conocimiento o información, sino la aplicación de este para elaborar conclusiones que expliquen los cambios ocurridos en la naturaleza. Lo importante en los exámenes del PISA, no es que los estudiantes repitan conceptos, teorías o fórmulas, sino que utilicen su conocimiento y muestren la comprensión de conceptos científicos para la explicar y argumentar hechos y situaciones que suceden en la vida real; las tareas requeridas son de razonamiento y comprensión, más que de retención y acumulación de hechos y principios.

Los conceptos encontrados en el análisis de los ítems de los exámenes del PISA, dan cuenta de que:

- El conocimiento evaluado por los exámenes del PISA es de tipo funcional.
- El razonamiento es utilizado para la solución de problemas de la vida cotidiana
- El rendimiento es valorado como la aplicación de conocimientos y aptitudes a situaciones auténticas.

Las matemáticas y ciencias como área de evaluación común en los exámenes del TIMSS, del LLECE y del PISA, permite comparar el contenido de los ítems de un examen de Matemáticas y Ciencias tanto para alumnos de cuarto grado, como para los que están por concluir la educación básica.

### **2.3 Contenido de los Ítems de matemáticas y ciencias de los exámenes internacionales para 4° grado y secundaria.**

En un intento por hacer una síntesis y comparación de la composición de los ítems de matemáticas y ciencias, que permita identificar la manera en que se evalúan temas y áreas comunes en las distintas pruebas internacionales, se elaboró el siguiente cuadro:



CUADRO 2.5

**CONTENIDO DE LOS ÍTEMS  
DE MATEMÁTICAS Y CIENCIAS DE LAS PRUEBAS INTERNACIONALES**

Examen	Área	Grado	Contenido	Procesos
TIMSS	M	4°	Números (Suma y resta)	Resolución de problemas de aritmética
			Medición (Sistema de medidas)	Uso Procedimientos: Observación, calculo y representación numérica
LLECE	M	4°	Numeración	Reconocimiento y establecimiento de relaciones entre cantidades Aplicación de procedimientos para establecer series numéricas
			Operaciones con números naturales	Relaciones entre datos y aplicar procedimientos para establecer estructuras matemáticas simples. Activación de procesos de razonamiento (Nivel III)
TIMSS	M	Fin Ed. Básica	Relaciones algebraicas	Uso de procedimientos: Lectura de patrones Encontrar relación entre representaciones gráficas
			Geometría (Figuras en tres dimensiones)	Procesos de representación y razonamiento
PISA	M	Fin Ed. Básica	Cambio y relaciones	Lectura de patrones Interpretación de información en palabras y símbolos Encontrar relaciones cuadráticas y lineales Vincular diferentes representaciones Explicar su respuesta mediante un razonamiento Comprender funciones matemáticas, mediante la comparación de una función lineal con una cuadrática. Reconocimiento y extracción de las matemáticas incluidas en una situación dada, para el planteamiento y resolución de problemas
			Geometría Espacio y forma	Reproducción, cálculo y representación Recuerdo de ciertas expresiones y procedimientos de medida relacionados con el espacio y la forma
TIMSS	C	Fin Ed. Básica	Ciencias de la vida: Nutrición	Comprensión de conceptos de nutrición
			Ciencias de la tierra	Comprensión de procesos de rotación para explicar un fenómeno natural
PISA	C	Fin Ed. Básica	Ciencias de la vida y la salud: Nutrición	Aplicación y comprensión de conceptos
			Ciencias de la tierra	Comprensión de conceptos



El cuadro establece que:

- Los temas numéricos de matemáticas en el 4<sup>a</sup> grado de la educación básica, se evalúan con ítems que comprenden uso de procedimientos aritméticos.
- Las relaciones algebraicas en alumnos que están por concluir la educación básica, se evalúan mediante ítems que contemplan la lectura de patrones y el establecimiento de relaciones cuadráticas y lineales.
- La geometría en cuanto a espacio y forma en exámenes para alumnos que se encuentran en los últimos grados de la educación básica se evalúa con ítems que consideran la representación como operación común .
- La evaluación de las ciencias al final de la escuela básica, se evalúa mediante ítems que requieren de una comprensión y aplicación de conceptos científicos más que de la reproducción de los mismos.

La tercera categoría establecida para el análisis de los exámenes de corte internacional, se refiere a la teoría de medición.



## 2.4. Modelo psicométrico en los exámenes internacionales.

Todo examen estandarizado, está sustentado en una teoría de medición o teoría del test.<sup>128</sup>

Como se ha venido señalando, dos son las teorías de medición aplicadas a la elaboración de pruebas estandarizadas: Teoría Clásica del Test (TC)<sup>129</sup> y Teoría de Respuesta al ítem (TRI).<sup>130</sup>

El análisis del diseño de los ítems en el capítulo anterior, presenta a la TRI, como la teoría en que se fundamenta la selección y calificación de ítems de las Pruebas Internacionales.

La Teoría de la Respuesta al Ítem, se concibe como un modelo alternativo psicométrico al que recurren las pruebas internacionales, porque permite obtener mediciones invariantes<sup>131</sup> y

---

<sup>128</sup> La teoría de la medición o teoría del test, es la metodología utilizada para estimar errores y mostrar fiabilidad; se trata de pruebas elaboradas con rigor. Una teoría del test proporciona modelos para hacer las puntuaciones de una prueba, construye 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.

<sup>129</sup> La teoría clásica de los test, es la teoría propuesta a principios de siglo por Spearman y consiste en que la puntuación empírica que obtiene un sujeto cuando se le administra un test (X), está en función del nivel real o verdadero en que el sujeto posee la característica o rasgo que se está evaluando (V) y el error de medida que siempre se introduce en cualquier proceso de medición (E); (es decir, cuántos contestaron bien el ítem). De ahí que no considere diferencias individuales ni niveles de dificultad entre los ítems. Los constructos o propiedades del ítem son tratados de forma generalizada.

<sup>130</sup> El propio nombre de Teoría de Respuesta al Ítem alude a que bajo la óptica de la TRI, la unidad básica de análisis es el ítem y no el test. En la Teoría Clásica, el índice de dificultad o su índice de discriminación, índices con los que se describe a los ítems, depende de la muestra de personas utilizadas para estimarlo. La TRI, permite realizar mediciones invariantes, con respecto al instrumento y a las personas seleccionadas y es más resistente a las variaciones en cuanto al número de sujetos y reactivos, lo que le da mayor coeficiente de consistencia interna. (Muñiz, 1997). En el modelo clásico, por su característica lineal, las personas examinadas se pueden ver afectadas por errores aleatorios de origen diverso. (Instrumento de medida, individuo y ambiente externo) (Bachoff, 2003: 58-61).

<sup>131</sup> Cuando hablamos de mediciones invariantes, nos referimos a la invarianza de parámetros, considerados como "el establecimiento de una medida constante tanto en los niveles de ítems como de los sujetos" (Martínez y Moreno 2002: 105) Es decir, la TRI permite obtener estimaciones de los parámetros de los ítems y estimaciones de la variable latente independientemente tanto de la muestra de examinados como del instrumento de medición (Muñiz, 1997: 53- 55).



representar con mayor precisión lo que se mide en las diferentes porciones de un examen. (Muñiz, 1997:15).

De acuerdo con Thorndike y Hagen (1978), en la teoría clásica, tres son los aspectos a tomar en cuenta para la construcción de los ítems de un examen: atributos, operaciones, y cantidades. Los *atributos* son considerados como las cualidades del objeto a medir, las *operaciones* vienen a ser los indicadores mediante los cuales se pueden percibir o manifestar y las *cantidades* consisten en traducir las observaciones en cuestiones cuantitativas. Es decir, que el proceso de medición se compone de: objetos que son los hechos o fenómenos empíricos que pueden identificarse de manera individual y concreta, de propiedades de esos objetos, que vienen a ser los constructos psicológicos y de medidas, que consisten en la asignación de una puntuación numérica a las respuestas dadas por los sujetos, en los ítems que componen una prueba psicológica (Martínez y Moreno, 2002).

En las pruebas internacionales, cada uno de los ítems utilizados presentan de manera explícita lo que se pretende medir, pues definen el objeto y sus atributos, al conceptualizar y especificar tanto los contenidos y rendimientos esperados, como los indicadores mediante los cuales se va a medir el objeto, es decir las operaciones y constructos<sup>132</sup> que vienen a ser las habilidades y procedimientos implícitos en los contenidos, como tareas a realizar; y el puntaje asignado a cada uno de los ítem.

Cabe destacar que el aprendizaje, no es algo tangible que se pueda medir de manera general; es necesario que los exámenes y ítems para tal efecto, consideren y especifiquen que es lo que se va a medir y definan las propiedades del objeto y a los operadores del mismo, mediante el establecimiento de constructos. Es decir que en todo modelo de medición, el qué, define el cómo de ésta.

Una característica importante de los exámenes internacionales es que en sus ítems se encuentra especificado tanto los atributos del objeto a medir, como las operaciones (indicadores) y propiedades (constructos) y las cantidades (puntaje).

---

<sup>132</sup> En el proceso de medición se denomina constructo a las propiedades de los objetos que se van a medir.



## 2.5. Conclusiones preliminares: Lo que se refiere de las Pruebas Internacionales para la elaboración de exámenes

En este contexto, lo que se refiere de las Pruebas Internacionales para la elaboración de exámenes, se enuncia a continuación:

- ↓ Definir el contenido de evaluación en estrecha relación con el tipo de ítems.
- ↓ Elaborar ítems centrados en la solución de problemas de la vida diaria que expresen una fuerte combinación de conocimiento y procesos de pensamiento<sup>133</sup> para evaluar la aplicación y uso del conocimiento, mas que la mera reproducción de datos, hechos o fórmulas.
- ↓ Incluir modelos y teorías actuales como la TRI para la elaboración de exámenes estandarizados.

Estos son los elementos más importantes a considerar para la elaboración de exámenes, dentro del concepto de examen que manejan las Pruebas Internacionales. La comparación con la experiencia mexicana se presenta en el siguiente capítulo.

---

<sup>133</sup> Para información sobre evaluación de procesos cognitivos, (véase Díaz Barriga Arceo y Hernández Rojas, 2002).

## CAPÍTULO TRES

### LA EXPERIENCIA MEXICANA EN PRUEBAS NACIONALES

---

El propósito del capítulo es ofrecer un panorama de la situación que guardan las pruebas nacionales en relación a las internacionales, mediante la descripción y análisis de los exámenes masivos de mayor cobertura. Se presentan los orígenes, propósitos, contenidos y ítems de cada experiencia, para posteriormente compararlas entre sí y con las internacionales.

#### **3.1. Exámenes masivos para la educación básica: Algunas pruebas que se aplican en los salones de clase de la escuela primaria mexicana.**

La experiencia de evaluación en México, iniciada desde hace tres décadas, ha tenido como propósitos: la verificación de objetivos y contenidos curriculares y la asignación de recursos. Su creación y aplicación, ha estado vinculada a considerar y conocer el logro obtenido por los alumnos en cuanto a lo establecido en los planes y programas de estudio; a comprobar la efectividad de los programas compensatorios y de estímulos; así como a la selección y diagnóstico de las posibilidades de éxito de los estudiantes; con un carácter meramente sumativo y fiscalizador, que presenta la enorme necesidad de una evaluación formativa que ayude a mejorar los procesos de enseñanza-aprendizaje en la educación básica.



La experiencia mexicana en cuanto a elaboración de pruebas, de acuerdo con los informes oficiales publicados en materia de evaluación,<sup>134</sup> se puede resumir en ocho experiencias,<sup>135</sup> ocurridas principalmente durante la década de los noventa, década que marca un auge tanto a nivel nacional como en la Región Latinoamericana y del Caribe, en materia de Sistemas Nacionales de Evaluación.<sup>136</sup>

Para efectos de posibilitar la comparación con las pruebas internacionales, hemos seleccionado sólo cuatro de los exámenes masivos nacionales de mayor cobertura.

Tres aplicados en la educación básica:

1. Instrumento de Diagnóstico para Alumnos de Nuevo Ingreso a Secundaria (IDANIS)
2. Pruebas del Sistema Nacional de Evaluación Educativa (SNEE).
3. Pruebas con base en estándares(PE) ahora Pruebas Nacionales (PN)

Uno más, dirigido a alumnos que concluyen la educación básica y están por ingresar a la educación media superior.

4. Examen Nacional de Ingreso a la Educación Media Superior  
(EXANI 1)

De los cuales presentamos sus principales características:

---

<sup>134</sup> Los informes consultados fueron: (SEP, 2002; SEP s/f; INEE, 2003 y 2004).

<sup>135</sup> Seis de estas experiencias han sido realizadas por la Dirección General de Evaluación (DGEE) de la SEP desde sus inicios en 1970 hasta el 2000, una corresponde a las Pruebas Nacionales (PN) aplicadas por el Instituto Nacional de Evaluación (INEE) en el ciclo escolar 2002 y las del Examen de Ingreso a la Educación Media Superior (EXANI-1) aplicadas por el Centro Nacional de Evaluación de la Educación Superior (CENEVAL).

<sup>136</sup> En Latinoamérica, los Sistemas de Nacionales de Medición, comenzaron a desarrollarse a partir de 1988. Argentina fue el primer país en constituir en 1993 un Sistema Nacional de Evaluación. (Véase Arancibia, 1997).



## 1. Instrumento de Diagnóstico para Alumnos de Nuevo Ingreso a Secundaria (IDANIS)

### ▪ Orígenes del IDANIS

El Instrumento de Diagnóstico para Alumnos de Nuevo Ingreso a Secundaria, ha estado a cargo de la Dirección General de Evaluación Educativa (DGEE), dependencia oficial de la Secretaría de Educación Pública (SEP), que durante 30 años (1972-2002), fue la instancia responsable de llevar a cabo los estudios de evaluación a nivel nacional.

Este examen se constituye como el primer estudio evaluativo de cobertura nacional, con pruebas a gran escala, surge a fines de 1989 y se continua aplicando hasta nuestros días.

### ▪ Propósito del IDANIS

El IDANIS tiene como propósito, determinar las posibilidades de éxito de los alumnos que ingresan a secundaria al medir habilidades verbales, matemáticas y de razonamiento verbal. En su diseño original, estaba fundamentado en la tecnología educativa, su propósito, constatar en qué medida el alumno había alcanzado los objetivos generales y particulares establecidos para la educación primaria. En el marco de la reforma curricular de los planes y programas de estudio para la educación básica a nivel primaria (1993); retoma los aportes de la psicología del aprendizaje en cuanto a teorías constructivistas y se transforma en una prueba centrada en habilidades <sup>137</sup> y en la medición del desarrollo de las capacidades de los estudiantes para aprender; su intención pasó de constatar el dominio de contenidos disciplinarios, a medir la aptitud para el aprendizaje.

---

<sup>137</sup> "Para efectos de su medición a través del IDANIS, las habilidades se caracterizan como recursos intelectuales que activan o regulan procesos más complejos al interior de las estrategias cognitivas de las cuales forman parte" (SEP 1999: 22).



▪ **Contenidos del IDANIS**

Tres son los tipos de habilidades que contempla la prueba del IDANIS:

- *Habilidad verbal*: Conjunto de recursos vinculados al conocimiento lingüístico que permiten procesar información lingüística para acceder a significados implícitos y explícitos.
  
- *Habilidad matemática*: Conjunto de recursos vinculados con el conocimiento matemático que mediante las estrategias cognitivas permite al sujeto procesar información.
  
- *Habilidad para el razonamiento abstracto*. Conjunto de recursos vinculados al conocimiento de la representación figurativa bidimensional que mediante estrategias cognitivas permiten procesar información gráfica para acceder a la interpretación de representaciones visuales diversas (cfr. SEP, 1999:22).

En el siguiente cuadro, se presentan los contenidos, procesos, tareas y tipo de ítem que conforman la prueba del IDANIS, con base en cada una de estas habilidades.



CUADRO 3.1

**IDANIS**  
**CONTENIDO GENÉRICO Y ESTRUCTURA DE LA PRUEBA**

<b>VARIABLE A MEDIR</b>	<b>CONTENIDO (Unidades de diagnóstico)</b>	<b>PROCESOS / HABILIDADES Que se ponen en acción</b>	<b>TAREAS</b>	<b>ESTRUCTURA DE LOS ITEMS/ O PREGUNTAS</b>
Habilidad verbal	Comprensión lectora	Habilidades para acceder al contenido de una comunicación escrita  Habilidades relacionadas con la construcción de ideas	Reconocer o construir el significado de vocablos Suprimir información trivial, redundante o ajena al interior de un texto, globalizar información y jerarquizarla, reconocer la estructura discursiva adoptada	Multireactivo, los ítems, dependen de un texto fuente
	Completar oraciones	Reconocimiento de la secuencia temporal, identificación de la relación lógica	Reintegrar un texto breve a partir de elementos subsistentes, identificación de la relación lógica, identificación de relaciones atributivas y reconocimiento de la concordancia gramatical.	Textos incompletos que caben dentro de la categoría del Test de Cloze (conformados por dos o más oraciones)
Habilidad Matemática	Aritmética	Traducción de situaciones verbales a su expresión numérica  Análisis de datos numéricos  Realización de cálculos básicos	Reconocer y relacionar los elementos de un problema que implique tanto la traducción de situaciones verbales a su expresión numérica, así como el análisis de datos numéricos y la realización de cálculos básicos.	Se muestra la estructura del problema y se exige al sustentante la identificación de los datos involucrados y la relación que guardan entre sí, así como la elección y realización de cálculos pertinentes



	Geometría	Expresión aritmética de relaciones espaciales de figuras y cuerpos geométricos	Interpretación de las relaciones espaciales de figuras y cuerpos representados en dos dimensiones	Se presentan figuras geométricas en las cuales el sustentante debe identificar relaciones de igualdad, desigualdad o proporción.
Habilidad para el razonamiento o abstracto	Series de figuras	Detección de reglas de un proceso que tiene lugar en el tiempo y el espacio	Solución de casos problema	Series de figuras en las cuales el sustentante debe reconocer procesos de cambio y evolución

Basado en SEP, 1999, pp. 28-29

Se afirma que el IDANIS es una prueba centrada en habilidades, pero nótese que éstas se encuentran expresadas en términos de contenidos curriculares; no hay delimitación específica de cada una, para su medición. Se determinan el tipo de éstas, pero no cuáles son; además, se consideran sólo como estrategias para resolver una tarea; lo que de entrada presenta una confusión para la definición del contenido de evaluación de esta prueba: *habilidades o estrategias*.<sup>138</sup>

#### □ Ítems del IDANIS<sup>139</sup>

El análisis de los ítems del IDANIS, se realiza, tomando como ejemplo un reactivo de cada uno de los contenidos que comprende.

<sup>138</sup> Como ya señaló en el capítulo anterior, la literatura existente en torno a habilidades y estrategias es muy variada y en muchos de los casos se tiende a confundir el término de estrategia con el de habilidad, e incluso a éste con el de capacidad; ante lo cual es importante insistir en que las pruebas que se dicen miden habilidades, comiencen por esclarecer muy bien el concepto para evitar confusiones en cuanto a lo que se mide.

<sup>139</sup> Consideramos conveniente resaltar que a falta de un documento oficial que muestre algunos ítems del IDANIS los ítems presentados, fueron seleccionados de una "Guía de estudio para el examen de ingreso a secundaria", proporcionada por un docente.



▪ Ítem de comprensión lectora para alumnos de 6° grado.

Para medir la *aptitud de comprensión lectora* se utilizan ítems como los siguientes:

**COMPRESIÓN DE LECTURA**

Lee con mucha atención el texto y procura comprenderlo bien. Después lee con cuidado cada una de las preguntas que le siguen, fijándote bien en lo que se te pregunta. A continuación estudia las opciones y elige aquella que responda correctamente a la pregunta, de acuerdo con lo que se dice en el texto.

Ejemplos de preguntas:

1. ¿Cuál de los siguientes enunciados confirma que Yaxchilán se encuentra cerca del río Usumacinta?
  - a) Se localiza Yaxchilán, a orillas del Río Usumacinta.
  - b) En la selva lacandona se localiza la ciudad de Yaxchilán.
  - c) Rodeada por bosques tropicales.
  - d) Forma parte de la zona de las lagunas de Montebello.
  - e) Se observan varias edificaciones monumentales.
  
2. ¿Qué lugar es uno de los más visitados en Chiapas?
  - a) Palenque
  - b) Tenam Puente
  - c) Chiapa de Corzo
  - d) Yaxchilán
  - e) Toniná.
  
3. ¿La idea que existe en el texto anterior es que Chiapas...?
  - a) Es la única entidad con zonas protegidas.
  - b) Cuenta con un territorio enorme.
  - c) Es un lugar bello que cuenta con zonas arqueológicas y naturales protegidas.
  - d) Es una entidad pobre pero bonita.
  - e) Es una entidad muy parecida a las de la República Mexicana.

Fuente: Guía de estudio.

En este reactivo se puede observar lo siguiente:

- La opción "c" de la pregunta No. 3, induce la respuesta de la pregunta No. 2. Una recomendación importante para la redacción de ítems de opción múltiple, es no trasladar opciones y mantenerlas independientes, en virtud de lo cual, la estructura de los ítems denota problemas importantes.



- En materia de contenido, el reactivo seleccionado muestra que las tareas para evaluar la comprensión lectora, están encaminadas básicamente a la obtención de información y a la identificación de las ideas dentro de un texto, tareas en las que van implícitas procesos muy básicos de pensamiento como la identificación, relación y generalización.

Los ítems que hemos presentado para la *evaluación de la comprensión lectora* en el IDANIS dan cuenta de que:

- ⇒ En el IDANIS, la comprensión lectora como contenido de evaluación, se basa en la mera obtención de información. No hay preguntas que promuevan la interpretación o reflexión sobre el texto.
- ⇒ Las tareas contemplan un tipo de lectura de carácter literal, que requiere de la mera localización de información y extracción de ideas centrales.
- ⇒ Los procesos implícitos en las tareas, residen en la identificación, localización y generalización.
- ⇒ La estructura de los ítems, presenta problemas de redacción.
- ⇒ Se trata de ítems un tanto abstractos y sin sentido, en donde los textos y situaciones presentadas se encuentran lejos de la realidad, necesidades e intereses de los estudiantes.

- **Ítems de habilidad verbal para alumnos de 6° grado**

Para medir la *habilidad o aptitud verbal*,<sup>140</sup> el IDANIS contempla ítems como los que a continuación se presentan:

---

<sup>140</sup> Nótese que en los ítems muestra de la guía, se maneja aptitud, lo que quiere decir que en cierta forma el término aptitud se considera como sinónimo de habilidad, de acuerdo con los fundamentos del IDANIS, este apartado correspondería a habilidad verbal.



## COMPLEMENTO DE ORACIONES

En esta sección de la evaluación se encuentran 10 oraciones; a cada una se le ha quitado una palabra, en su lugar se ha puesto una línea. Debajo de cada oración aparecen cinco palabras. Lee con cuidado cada oración y elige la palabra que complete apropiadamente el sentido de la misma. Después busca en tu hoja de respuesta el número de la pregunta que estés contestando y llena el espacio correspondiente a la opción que hayas elegido.

EJEMPLOS.

1. El payaso es el personaje del circo más \_\_\_\_\_
  - a) Triste.
  - b) Contento.
  - c) Distraído.
  - d) Simpático.
  - e) Inteligente.
2. El pan está \_\_\_\_\_ porque está recién hecho.
  - a) Caliente.
  - b) Dulce.
  - c) Grande.
  - d) Sabroso.
  - e) Oloroso.
3. El disco se vendió mucho, la voz de la cantante es \_\_\_\_\_
  - a) Ronca.
  - b) Aguda.
  - c) Agradable.
  - d) Espantosa.
  - e) Desentonada.

Fuente: Guía de estudio.

Este es un ejemplo en el que se pueden identificar problemas de estructura y contenido en la evaluación de la "habilidad verbal":

- o Nótese que en la pregunta 1, cualquiera de las opciones de respuesta es válida, en la preguntas 2, la opción incorrecta es muy obvia. De acuerdo con las recomendaciones de autores especialistas en materia de elaboración de pruebas con formato de opción múltiple, (véase Haladyna, Haladyna y Merino, s.f.). Los ítems seleccionados, presentan problemas importantes de constitución. La redacción de pruebas objetivas debe cuidar que sólo una de las respuestas sea válida y evitar utilizar pistas que indiquen la opción correcta.



- El contenido del reactivo está centrado en operaciones muy elementales de pensamiento; que si bien son básicas, no son suficientes para determinar una "aptitud o habilidad verbal", sobre todo si ésta se concibe como los procesos lingüísticos para la obtención de un significado.
  - Las tareas, consisten en la identificación, relación y discriminación de estructuras lingüísticas, que poco tienen que ver con una habilidad verbal.
  - La constitución de los ítems utilizados en el IDANIS para la evaluación de la "aptitud verbal," no favorece el empleo del lenguaje como proceso de comunicación.
- **Ítems de "habilidad matemática" para alumnos de 6° grado**

Para la evaluación de la "habilidad matemática", se incluyen dos campos temáticos: aritmética y geometría.<sup>141</sup>

#### ÍTEM DE ARITMÉTICA

1. Un Kilogramo de huevo costaba \$8.50 y su precio aumentó a \$10.50 ¿Cuánto pago más ahora por seis Kilogramos de huevo?
  - a) \$ 12.00
  - b) \$ 51.00
  - c) \$ 63.00
  - d) \$ 2.00
  - e) \$10.00
  
2. ¿Qué número multiplicado por 6 da el doble de 60?
  - a) 50.
  - b) 40.
  - c) 30.
  - d) 20.
  - e) 10.
  
3. Susana pagó 18 pesos por media docena de manzanas. ¿Cuánto cuestan 11 manzanas?

<sup>141</sup> Nótese que aquí también hay una confusión de términos y conceptos, pues se considera habilidad matemática como contenidos específicos del área de matemáticas. Es decir, notamos una falta de clarificación en el uso de los términos habilidad y su equiparación con campo temático.



- a) 39 pesos
- b) 36 pesos
- c) 33 pesos
- d) 30 pesos
- e) 27 pesos

Fuente: Guía de estudio

El análisis de este reactivo lleva a decir que:

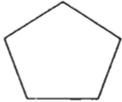
- Se trata de un reactivo que implica la realización de cálculos descontextualizados de la vida cotidiana del alumno, no se incluye una situación real; el comprar huevos y manzanas no es una actividad primordial de quien está por ingresar a la educación secundaria; además en la pregunta No. 2, la tarea se presenta abstracta y sin sentido.
- El reactivo no incluye la evaluación de recursos que permitan procesar la información, lo único que se requiere es la realización de operaciones que no conllevan a la transferencia, o a un razonamiento lógico y sistemático que favorezca el uso de las matemáticas para resolver problemas cotidianos mediante el empleo de ciertos procedimientos matemáticos. Esto presenta una evidente desvinculación entre la definición de habilidad matemática y el tipo de tareas que se emplean para evaluarla.

Otro de los campos temáticos incluidos, es Geometría, en donde se utilizan ítems como los siguientes:



### ÍTEM DE GEOMETRÍA

1. Observa las siguientes figuras: Considera el dato que se muestra y que el perímetro de una figura es igual a la suma de sus lados. ¿Cuánto mide el lado del cuadrado si el perímetro de las dos figuras miden lo mismo?



8 cm.



- a) 20 cm.
- b) 10 cm.
- c) 8 cm.
- d) 4 cm.
- e) 15 cm.

2. Observa las siguientes figuras: El cuadrado grande mide 12 cm., por lado y el pequeño mide la mitad por cada lado. ¿Cuál es el área del pequeño sabiendo que la superficie del cuadrado es igual al producto de la base por la altura?

12cm.



- a) 24 cm.
- b) 36 cm. Cuadrados
- c) 36 cm.
- d) 24 cm cuadrados
- e) 87 cm.

Fuente: Guía de estudio



En este ítem tenemos que:

- Las situaciones presentadas requieren de la mera ejercitación y práctica de cálculos para obtener el área de figuras, no se incluyen situaciones reales.
- Las tareas, requieren de procesos muy básicos de pensamiento, como la observación y comparación, que si bien son importantes, no son suficientes para favorecer un conocimiento matemático.

Los ejemplos de los ítems que el IDANIS emplea para evaluar la *habilidad matemática* dan cuenta de que:

- ⇒ El contenido de los ítems denota que habilidad matemática es igual a reproducción y práctica; lo que conlleva un enfoque tradicional de aprendizaje y evaluación de la matemática, en donde la activación de procedimientos aritméticos y el ejercicio de un pensamiento lógico-matemático, es inexistente.
- ⇒ Los procesos implícitos en los ítems, residen básicamente en la operación y el cálculo.
- ⇒ Las tareas de los ítems consisten en meros ejercicios aritméticos que se presentan como el fin y el medio de la evaluación matemática; lo que ofrece una visión tradicional del ejercicio didáctico trabajada por Snyders desde los años 70, quien plantea: "La educación tradicional corre el riesgo de caer en una simple sucesión de ejercicios; el ejercicio que no era más que un medio para llegar al modelo, va a convertirse en un fin en sí mismo y a ocupar todo el sitio"(Snyders, 1972).
- ⇒ El tipo de ítems presenta situaciones rutinarias y abstractas, lejanas de la realidad y necesidades de un alumno de sexto grado de educación primaria; lo que dificulta la transferencia y uso del conocimiento a situaciones matemáticas de la vida cotidiana.



▪ **Ítems de razonamiento abstracto para alumnos de 6° grado**

El razonamiento abstracto, otra dimensión incluida en la prueba del IDANIS, se evalúa mediante dos elementos base: *Series de números* y *Series de figuras*.

Algunos ejemplos de los ítem utilizados:

**SERIES DE NÚMEROS**

En esta sección de la evaluación encontrarás diez series incompletas de números. Analiza la secuencia de cada una de ellas y elige la opción que contenga el(los) número(s) que la complete(n)

8, 16, 24, 32, 40...

- A) 44
- B) 48.
- C) 52.
- D) 56.
- E) 60.

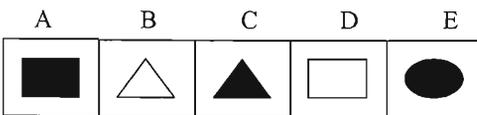
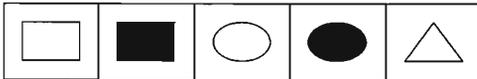
880, 440, 220, 110...

- A) 95.
- B) 85.
- C) 75.
- D) 65.
- E) 55.

Fuente: Guía de estudio

**SERIES DE FIGURAS**

En esta parte de la evaluación encontrarás series de figuras y, delante de cada una, cinco cuadros. Cada serie de figuras es una secuencia que se continúa con una de las figuras contenidas en los cuadros. Analiza cada serie y elige el cuadro que contiene la figura que completa correctamente la serie.



Fuente: Guía de estudio



En este reactivo podemos observar que el contenido denominado como de *razonamiento abstracto*, presenta a esta operación mental como un proceso de percepción y análisis para inferir el número o figura que complete correctamente las secuencias presentadas; procesos, que si bien son importantes, el razonamiento es una operación mental que va más allá de la mera identificación e interpretación de representaciones visuales en figuras bidimensionales. El razonamiento implica: asociar, integrar y aportar un enfoque y aplicación diferente. El análisis de las características de los ítems seleccionados, lleva a decir que el IDANIS es una prueba que presenta problemas de estructura y contenido como los siguientes:

- ▶ El contenido de evaluación, referente a estrategias y habilidades, se centra en estrategias de repetición para la realización de tareas que demandan un aprendizaje básico (véase Weinstein y Mayer, 1986). no se encuentran habilidades para la regulación de procesos complejos, conforme a lo que el IDANIS contempla en su definición.
- ▶ El tipo de ítems, incluye tareas centradas en la mera reproducción de ejercicios ajenos a situaciones cotidianas de la vida del estudiante; no contemplan la evaluación de estrategias de elaboración, organización y control, que permitan detectar la capacidad del alumno para integrar y aplicar un conocimiento.
- ▶ Las tareas de evaluación, ofrecen una visión tradicional del ejercicio didáctico, se presentan como un fin en sí mismo de enseñanza, aprendizaje y por tanto de evaluación; no como un medio que permita evaluar las capacidades de los estudiantes para el aprendizaje, que en sí vendría a ser uno de los objetivos de la aplicación del IDANIS.
- ▶ No existe vinculación entre la definición del contenido a evaluar y el tipo de ítems y tareas que se emplean.
- ▶ En el diseño de sus ítems se observan serios problemas de estructura y redacción.



## 2. Pruebas del Sistema Nacional de Evaluación Educativa (SNEE)

### ▫ Origen de las pruebas del SNEE

En el año de 1994, como respuesta al diagnóstico elaborado por el Programa de Desarrollo Educativo 1995-2000, que plantea la insuficiencia de instrumentos útiles para evaluar sistemáticamente el desempeño escolar,<sup>142</sup> la Dirección General de Evaluación Educativa (DGEE), crea el Sistema Nacional de Evaluación Educativa (SNEE), encargado de elaborar y aplicar pruebas a nivel nacional para la educación básica. Se trata de pruebas de aprovechamiento escolar, cuya elaboración y aplicación se dio a partir del ciclo escolar 1995-1996, hasta el ciclo 2002- 2003. La población a la que se aplicó contempla alumnos de 3° a 6° de educación primaria y de 1°, 2° y 3° de educación secundaria.

### ▫ Propósitos de las pruebas del SNEE

Entre los objetivos del SNEE, se encuentran: detectar posibles deficiencias en el proceso enseñanza-aprendizaje, abatir las desigualdades de los servicios educativos identificando los factores que causan impacto en la enseñanza, apoyar la toma de decisiones y asignar recursos según resultados de evaluación.

### ▫ Contenidos de las pruebas del SNEE para el 3° grado de educación primaria:

Los exámenes para la educación primaria, constan de un sólo instrumento en donde se incluyen ítems de las diferentes áreas. En el caso de la educación secundaria, existe un

---

<sup>142</sup> "La educación carece de suficientes instrumentos útiles para evaluar sistemáticamente el desempeño escolar y adecuar las políticas y decisiones a sus resultados[...].todavía no se encuentra con un sistema de evaluación que reporte información periódica" (Poder Ejecutivo Federal, 1996: 29).



instrumento para cada una de las áreas evaluadas que son: Biología, Civismo, Español, Geografía Historia, Inglés, Matemáticas, Francés, Física y Química.<sup>143</sup>

Para efectos de una posterior comparación con las pruebas del LLECE, en el siguiente cuadro se presentan los contenidos de evaluación de la prueba para el Tercer Grado aplicada en el ciclo 1999-2000.

CUADRO 3.2

**PRUEBA DEL SNEE PARA EL 3º GRADO DE EDUCACIÓN BÁSICA**

ÁREA	CONTENIDOS	TAREAS	PROCESOS
Español	Lectura Sinónimos Antónimos Orden alfabético Sujeto Signos de admiración Tipos de textos Usos de C, K y Q Datos de una ficha bibliográfica Homónimos Acento ortográfico Tipos de textos Tiempos verbales Género y número Escritura correcta de palabras	A partir de un texto identificar: Acentos y verbos Sujetos Uso de sinónimos y antónimos Género y número Orden alfabético los personajes Significado de palabras Correcta o incorrecta escritura de palabras Ideas principales Palabras clave Personajes Acontecimientos Idea General Tipo de texto Obtener información del texto Datos de una ficha bibliográfica Partes de un libro Interpretación de frases	Localización Síntesis Adaptación Identificación Interpretación Comprensión Discriminación Análisis

<sup>143</sup> En la páginas del SNEE existe un banco de instrumentos que contiene las pruebas de aprovechamiento escolar para cada uno de los ciclos en los diferentes grados y niveles a los que se aplicó. Ver [http://snee.sep.gob.mx/BROW-AES/SubMnPruebas\\_AprovEsc.htm](http://snee.sep.gob.mx/BROW-AES/SubMnPruebas_AprovEsc.htm)



Matemáticas	Equivalencia y división de fracciones Perímetro Problemas de suma y multiplicación La hora Perspectivas Cantidades Planos, Problemas de reparto Problemas de suma y resta Lectura de tablas y gráficas de barras Comprensión de datos de un problema Obtención de superficies Uso de planos y trayectorias	Ubicar: Elementos que integran la fórmula del perímetro, Datos para resolver un problema, Cuántas veces cabe una figura pequeña en otra más grande Fracciones equivalentes a $\frac{1}{2}$ en un dibujo, Resolver problemas que requieran suma, resta y multiplicación de cantidades Comparar cantidades y fracciones Interpretar un conjunto de datos de una situación Leer mapas y la hora Repartir las cantidades representadas en un dibujo para resolver un problema Obtener información de una tabla Analizar los datos de una situación o hecho para relacionarlos con las preguntas planteadas Hacer desplazamientos en un plano Dividir objetos en fracciones de acuerdo a situación de reparto y ubicar trayectorias.	Identificación Cálculo Observación Reparto Representación Percepción analítica Comprensión y orientación espacial
-------------	--	--	---

□ Ítems de las pruebas del SNEE para el 3° grado de primaria

Las pruebas del SNEE, incluyen ítems de distintas áreas en un solo instrumento, con la intención de analizar el contenido y forma de evaluación del español y las matemáticas en el tercer grado de educación primaria, hemos seleccionado algunos ejemplos de ítems referentes sólo al español y las matemáticas,<sup>144</sup> los cuales se presentan y analizan por separado aunque pertenezcan a un sólo instrumento:

<sup>144</sup> El instrumento completo se encuentra en:  
[http://snee.sep.gob.mx/Pruebas/Aprovechamiento\\_Escolar/primaria/00/aep3\\_00q.pdf](http://snee.sep.gob.mx/Pruebas/Aprovechamiento_Escolar/primaria/00/aep3_00q.pdf)



▪ Ítems de la prueba de español\*

MUNECO + ROSCA = TAMALES

1. Según el diccionario, una tradición es la transmisión de historias, noticias y costumbres de generación en generación.
  2. Seguramente (y tal vez sin que te des cuenta) formas parte de la cadena que hace subsistir en alguna tradición. Por ejemplo, si eres de los que el pasado 6 de enero comieron rosca de reyes, permíteme decirte que estás continuando un ritual que llegó a México hace unos 400 años y existen bastantes posibilidades de que si algún día decides tener hijos, ellos se unirán como un eslabón más.
  3. ¿Sabías que antes, en lugar de muñequito se ponía una haba en la rosca? El objetivo era el mismo que ahora, quien encontraba a la semilla apadrinaba al niño Jesús el día de la Candelaria –2 de febrero- y en consecuencia tenía que hacerse cargo de organizar la fiesta con una muy buena dotación de tamales.
  4. Más tarde y en otras regiones, también se acostumbró incluir un dedal y una aguja; si alguna muchacha encontraba la aguja, esto significaba que se casaría pronto, en cambio la desafortunada que encontraba el dedal, prácticamente quedaba condenada a la eterna soltería.
- 
1. De acuerdo con el texto, ¿qué es lo que se trasmite de generación en generación?
    - a) La tradición
    - b) Las roscas
    - c) El muñequito
    - d) Los dedales
  2. En el segundo párrafo, ¿cuál es el sinónimo de la palabra ritual?
    - a) Obligación
    - b) Ceremonia
    - c) Formalidad
    - d) Firmeza
  3. ¿Cuándo llegó la tradición de la rosca a nuestro país?
    - a) El 6 de enero
    - b) El 2 de febrero
    - c) Hace 40 años
    - d) Hace 400 años
  4. ¿Qué opción resume las ideas del tercer párrafo?
    - a) Se sigue incluyendo una haba en la rosca como antes
    - b) Quien encontraba la semilla apadrinaba al niño Jesús el día de la Candelaria.
    - c) El día de la candelaria es el 2 de febrero
    - d) Para organizar una fiesta se debe tener una buena dotación de tamales

Fuente: [http://snee.sep.qob.mx/Pruebas/Aprovechamiento Escolar/primaria/00/aep3\\_00q.pdf](http://snee.sep.qob.mx/Pruebas/Aprovechamiento_Escolar/primaria/00/aep3_00q.pdf) p.1



En el ejemplo presentado se puede observar que:

- ⇒ El contenido de la comprensión lectora, como tema de español, se evalúa con ítems enfocados a la obtención de información y a la extracción de ideas de un texto, lo que presenta una concepción de lectura como información y extracción de ideas.
- ⇒ Los contenidos para evaluar el español se presentan como un listado de temas curriculares contemplados en los Planes y Programas de Estudio para la Educación Primaria.
- ⇒ Los procesos implícitos en las tareas, tienen que ver con procesos básicos de pensamiento como la identificación y el reconocimiento, que no contemplan la reflexión, evaluación, elaboración o inferencia.
- ⇒ Los ítems para evaluar los contenidos curriculares propios del área de español, residen en multireactivos referentes a la presentación de un texto.

▪ **Ítems de las pruebas de matemáticas**

1. Si una persona gasta 90 litros de agua bañándose en regadera, ¿cuántos litros de agua gastarán 5 personas al bañarse en regadera?
  - a) 18
  - b) 85
  - c) 95
  - d) 450

2. En una florería tienen la siguiente tabla:

Rosas	
Cantidad	Precio
1 Docena	\$ 15 pesos
3 Docenas	\$ 45 pesos

¿Cuánto se debe pagar por 4 docenas?

- a) \$ 30.00
- b) \$ 35.00
- c) \$ 50.00
- d) \$ 60.00



Los ejemplos de los ítems de matemáticas en estas pruebas, dan cuenta de que:

- ⇒ El contenido de la evaluación matemática consiste en una serie de temas curriculares contemplados en los planes y programas de estudio de la educación primaria.
- ⇒ Las tareas para la evaluación matemática están centradas en la realización de ejercicios aritméticos, acordes con la visión tradicional que Snyder plantea del ejercicio didáctico como fin y medio de la evaluación y el aprendizaje.
- ⇒ Los procesos implícitos en las tareas, consisten en procesos básicos de pensamiento como la identificación, representación, análisis y relación que tienen que ver más con el cálculo, que con un pensamiento lógico matemático.
- ⇒ Se trata de ítems descontextualizados, lejanos a la realidad y necesidades de un alumno de tercer grado de educación primaria.

El análisis de las características de los ítems seleccionados, de las pruebas del SNEE, permiten enunciar las siguientes características:

- ▶ Su contenido de evaluación está centrado en temas curriculares, que da la idea de conocimiento como acumulación de información.
- ▶ Los procesos, implícitos en las tareas y ítems son procesos básicos, que aunque importantes en todo proceso de pensamiento, no promueven el uso de la lengua y las matemáticas como medio de comunicación y desarrollo de habilidades mentales superiores como la reflexión, la inferencia y el razonamiento matemático, la reflexión, integración y aplicación del conocimiento a situaciones cotidianas de la vida de los alumnos.
- ▶ El tipo de tareas de sus ítems están centradas en el ejercicio como práctica y no como aplicación.
- ▶ Los ítems contemplan situaciones descontextualizadas que dificultan la aplicación y uso de lo aprendido a realidades cercanas de los alumnos.
- ▶ Se desconocen las normas y procedimientos para su elaboración.



### 3. Pruebas con base en estándares nacionales (PE), ahora Pruebas Nacionales

#### □ Orígenes de las Pruebas Nacionales:

En el Sistema Educativo Mexicano, el establecimiento de estándares nacionales y su medición a través de pruebas, tuvo sus orígenes con la Reforma Curricular de 1993 en materia de Planes y Programas de estudio para la educación primaria, en la cual se presentan las "habilidades"<sup>145</sup> y conocimientos que se espera dominen los alumnos en cada materia y grado. Los antecedentes de las Pruebas de Estándares los podemos encontrar en las pruebas del proyecto: *Evaluación de la Educación Primaria (EVEP)*, aplicadas desde 1996 con el fin de obtener información representativa a nivel nacional del estado que guardaba la educación básica.

Conforme a informes publicados, (Véase SEP, 2000:152) la SEP en el año de 1997, comienza el desarrollo de instrumentos de medición de logro educativo sustentados en estándares nacionales.<sup>146</sup>

En 1998, se da por primera vez, la elaboración y aplicación de pruebas con base en estándares nacionales<sup>147</sup> (PE) para medir las habilidades que se señalan en el Plan y Programas de Estudio para la Educación Básica de 1993.

---

<sup>145</sup> El Plan y Programa de Estudios, 1993 para la Educación Primaria considera como desarrollo de habilidades intelectuales: la lectura, la escritura, la expresión oral, la búsqueda y selección de información, la aplicación de las matemáticas a la realidad. (cfr .SEP, 1993).

<sup>146</sup> En los informes oficiales de la SEP se define a los estándares como: "las habilidades que debemos esperar hayan adquirido los alumnos en determinados momentos de su formación, conforme a los propósitos que persigue la educación básica en nuestro país y cómo se establecen en los planes y programas de estudio". (SEP, 2000:152) Nótese que aquí el concepto de estándares está estrechamente relacionado a las habilidades mencionadas en el Plan de Estudios.

<sup>147</sup>El concepto de estándares, referente a meta o a confirmación del logro, no es un concepto nacional, su uso proviene del modelo utilizado por Estados Unidos. Además la incorporación de estándares al campo educativo proviene del ámbito de la producción; ya que éstos surgen como resultado de la proliferación dada en materia de producción y en la educación se adoptan a partir de la proliferación dada en las escuelas. Sin embargo conviene señalar que para poder estandarizar la producción, se requiere contar con las mismas condiciones y en México no todas las escuelas y mucho menos los docentes cuentan con condiciones



Actualmente el Instituto Nacional de Evaluación Educativa (INEE),<sup>148</sup> es quien retoma la elaboración y aplicación de las pruebas con base en estándares, llamadas ahora Pruebas Nacionales (PN) cuya primera aplicación tuvo lugar a fines del ciclo escolar 2002-2003 (véase INEE 2003).

□ **Propósito de las Pruebas Nacionales:**

Documentos publicados por el SNEE establecen que “el objetivo de la evaluación con base en estándares nacionales, es medir el logro académico de los estudiantes, en función de algunos aspectos que el currículo actual enfatiza como es el desarrollo de habilidades fundamentales de lectura y matemáticas para todos los grados de primaria. Se pretende medir por un lado, cuántos alumnos alcanzan los estándares de habilidades identificados en el currículo y por otro, intentan decir algo acerca de qué habilidades poseen los alumnos y qué les falta por desarrollar para alcanzar dichos estándares”(Martínez y Schmelks, s.f).

Las Pruebas Nacionales tienen la misma finalidad que las llamadas Pruebas de Estándares: medir las habilidades de comprensión lectora y matemáticas incluidas en los planes y programas de estudio para la educación básica 1993.

---

iguales. De ahí que consideramos que la educación y la evaluación o medición a partir de estándares todavía es algo difícil de alcanzar.

<sup>148</sup> Creado de acuerdo por lo estipulado en el Plan Nacional de Desarrollo 2001-2006, por decreto presidencial el 8 de agosto de del 2002. El texto completo del decreto por el que se crea el INEE, publicado en el diario oficial se puede consultar en: Poder Ejecutivo Federal (2002), “Decreto por el que se crea el Instituto Nacional para la Evaluación de la Educación”, en *Diario Oficial de la Federación*, 8 de agosto, México Documento en red consultado en: [http://www.gobernacion.gob.mx/dof/2002/agosto/dof\\_08-08-2002.pdf](http://www.gobernacion.gob.mx/dof/2002/agosto/dof_08-08-2002.pdf). O en la página del INEE en la siguiente dirección electrónica [http://capacitacion.ilce.edu.mx/inee/pdf/Decreto de creacion.pdf](http://capacitacion.ilce.edu.mx/inee/pdf/Decreto%20de%20creacion.pdf)



□ **Contenidos de las pruebas nacionales para la educación primaria:**

El contenido de evaluación de las Pruebas de Estándares, ahora Pruebas Nacionales, está determinado por lo que contemplan los planes y programas de estudio vigentes de educación primaria, con respecto a las habilidades de:

- **Comprensión de la lectura**
- **Matemáticas básicas**

Las *habilidades básicas* que del Plan y Programa de Estudios en relación al área de **español** consisten en:

- El dominio de la lectura, la escritura y la expresión oral, el propósito central es que los niños desarrollen su capacidad de comunicación en la lengua hablada y escrita. (cfr. SEP, 1993).

En el área de **matemáticas**, la prioridad reside en desarrollar:

- Habilidad para la resolución de problemas.
- El razonamiento matemático a partir de situaciones prácticas.

▪ **Evaluación de la comprensión lectora en las Pruebas Nacionales:**

Para especificar lo que se mide en las Pruebas Nacionales, el INEE retoma lo establecido en los Planes y Programas de Estudio de la Educación Primaria y presenta lo que se entiende por *comprensión de lectura*, una de las principales habilidades a desarrollar según lo estipulado en los documentos curriculares de la educación primaria.<sup>149</sup>

---

<sup>149</sup> "El nuevo plan de estudios y los programas de asignaturas que lo integran, tienen como propósito organizar la enseñanza y el aprendizaje de contenidos básicos para asegurar 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 las matemáticas a la realidad) que les permitan aprender permanentemente y con independencia, así como actuar con eficacia e iniciativa en las cuestiones prácticas de la vida cotidiana." (SEP, 1993:13)



En documentos publicados por el INEE se puede encontrar que la evaluación de la *comprensión lectora* implica relacionar la información que proviene del texto con los conocimientos previos del lector, para construir así el sentido de lo que el autor quiso comunicar; a partir de cuatro factores: lo que sabe, lo que lee, el contexto y los propósitos que lo motivaron a leer.

La comprensión implica, no solamente leer lo explícito, sino también:

- Leer entre líneas
- Reconocer la ironía, la metáfora y el humor
- Inferir y sacar conclusiones
- Identificar evidencia
- Detectar peculiaridades y sutilezas del lenguaje
- Evaluar la información leída
- Reflexionar sobre ella
- Contrastar el conocimiento con otra información proveniente de otras fuentes;
- Reconocer las maneras en que los textos están contruidos

Para evaluar la comprensión lectora se toman en cuenta dos *aspectos* fundamentales:

- Las habilidades que subyacen a la comprensión de lectura
- El contenido o tipos de texto utilizado (cfr. Quezada, 2004: 20)



CUADRO 3.3

**EVALUACIÓN DE LA COMPRENSIÓN LECTORA EN LAS PRUEBAS NACIONALES**

CONCEPTO	TAREAS	ASPECTOS QUE COMPRENDE
<p>Relacionar la información que proviene del texto con los conocimientos previos del lector, para construir así el sentido de lo que el autor quiso comunicar; a partir de cuatro factores:</p> <ul style="list-style-type: none"> <li>• Lo que sabe</li> <li>• Lo que se lee</li> <li>• El contexto</li> <li>• Los propósitos que motivaron a leer</li> </ul>	<p>Leer entre líneas Reconocer la ironía, la metáfora y el humor Inferir y sacar conclusiones Identificar evidencia Detectar peculiaridades y sutilezas del lenguaje Evaluar la información leída Reflexionar sobre ella Contrastar el conocimiento con otra información proveniente de otras fuentes; Reconocer las maneras en que los textos están contruidos</p>	<p>Las habilidades que subyacen a la comprensión de lectura</p> <p>El contenido o tipos de texto utilizado</p>

Para medir las habilidades de comprensión lectora, el INEE, las clasificó de acuerdo a tres criterios y posteriormente las relacionó con lo presentado en los planes y programas de estudio tanto de primaria como de secundaria.

CUADRO 3.4

**HABILIDADES DE COMPRENSIÓN LECTORA PARA LA EDUCACIÓN PRIMARIA DE LAS PRUEBAS NACIONALES**

HABILIDADES (INEE)	PLAN DE ESTUDIOS PARA LA EDUCACIÓN PRIMARIA
<p>1. Habilidades que permiten reconocer las diferencias entre diversos tipos de texto, y facilitar así, la anticipación del contenido que se va a leer.</p>	<ul style="list-style-type: none"> <li>□ Predecir el contenido de los textos a partir de sus propiedades gráficas.</li> <li>□ Utilizar la organización gráfica y tipográfica para identificar el tópico, subtópico, tipo de texto, función y portador.</li> <li>□ Relacionar las propiedades lingüísticas relevantes para predecir el tipo de texto.</li> </ul>



<p>2. Habilidades que permiten construir el sentido del texto que se está leyendo.</p>	<ul style="list-style-type: none"><li>□ Referencia: Construir el significado que las palabras y expresiones adquieren dentro del texto.</li><li>□ Co-referencia o referencia textual: relacionar un pronombre, sinónimo o expresión con la información a la cual sustituye y aparece en otra parte de la lectura.</li><li>□ Interpretación de la información explícita en el texto: reconocer ideas principales expresadas con los mismos términos con que figuran en un texto, o expresadas en paráfrasis o a través de la reorganización de la información.</li><li>□ Interpretación de la información no presente, pero sugerida en el texto: reconstruir las intenciones, motivaciones y opiniones de personajes o autores, o extraer conclusiones a partir de relaciones lógicas explícitas en un texto.</li><li>□ Integración de información relacionada con la interpretación del texto como un todo coherente: integrar en un enunciado temático, cuadro sinóptico o diagrama, las ideas principales de la totalidad de un texto informativo.</li></ul>
<p>3. Habilidades que permiten valorar críticamente un texto.</p>	<ul style="list-style-type: none"><li>□ Advertir la perspectiva del autor: determinar la actitud y la posición del autor en relación con lo que escribe.</li><li>□ Reflexionar sobre la forma de un texto: distinguir matices del lenguaje que modifican la interpretación en textos narrativos, apelativos y argumentativos.</li><li>□ Interpretar, cuestionar y evaluar un mensaje para construir una opinión.</li></ul>

Adaptado de Quezada., 2004, p. 21

Se definen también los “niveles de desempeño”<sup>150</sup> y se establecen cuatro niveles de competencia para cada grado y nivel de la educación básica.<sup>151</sup>

En el siguiente cuadro se presentan los desempeños y niveles de competencia establecidos para el sexto grado, considerando que los únicos ítems publicados hasta el momento por el INEE, que permitan detectar como se comportan estos desempeños en los ítems correspondientes , pertenecen al sexto grado de educación primaria.

<sup>150</sup> Término que también es utilizado en las pruebas del LLECE.

<sup>151</sup> Nótese que el término competencia se equipara con el de desempeño.



CUADRO 3.5

**NIVELES DE COMPETENCIA EVALUACIÓN DE LA COMPRENSIÓN LECTORA  
6º GRADO DE EDUCACIÓN PRIMARIA PRUEBAS NACIONALES (INEE)**

Niveles	Definición	Desempeños
IV	Nivel más alto: Cumple perfectamente los objetivos curriculares. Los alumnos se consideran buenos lectores.	<ul style="list-style-type: none"><li>• Construir el significado de términos especializados, relacionando diferentes fragmentos del texto.</li><li>• Predecir el tipo de expresiones que pueden encontrarse al interior de un texto, las cuales respeten el género discursivo y la función del texto.</li><li>• Seleccionar y jerarquizar la información explícita para comprender el sentido de un texto.</li><li>• Interpretar información no presente, pero sugerida en el texto, para reconstruir las intenciones del autor al escribir el texto. Reconstruir relaciones de causa/efecto y de problema/solución.</li><li>• Integrar en un enunciado temático las ideas principales de la totalidad de un texto informativo.</li><li>• Comprender el mensaje más importante y el tema central de un texto literario.</li></ul>
III	Expresión parcial suficiente.	<ul style="list-style-type: none"><li>• Interpretar información explícita en <i>textos discontinuos</i>, localizando, relacionado y/o comparando más de tres datos, en dos o más textos, para llegar a una conclusión.</li><li>• Interpretar los aspectos gráficos y las propiedades lingüísticas de los textos para identificar portador, tópico y tipo de información que contiene un texto o para localizar información específica, en textos de uso poco frecuente dentro de la escuela.</li><li>• Interpretar aspectos gráficos del texto y sus propiedades lingüísticas para anticipar el tema global en textos discontinuos.</li><li>• Relacionar información de un texto discontinuo con información externa para llegar a una conclusión.</li><li>• Construir el significado que un término de uso poco frecuente adquiere dentro de un texto particular.</li><li>• Reconocer ideas importantes expresadas a través de una reorganización de la información contenida en el texto.</li><li>• Extraer conclusiones a partir de relaciones lógicas explícitas en textos informativos.</li><li>• Relacionar la esencia de un texto con un título pertinente.</li><li>• Valorar críticamente un texto para determinar.</li></ul>



II	Expresión parcial insuficiente.	<ul style="list-style-type: none"><li>● Interpretar información explícita en <i>textos discontinuos</i>, por ejemplo, pueden localizar información relacionando o comparando no más de tres datos, que aparecen en uno o dos textos, para llegar a una conclusión.</li><li>● Construir el significado que una expresión de uso frecuente adquiere dentro del texto.</li><li>● Interpretar información explícita en los textos para reconocer ideas principales expresadas en los mismos términos que aparecen en los textos o que aparecen en paráfrasis.</li></ul>
I	Nivel más bajo. Indica ausencia de competencia. El alumno no logra los objetivos establecidos en los Planes de Estudio.	<ul style="list-style-type: none"><li>● Interpretar los aspectos gráficos y las propiedades lingüísticas de los textos para identificar portador, tópico y tipo de información que contiene un texto.</li><li>● Localizar información específica en <i>textos discontinuos</i> de uso común dentro de la escuela, tales como mapas o diagrama del sistema planetario.</li><li>● Relacionar un pronombre con la palabra a la que sustituye.</li><li>● Interpretar información específica explícita en un texto, por ejemplo, reconocer una idea que engloba, en un campo semántico, un conjunto de elementos que aparecen en el texto.</li></ul>

Adaptado de Quezada, 2004, p. 28

En la selección de textos para evaluar la comprensión lectora, se consideró que los relatos utilizados concordaran con el tipo de textos manejados en los planes y programas de estudio y que representaran “situaciones auténticas;” es decir que sean semejantes a los que se pueden encontrar en la vida real.

Se utilizaron dos tipos de textos: *textos continuos* y *textos discontinuos*<sup>152</sup>

Entendiendo por *textos discontinuos* aquellos que presentan información de diversa manera como: formatos, gráficas y mapas; para evaluar y que permiten evaluar habilidades relacionadas con el reconocimiento de los distintos tipos de texto y por *textos continuos*, los compuestos por enunciados distribuidos en párrafos; utilizados evaluar las habilidades para construir el sentido del texto y valorar críticamente lo que se lee.

A fines del 2004, el INEE en su informe sobre *La Calidad de la Educación Básica en México* (véase INEE 2004), publica algunos ítems de las llamadas Pruebas Nacionales.<sup>153</sup>

<sup>152</sup> Nótese que esta denominación es la misma que se da a los tipos de textos manejados en las pruebas del PISA.



▪ Ítems de comprensión lectora para el 6° grado

Los ítems utilizados para la evaluación de la lectura se presentan a continuación:

Los dos ítems publicados por el INEE de las pruebas de comprensión lectora para el 6° grado de la educación primaria fueron los siguientes:

**PRUEBAS NACIONALES DEL INEE (2003-2004).  
ÍTEM DE COMPRENSIÓN LECTORA 6° GRADO**

Lee con cuidado el siguiente texto y después contesta las preguntas de la 67 a la 78. Vuelve a consultarlo siempre que lo necesites.

**Reactivo fácil**

Susana Elizondo. *Buenos Aires, 4 de julio de 1999*

Una botella con un mensaje de amor escrito hace 85 años por un soldado británico fue hallada el pasado 26 de julio en el estuario del río Tamesis, por un pescador que navegaba por la costa sudoeste de Inglaterra.

Doce días antes de morir, el 16 de diciembre de 1914 durante la Primera Guerra Mundial, el soldado Thomas Hughes le escribió a su esposa Elizabeth un mensaje que concluía: "Adiós por ahora, querida. Tu marido", y lo echó al río cuando se dirigía a Francia a luchar contra los alemanes".

La esposa de Hughes murió en 1979 pero su hija, Emily Crowhurst, se mostró muy emocionada al recibir la noticia del hallazgo. "Solamente conservo unas pocas cartas, libros y ropa de mi padre, pero me he pasado toda la vida tratando de saber más de su vida", manifestó Emily, quien actualmente reside en Nueva Zelanda.

Hughes, quien tenía 26 años al morir en su primer día en las trincheras, se despidió de Emily el 6 de noviembre de 1914 cuando ella tenía dos años. Sin embargo, Emily todavía guarda ese recuerdo y el de la tristeza de su madre tras la pérdida de su marido.

74. ¿ Para qué Thomás Hughes le escribió un mensaje a su esposa?

- Para despedirse de ella

<sup>153</sup> Los ítems publicados por el INEE en su 2° informe de actividades (véase INEE 2004), tiene la intención de explicar el sentido de los puntajes, más no de analizar el contenido y forma de evaluación. Anteriormente, los únicos ejemplos de ítems muestra de las Pruebas Nacionales fueron los publicados por el periódico Reforma, a propósito de la difusión de resultados y de la discusión en torno de la posibilidad de una comparabilidad entre los mismos. En mayo del 2004, a raíz de la noticia de que los resultados obtenidos en las Pruebas de Estándares Nacionales de Lectura y Matemáticas aplicadas en el 2003, son inferiores a los obtenidos en el 2002, el periódico Reforma publicó las habilidades de lectura para el sexto grado, consideradas en cada uno de los cuatro niveles de competencia en que se ubican a los alumnos a partir de sus aciertos y errores; así como una muestra de las preguntas de la prueba de Matemáticas para sexto grado (véase Del Valle, 2004).



- Para encargarle a su hija
- Para avisarle a dónde se dirigía
- Para avisarle que pronto regresaría

Tomado de INEE, 2004, p. 74

En este reactivo se puede decir que:

- o La tarea principal reside en inferir y sacar conclusiones mediante una cierta interpretación de lo que dice el texto.
- o El contenido de las tareas y procesos implícitos en este ítem, no contempla ninguno de los elementos que el INEE presenta como *comprensión de la lectura*.
- o Se trata de un texto en el que difícilmente el lector podrá relacionar lo que se dice con sus conocimientos previos y contexto, la situación que presenta está completamente alejada de la realidad e intereses de un alumno de 6° grado de primaria.



Reactivo difícil:

**D**

¿SABÍAS QUE...?



Desde que la antigua Unión Soviética lanzó en 1957 al primer satélite espacial, el Sputnik I, se han lanzado al espacio más de 4000 satélites artificiales, como otros dos millones el Marisat II y el Marisat III. En términos generales, los satélites han proporcionado muchísima información a los científicos, además de jugar un papel muy importante en las telecomunicaciones.

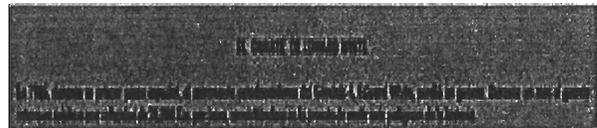
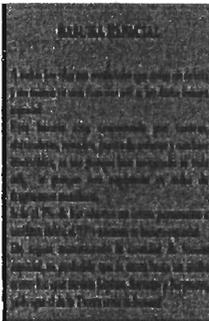
Desafortunadamente, para servir los satélites al espacio y poderlos poner en operación, se requiere de cables y otras maquinarias que son desechados una vez que cumplen su misión. Entre esas residuos espaciales están los satélites alrededor de la tierra. Actualmente hay

cerca de 9,000 objetos espaciales de 10 cm de diámetro y millones de objetos del tamaño de una bola que siguen a velocidades increíblemente altas.

El problema principal de estos residuos espaciales es que pueden chocar contra los satélites y causarles daños irreparables. Cuando los objetos son grandes, espaciales de 10 cm, se pueden detectar con anticipación y maniobrar los satélites para evitar colisiones. Cuando los objetos son pequeños, menores a 1 cm, se pisan los datos recabados a los satélites espaciales con un bombardeo muy silencioso. El mayor problema lo representan los fragmentos metálicos de

hierro 1 y 10 cm, porque es muy difícil detectarlos y evitarlos en su trayectoria de movimiento. Esto puede causar el impacto de estos objetos.

Existen algunas propuestas específicas encaminadas a solucionar el problema de los residuos espaciales. Entre ellas está el Proyecto Orion, que utilizará rayos láser lanzados desde la Tierra para desviar la trayectoria de los objetos en órbita, bien para hacerlos caer a la atmósfera y desintegrarlos, o bien para enviarlos a órbitas de mayor altura donde no representen peligro alguno para los satélites en uso.



102. El texto "Sabías qué...?" se trata de:

- cómo se generan los residuos espaciales, los problemas que causan y los distintos modos que se tiene para resolverlos.
- cómo los científicos detectan los objetos en órbita, los clasifican en chicos, medianos y grandes, y los desintegran.
- cómo la antigua Unión Soviética ganó la carrera espacial al ser el primer país en lanzar el satélite Sputnik

Tomado de INEE, 2004, p. 75

Este reactivo presenta las siguientes características:

- o Reactivo de comprensión global, que requiere de una evaluación de la información que se presenta.



- Se trata de un texto científico que puede resultar de interés para los alumnos por el tema que aborda, aunque convendría saber que tan familiarizados están los estudiantes con el contenido del mismo, pues el desconocer algunos de los términos manejados, puede dificultar el establecer las relaciones necesarias entre lo que se lee, lo que sabe, lo que el texto dice y los propósitos que motivan la lectura, que en sí es lo que el INEE, considera como comprensión lectora.

El análisis de los ítems seleccionados para la evaluación de la comprensión lectora en las pruebas nacionales denota que:

- ⇒ Existe un desfase entre lo que se considera como comprensión lectora y el contenido de los ítems; en virtud de que las situaciones de lectura que presentan, están totalmente ajenas a la realidad y contexto de un alumno de sexto grado, lo que impide que el estudiante pueda relacionar la información del texto con su conocimientos previos y darle sentido a lo que el autor quiso comunicar.
- ⇒ Aunque las tareas de comprensión lectora, contemplan procesos de interpretación, inferencia y evaluación, el tipo de ítems dificulta ver a la lectura como un proceso de comunicación.
- ⇒ El tipo de ítems a los que tuvimos acceso, no contempla capacidades de comunicación de la lengua hablada o escrita, lo que muestra un desfase entre el propósito de lectura y ítems para evaluarla.

#### ▪ Evaluación de las matemáticas en las Pruebas Nacionales

Las Pruebas Nacionales del INEE, conciben a las matemáticas conforme a lo establecido en los Planes y Programas de estudio para la Educación Primaria (Quezada, 2004: 62); y señalan que las matemáticas en la educación primaria tienen como propósitos desarrollar:

- La capacidad de utilizar las matemáticas como una herramienta para reconocer, plantear y resolver problemas.
- La capacidad de anticipar y verificar resultados.



- La capacidad de comunicar e interpretar información matemática.
- La imaginación espacial.
- La habilidad para estimar resultados de cálculos y mediciones.
- La destreza en el uso de ciertos instrumentos de medición, dibujo y cálculo.
- El pensamiento abstracto por medio de distintas formas de razonamiento, entre otras, la sistematización y generalización de procedimientos y estrategias.

Para la evaluación de las matemáticas el INEE, con la ayuda de la Dirección de Materiales y Métodos Educativos de la SEP, se dio a la tarea de hacer una revisión de las habilidades matemáticas y secuencias conceptuales especificadas en los planes y programas de estudio y demás materiales de apoyo de primaria y secundaria. Con base en esto, decidió elaborar instrumentos de medición orientados a las habilidades de pensamiento, implícitas en los planes de estudio de primaria y secundaria.

En las Pruebas de Estándares, se establecen seis de las siete habilidades matemáticas comprendidas en el Plan y Programa de Educación Primaria.

Estas son:

- Comunicar e interpretar ideas, cantidades o números mediante símbolos, signos, diagramas, tablas o gráficas.
- Imaginar: construir ideas, evocando imágenes de objetos, sucesos, relaciones, atributos o procesos.
- Generalizar: analizar regularidades para encontrar patrones.
- Medir: Establecer relaciones entre magnitudes para encontrar una medida.
- Operar: Establecer relaciones entre números o expresiones para producir un resultado.
- Inferir: establecer relaciones causa-efecto, o entre datos explícitos o implícitos que aparecen en un texto, ilustración, tabla, gráfica o diagrama para encontrar un resultado (cfr. Martínez y Schmelks, s/l. ).

A partir de éstas el INEE establece "seis habilidades medibles" que se aplican en la resolución de problemas:



- Operar
- Medir
- Comunicar
- Imaginar
- Generalizar
- Inferir

*Operar:* Descubrir o establecer relaciones entre números, entre unidades de medida, entre figuras o cuerpos o entre las partes de un instrumento físico o gráfico con la finalidad de comprender los procesos que subyacen en las relaciones y obtener resultados de manera más eficaz.

*Medir:* Capacidad que muestran los sujetos para establecer la magnitud de un fenómeno cotidiano (espacial, de tiempo, de masa o peso, de temperatura, etc. ) considerando sus distintas unidades. Además de las comparaciones directas para las distintas clases de características involucradas en determinadas.

*Imaginar:* Capacidad del sujeto para describir la situación resultante de una o varias transformaciones espaciales sobre un objeto o situación inicialmente dada y emitir un juicio o resultado, producto de las transformaciones mencionadas.

*Comunicar – interpretar:* Capacidad que muestran los sujetos para traducir situaciones matemáticas entre los distintos lenguajes que se utilizan en las mismas; sean éstos: numérico, coloquial, algebraico, geométrico o figurativo en general, etc. Por otra parte, también se considera aquí la lectura comprensiva de las situaciones expresadas en los lenguajes arriba mencionados, así como la descripción que los sujetos puedan hacer de una situación determinada. Es claro que en este último caso se observa la precisión y concisión de dicha tarea.

*Generalizar:* Capacidad de los sujetos para descubrir los elementos fundamentales de una situación determinada y ampliarlos por extensión o desarrollo de la situación inicial propuesta.



*Inferir*: Manipulación que el sujeto realiza con la información de una situación determinada para deducir información nueva consistente con la inicial (Quezada, 2004: 63-64).

Nótese que tres de las habilidades señaladas, se definen como *capacidades*, el considerar el carácter innato de las capacidades y a las habilidades como capacidades desarrolladas, se puede decir que las "habilidades" establecidas para la evaluación matemática, tendrán que ser capacidades altamente desarrolladas en los Planes y Programas de Estudio de la Educación Básica.

Al igual que en la comprensión lectora se cuidó que los problemas presentados estuvieran acordes con los documentos curriculares vigentes; con base en los cuales se hizo la distribución y relación entre: No. de reactivo, habilidad, tarea y contenido curricular. (véase Quezada, 2004).

Una vez establecida la relación entre habilidades, tareas y contenidos curriculares a medir, se establecieron cuatro "*niveles de desempeño*" que permiten ubicar a los alumnos y alumnas de acuerdo a lo que saben y en relación al contenido de las pruebas.<sup>154</sup>

En el siguiente cuadro se sintetizan la descripción de los niveles de desempeño en matemáticas que el INEE establece para el 6° grado de educación primaria.

---

<sup>154</sup>Los niveles de desempeño establecidos por el INEE, tienen la intención de que los resultados no se limiten a "aprobar" o "reprobar" a un alumno; sino que constituyen una pauta para interpretar el logro alcanzado en las pruebas nacionales del ciclo 2002-2003, para ubicar a los estudiantes en relación al contenido de las pruebas y saber qué es lo que pueden hacer los alumnos y qué es lo que no todavía no pueden hacer, pero que se espera que hagan. El manejo que el INEE le da a los niveles de desempeño, permite "interpretar cada nivel como la descripción de un estado transitorio que podría cambiar a causa de la acción docente, la intermediación de la escuela, la influencia del medio familiar o social, así como de la voluntad de los propios alumnos" (Quezada, 2004:79). Nótese que es un enfoque diferente al planteado por el LLECE.



CUADRO 3.6

**NIVELES DE DESEMPEÑO PARA LA EVALUACIÓN MATEMÁTICA  
EN EL 6° GRADO DE EDUCACIÓN PRIMARIA  
PRUEBAS NACIONALES (INEE)**

IV	<ul style="list-style-type: none"><li>• Realizan las tareas ya mencionadas en los niveles anteriores pero en ejercicios con mayor dificultad: resuelven problemas que implican situaciones proporcionales.</li><li>• Realizan transformaciones de unidades en sistemas no decimales.</li><li>• Muestran un buen manejo del concepto de área.</li><li>• Manejan correctamente el concepto de unidad de medición y están conscientes de su arbitrariedad.</li><li>• Realizan rotaciones y traslaciones imaginarias de figuras con fines específicos, descomponen figuras complejas en sus elementos más simples y también realizan la operación inversa; es decir, estructuran figuras sencillas para generar otras más complejas con propiedades específicas.</li><li>• En cuanto a sus razonamientos, establecen desarrollos lógicos de hasta tres etapas o pasos inferenciales.</li><li>• Establecen el orden tricotómico en los números racionales dentro de su expresión decimal y en particular realizan inferencias a partir de las características tricotómicas de las mediciones.</li><li>• Sólo hasta este nivel se observa que los estudiantes estiman el rango de valores que puede mostrar cierto resultado y también estiman el «tamaño» de una cantidad asociada de los cuerpos materiales.</li><li>• También se observa, aunque es poca la evidencia, que manejan el orden de operaciones aritméticas cuando los procesos no son conmutativos.</li></ul>
III	<ul style="list-style-type: none"><li>• Aplican correctamente relaciones de proporcionalidad aun cuando las cantidades que en ellas intervienen sean fraccionarias.</li><li>• Manejan bien las equivalencias entre fracciones.</li><li>• Resuelven problemas asociados y también aquellos que requieren además de transformaciones de unidades.</li><li>• Respecto a la lectura e interpretación de gráficas, el conocimiento de términos propios de las matemáticas permite un desempeño más preciso que en los niveles anteriores, y no sólo en aspectos de lectura e interpretación de gráficas.</li><li>• En general, parece que hay cierta correlación positiva entre el conocimiento de términos matemáticos y la sistematización con la lectura de imágenes complejas, que no necesariamente son gráficas.</li><li>• Se observa que los alumnos en este Nivel, traducen planteamientos propuestos en términos de figuras.</li><li>• Leen y consideran todos los elementos relevantes de una figura para emitir un juicio acerca de ella.</li><li>• Pueden encontrar los patrones aditivos en secuencias de figuras – sean éstos constantes o variables-, o interpretar representaciones no ortodoxas de elementos numéricos.</li></ul>



	<ul style="list-style-type: none"><li>• Los efectos de buenas lecturas de figuras en los aspectos geométricos no pasan desapercibidos, pues se observa para representaciones planas de figuras, estiman longitudes en las figuras que se les presentan; o bien, las descomponen en sus elementos rectangulares y triangulares, para realizar cálculos de perímetros o de superficies, sea por superposición o aplicando los algoritmos tradicionales correspondientes.</li><li>• Reconocen y aplican las propiedades de los segmentos de un círculo para resolver ejercicios con estos elementos.</li><li>• Traducen a diferentes representaciones con suficiente soltura, es decir, interpretan diagramas sencillos de la vida cotidiana, aplican sus conocimientos de geometría para realizar inferencias en situaciones mecánicas o transformaciones evolutivas en la propia geometría.</li><li>• Descartan información no relevante en expresiones del lenguaje cotidiano, en expresiones numéricas y en las figuras en general. Pueden inclusive, considerar expresiones combinadas como las manipulaciones de las cotas de las figuras correspondientes, o tomar en cuenta dos condicionantes para obtener la conclusión requerida.</li></ul>
II	<ul style="list-style-type: none"><li>• Manejan adecuadamente las proporciones y aplican estos conceptos a situaciones que involucran áreas y cantidades, asociando tamaños de superficies a las cantidades correspondientes.</li><li>• Encuentran el factor de proporcionalidad existente entre varios números y resuelven problemas asociados a figuras semejantes, lo cual significa que aplican sus conocimientos a situaciones cotidianas como pueden ser proporcionalidades en situaciones comerciales o lectura de gráficas de barras, o de pastel.</li><li>• El manejo que poseen estos alumnos de las proporciones y sus elementos, representan una buena base para iniciarse en transformaciones algebraicas, ya que explícita o implícitamente, hacen uso de la llamada «ley de cancelación del producto».</li><li>• El nivel de abstracción mostrado por los alumnos de este nivel les permite realizar traducciones de códigos para operar con ellos. Esto también resulta básico en su formación pre-algebraica.</li><li>• Respecto a los conceptos de probabilidad, comparan la magnitud de la probabilidad de eventos a partir de situaciones gráficas; es decir, nuevamente aplica sus conocimientos de proporcionalidad y los aplica en situaciones de distinta índole.</li><li>• Sin embargo, se observa que a los alumnos les hace falta ampliar el léxico propio de las matemáticas para describir y comprender descripciones con la precisión requerida.</li><li>• En el aspecto espacial, la descripción poco precisa de lo observado conlleva poca sistematización en la revisión de las gráficas o figuras en general y con ello aspectos o características no consideradas.</li></ul>
I	<ul style="list-style-type: none"><li>• Poseen un buen manejo de las razones y proporciones, concepto que pueden traducir a la representación como la división.</li><li>• Tienen problemas para manejar el concepto de la división como operación inversa a la multiplicación.</li><li>• Durante la resolución de problemas que implican aplicaciones de ideas de proporcionalidad, sólo atienden a aquellas que involucran cantidades enteras; cuando</li></ul>



	<p>las proporciones involucran números fraccionarios o racionales, los alumnos sólo consideran la parte entera de los números.</p> <ul style="list-style-type: none"><li>• Otra de las dificultades con las cuales se enfrentan los alumnos de este nivel, es que operan el algoritmo de la proporcionalidad o regla de tres simple de forma equivocada, lo cual implica: que no estiman resultados, como lo sugiere el programa y tampoco tienen claridad en lo que están haciendo.</li><li>• En cuanto a la lectura e interpretación de gráficas, el error más frecuente que cometen los alumnos es no terminar el proceso, y quedarse únicamente en alguna de las etapas intermedias.</li><li>• Con respecto a la elaboración de gráficas, correspondiendo con el mal manejo que presentan en el tema de las proporciones, los alumnos establecen el orden según el tamaño de las diferentes secciones de una gráfica, pero no así las proporciones que guardan entre ellas.</li><li>• En los sólidos y en figuras planas en general, las relaciones no proporcionales, como son las correspondientes a las dimensiones lineales con sus respectivas áreas, o volúmenes, no las manejan. Los alumnos sólo atienden a la forma lineal y como se observó, aún en ese aspecto, no muestran el dominio esperado.</li></ul>
--	---

Adaptado de Quezada, 2004, p. 85 y 86

Para asegurar el apego del currículum a lo que se está evaluando,<sup>155</sup> se establecieron ítems<sup>156</sup> para cada una de las habilidades propuestas, las cuales se operacionalizaron en tareas y se vincularon con los contenidos curriculares de los planes y programas de estudio.

▪ **Ítems de matemáticas para el 6° grado.**

De la prueba de matemáticas para el 6° grado se publicaron los siguientes ítems:

<sup>155</sup> Nótese que aunque el INEE presenta cuadros para vincular los ítems, las habilidades y las tareas con los contenidos curriculares del Plan y Programa de estudio, no da muestra de los ítems utilizados. (véase Quezada, 2004:30-53).

<sup>156</sup> Cabe señalar que de las primeras pruebas nacionales aplicadas a fines del ciclo escolar 2002-2003 no hay muestras de estos ítems. Solamente se mencionan y numeran. De ahí que se reconoce que el INEE se ha esmerado por presentar la relación existente entre reactivo, tarea y contenido curricular, este análisis quedaba incompleto al no presentar una muestra de los ítems utilizados para poder identificar en estos la tarea y habilidad señalada.



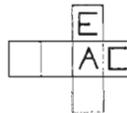
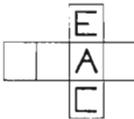
**PRUEBAS NACIONALES DEL INEE (2003-2004)**  
**ÍTEM DE MATEMÁTICAS 6° GRADO**

**Reactivo fácil**

29. Observa el dado que sólo tiene tres letras:



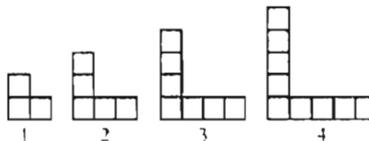
¿Con cuál de las plantillas se armó el dado?



Tomado de Quezada., 2004, p. 85 y 86

**Reactivo difícil:**

88. Observa la secuencia de figuras. Si la secuencia continua, ¿Cuántos cuadritos tendrá la figura número 8?



- 13
- 15
- 17
- 19

Tomado de Quezada., 2004, p. 77



En este reactivo, se puede observar los siguiente:

- La pregunta No. 29, denominado por el INEE como "reactivo fácil",<sup>157</sup> es un ítem cuyo contenido está relacionado con la habilidad de imaginar, implica cierta representación mental y capacidad de discriminación, además de un conocimiento de las características de las figuras tridimensionales. Sin embargo la situación que presenta, está ajena a las realidades cotidianas de un alumno de sexto grado de primaria; se trata de un reactivo un tanto abstracto, de cierto grado de dificultad, por las capacidades que se tienen que poner en juego, pero que no requiere de la transferencia.
- El reactivo de la pregunta No. 88, denominado como "difícil," contempla habilidades de imaginación, generalización e inferencia necesarias para encontrar la relación numérica con respecto al número de cuadros que existe en cada figura. Se trata también de una tarea descontextualizada que si bien comprende la activación de un mayor número de procesos mentales, la respuesta se puede encontrar fácilmente al contar los cuadros de cada figura y reparar la diferencia entre ellos.

Al revisar los procesos mentales implicados en cada tarea, notamos una confusión mayor con respecto a lo que se entiende por reactivo "fácil" o reactivo "difícil", que obedece a la falta de criterios claros para su elaboración.

El análisis de las características de los ítems publicados por el INEE para la evaluación de las matemáticas en sexto grado de la educación primaria, permite enunciar los siguiente:

- ⇒ En cuanto al contenido de evaluación, los ítems que se publican, no contemplan habilidades para la resolución de problemas y para el desarrollo del razonamiento matemático en situaciones prácticas; lo que refleja un desfase entre el propósito y contenido de la evaluación matemática y el tipo de ítems con que ésta se realiza.

---

<sup>157</sup> La denominación "reactivo fácil" queda un tanto ambigua puesto que aunque puede corresponder a una escala, no hay criterios establecidos para comprender que es lo que se entiende por reactivo "fácil" o "difícil".



- ⇒ En relación a los procesos, los niveles de desempeño que se describen no incluyen el comportamiento de las habilidades establecidas por el INEE para la evaluación matemática.
  - ⇒ De acuerdo a las tareas asignadas en cada reactivo, hay cierta incongruencia entre denominación del reactivo conforme a su grado de complejidad y las situaciones y procesos mentales que comprende la tarea a realizar.
  - ⇒ Se trata de ítems descontextualizados, lejanos a situaciones reales que se le pueden presentar a un alumno de sexto grado de primaria, lo que impide el uso y aplicación de las matemáticas para resolver problemas, uno de los principales objetivos establecidos en los planes y programas de estudio de la educación primaria.
- **Escalas y normas técnicas de las pruebas nacionales:**

Las Pruebas Nacionales se estructuran como pruebas referidas a criterios; es decir no están constituidas con base a una norma; de lo que se trata es de medir la habilidad real de un alumno y no la distancia de éstos con respecto a la media poblacional.

Las pruebas se califican conforme a puntajes obtenidos por los alumnos, establecidos conforme a escalas determinadas a partir del número de aciertos y errores. La media teórica es de 500 puntos, lo que indica que los valores extremos en cuanto a puntajes que pueden alcanzar los alumnos va de 200 como mínimo a 800 como máximo. "La escala se construye mediante procedimientos estadísticos basados en Teoría de Respuesta al Ítem;<sup>158</sup> se utilizó un modelo de un solo parámetro, o modelo de Rasch"(INEE, 2004: 57).

La revisión y análisis de las recientes Pruebas Nacionales aplicadas por el INEE y de los ítems muestra publicados en los informes de este instituto, da cuenta de que:

---

<sup>158</sup> Conviene señalar que la inclusión de la Teoría de respuesta al ítem en las pruebas nacionales es de reciente incorporación.



- ▶ Aunque a la evaluación de la *comprensión lectora y las matemáticas*, le precede una definición del contenido a evaluar y las habilidades y tareas que se comprenden se especifican conforme a niveles de desempeño; No existe una conceptualización y organización propia para la prueba, sino que lo que el INEE plantea, es retomado de los Planes y Programas de estudio para la Educación Primaria.
- ▶ La denominaciones que emplea para presentar la organización de su contenido de evaluación, es retomado del marco de evaluación de las Pruebas Internacionales.
- ▶ En la evaluación de la comprensión lectora, no existe una vinculación, entre lo que se establece como comprensión lectora y el tipo de ítems con que ésta se evalúa.
- ▶ En la evaluación matemática, las situaciones y procesos implícitos en los ítems, no contemplan el propósito de la educación matemática establecido en los planes y programas de estudio para la educación primaria.
- ▶ Las situaciones incluidas en los ítems no favorecen el uso de la lectura y las matemáticas en situaciones de la vida cotidiana.

#### **4. Examen Nacional de Ingreso a la Educación Media Superior para alumnos que concluyen la educación básica. (EXANI-1)**

##### **□ Orígenes del EXANI-1**

El Examen Nacional de Ingreso a la Educación Media Superior (EXANI-1), creado fundamentalmente por la Coordinación Nacional para la Planeación de la Educación Superior (CONPES), cuyos orígenes datan del año 1996,<sup>159</sup> se estructura y elabora por el Centro Nacional de Evaluación de la Educación Superior (CENEVAL) con base en las normas,

---

<sup>159</sup> El Concurso de Selección para la Educación Media Superior, convocado en el año de 1996 por la Comisión Metropolitana de Instituciones Públicas de Educación Media Superior, se sitúa como un antecedente importante del EXAN-1 y se establece como respuesta para atender a las problemáticas originadas por el incremento de la demanda de la educación media superior que las instituciones por sí solas no podían resolver. (Véase Comisión Metropolitana de Instituciones Públicas de Educación Media Superior. *Concurso de Selección para la Educación Media Superior 1996*).



políticas y criterios que establece su consejo técnico. (Véase <http://www.cenevalac.edu.mx/portalceneval/index.php?q=info.fichas.ficha54>)

□ **Propósito del EXANI- 1:**

El EXANI -1, examen dirigido a jóvenes que actualmente se encuentran cursando el tercer grado de secundaria, con fines meramente de diagnóstico; de acuerdo a lo señalado por CENEVAL, es un examen para explorar lo que se sabe y cómo se usa lo que se sabe, es decir las habilidades<sup>160</sup> que deben haber desarrollado los estudiantes durante los estudios de secundaria y que son indispensables para enfrentar con éxito los aprendizajes que se ofrecen en la educación media superior. Los resultados del EXANI-1, permiten identificar a los aspirantes con mayores probabilidades de éxito en la educación media superior (CENEVAL, 2004).

▪ **Contenido del EXANI- 1:**

El EXANI-1, es una prueba de razonamiento y conocimientos básicos en donde el tipo de preguntas que incluye están organizadas de acuerdo a dos campos de evaluación:

- Habilidades intelectuales básicas
- Conocimientos disciplinarios.

---

<sup>160</sup> Nótese que aquí se maneja habilidades como uso de conocimientos.



CUADRO 3.7

**EXANI-1  
CONTENIDO DE EVALUACIÓN**

Habilidades intelectuales básicas	Habilidad verbal	Comprensión de lectura Sinónimos Antónimos Analogías
	Habilidad matemática	Sucesiones numéricas Series espaciales Imaginación espacial Problemas de razonamiento
Conocimientos disciplinarios	Español Historia Geografía Formación cívica Ética Matemáticas Física Química Biología	

Adaptado de: CENEVAL, 2004. <http://www.cenevalac.edu.mx/portalceneval/index.php?q=docs.desp&ndf=49>

□ Ítems del EXANI-1

Para el análisis de los ítems se seleccionaron ítems referentes a la “Habilidad verbal”, “Habilidad matemática” y de algunos “Conocimientos disciplinarios.” Éstos fueron extraídos, de los publicados tanto en la página electrónica del CENEVAL, como en la guía y CD.<sup>161</sup>

<sup>161</sup> Véase: <http://www.cenevalac.edu.mx/portalceneval/index.php?q=docs.desp&ndf=49> ; CENEVAL, 2004 y Guía Interactiva Examen Nacional de Ingreso a la Educación Media Superior Documento en CD. Documentos elaborados como ayuda a los estudiantes que presentan estos exámenes.



- Ítems de comprensión lectora para alumnos que concluyen la educación básica:

Lee con atención el texto<sup>162</sup> y contesta las preguntas 1 a 4:

En Siberia, la palabra ukok quiere decir el fin de todas las cosas. Así que no es extraño que éste sea el nombre elegido para una inmensa praderasituada a 2,000 metros de altitud al sur de la región. Hace algunos millaresde años, Ukok se extendía hasta más allá de donde alcanza la vista humana sin que pudiera encontrarse una frontera. Ahora, la mesetasiberiana ha quedado incrustada como una cuña entre cuatro países distintos: China, Mongolia, Kazajstán y Rusia.

Aquí, el verano transcurre ajeno a cualquier presencia humana, sólo visitado por las continuas tormentas de granizo. Pero, durante el invierno, llegan hasta el lugar docenas de pastores en busca de comida para sus bestias. Animales y hombres tienen que prepararse para pasar meses enteros a temperaturas de -30° y protegerse contra los temidos vientos que levantan la nieve del suelo y la lanzan a gran velocidad contra todo lo que hallan a su paso. Hace 2,500 años, sobre estas mismas nieves invernales, y sobre sus sucesoras flores amarillas de primavera, cabalgaba un pueblo preparado para la guerra, la caza y el pastoreo como ningún otro en la región: los pazyryk, una estirpe de valientes jinetes que vivían en la estepa. Este clan estaba emparentado con los escitas, antiguo pueblo seminómada que provenía de las regiones asiáticas donde actualmente se encuentra Irán. Hoy, los miembros de un equipo científico del Instituto Ruso de Arqueología y Etnografía de Novosibirsk desenterran uno a uno, fascinados y con mano trémula, los vestigios más sorprendentes de aquella civilización. Entre ellos, las dos joyas más valiosas son las momias de una dama de la alta sociedad pazyryk (siglo VI antes de Cristo) y de un joven jinete enterrado junto a su caballo (siglo III o II antes de Cristo). Este último, que es el hallazgo más reciente, permite a los expertos albergar esperanzas sobre la posibilidad de que haya más restos momificados escondidos en las frías tierras de Ukok. La momia del joven escita fue encontrada con sus vestimentas de caballero -gorro, abrigo de marmota, pantalones rojos y botas de fieltro casi intactas. Su cabeza estaba adornada por dos largas trenzas pelirrojas que, sorprendentemente, han desafiado incólumes el paso de más de dos mil años bajo tierra, y en la piel de la espalda y el pecho tenía tatuados dibujos de un gran alce.

Este último dato contradice algunas teorías históricas sobre el papel de los tatuajes en las antiguas comunidades euroasiáticas. Tradicionalmente se había creído que los escitas y otros pueblos utilizaban los adornos en la piel como símbolo de nobleza o posición jerárquica elevada. Sin embargo, este joven recientemente desenterrado era un simple guerrero nómada sin abolengo. ¿Por qué estaba también tatuado? El estudio de posibles futuros restos podrá, quizás, dar una respuesta a esta pregunta.

Mientras tanto, su momia ya ha sido trasladada a Moscú, al mismo laboratorio en donde se conserva el cuerpo embalsamado de Lenin, para que sus técnicos estudien cómo mantener los restos de este hombre fuerte y alto que tenía 30 años de edad cuando murió.

"La dama de hielo", en *Muy interesante*. Año XIII, No. 5, México, pp. 39-40.

<sup>162</sup> El texto utilizado se denomina: "La dama de hielo", publicado en *Muy interesante*, Año XIII., No. 5, México, pp. 33-40



1. Las características más sobresalientes de la momia de guerrero son:
  - a) el cabello y el tatuaje
  - b) la vestimenta y la piel
  - c) la vestimenta y las joyas
  - d) el gorro y las botas
  - e) el abrigo y los pantalones rojos
2. De la región de Irán provienen los:
  - a) Escitas
  - b) Rusos
  - c) Nómadas
  - d) Pazyrik
  - e) Euroasiáticos
3. Los párrafos cuatro y seis se relacionan porque hablan sobre:
  - a) Las costumbres guerreras de los escitas
  - b) Los descubrimientos arqueológicos de Ukok
  - c) Los tatuajes encontrados en la momia
  - d) Las condiciones de vida de los escitas
  - e) Las características físicas de los nómadas
4. La idea principal del párrafo seis se refiere a:
  - a) La descripción de l guerrero
  - b) La vida del guerrero
  - c) La muerte del guerrero
  - d) La descripción de la dama
  - e) La narración de un descubrimiento

Tomado de: CENEVAL, <http://www.cenevalac.edu.mx/portalceneval/index.php?q=docs.desp&ndf=49> p. 42, 43

El análisis de estos ítems permite decir que la evaluación de la comprensión lectora en el EXANI-1 presenta las siguientes características:

- Se trata de textos encaminados a identificar información, más que a darle un uso personal y social a la misma. Lo que hace pensar en un concepto la lectura un medio para adquirir información.
- Los procesos implícitos en las tareas se refieren a la identificación, comparación, análisis-síntesis e interpretación. No existen ítems de reflexión.
- Las tareas para evaluar la comprensión lectora se centran en la localización de información, relación y ubicación de la idea principal en un párrafo.
- En los ítems se utilizan textos lejanos a la realidad de los estudiantes.

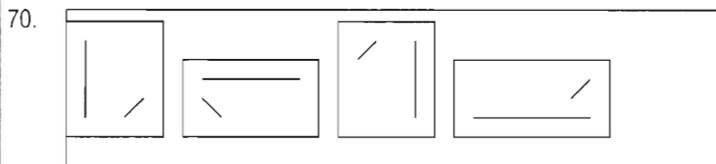


- Ítems de habilidad matemática para alumnos que concluyen la educación básica

En la pregunta 65 a 68, selecciona la opción que contenga el término que sigue en la sucesión presentada.

65.
- a. 10
  - b. 8
  - c. 7
  - d. 6
  - e. 4

En las preguntas 69 a 72, elige la opción que complete la serie presentada.



- a)
- b)
- c)
- d)
- e)

Tomado de CENEVAL, <http://www.cenevalac.edu.mx/portalceneval/index.php?q=docs.desp&ndf=49> p. 63



En estos ítems se puede observar que en el EXANI-1, la evaluación de las matemáticas se realiza bajo las siguientes circunstancias:

- Las tareas se refieren a encontrar la relación numérica o gráfica de la secuencia presentada.
- Los procesos implícitos en las tareas comprenden procesos de ordenamiento, percepción y análisis.
- Con base en los procesos incluidos en las tareas se puede decir que la habilidad matemática se considera como el ejercicio<sup>163</sup> de operaciones mentales encaminadas al tratamiento y manejo de la información.
- El tipo de ítems presenta situaciones un tanto abstractas, lejanas de la realidad y necesidades cotidianas de los estudiantes.

▪ **Ítems para evaluar conocimientos disciplinarios en alumnos que concluyeron la educación básica:**

Para evaluar los **conocimientos disciplinarios**, se utilizan ítems como los siguientes:

**Ítem de español**

¿En qué palabra está subrayada la sílaba tónica?

- a) Lapicero
- b) Computadora
- c) Excelencia
- d) Escritura
- e) Moneda

¿Cuál es la categoría gramatical de las palabras escritas en negritas en el siguiente fragmento?

La contemplación amorosa y devota de las maravillas del suelo nos dejó alguna vez impresiones encantadoras y gratas, que conservamos con cariño durante nuestra vida.

Tomado de: CENEVAL, <http://www.cenevalac.edu.mx/portalceneval/index.php?q=docs.desp&ndf=49> p. 48

---

<sup>163</sup> Referimos ejercicio desde la noción de ejercicio didáctico de Snyders, es decir, como simple práctica.



### Ítem de geografía

Las capas de la atmósfera ordenadas de abajo hacia arriba son:

- A) ionosfera, troposfera, estratosfera
- B) estratosfera, troposfera, ionosfera
- C) troposfera, estratosfera, ionosfera
- D) ionosfera, estratosfera, troposfera
- E) estratosfera, ionosfera, troposfera

Tomado de: CENEVAL, <http://www.cenevalac.edu.mx/portalceneval/index.php?q=docs.desp&ndf=49> EXANI p-1, p.23

### Ítem de química

Relaciona cada unidad de medida de la columna izquierda con su correspondiente magnitud en la columna derecha:

#### UNIDADES

- 1. coulomb
- 2. volt
- 3. ampere
- 4. ohm
- 5. watt

#### CONCEPTOS

- a) Corriente eléctrica
- b) Diferencia de potencial
- c) potencia
- d) Carga eléctrica
- e) Resistencia

- a) 1a, 2b, 3c, 4d, 5e
- b) 1b, 2d, 3c, 4e, 5a
- c) 1e, 2c, 3d, 4a, 5b
- d) 1d, 2b, 3a, 4e, 5c
- e) 1d, 2c, 3e, 4a, 5b

Tomado de: CENEVAL, <http://www.cenevalac.edu.mx/portalceneval/index.php?q=docs.desp&ndf=49> EXANI P-1, p.25.

En todos estos ítems, llamados de "conocimiento", el único proceso que se requiere es el de memoria, para la reproducción de un conocimiento factual. No hay ítems de aplicación de



conocimiento; lo que hace pensar en un concepto de conocimiento como mera acumulación de conceptos, hechos y principios.

#### □ Normas técnicas para la elaboración del EXANI-1

Las normas técnicas para la elaboración del EXANI-1, del CENEVAL, se obtuvieron de la Guía existente (CENEVAL, 2004)<sup>164</sup> de donde se pudieron extraer las siguientes características:

- El EXANI-1, incluye preguntas cuya gama de dificultad va desde las muy fáciles (aquellas que la gran mayoría de los sustentantes pueden responder correctamente), hasta las muy difíciles (aquellas que sólo los sustentantes muy capaces, pueden responder correctamente). Se puede decir que es una prueba cuyo nivel de dificultad no está centrado en los ítems, sino en el % de alumnos que pueden contestar correctamente cada reactivo de la prueba, lo que hace pensar que está sustentado en una Teoría Clásica de los Test (TC).
- El formato de los ítems es de opción múltiple, el cual sólo exige que se identifique la respuesta correcta, sin necesidad de construirla.

La presentación y análisis de los ítems que hemos seleccionado de las pruebas del EXANI-1, da cuenta de que:

- ▶ Aunque se dice que el EXANI-1 es un examen para explorar lo que se sabe y cómo se usa lo que se sabe en términos de habilidades, no hay una definición de lo que se entiende por éstas.
- ▶ El tipo de ítems utilizados para medir la llamada “habilidad verbal” o “habilidad matemática” y los procesos mentales implícitos en estos, hacen pensar que en las

---

<sup>164</sup> Cabe aclarar que se recurrió a la Guía del EXANI-1, para analizar las normas técnicas para la elaboración del EXANI-1 porque a diferencia de las pruebas de organismos internacionales, no se han encontrado documentos de acceso público del CENEVAL, en donde se presenten las técnicas para la elaboración de los ítems.



pruebas del EXANI- 1, la comprensión lectora reside en un mero proceso de identificación, comprensión, análisis, síntesis e integración. La habilidad matemática es trabajada como un ejercicio de ordenamiento, percepción y análisis. El conocimiento es conceptualizado como mera acumulación de información y no como una aplicación de éste a hechos o situaciones cotidianas.

- Los procesos implícitos en los ítems utilizados, quedan en un mero nivel de conocimiento, que si recordamos la taxonomía de Blomm,<sup>165</sup> éste es el nivel más bajo y se refiere a la operación mental con la cuál recordamos hechos, eventos, símbolos, fórmulas y teorías que han sido consideradas anteriormente en la memoria; las tareas o acciones implícitas en estas residen en: definir, repetir, memorizar, registrar, nombrar, relatar, subrayar, identificar; las cuales coinciden plenamente con lo requerido en los exámenes del CENEVAL.

La revisión y análisis del origen, propósitos, contenidos y ítems de cuatro de los exámenes masivos aplicados a nivel nacional, permite caracterizar con base en sus áreas de evaluación común, la situación mexicana en cuanto a elaboración de pruebas, para posteriormente compararla con la de la experiencia internacional.

---

<sup>165</sup> Aunque la taxonomía de Bloom ya está superada en cuanto a objetivos de aprendizaje, es importante destacar que los procesos implícitos en las tareas de muchos de los ítems utilizados, coincide con la jerarquización de los niveles cognitivos que éste hace: Conocimiento, comprensión, aplicación, análisis síntesis y evaluación (véase Bloom, 1981).



### Características de los exámenes masivos nacionales

El siguiente cuadro ofrece un panorama del contenido y estructura de los exámenes masivos en México:

CUADRO 3.8

**CONTENIDO Y ESTRUCTURA DE LOS EXÁMENES MASIVOS EN MÉXICO**

AREA	CARACT.	EXÁMENES			
		IDANIS	SNEE	PN	EXANI-1
Comprensión lectora	Contenido	Obtención de información	Obtención de información	Desfasado de los propósitos.	Medio para adquirir información
	Procesos	Identificación Relación Generalización	Identificación Relación Generalización	Inferencia Interpretación Evaluación	Identificación Comparación Análisis Síntesis Interpretación
	Tareas	Localización de información e ideas	Localización de información e ideas.	No favorecen procesos de comunicación oral y escrita	Localización de información Relación Ubicación
	Tipo de ítems	De carácter literal	Multiítems	Lejanos a la realidad de los alumnos	Lejanos de las realidades de los estudiantes
Habilidad Matemática	Contenido	Reproducción y práctica de temas curriculares	Reproducción y práctica de temas curriculares	Desfasado de los propósitos	Tratamiento y manejo de la información
	Procesos	Cálculos descontextualizados Observación Comparación	Cálculo	Centrados en la imaginación, generalización e inferencia	Ordenamiento Percepción Análisis
	Tareas	Operaciones Ejercicios	Ejercicios como práctica	Incongruentes en cuanto a su grado de dificultad	De relación numéricas y gráfica
	Tipo de ítems	Descontextualizados	Descontextualizados.	Descontextualizados	Situaciones abstractas y rutinarias



La información presentada en el cuadro permite observar y comparar las características de los exámenes nacionales y enunciar el contenido y estructura de las pruebas en las áreas de evaluación común:

En cuanto la comprensión lectora se puede decir que:

- La comprensión lectora se concibe como un medio para obtener información.
- Los procesos implícitos en la evaluación de la comprensión lectora, se quedan en un nivel de lectura literal y de interpretación, no llega a la reflexión.
- Las tareas para evaluar la comprensión lectora están enfocadas a la localización de información.
- Las situaciones que se presentan en los ítems, no corresponden a necesidades e intereses reales de los alumnos de educación básica.

La evaluación de la comprensión matemática consiste en:

- Un contenido centrado en la reproducción de contenidos curriculares y en el manejo y tratamiento de la información.
- Procesos centrados en el cálculo y organización de información numérica y gráfica.
- Tareas enfocadas al ejercicio como práctica y no como aplicación.
- Ítems abstractos y descontextualizados.

Estas son el tipo de pruebas que los estudiantes mexicanos están acostumbrados a contestar; aunque se afirme que el IDANIS, de la DGEE o bien, las ahora Pruebas Nacionales del INEE o el EXANI-1 del CENEVAL; son pruebas que miden habilidades, las habilidades o procesos implícitos en las tareas de los ítems, están encaminados a la activación de la memoria para la reproducción de un conocimiento; se desconoce el ejercicio de procesos superiores de pensamiento encaminados al uso, transferencia, o aplicación del conocimiento como ocurre en las pruebas internacionales.



### 3.3. Como conclusión inicial y a manera de síntesis:

El análisis, caracterización y comparación de los exámenes masivos nacionales, da cuenta de importantes problemas de contenido y estructura en materia de elaboración de pruebas en la experiencia mexicana, de entre los cuales se destacan los siguientes:

- ❖ El contenido de evaluación está centrado en los planes y programas de estudio de la educación básica, no existe un contenido propio.
- ❖ La mayor parte de los contenidos a evaluar son de tipo declarativo, lo que hace escasa la evaluación de procedimientos.
- ❖ Los conocimientos evaluados son presentados como un listado de temas, lo que revela cierto enciclopedismo.
- ❖ No hay relación entre la definición del objeto de evaluación y la organización del contenido a evaluar en sus ítems.
- ❖ La conceptualización sobre las habilidades o contenidos que se evalúan es inexistente.
- ❖ Las "habilidades" se evalúan de manera independiente a los contenidos," lo que hace pensar que no se consideran como parte de un proceso de conocimiento.
- ❖ Se refieren habilidades como contenidos, lo que denota no sólo una ausencia de clarificación del concepto, sino una confusión en el uso de los mismos.
- ❖ Se nota la ausencia de una fundamentación que justifique el tipo de ítems utilizados y la forma en como éstos se estructuran e integran en la prueba.<sup>166</sup>
- ❖ La evaluación de la comprensión lectora se limita a la solicitud de información, no llega al manejo de significados.
- ❖ La evaluación del español continua con un enfoque gramatical, más que funcional.

---

<sup>166</sup> El argumento que se les da a los estudiantes para explicarles la estructura del examen es: "el EXANI-1 evalúa sólo los conocimientos y habilidades que van a ser indispensables para que prograses en los estudios del nivel superior, y que debes haber aprendido gracias al trabajo regular realizado en la escuela secundaria." (CENEVAL, 2004:13)



- ❖ La evaluación de las matemáticas reside, la mayoría de las veces, en tareas de cálculo a manera de práctica y realización de ejercicios descontextualizadas, para las cuales generalmente no se requiere el uso del pensamiento lógico.
- ❖ Las tareas incluidas en los ítems se centran en ejercicios abstractos que adolecen de una transferencia.
- ❖ La constitución de sus ítems presenta problemas de redacción y elaboración, lo que les resta confiabilidad y validez.
- ❖ El tipo de ítem al que se recurre se centra en ítems con formato de opción múltiple, no ha hay la demanda de respuestas construidas por parte de los estudiantes a manera de argumentación.
- ❖ Aunque no es del dominio público las normas y procedimientos para su elaboración, se denota la primacía de la Teoría Clásica del Test; aún no se incorporan en todas las pruebas las alternativas psicométricas existentes.
- ❖ La validación de los ítems continua estando en función de los sujetos en los que se pilotea y no de las variables que éstos miden.
- ❖ Aunque el INEE, inicia con la presentación de resultados por niveles de desempeño, no se observa un análisis cognitivo de tareas; para jerarquizarlos; sino que la clasificación se realiza más bien por niveles de dificultad.
- ❖ La difusión de lo que se evalúa y cómo se evalúa a nivel nacional, aún es escasa, pues es difícil acceder a documentos públicos que presenten y difundan el contenido y estructura de las pruebas.

La revisión de los documentos existentes con respecto a la evaluación del aprendizaje en México, permite hacer los siguientes señalamientos:

- ❖ Se puede caracterizar a los noventa como la década de la proliferación de exámenes a gran escala, más no de la evaluación.
- ❖ El origen de la "evaluación" del aprendizaje en México, ha estado vinculado a una política educativa de programas compensatorios y de estímulos.



- ❖ La idea que prevalece en los programas de evaluación en México, ha residido en la mera selección, jerarquización y comprobación, más que la de mejora. Algunas acciones como las del EXANI-1, surgen como respuesta a problemáticas de cobertura y recepción que las instituciones no han podido resolver.
- ❖ La vinculación evaluación-enseñanza-aprendizaje, continua sin resolverse.
- ❖ El tratamiento pedagógico de la evaluación nivel nacional, en las pruebas aplicadas a gran escala, continua como tarea pendiente.
- ❖ Existe una abundante enumeración de acciones, pero sin la construcción y sobre todo difusión de un marco teórico-conceptual que las fundamente.
- ❖ El INEE, comienza en fechas recientes con la elaboración de documentos que clarifiquen el qué y el cómo de lo que se pretende evaluar, pero se desconoce algunos ítems muestra que permitan constatar los lineamientos y características enunciadas.
- ❖ Prevalece una escasa, o más bien nula, formación de docentes, en materia de evaluación.

Una comparación entre las características de las Pruebas Internacionales y las Pruebas Nacionales, permitirá encontrar ejes de acción que ayude a mejorar la elaboración de exámenes masivos en México.

## CAPÍTULO CUATRO

### PRUEBAS INTERNACIONALES Y PRUEBAS NACIONALES: CONTENIDOS Y TEORÍAS

---

La descripción y análisis del contenido, estructura y marco teórico de los exámenes masivos, realizados tanto en el ámbito internacional como nacional, permite ofrecer en este capítulo una comparación que proporcione elementos de corte teórico-metodológico importantes a considerar en materia de elaboración de exámenes tanto en el plano internacional como en el nacional.

Retomamos las categorías establecidas en el capítulo dos, con la intención de reconocer la situación de las pruebas nacionales y su relación con las internacionales aplicadas a poblaciones de edad semejante, en los tres ejes de análisis y comparación planteados para este estudio.

1. Concepto de contenido.
2. Tipo de ítem.
3. Modelo psicométrico.

#### **4.1. Concepto de contenido en los exámenes masivos**

La existencia de una posible Concepto de contenido en exámenes masivos la podemos observar a partir de la relación existente entre conceptualización, dimensiones, campos temáticos y procesos del contenido de evaluación.

Los criterios para la realización de las comparaciones pertinentes giraron en torno a la edad y áreas de evaluación común.



### a. Exámenes masivos para alumnos de ocho y nueve años de edad

Como exámenes masivos para alumnos ocho y nueve años, en el ámbito internacional y nacional, se refieren los realizados en el TIMSS, el LLECE y los del SNEE. Analizamos y comparamos en cada uno de éstos, en sus áreas comunes, para dar cuenta de las semejanzas y diferencias encontradas en cuanto a conceptos, campos temáticos y procesos del contenido de su evaluación.

#### □ Contenido de la evaluación matemática

El siguiente cuadro retoma y conjunta el contenido de la evaluación matemática para alumnos de ocho y nueve años en sus diferentes aspectos.

CUADRO 4.1

#### CONTENIDO DE LA EVALUACIÓN MATEMÁTICA EN EXÁMENES MASIVOS PARA ALUMNOS DE OCHO Y NUEVE AÑOS

EJES	EXÁMENES MASIVOS		
	TIMSS	LLECE	SNEE
<b>Concepto</b>	Pensamiento lógico y sistemático para la transferencia de conocimientos y destrezas	Uso de procedimientos matemáticos denominada competencia matemática.	No se define
<b>Dimensiones</b>	Conocimientos y comportamientos	Habilidades y destrezas	Temas del plan y programa de estudiados para la evaluación matemática.
<b>Campos temáticos</b>	Números Álgebra Medición Geometría Datos	Numeración Operación con números naturales Fracciones comunes Geometría Medición	Equivalencia y división de fracciones Perímetro Problemas de suma y multiplicación La hora Perspectivas Cantidades Planos, Problemas de reparto Problemas de suma y resta Lectura de tablas y



			gráficas de barras Comprensión de datos de un problema Obtención de superficies Uso de planos y trayectorias
<b>Procesos</b>	Conocimientos de hechos y principios Utilización de conceptos Resolución de problemas Razonamiento	Interpretación de gráficos Reconocimiento de patrones Manejo de probabilidades Establecimiento de relaciones entre datos	Identificación Representación Análisis Relación Cálculo

Algunas semejanzas y diferencias que se perciben en el contenido de la evaluación matemática para niños de ocho y nueve años:

- La evaluación de las matemáticas es conceptualizada en las pruebas internacionales del TIMSS y LLECE, como la activación de un pensamiento lógico y sistemático; se evalúa mediante uso de procesos cognitivos y procedimientos matemáticos para la aplicación del conocimiento. En las pruebas del SNEE, no se define, sólo se encuentran reactivos de diversos temas.
- Las dimensiones de las pruebas internacionales del TIMSS y LLECE, incluyen conocimientos y comportamientos<sup>167</sup> o habilidades. En el TIMSS, la organización de su contenido, aborda tanto conocimientos factuales<sup>168</sup> como una comprensión conceptual que exige el razonamiento y análisis de diversos conceptos propios de cada disciplina. El estudio del LLECE está enfocado a detectar las destrezas adquiridas por los educandos en cuanto a comprensión lectora y pensamiento lógico-matemático. Las pruebas del SNEE, no establecen dimensiones y sólo se enfocan a los contenidos del plan de estudios.

<sup>167</sup> Aunque todavía no está resuelto el debate en cuanto a la conceptualización y diferencia entre capacidad, habilidad y competencias, algunos autores, utilizan el término comportamientos para definir a una competencia (Rey, 1996).

Perrenoud (1999), establece que una competencia exige la activación tanto de conocimientos como de operaciones mentales complejas y esquemas

<sup>168</sup> En el marco de la Reforma Española, se denominan *conocimientos factuales* al manejo de hechos y conceptos (véase Coll, Pozo, Sarabia, y Valls, 1992).



- Existen campos temáticos comunes en las pruebas internacionales del TIMSS y LLECE para alumnos de ocho y nueve años: *números, geometría y medición*. En las pruebas del SNEE, no existen campos temáticos, lo que contienen son reactivos pertenecientes a diversos temas, de entre los que sobresalen la resolución de problemas aritméticos. Sólo lo referente a *fracciones*, aparece como un tema común tanto en las pruebas del LLECE, como en las del SNEE.
- Los procesos implícitos en las pruebas internacionales del TIMSS y LLECE, están vinculados con el *razonamiento y la transferencia*. En las pruebas del SNEE, están más enfocados al *cálculo*, sólo el establecimiento de *relaciones*, aparece como un tema común con las pruebas del LLECE.

Las diferencias y semejanzas encontradas; marcan importantes distancias de enfoque y organización entre exámenes masivos internacionales y nacionales. Mientras en el plano internacional las pruebas para la evaluación matemática en niños de ocho y nueve años, establecen campos precisos y genéricos de evaluación que comprenden tanto conocimientos como habilidades y procedimientos vinculados al pensamiento lógico-sistemático, para la aplicación de un conocimiento; en México los exámenes siguen centrados en diversos contenidos del plan y programa de estudios, sin ninguna agrupación, requieren sólo del ejercicio de operaciones muy básicas de pensamiento.



□ Contenido de evaluación en la comprensión lectora

CUADRO 4.2

**EVALUACIÓN DE LA COMPRENSIÓN LECTORA EN EXÁMENES MASIVOS  
PARA ALUMNOS DE TERCERO Y CUARTO PRIMARIA**

	LLECE	SNEE
<b>Concepto</b>	Posibilidades que tiene un estudiante para comprender, organizar, negociar y producir actos significativos a través de distintos sistemas de signos lingüísticos y no lingüísticos de comunicación, considerado como competencia comunicativa.	No se especifica, pero los reactivos están enfocados a la obtención de información y extracción de ideas dentro de un texto.
<b>Dimensiones</b>	Lectura literal primaria Lectura a manera de paráfrasis Lectura de carácter inferencial.	No se especifican
<b>Campos temáticos</b>	Estrategias de pensamiento	Reactivos con temas de español del Plan y Programa de estudios de la educación primaria
<b>Procesos</b>	Posibilidades interpretativas para construir o develar el significado de un texto, formulando hipótesis globales de lectura	Identificación Reconocimiento

El cuadro presenta diferencias significativas, como las siguientes:

- La evaluación de la comprensión lectora en las pruebas del LLECE, es concebida como producción de significados. Los reactivos de las pruebas del SNEE, están enfocados a la obtención de información y extracción de ideas en un texto.
- Las pruebas del LLECE organizan el contenido de su evaluación en tres dimensiones que van desde la lectura literal, hasta la inferencial.
- Los temas de las pruebas del LLECE están centrados en estrategias de pensamiento; mientras que los de las pruebas del SNEE, se enfocan a los contenidos del plan y programa de estudios.



- Los procesos implícitos en las pruebas del LLECE, se enfocan a posibilidades interpretativas, que permitan recuperar el significado de un texto; mientras que las pruebas del SNEE, sólo contemplan la identificación y el reconocimiento.

Estas diferencias hacen pensar que la situación que guardan los exámenes masivos en México, con respecto a la evaluación de la comprensión lectora en niños de ocho y nueve años, dista mucho del contenido que se trabaja en el plano internacional. Mientras que el LLECE, elabora y aplica pruebas de comprensión lectora que privilegian e incluyen procesos para la recuperación del significado de un texto, en tres dimensiones; las pruebas mexicanas aún conservan un enfoque centrado en la obtención de información y en microprocesos, insuficientes para una comprensión total, que requiere de procesos más complejos o macroprocesos vinculados a una macroestructura.<sup>169</sup>

En las pruebas del SNEE, se promulga una evaluación de la comprensión lectora, pero más bien se incluyen dentro de ésta, la evaluación de contenidos gramaticales; la comprensión lectora, "es una actividad constructiva que implica la interacción entre las características del lector y del texto, dentro de un contexto determinado" (Díaz Barriga Arceo y Hernández, 2002:275) Es una actividad constructiva porque no se trata de una transposición unidireccional de los mensajes comunicados en el texto, sino que a partir de los significados sugeridos, el lector tendrá que utilizar sus recursos cognitivos pertinentes tales como habilidades psicolingüísticas, esquemas y demás habilidades para construir una representación fidedigna. La interacción entre las características del lector (intereses, actitudes, conocimientos previos), las del texto (intenciones presentadas explícita o implícitamente por el autor) y el contexto (demandas específicas de la situación real es

---

<sup>169</sup> De acuerdo con lo planteado por Díaz Barriga Arceo . y Hernández (2002), para comprender un texto se requiere que el lector realice actividades de micro y macroprocesamiento...Los microprocesos involucrados en los niveles inferiores de la comprensión son: reconocimiento de palabras escritas en los enunciados, identificación de ideas...Las actividades de macroprocesamiento son de naturaleza abstracta y tienen que ver con la representación semántica y global del texto... Algunos de los macroprocesos más relevantes son: supresión, generalización, construcción, jerarquización de ideas y coherencia global (Cfr. Díaz Barriga Arceo y Hernández, 2002:277).



básica para la interpretación) Y es una actividad estratégica porque requiere de la organización y uso de sus recursos cognitivos. (Díaz Barriga Arceo y Hernández, 2002:276)

### b. Exámenes masivos para alumnos que concluyen la educación básica.

Como exámenes masivos para estudiantes que concluyen la educación básica, a nivel internacional y nacional, se ubican los realizados por el PISA y el EXANI-1. Analizamos y comparamos cada uno de éstos en sus áreas comunes, para dar cuenta de las semejanzas y diferencias encontradas en su contenido.

#### □ Contenido de la evaluación matemática

CUADRO 4.3

### CONTENIDO DE LA EVALUACIÓN MATEMÁTICA EN EXÁMENES MASIVOS PARA ALUMNOS QUE CONCLUYEN LA EDUCACIÓN BÁSICA

	TIMSS	PISA	EXANI-1
<b>Concepto</b>	Pensamiento lógico y sistemático para la transferencia de conocimientos y destrezas	Aptitud para plantear, formular, resolver e interpretar problemas a través de las matemáticas en diferentes situaciones y contextos (OCDE, 2004:27)	No se define
<b>Dimensiones</b>	Conocimientos y comportamientos	Grandes ideas (Contenidos) Competencias (procesos y procedimientos) Situaciones (OCDE, 2000: 25, 26)	Conocimientos disciplinarios y habilidades de sucesiones numéricas, series espaciales, imaginación espacial y problemas de razonamiento. (GENEVAL, 2004: 14)
<b>Contenidos</b>	Números Álgebra Medición Geometría Datos	Cantidades Espacios y formas Cambios Relaciones Incertidumbre (OCDE 2000: 24)	Listado de diversos temas del plan de estudios



<b>Procesos</b>	Conocimientos de hechos y principios Utilización de conceptos Resolución de problemas Razonamiento	Reproducción Conexiones Reflexión Tareas auténticas (OCDE, 2000: 76)	Sucesiones numéricas Series espaciales Imaginación espacial Problemas de razonamiento (CENEVAL, 2004:14)
-----------------	---	---	---

Las semejanzas y diferencias encontradas en el contenido de los exámenes de matemáticas para alumnos que concluyen la educación básica, residen en que:

- Las pruebas internacionales del TIMSS y PISA, manejan un concepto funcional y cognitivo de las matemáticas al conceptualizarlas como la activación de procesos de identificación, lógica, sistematización, comprensión, práctica y transferencia; en el EXANI-1, no se define.
- El TIMSS y el PISA, contemplan dos dimensiones en su contenido de evaluación: Conocimiento y habilidades. (procesos, comportamientos o competencias). En el EXANI-1, no se contempla.
- Los exámenes masivos de corte internacional, establecen cinco campos temáticos, de entre los cuales el manejo de cantidades y la geometría (espacio y forma), aparecen como temas comunes. En el EXANI-1- sólo se presentan tareas encaminadas al tratamiento y manejo de la información.
- En los estudios internacionales, los procesos que se incluyen en la evaluación de las matemáticas son de distinta índole. En el TIMSS, se refieren más a procesos curriculares, vinculados con la aplicación de un conocimiento matemático; el PISA comprende proceso de corte más genérico, funcional y cognitivo. El EXANI-1, sólo se refiere a cuestiones de ordenamiento, percepción y análisis.

Estas diferencias dan cuenta de que mientras los exámenes masivos en el ámbito internacional presentan cierta relación entre conceptualización y contenido de la evaluación matemática que induce a un concepto de contenido, enfocada a la aplicación del conocimiento matemático: el EXANI-1, no ofrece ningún planteamiento teórico y su evaluación se centra a la mera resolución de ejercicios matemáticos.



□ Contenido de evaluación en la comprensión lectora

CUADRO 4.4

**EVALUACIÓN DE LA COMPRENSIÓN LECTORA EN EXÁMENES MASIVOS  
PARA ALUMNOS QUE CONCLUYEN LA EDUCACIÓN BÁSICA**

	<b>PISA</b>	<b>EXANI-1</b>
<b>Concepto</b>	Capacidad de comprender, emplear y reflexionar sobre textos escritos con el fin de lograr las metas individuales, desarrollar el conocimiento potencial y participar eficazmente en la sociedad(OCDE, 2002:22).	La comprensión de textos es una actividad mental que requiere de concentración total para su realización eficiente. (CENEVAL, 2004:37)
<b>Dimensiones</b>	Forma y estructura del material de lectura. Tipo de reactivo. Propósito y uso del texto. (OCDE, 2002: 36-38).	Comprende el reconocimiento del significado de los vocablos, el descubrimiento de las relaciones entre las frases y las oraciones, así como de los distintos párrafos que integran el texto. (CENEVAL, 2004:37)
<b>Tareas</b>	Obtención de información Interpretación de textos Reflexión y evaluación sobre el contenido y la forma del texto. Empleo privado. Uso público. Ocupacionales. Educativas. (OCDE, 2000: 44-51).	No se especifican, pero las tareas residen en: Localización de la información. Relación y ubicación de la idea principal en un párrafo. Decir de que trata un texto o a qué conclusión conduce.
<b>Procesos</b>	Comprensión global Obtener información Interpretación de textos Reflexión y evaluación sobre el contenido y forma del texto (OCDE, 2000:52-58).	Identificación Comparación Análisis Síntesis Interpretación



El cuadro presenta diferencias significativas de entre las cuales se encuentran:

- El concepto de lectura en el proyecto PISA, presenta un enfoque más funcional, en el EXANI-1, no sólo no se define, sino que los reactivos, presentan a la lectura como adquisición de información.
- El PISA, contempla la lectura en tres dimensiones, el EXANI- no lo define.
- Los campos temáticos en el PISA, están definidos de acuerdo a las dimensiones. En el EXANI-1, no se especifican, pero las tareas se centran en localizar información e ideas.
- En el PISA, se especifican los procesos a evaluar en la comprensión de la lectura, los cuales van desde la comprensión, como el proceso básico, hasta la reflexión. El EXANI-1 incluye procesos que van desde la identificación, pero que llegan sólo a la interpretación.

Las diferencias encontradas dan cuenta de que, PISA plantea una evaluación de la lectura, definiendo en primer lugar qué es lo que se entiende por lectura y establece dimensiones, campos temáticos y procesos acordes con esta definición, conformando así una Concepto de contenido de evaluación, que el EXANI-1 no contempla.

El contenido de evaluación de las matemáticas y la comprensión lectora, trae consigo conceptos implícitos que conviene comparar:

### **c. Conceptos estructurantes en los exámenes masivos**

En capítulos anteriores encontramos como conceptos estructurantes dentro de la Concepto de contenido de las pruebas internacionales, los términos de *habilidad, destreza, competencia y comprensión*.



En las pruebas del TIMSS *la habilidad*, se equipara con comportamientos, la *destreza*, se maneja a la par del concepto de habilidad, la *competencia* se refiere a las destrezas implícitas y la comprensión alude al uso y aplicación del conceptos.

En el LLECE, la habilidad se refiere a la capacidad y aptitud para realizar un acto físico o mental y está muy relacionada con el poder hacer algo, la destreza implica la realización de una tarea con maestría y eficacia, la competencia, uno de los conceptos más trabajados, representa la aplicación de un saber, es decir la posibilidad de una acción a partir de un saber inmerso mediante la construcción y apropiación de instrumentos que posibilitan la transformación de la experiencia humana. Y la comprensión se refiere a poner en juego tanto los conocimientos previos como las representaciones y saberes. El PISA entiende la habilidad como destreza y a ésta como ciertos procedimientos agrupados en competencias, las que define como conjunto de conocimientos y destrezas para la vida adulta aplicadas a contenidos y situaciones de la vida diaria. A la comprensión la considera importante para hacer predicciones y explicar relaciones.

En los exámenes masivos de la experiencia mexicana, solamente el IDANIS y las Pruebas Nacionales (PN) del INEE, cuentan con documentos en donde se puede encontrar cierta conceptualización de los términos que incluyen; encontrándose dentro de éstos, el de habilidad como un concepto común, tanto en estas pruebas, como en las del CENEVAL. Sin embargo no se tiene claridad del concepto, en el IDANIS, se le da un manejo superficial y confuso al concepto, porque aunque se afirma que no es una prueba centrada en contenidos curriculares; la habilidad verbal se diagnostica a partir de la comprensión lectora y el completar oraciones; la habilidad matemática está en función de los problemas aritméticos y la geometría y la habilidad para el razonamiento abstracto, está vinculada con la representación gráfica; lo que denota cierta confusión entre contenido y habilidad, aunque se especifiquen los procesos implícitos en cada una de éstas. Se le concibe además como un recurso subordinado a las estrategias y vinculado al conocimiento para procesar información, ya sea lingüística, matemática o gráfica.



En las Pruebas Nacionales (PN), se retoman de los objetivos establecidos en los Planes y Programas de estudio y en el EXANI-1, se presentan como aplicación de conocimientos, pero no se definen.

Comparamos ahora cómo se definen éstos conceptos tanto en el plano internacional como nacional.

CUADRO 4.5

**CONCEPTOS ESTRUCTURANTES DE LOS EXÁMENES MASIVOS**

	PRUEBAS INTERNACIONALES			PRUEBAS NACIONALES			
	TIMSS	LLECE	PISA	IDANIS	SNEE	PN	EXANI-1
Habilidad	Comportamientos	Capacidad y aptitud para realizar algo	Destrezas	Recursos vinculados al conocimiento o para procesar información		Se usa el término para referir habilidades de matemáticas del Plan de estudios	No se define pero se utiliza como aplicación de conocimientos
Destreza	Habilidad	Realización de una tarea con maestría y eficacia	Procedimientos agrupados en competencias				
Competencia	Destrezas implícitas	Aplicación de un saber	Conocimientos y destrezas aplicadas a la vida diaria				
Comprensión	Uso y aplicación de conceptos		Como medio para explicar relaciones y hacer predicciones				



El comparar la información concentrada en el cuadro, permite enunciar los siguiente:

- La habilidad, en las Pruebas Internacionales, es un concepto que destaca y predomina en los tres estudios con diferentes denominaciones: comportamientos, capacidades o destreza,<sup>170</sup> para referir la activación de operaciones cognitivas en la aplicación de un conocimiento. En las pruebas mexicanas, se le incluye en tres de ellas, (IDANIS, PN y EXANI- 1). El IDANIS, es la única prueba en donde se le define como recurso vinculado al conocimiento para el procesamiento de la información. En las pruebas nacionales se le refiere como contenidos del Plan y Programa de Estudio y en el EXANI-1, sólo se le menciona como aplicación de un conocimiento.
- El término destreza, sólo es incluido en los exámenes masivos internacionales, como manejo de procedimientos que permiten realizar adecuada y eficazmente una tarea.
- Competencia, es un concepto que incorporan las pruebas del LLECE y el PISA, para referirse a la aplicación de un saber y que en los exámenes masivos de la experiencia mexicana todavía no se contempla.
- La comprensión, concepto al que hacen referencia básicamente las pruebas del TIMSS, como aplicación de un conocimiento; que en las pruebas del LLECE y PISA, es lo que se considera como competencia, no se integra en los exámenes masivos de la experiencia mexicana.

La comparación analítica entre los exámenes masivos internacionales y nacionales en cuanto a su contenido y conceptos estructurantes, da cuenta de:

- ✓ En los exámenes masivos internacionales se puede hablar de la existencia de una teoría de contenido, dada la relación encontrada entre conceptos, dimensiones, campos temáticos y procesos incluidos .

---

<sup>170</sup> Aunque al término destreza se le atribuye un origen meramente físico o motor, algunos autores (Hontangas, 1994) destacan que "toda destreza humana implica la coordinación de percepción y acción y, consecuentemente, necesita la contribución de procesos preceptuales, decisiones cognitivas, control motor e incluso de procesos sociales." (citado por Agut Nieto y Grau Gumbau, 2001:13-24).



- ✓ En los exámenes masivos nacionales, no existe una teoría del contenido, el contenido de su evaluación, es retomado del plan y programa de estudio y los procesos inmersos en estos contenidos se quedan en el nivel de mero conocimiento, no llegan a la aplicación, transferencia, reflexión o evaluación.
- ✓ En las pruebas internacionales, el debate entre habilidad y destreza como objeto y contenido de evaluación, continua sin ser resuelto; se utilizan los términos de manera indistinta; inclusive se llega a equiparar con el de competencia; concepto polisémico que presenta diversas aseveraciones.<sup>171</sup> De acuerdo al análisis realizado, en las pruebas internacionales del LLECE y PISA aparece como integración de conocimientos y destrezas aplicadas. Sin embargo, la incursión y definición de dichos términos y conceptos en los documentos de evaluación en el plano internacional, plantea un nuevo marco teórico y enfoque en materia de evaluación y por ende de aprendizaje.
- ✓ En la experiencia mexicana el término de habilidad, es el único de los conceptos estructurantes, que se incorpora, pero de manera confusa en dos de los exámenes referidos: IDANIS y PN. En el IDANIS, aparece en estrecha vinculación con el de estrategia, es decir, no hay una diferenciación entre estrategia y habilidad. En las Pruebas Nacionales (PN), las habilidades se definen a partir de los objetivos y contenidos curriculares de los planes y programas de estudio; no en cuanto a las operaciones o procesos cognitivos del aprendizaje.
- ✓ Al comparar, los conceptos estructurantes de las experiencias de evaluación y medición de corte nacional, con los del plano internacional, se percibe una diferencia significativa. Mientras en el plano internacional, se está hablando ya de una evaluación en términos de competencia, que demanda la activación de procesos complejos de pensamiento, para la integración y aplicación de saberes y procedimientos dentro de un contexto real, de acuerdo a diferentes niveles de

---

<sup>171</sup>Como diría Perrenoud: "La noción de competencia sigue siendo un desafío en las ciencias cognitivas, así como en didáctica" (Perrenoud, 1999) *Construir competencias desde la escuela*. Chile: Dolmen Ediciones.



desempeño; en el ámbito nacional, el contenido de evaluación se reduce a la verificación del currículum preescrito; se habla de evaluación de habilidades, pero éstas son muy elementales y están supeditados a lo establecido en los planes y programas de estudio.

- ✓ En los exámenes masivos de corte nacional; no nada más se percibe ausencia de una clarificación en cuanto al contenido y concepto de evaluación; la inclusión de conceptos como el de habilidad, dista mucho del contenido de la enseñanza y del aprendizaje en la escuela básica. Aunque los planes y programas de estudio planteen el desarrollo de habilidades como uno de los principales objetivos, la realidad es que los contenidos que se privilegian en los procesos enseñanza-aprendizaje de la escuela primaria, se reducen a la enseñanza de hechos y conceptos.

**d. Algunas reflexiones a propósito del concepto de contenido en los exámenes masivos.**

Conviene aclarar el uso del término habilidad como contenido de evaluación de los exámenes masivos tanto en el plano internacional, como nacional. No queda claro si lo que se mide es una habilidad, entendida como destreza, como finalidad educativa, como contenido, como estrategia, o como capacidad.

En la concepción tradicional sobre enseñanza y aprendizaje, los contenidos son saberes constituidos, sistemáticamente planificados con fines de transmisión, incorporación y acumulación de los conocimientos más importantes para una determinada sociedad. La concepción cognitiva y constructivista minimiza y relativiza la importancia de los contenidos y se centra en el desarrollo de las capacidades cognitivas de los alumnos, rechazando y cuestionando la importancia que se le da a los contenidos en la postura llamada comúnmente como tradicional.



Existen ciertas posturas como la de la Reforma Española, que tienden a incluir todos estos aspectos en los objetivos curriculares.<sup>172</sup> El aporte de la Reforma Española y en concreto, de la propuesta de Coll, es integrar ambos conceptos como contenidos de aprendizaje y romper con el antagonismo de dichas posturas al conceder una importancia considerable al aprendizaje de ciertos contenidos, pero considerándolos no como un fin en sí mismo, sino como un medio para el desarrollo de las capacidades cognitivas<sup>173</sup> de los estudiantes; destacando así la necesidad de una actividad constructiva del alumno en el aprendizaje a través de la enseñanza no sólo de hechos o principios sino de procedimientos y actitudes. De ahí que dentro la propuesta de la Reforma Española, los contenidos se conciben en una acepción mucho más amplia y se definen como: "conjunto de saberes o formas culturales cuya asimilación y apropiación por los alumnos se considera esencial para su desarrollo y socialización" (Coll, Pozo, Sarabia, y Valls, 1992:13).

Con respecto a la habilidad, o a las llamadas habilidades del pensamiento éstas son concebidas dentro de la teoría constructivista como la práctica controlada y consciente de un proceso mental; entendiéndose como proceso a los operadores intelectuales que actúan sobre los conocimientos para transformarlos y generar nuevas estructuras. (De Sánchez, 1991). Algunos procesos considerados como básicos son: la observación, la comparación y la clasificación. Estos procesos a su vez pueden descomponerse en procedimientos para generar estructuras mentales de tipo procedimental; que es a lo que César Coll ha llamado contenidos procedimentales.<sup>174</sup>

---

<sup>172</sup> Con la finalidad de no privilegiar o descartar la enseñanza y aprendizaje de contenidos, o bien de habilidades y/o estrategias cognitivas, Coll ha caracterizado y clasificado a los contenidos curriculares en tres tipos: Hechos y Conceptos, Procedimientos, Actitudes y Valores (véase: Coll, C. y Martín E., 1994).

<sup>173</sup> En su más reciente libro Coll, considera el desarrollo de las capacidades como finalidades educativas de los procesos enseñanza-aprendizaje y afirma que: "Las capacidades son la meta de la enseñanza, la guía de la práctica en el aula, el referente del éxito o fracaso de las instituciones escolares" (véase. Coll, y Martín, 2003).

<sup>174</sup> El Diseño Curricular Base de la Reforma Española (DCB), considera a los procedimientos como "conjunto de acciones ordenadas, orientadas a la consecución de una meta". Lo que se pretende es que el alumno aprenda a llevar a cabo determinadas acciones de manera ordenada para alcanzar una meta. (véase Coll, Pozo, Sarabia, y Valls, 1992).



En cuanto a las estrategias, Nickerson las concibe como mecanismos a través de los cuales se pueden relacionar los procesos y las estructuras y dependen del tipo de situación y de tarea; de ahí que se refieran al saber qué hacer y cuándo hacerlo, es decir a la clase de operaciones mentales que se es capaz de aplicar ante diferentes situaciones de aprendizaje (Nickerson, Perkins, y Smith, 1994).

La necesidad de clarificar qué se entiende por el contenido de evaluación, radica en que, un elemento básico en todo proceso de medición es la definición del atributo u objeto a medir, (véase Thorndike y Hagen 1978) de ahí la importancia que si se dice que es una prueba centrada en habilidades, se especifique y defina qué se entiende por dicho término; es decir, si es un "contenido procedimental", una "estrategia" o una "capacidad desarrollada"; una clarificación del concepto, permitiría a los docentes trabajar mucho más y mejor, dichos elementos en el proceso enseñanza-aprendizaje.

La inclusión de términos como habilidades y competencias en el contenido de evaluación de los exámenes masivos tanto en el plano internacional como en algunas de las experiencias nacionales, induce a cuestionamientos como los siguientes:

- ⇒ ¿Qué tanto se promueve la enseñanza y el aprendizaje de procedimientos, estrategias habilidades o competencias, dentro de la enseñanza básica?
- ⇒ ¿Realmente el proceso enseñanza-aprendizaje, consiste en una práctica constante de acciones ordenadas a sistematizar y desarrollar procesos mentales de orden superior para la aplicación de un conocimiento en miras a resolver un problema o alcanzar una meta?
- ⇒ ¿A los alumnos en la escuela primaria o secundaria se les enseña cómo utilizar el conocimiento trabajado?
- ⇒ ¿Los exámenes elaborados por los docentes, son exámenes que demandan la aplicación del conocimiento e incluyen la evaluación de algunas habilidades mentales, o sólo se enfocan a la reproducción de información ?

Tales cuestionamientos inducen a pensar que la mayoría de los estudiantes, en especial los alumnos de la escuela básica en México, están muy poco familiarizados con pruebas



centradas en habilidades o “competencias;” la mayoría de los exámenes que ellos han contestado son pruebas que demandan la memorización de conceptos, hechos y principios y no el uso de procesos cognitivos de mayor nivel.

La comparación del tipo de ítems que contienen los exámenes masivos tanto en el plano internacional como nacional, permitirá corroborar que las pruebas mexicanas no sólo difieren de las internacionales, en cuanto al contenido de su evaluación, sino también al tipo de reactivo de sus exámenes.

#### **4.2 Tipo de ítems en los exámenes masivos**

El tipo de ítem de los exámenes masivos, lo analizamos y comparamos en dos aspectos: composición y estructura.

##### **a. Composición de los ítems de matemáticas**

Para referir la composición de los ítems retomamos contenidos, tareas y procesos implícitos en un área de evaluación común como es las matemáticas.

El siguiente cuadro concentra los contenidos, tareas y procesos implícitos en los ítems de matemáticas de los exámenes masivos de corte internacional y nacional.



CUADRO 4.6

**COMPOSICIÓN DE LOS ÍTEMS DE MATEMÁTICAS EN LOS EXÁMENES MASIVOS**

	PRUEBAS INTERNACIONALES			PRUEBAS NACIONALES			
	TIMSS	LLECE	PISA	IDANIS	SNEE	PN	EXANI-1
<b>Contenidos</b>	Números Álgebra Medición Geometría Datos	Numeración Operaciones con números naturales Fracciones comunes Geometría Medición	Cantidades Espacios y formas Cambios Relaciones Incertidumbre	Aritmética Geometría	Temas del plan y programa de estudios	De acuerdo con lo establecido en los planes y programas de estudio	Sucesiones numéricas Series espaciales Imaginación espacial Problemas de razonamiento
<b>Procesos</b>	Aplicación y transferencias de conocimientos y procedimientos	Uso de procedimientos: como: Reconocer y establecer relaciones entre cantidades	Cálculos simples Reunión de ideas y procedimientos Razonamiento matemático Generalización	Cálculos Observación y comparación	Procesos básicos de pensamiento como la identificación y el reconocimiento	Operar Medir Comunicar Imaginar Generalizar Inferir	Ordenamiento Percepción Análisis
<b>Tareas</b>	Resolver problemas reales	Modelar situaciones: con fines de matematización	Contextos auténticos Problemas donde se empleen las matemáticas	Realización de operaciones y ejercicios	Realización de ejercicios aritméticos	Ajenas a Situaciones reales	Encontrar relaciones numéricas o gráficas

A partir de la información presentada en el cuadro, podemos citar algunas semejanzas y diferencias encontradas:

- ⊗ El contenido de los ítems de las pruebas internacionales, está determinado por cinco campos temáticos. En las pruebas nacionales se refieren a temas y "habilidades" del Plan y programa de estudios correspondientes.
- ⊗ Los procesos implícitos en los ítems de las pruebas internacionales, están vinculados con el uso de procedimientos para la aplicación del conocimiento. En las pruebas nacionales, se trata de procesos básicos del pensamiento, donde prevalece el cálculo y la operacionalización, pero no el razonamiento.



- ⊗ Las tareas de los reactivos en las pruebas internacionales, se refieren a problemas reales que el alumno tendrá que resolver mediante una modelización matemática. En las pruebas nacionales están centrados en la práctica y ejercitación que poco o nada tienen que ver con situaciones de la vida diaria.

Las diferencias encontradas, permiten afirmar que la composición de los ítems de las pruebas internacionales, dista mucho de los ítems que se utilizan a nivel nacional; mientras que en el plano internacional, comprenden tareas auténticas que demandan la aplicación del conocimiento mediante la activación de procesos de orden superior; en México, los exámenes masivos conservan la realización de ejercicios descontextualizados ajenos a situaciones reales, con fines de práctica y reproducción del conocimiento.

La constitución de los ítems en los exámenes masivos, presenta también estructuras y formatos diferentes tanto en el plano internacional como nacional.

## b. Estructura y formato de los ítems

La revisión y análisis de los reactivos muestra, publicados tanto en las pruebas internacionales como en algunas de las nacionales, lleva a decir que:

- ⇒ Los ítems de las pruebas internacionales, conservan un formato de opción múltiple; en las pruebas del TIMSS y en el PISA. Se pueden encontrar además, ítems de respuesta abierta y de respuesta construida;<sup>175</sup> lo que resulta importante de considerar, sobre todo porque se trata de exámenes masivos que presentan criterios para calificar ítems de respuesta abierta en pruebas a gran escala, situación que no parecía muy factible en este tipo de exámenes.

---

<sup>175</sup> Este es un tipo de ítems, en el que al alumno se le presenta un problema o pregunta específica, para que produzca con sus propias palabras la respuesta. Recordemos que de acuerdo a las consideraciones técnicas dadas por Thorndike y Hagen, para elegir los tipos de ítem a utilizar en una pruebas, vemos que los ítem de respuesta abierta y de respuesta construida son los más adecuados para medir la capacidad de utilización de los conocimientos en la solución de un problema o la capacidad de interpretar o analizar o evaluar (Véase Thorndike y Hagen, 1978).



En las pruebas de la experiencia nacional, prevalecen los ítems con un formato de opción múltiple,<sup>176</sup> y como se señaló oportunamente en una de éstas (IDANIS), se encontraron problemas de redacción.

De acuerdo con Thorndike y Hagen (1978), el formato de opción múltiple, opera dentro de situaciones completamente estructuradas, en donde no se puede medir la originalidad de la respuesta o la capacidad de reunir varias ideas o conceptos de una o más áreas de contenido hasta formar un todo organizado. Lo único que se requiere, es saber elegir la respuesta dentro de un número limitado de opciones. De ahí que el ítem objetivo no es muy conveniente para tratar cuestiones controvertidas o complejas.

La importancia de considerar el formato en la estructura de los ítems de los exámenes masivos, radica en que, según Thorndike y Hagen (1978), existen ciertas consideraciones teóricas para elegir el tipo de ítem en una prueba, entre las que se encuentran, la capacidad del ítem para provocar la conducta que se trata de medir (Thorndike y Hagen, 1978: 60-78); de ahí que la elección del formato, dependa de la finalidad de la pruebas y de lo que se quiera medir. Además, considerando que en materia de medición, la calidad técnica de un instrumento reside en dos propiedades básicas referentes a la confiabilidad<sup>177</sup> y validez,<sup>178</sup> que consiste en la correcta representación del constructo,<sup>179</sup> cabría preguntarse si de acuerdo con los constructos establecidos en las pruebas del IDANIS, en las llamadas ahora Pruebas Nacionales del INEE y en el EXANI-1 los formatos de opción múltiple, son los más adecuados para medir lo que se afirma que se quiere medir.

---

<sup>176</sup> Este formato se establece como de preferencia a partir de la aparición del Stanford Achievement Test en los Estados Unidos (1923) y consiste en un enunciado parcial o en una pregunta seguida de opciones. (Cfr. Haladyna, Haladyna y Merino, s.f.).

<sup>177</sup> La confiabilidad significa que aplicada en las mismas circunstancias, a los mismos sujetos produzca los mismos resultados.

<sup>178</sup> Conviene recordar que la validez de una prueba en el plano estadístico, también se refiere al índice de dificultad del reactivo, el cual se obtiene conforme a la relación inversa de los que no contestaron.

<sup>179</sup> Los constructos son entonces conceptualizaciones, construcciones mentales que usan los científicos sociales para lograr describir o explicar aspectos que se desean estudiar en los individuos. El constructo está asociado a ciertos referentes teóricos que se emplean para definirlo y describirlo. Al hablar, por ejemplo, de medir constructos como inteligencia, actitud hacia la matemática, aptitud académica, grado de socialización o nivel de motivación en una persona, los desafíos para la medición se presentan con su definición, para luego abarcar todo su proceso de "operacionalización" (Montero, 2003).



### 4.3 Modelo psicométrico en los exámenes masivos

Todo test<sup>180</sup> lleva consigo una teoría, implícita en el proceso mismo de medición, que presenta ciertos retos desde el momento en que se intenta definir las características o fenómenos a medir.

Como ya se ha mencionado, en todo proceso de medición, los objetos o atributos deben estar claramente definidos.<sup>181</sup> Una vez definido el objeto a medir, el segundo paso consiste en determinar las operaciones mediante las cuales el atributo puede manifestarse y percibirse; es decir, las propiedades observadas en los objetos, se expresan en constructos, mediante los cuales se operacionaliza el objeto, porque no se miden objetos tangibles, sino cualidades de éste, a través de los llamados indicadores<sup>182</sup> y finalmente, se utilizan teorías, métodos y técnicas definidas y se establecen medidas estadísticas para los indicadores. Estas técnicas, métodos y modelos, se basan en enfoques cuantitativos y utilizan conceptos, procedimientos y medidas derivadas de la estadística y la matemática.

A continuación se presenta, cómo se asumen estos elementos tanto en las pruebas internacionales, como nacionales.

⇒ Con respecto al atributo, u objeto de medición, en los exámenes masivos de corte internacional, existe una clara definición del objeto que se va a medir. El TIMSS, habla de dominios de contenidos y dominios cognitivos como objetos de evaluación, el LLECE define su contenido a partir de niveles de desempeño y el PISA lo presenta como conocimientos relevantes (grandes ideas) y destrezas para la vida (competencias). En los exámenes masivos de la experiencia mexicana, el IDANIS señala su objeto como habilidad verbal, habilidad matemática y habilidad para el razonamiento abstracto, las pruebas del SNEE no lo contemplan, las recientes Pruebas Nacionales, retoman su atributo de los Planes y programas de estudio y en

---

<sup>180</sup> Entendido este término como forma de estandarizar un instrumento de medición.

<sup>181</sup> En el plano psicológico de la medición, la definición del atributo se refiere a señalar o definir las cualidades del objeto a medir.

<sup>182</sup> Operacionalizar un constructo significa representarlo por medio de indicadores. Los indicadores se refieren usualmente a hechos, conductas o acciones observables (Montero, 2003).



el EXANI-1 se presenta como habilidades básicas y conocimientos disciplinarios, aunque no se define.

- ⇒ En relación a las propiedades u operaciones a medir, el TIMSS, habla de aplicación de conocimientos y destrezas, el LLECE de competencia matemática y competencia comunicativa y el PISA presenta claramente lo que entiende por aptitud<sup>183</sup> lectora, por aptitud matemática y por aptitud científica<sup>184</sup> y define los conceptos, procesos y situaciones que comprende. En los exámenes de la experiencia mexicana, sólo en el IDANIS se encuentra una definición de las propiedades a medir, en las recientes pruebas nacionales del INEE, hay un intento por presentar y operacionalizar las propiedades del objeto a través de desempeños, pero estos son retomados de lo establecido en los planes y programas de estudio.
- ⇒ En cuanto a las operaciones o medidas estadísticas. hemos encontrado que las pruebas internacionales ofrecen aportaciones importantes, referentes a la elaboración de guías codificadas (TIMSS), criterios y muestras de respuestas y palabras clave (PISA) para calificar preguntas abiertas; así como el establecimiento de puntajes estandarizados y escalas de calificación con puntajes promedio por niveles de desempeño (TIMSS, LLECE y PISA). En las pruebas mexicanas, no se conocen las medidas estadísticas para su calificación. Sólo en el EXANI-1 se habla del establecimiento de preguntas conforme a su grado de dificultad, de acuerdo con la Teoría Clásica del Test.

El modelo psicométrico de medida que prevalece en los exámenes masivos de corte internacional es el de la Teoría de Respuesta al ítem, en donde la validación estadística del instrumento consiste en el análisis detallado de las propiedades del ítem y no en una perspectiva global como en la Teoría Clásica. La Teoría de Respuesta al ítem en que incursionan los instrumentos de evaluación internacionales, consiste en la relación

---

<sup>183</sup> Conviene destacar el hecho de que el PISA, emplee el término "aptitud" para definir al objeto de medida, en lugar de habilidad, le da una connotación diferente. Según algunos autores las aptitudes se distinguen de las habilidades, por su carácter permanente e innato y porque estas se infieren a partir de la conducta (véase Agut Nieto y Grau, 2001).

<sup>184</sup> Lo que se entiende por cada uno de estos constructos ha sido definido en la descripción y análisis de cada uno de los instrumentos.



matemática que se establece entre la variable a medir en cada uno de los times (es decir, los niveles de constructo) y las probabilidades de que se de respuesta a éstos; lo cual es representada gráficamente por una curva. Mientras que a nivel nacional prevalece todavía el modelo de Teoría Clásica del Test, en donde el grado de dificultad de sus ítems, está determinado por el % de alumnos que contesten adecuadamente cada ítem y no por las propiedades de éste.

El siguiente cuadro facilita la comparación de los elementos y teorías de la medida con que cuentan los exámenes masivos tanto de la experiencia internacional, como nacional.

CUADRO 4.7

		PRUEBAS INTERNACIONALES			PRUEBAS NACIONALES		
		TIMSS	LLECE	PISA	IDANIS	SNEE	PN
<b>Objeto o atributo</b>	Dominios de contenidos y dominios cognitivos	Niveles de desempeño	Conocimientos relevantes y destrezas para la vida.	Habilidad verbal, matemática y para el razonamiento abstracto	No lo define	Del Plan y programa de estudios	Habilidades intelectuales básicas y conocimientos disciplinares
<b>Operaciones o propiedades</b>	Aplicación de conocimientos y destrezas	Competencia matemática y comunicativa	Definición de aptitud para la lectura, las matemáticas y las ciencias a través de conceptos, procesos y situaciones	Procesos que componen cada habilidad	No se especifican.	Clasificación de habilidades y desempeños	No se definen
<b>Medidas estadísticas</b>	Guías codificadas para preguntas abiertas Puntajes promedio	Por criterios conforme a ciertos niveles de desempeño	Puntajes de calificación por niveles de desempeño o Criterios y respuestas	No se especifica	No se especifica	No se especifica	Por el nivel de dificultad



			tipo para preguntas abiertas				
Teoría de test	Teoría de Respuesta al Ítem	Teoría de Respuesta al Ítem y Modelo de Rasch	Teoría de Respuesta al Ítem	Modelo de Rasch	No se especifica	No se especifica	Teoría clásica del test

La comparación de los elementos y modelos de medición empleados en las pruebas internacionales y en las nacionales, dan cuenta de que mientras los exámenes masivos a nivel internacional especifican claramente los elementos de medición y modelo teórico utilizados y presentan importantes aportaciones al respecto; en México, no solamente no todos los exámenes masivos, cuentan con una especificación de la totalidad de elementos correspondientes, sino que, aún conservan la Teoría clásica del test para validar sus ítems; método tradicional, que por mucho tiempo ha sido el más utilizado para el análisis de calidad técnica de instrumentos de medición psicométricos. Sólo el IDANIS, incluye el modelo de Rasch,<sup>185</sup> que es un modelo propio de las Teoría de Respuesta al Ítem (TRI).

La mayoría de los exámenes masivos elaboradas y aplicadas a nivel nacional, incluyendo el EXANI-1, conservan el modelo de Teoría Clásica del Test, el cual es un modelo unitario que no considera diferencias individuales; es decir, los constructos son tratados de forma generalizada y en donde el nivel de dificultad,<sup>186</sup> depende del nivel de ejecución de los sujetos. En cambio la Teoría de Respuesta al Ítem (TRI),<sup>187</sup> de acuerdo con el principio de

<sup>185</sup> El modelo de Rasch es uniparamétrico, es decir, utiliza una sola escala para medir tanto el desempeño de la persona en cuanto conocimientos y habilidades, como la calidad de un reactivo, en cuanto al número de personas que lo contestan bien. La ventaja de esto es que ambos, es que tanto los reactivos como las personas se pueden comparar en una escala única.

<sup>186</sup> Entiéndase este como la proporción de personas que responden correctamente un reactivo.

<sup>187</sup> Conviene recordar que como se ha venido señalando, el modelo psicométrico de TRI es una conceptualización, en donde partiendo de ciertos conceptos básicos de medición y usando las herramientas de la estadística y la matemática; los parámetros estimados por el modelo permiten evaluar la calidad técnica de cada uno de los ítems o reactivos por separado y del instrumento como un todo y a la vez estimar el nivel que cada examinado presenta en el constructo de interés. En un modelo de TRI se asume que hay una variable latente o constructo que no es observable directamente y que se desea estimar para cada examinado a partir de las respuestas suministradas por éste en el instrumento de medición. Además se asume que para cada ítem o pregunta el comportamiento de las respuestas dadas por los examinados puede ser modelado mediante una función matemática que se denomina Curva Característica del Ítem o CCI. Aunque los desarrollos básicos de los modelos de TRI se generaron con el trabajo pionero de psicómetras como A. Birnbaum, G. Rasch, F.M. Lord y M.R. Novick que datan de los años 50, no es sino hasta los años 70 que estos enfoques comienzan a divulgarse y a aplicarse en los Estados Unidos y otros países industrializados. En América Latina,



invarianza de parámetros; el nivel de dificultad, depende de la dificultad del ítem y no de las capacidades del sujeto.

#### 4.4 Par concluir: Unas comparaciones finales:

El análisis y comparación de los sustentos teóricos-metodológicos con respecto al cómo se mide en el plano internacional y lo que prevalece y existe a nivel nacional en materia de instrumentos para evaluar y medir el aprendizaje en la educación básica, dan cuenta de que:

- Los exámenes que circulan en el plano internacional incorporan y presentan una Concepto de contenido en su marco de evaluación, misma que se puede identificar en la conceptualización que hacen de su objeto de evaluación y la vinculación que existe con la manera en como organizan su contenido. La situación de los exámenes en México, dista mucho de contar con un concepto de contenido.
- Las pruebas internacionales están estructuradas en torno a conceptos tales como habilidades, competencias, conocimientos y aptitudes para la vida. No son pruebas elaboradas conforme a un plan de estudios, como en las pruebas de la experiencia mexicana, sino a las capacidades adquiridos por los estudiantes para aplicar el conocimiento y desempeñarse adecuadamente en situaciones de la vida cotidiana. El contenido de evaluación en el marco internacional está ubicado con un enfoque más funcional que nominal como es el caso de las pruebas mexicanas. He aquí una de las diferencias más importantes.
- En los instrumentos de evaluación internacionales, los procesos y contenidos implícitos en los ítems, se refieren a procesos complejos de pensamiento para la aplicación de procedimientos. Una aportación importante en el tipo de ítems de los exámenes masivos de corte internacional es que incluyen preguntas abiertas y al mismo tiempo proponen métodos para calificarlas; lo que revela una nueva forma de

---

debido a la escasa formación en materia de métodos cuantitativos de evaluación, es muy reciente la incursión del modelo de Rasch y de la Teoría de Respuesta al ítem como modelo de medición. Argentina y Colombia se pueden ubicar como los países pioneros en el empleo de este modelo (véase Montero, 2003).



elaboración de ítems en pruebas a gran escala. El tipo de ítem de los exámenes nacionales se caracteriza por ser ítems de opción múltiple en su totalidad, centrados en la realización de ejercicios que demandan el uso de procesos muy básicos como la identificación y el recuerdo, con fines de mecanización o reproducción de conceptos o hechos.

- Los exámenes internacionales, presentan cambios importantes en los atributos y operaciones a medir; los cuales están perfectamente definidos e incluyen nuevos objetos y propiedades de evaluación, tales como, pensamiento lógico y sistemático, competencia comunicativa, competencia matemática y reflexión sobre los textos. Los exámenes en México, definen su atributo a partir de los planes y programas de estudio solamente.
- Los sustentos teóricos de las pruebas internacionales, nos presentan a la evaluación desde otra mirada, pues con éstas, el campo de la evaluación y medición, no sólo se abrió sino que también surgieron nuevas teorías. Las pruebas aplicadas por agencias y Organismos Internacionales suprimen la Teoría Clásica del Test, e incorporan las tendencias más recientes en materia de medición, como la Teoría de Respuesta al Ítem.

Hasta aquí queda claro que las agencias internacionales están promoviendo y aplicando nuevas teorías y formas para medir y evaluar el aprendizaje escolar. México se incorporó de manera abrupta a la evaluación, pero sin una inclusión de las tendencias y teorías actuales. Es hasta las recientes Pruebas Nacionales aplicadas por el INEE que comienza a darse y publicarse una clarificación del objeto y de los constructos a medir; sin embargo éste se define todavía en torno a los planes de estudio. En los exámenes masivos aplicados en México, prevalece la idea de aprendizaje como acumulación de contenidos, que reduce a la evaluación a una mera reproducción del conocimiento de los mismos; mientras que en las pruebas internacionales subyace una idea de aprendizaje más funcional y sobre todo integrador, pues éste se concibe como la vinculación entre el conocimiento y su aplicación a una situación real y cotidiana.



Detrás de toda práctica educativa y dentro de ésta, la evaluación del aprendizaje no es la excepción, siempre hay teorías o supuestos que se asumen, la intención de estos capítulos fue ponerlos de manifiesto y compararlos, para situar el estado que guardan los exámenes masivos a nivel nacional dentro del plano internacional.

### **Tareas pendientes**

La comparación realizada en cuanto a contenido y estructura de los exámenes masivos a nivel internacional y la situación que guardan los exámenes realizados en México, permite identificar algunas tareas pendientes en materia de elaboración de pruebas.

1. Pruebas como el IDANIS, las Nacionales (PN) y el EXANI-1, que de acuerdo con los documentos que las sustentan, se afirma evalúan habilidades, es importante que empiecen por definir de manera independiente al plan de estudios, qué entienden por habilidades y en qué consisten las habilidades que se pretende medir; es decir que presenten una definición propia de los atributos y constructos y una organización del contenido acorde las definiciones elaboradas, para comenzar a conformar una Concepto de contenido que las sustente.
2. Los exámenes masivos a nivel nacional, deberán presentar conceptos estructurantes propios, diferentes a los de los planes y programas de estudio que ofrezcan un nuevo concepto y teoría de evaluación y por ende de aprendizaje.
3. Los exámenes nacionales tendrán que diversificar el formato de sus reactivos e incorporar procesos cognitivos que no sólo demanden reproducción de información o práctica de algún ejercicio; sino procesos de pensamiento más complejos para una aplicación y reflexión de lo aprendido
4. Las pruebas aplicadas en México requieren de un sustento teórico-metodológico, acorde con las tendencias actuales a nivel internacional, que incorpore modelos y elementos recientes en materia de medición psicométrica.



5. Las pruebas nacionales, podrán que incluir nuevos enfoques que retomen los avances existentes con respecto a la teoría cognitiva para la elaboración de test, con el fin de identificar los procesos y actitudes necesarios para el aprendizaje, tales como las ideas preconcebidas, las estrategias y los recursos cognitivos de los estudiantes(véase Wittrock y Baker 1998).

Este análisis comparativo, deja muchas tareas pendientes en materia de elaboración de exámenes masivos, que es importante retomar y comenzar a trabajar, si es que se aspira a contar con un sistema de evaluación y por ende un sistema educativo, acorde las tendencias teórico-metodológicas que a nivel internacional se están operando en materia de evaluación del aprendizaje.

## CONCLUSIONES

---

El análisis realizado del contenido y estructura de exámenes masivos aplicados a nivel internacional durante la última década del siglo XX y la comparación establecida con las experiencias nacionales, permite a partir de las interrogantes planteadas al inicio de esta investigación, dar cuenta de algunas aproximaciones conceptuales y metodológicas en materia de evaluación del aprendizaje.

### **i. Una respuesta a las interrogantes planteadas:**

#### **1. *¿Bajo qué orientaciones se organiza el contenido de evaluación en el plano internacional?***

Contenidos curriculares básicos, estándares y preparación de los jóvenes para la vida adulta, son los conceptos base para la organización del contenido de evaluación en las pruebas internacionales; además de que se abordan los factores tanto sociales, como escolares y familiares que influyen en los resultados obtenidos. Lo que para México sigue siendo un reto.

#### **2. *¿Qué concepto de contenido de evaluación está implícito en las pruebas internacionales?***

Los exámenes internacionales evalúan no sólo conocimientos, sino también capacidades y/o aptitudes para la aplicación de esos conocimientos a situaciones cotidianas. Plantean un concepto de contenido de corte más funcional y procedimental. Presentan pruebas que descartan el manejo enciclopédico del contenido curricular, para dar paso a la evaluación de contenidos básicos y procesos superiores de pensamiento, en contextos reales; que en México todavía no han sido generalizados ni trabajados a profundidad.



**3. ¿Cuáles son los fundamentos teórico-metodológicos que sustentan a los exámenes internacionales?**

Las pruebas internacionales presentan una conceptualización de su contenido, que dista de lo establecido en planes y programas de estudio y actúa como el eje organizador de las mismas.

Los exámenes están estructurados con ítems contextualizados que requieren la aplicación de un conocimiento mediante la activación de procesos mentales de orden superior, como el razonamiento, la reflexión, el pensamiento lógico, la interpretación y la transferencia.

En materia de medición, recurren a modelos psicométricos alternativos, como la TRI en donde los ítems son validados y seleccionados de acuerdo a lo que miden y a la probabilidad de que sean contestados correctamente.

Se trata de pruebas criterioles, estructuradas con base en niveles de desempeño que procuran una evaluación realizada a través de situaciones reales.

**4. ¿Cuáles son los conceptos que subyacen en los exámenes internacionales?**

Los exámenes masivos a nivel internacional, están estructurados con base en un concepto de competencia, considerada como habilidad y/o aptitud para la aplicación de un conocimiento; y aunque carecen de una definición precisa, presentan un idea más funcional de aprendizaje y evaluación.

**5. ¿Qué se puede apreciar de la comparación entre las experiencias internacionales y las nacionales, en materia de tendencias y contenidos de evaluación?**

La comparación de las experiencias internacionales en relación a las nacionales en materia de evaluación del aprendizaje permite identificar diferencias significativas en los siguientes puntos:



- 1) Los estudios de evaluación internacional organizan su contenido conforme a la definición que de éste hacen.
- 2) Las pruebas a nivel nacional, no ofrecen un marco de evaluación propio que explique su contenido y estructura. Éste se da con base en los Planes y Programa de estudios y la Taxonomía de Bloom (1971).
- 3) El concepto de contenido de evaluación en los estudios de evaluación internacional consiste en comprobar y comparar el rendimiento escolar en cuanto a comprensión y aplicación de los conocimientos.
- 4) Las pruebas realizadas en México continúan con contenidos centrados en la reproducción de los temas del plan de estudios.
- 5) El tipo de ítem de los exámenes internacionales, evalúa capacidades cognitivas para la aplicación del conocimiento a situaciones de la vida cotidiana, mediante el uso de procedimientos.
- 6) Los ítems de los exámenes nacionales, requieren de operaciones básicas de pensamiento centradas en la identificación y el recuerdo, para la realización de ejercicios rutinarios.
- 7) Las pruebas internacionales incorporan tendencias recientes en materia de evaluación del aprendizaje como es la evaluación criterial, por desempeños y contextualizada.
- 8) Las pruebas de la experiencia mexicana, conservan los modelos de la TC .

**6. ¿Cuál es el impacto que el contenido de evaluación de las pruebas internacionales tiene en materia de currículo y evaluación?**

El concepto de contenidos a los que se recurre en el plano internacional, para la evaluación del aprendizaje, obliga a reparar en el tipo de contenidos trabajados en la escuela básica y en los factores asociados en materia de desarrollo curricular; porque los estudios de evaluación internacionales sobre el aprovechamiento escolar, no sólo contemplan lo relacionado al rendimiento en



términos de resultados; sino que además ofrecen una evaluación de los factores asociados al rendimiento y desempeño escolar, que incluye lo referente a procesos y situaciones que se dan en la práctica educativa. Dado las características conceptuales de los contenidos en los estudios de evaluación internacional, se puede decir que éstos retoman los planteamientos curriculares de John Dewey referentes a contenidos inmersos en la experiencia o la práctica y de Franklin Bobbit, quien desde 1918, planteaba la necesidad de articular el currículum al mundo laboral.

### 7. *¿Qué necesidades surgen en materia de docencia ?*

En materia de docencia los conceptos implícitos en los exámenes internacionales plantean la necesidad de encontrar formas de enseñanza más vivenciales, centradas en la experiencia y aplicación de un conocimiento, mediante el ejercicio y desarrollo de operaciones mentales de orden superior.

#### ii. **Un concepto de examen diferente:**

Las tendencias que en materia de elaboración de exámenes masivos, se han encontrado tanto en el plano internacional como nacional, presentan un importante cambio en la concepción de examen, que los especialistas en elaboración de exámenes a nivel nacional podrían considerar:

- ◆ Los exámenes a nivel internacional presentan conceptos de contenido, aprendizaje y evaluación diferentes; el contenido que se evalúa está centrado más en habilidades para la aplicación del conocimiento, que en la reproducción de temas curriculares; lo que requiere de un aprendizaje de corte más utilitario y una idea de evaluación más funcional, centrada no en lo que se puede acumular sino en lo que se sabe hacer. Lo que invita a los especialistas en evaluación y currículo, a repensar el concepto de contenido implícito tanto en las pruebas nacionales como en los planes y programas de estudio de la educación básica en el país y a elaborar pruebas que demanden la





aplicación de conocimientos básicos y no la repetición de un sin fin de temas curriculares.

- ◆ Lo importante ahora en el plano internacional, no es medir a través de resolución de ejercicios descontextualizados, lo que se sabe o lo que se puede retener, sino valorar las capacidades de los estudiantes para responder a situaciones cotidianas. Lo que hace necesario que en México se trabaje cada vez más por la elaboración de exámenes con reactivos que reflejen situaciones reales.
- ◆ Los exámenes internacionales, incluyen en el diseño de sus ítems, tareas referentes a problemas la vida diaria, que requieren la activación de procesos cognitivos de nivel superior como el razonamiento, la transferencia y la reflexión, para la aplicación e integración del conocimiento adquirido. De ahí la necesidad de incluir en los reactivos de las pruebas mexicanas tareas más de corte cognitivo que de mecanización y memoria.
- ◆ Las pruebas internacionales incorporan modelos psicométricos recientes como el del TRI, en donde los ítem se seleccionan no por el número de estudiantes que lo contestaron correctamente según los planteamientos de la Teoría Clásica del Test, sino con base en una jerarquización lógica conceptual y cognitiva de sus propiedades que ayudan a identificar niveles de desempeño.

Los exámenes a gran escala adoptan la evaluación por niveles de desempeño y transforman los sistemas de calificación de una prueba, al no limitarse al lector óptico y elaborar guías y códigos para calificar preguntas abiertas. Ante lo cual, es importante que las pruebas elaboradas en México por los institutos de evaluación, contemplen las alternativas teórico, metodológicas que existen en materia de elaboración y calificación de exámenes y no se limiten a los procedimientos clásicos.



**iii. Lo que se puede apreciar:**

Las comparaciones establecidas, permiten apreciar que los resultados obtenidos por México en los estudios evaluativos de corte internacional en los que ha participado a partir de 1995, obedecen a diferencias significativas en los siguientes puntos:

- I. **Conceptualización sobre el contenido y tipo de ítems en exámenes internacionales:** La comparación entre lo que se evalúa y cómo se evalúa en el nivel internacional y nacional, revela que mientras el contenido de las pruebas nacionales gira en torno a la evaluación de *competencias* entendidas como la aplicación del conocimiento adquirido, mediante la activación de operaciones mentales de orden superior; la mayoría de las pruebas nacionales, se enfoca a la mera reproducción de información. Operaciones como el razonamiento, la interpretación y la transferencia, implícitas en las tareas de evaluación de los estudios internacionales, no son ejercitadas en las pruebas mexicanas, en donde la memoria y el recuerdo para identificar la respuesta correcta, es lo fundamental. Los ítems de las pruebas internacionales en los que los estudiantes tienen que construir y justificar su respuesta; son inexistentes en nuestros exámenes masivos, compuestos en su totalidad por ítems de opción múltiple. La construcción y selección de reactivos con base en modelos psicométricos que consideran las diferencias individuales en el desempeño de los estudiantes en términos de probabilidad y las diferentes propiedades de los ítem, no es considerado en las experiencias nacionales de evaluación, en donde todavía se recurre a un modelo psicométrico clásico, sin una definición clara y propia de sus elementos; lo que impide la detección de las fortalezas y limitaciones de capacidades cognitivas de los estudiantes.
- II. **Formas de enseñanza y evaluación:** Las pruebas internacionales, evalúan un aprendizaje de conceptos, hechos, principios y procedimientos, pero aplicados a escenarios reales; lo que requiere de formas de enseñanza, centradas en el desarrollo de capacidades mentales que enseñen al alumno no sólo a adquirir un conocimiento



sino a reflexionar sobre el mismo y a transferir lo que va aprendiendo a situaciones prácticas. Los estudiantes mexicanos están familiarizados con formas de enseñanza centradas en la acumulación y transmisión de información y en exámenes que miden la capacidad para retener y reproducir esa información.

III. **Enfoques curriculares:** El contenido y estructura de las pruebas internacionales, requieren de contenidos curriculares de tipo conceptual, pero sobre todo, procedimental y cognitivo. Los planes y programas de estudio para la educación básica, en el Sistema Educativo Mexicano, señalan algunas habilidades a desarrollar, pero a manera de grandes objetivos, no como contenidos y procedimientos que hay que enseñar; pese a ésto, se sigue privilegiando un currículo enciclopédico y sin sentido.

IV. **Prácticas socioculturales:** Muchas de las tareas de evaluación de las pruebas internacionales, implican un tipo de prácticas socioculturales un tanto alejadas de la cultura nacional. Los jóvenes en México no tienen incorporado como parte de sus prácticas culturales, el leer sin un fin obligatorio y el reflexionar sobre el contenido y forma de lo que leen, para tomar postura y hacer uso de la lectura para diferentes fines no sólo los escolares, aplicándola en situaciones cotidianas, como es la lectura de una garantía o de un esquema. Los estudiantes mexicanos, no están acostumbrados a pensar matemáticamente o como declaran las pruebas internacionales, a "matematizar el mundo" mediante la construcción y explicación de modelos. Las matemáticas son asumidas como algo complejo y generador de angustia, que se resuelve cuando se domina el mecanismo. El conocimiento científico es considerado como algo abstracto y sin sentido reservado para "sabios" o gente especializada, no como una explicación de lo que ocurre en la persona y en el medio en el que se vive.

Las apreciaciones generadas, permiten comprender el porqué de los resultados obtenidos y enunciar algunas aportaciones:



#### iv. Algunas aportaciones y sugerencias:

El estudio de las pruebas elaboradas y aplicadas por Organismos Internacionales y la situación que guarda México respecto a éstas, presenta aspectos importantes a considerar en materia de currículum y evaluación:

- El contenido y estructura de los exámenes internacionales, lleva a pensar en la necesidad de enseñar y evaluar no sólo contenidos, sino procesos y procedimientos, que favorezcan el uso del conocimiento adquirido en la resolución de los problemas de la vida cotidiana. Es decir, una tarea importante en términos de contenidos curriculares, reside en enseñar no sólo lo estipulado en el plan de estudios, sino su aplicación y transferencia.
- Con miras a darle un sentido más pedagógico a la evaluación es importante, no quedarse con los resultados obtenidos, sino reparar y analizar el contenido de un examen y los conceptos de contenido, aprendizaje y evaluación implícitos en éste.
- Las funciones formativas de la evaluación, sólo se podrán realizar cuando los resultados no queden en el número, calificación o puntaje alcanzado; es necesario pasar al terreno de la comprensión y analizar el por qué de los resultados obtenidos, para darle un sentido formativo y no sólo informativo a la evaluación.
- Es importante complementar y diversificar la elaboración y análisis de los instrumentos de evaluación e incorporar cuestionarios, que permitan realizar estudios de contexto, para conocer no meros resultados, sino procesos y significaciones.
- La evaluación del aprendizaje debe ser abordada con un enfoque teórico y metodológico serio y no reducirla al plano técnico centrado en la medición.

Queda por señalar que los hallazgos generados hasta el momento, genera nuevos debates y cuestionamientos para la realización de trabajos posteriores, que promuevan una reflexión



pedagógica con docentes en torno a las implicaciones didácticas y curriculares que se derivan del contenido y estructura de los estudios de evaluación internacional.

#### **v. Una puerta a nuevos debates y cuestionamientos**

El penetrar en el ámbito de las evaluaciones internacionales, abre la puerta a debates y cuestionamientos en muchos términos: políticos, económicos, socioeducativos, culturales, curriculares, didácticos y de la evaluación misma; sin olvidar que la intención de esta investigación es presentar un análisis pedagógico de las pruebas internacionales y de su relación con las nacionales, no se puede dejar de advertir la discusión que la adopción de estos estudios ha generado entre la sociedad mexicana en muchos aspectos; así como las limitaciones propias encontradas en las pruebas internacionales, pues aunque ofrecen un concepto de examen diferentes, se advierten factores y principios que vale la pena cuestionar, con el fin de repensar y recuperar el sentido pedagógico de la evaluación.

⇒ Políticamente, no se puede dejar de observar cómo la rendición de cuentas viene a ser el eje rector de la elaboración y aplicación de estos estudios, lo que denota una vez más, el uso de la evaluación como medida política más que educativa; ¿serán éstas las orientaciones que la evaluación debe conservar en la sociedad de hoy?. Ante la rápida inserción de la adopción de la evaluación como medida política, vale la pena retomar los controversiales cuestionamientos que se le han formulado a la evaluación desde su nacimiento y proliferación y reparar en si los actuales sistemas educativos de México y el mundo: ¿deben continuar con la adopción de la evaluación como medio control y regulación de un sistema?, ¿qué resultados se obtienen con esto?, ¿la evaluación en los sistemas educativos se utiliza para juzgar, relegar, excluir, informar, conocer, o mejorar?, ¿cuál es o debe ser el sentido de la evaluación para las sociedades del mundo



contemporáneo?, ¿cómo superar el enfoque fiscalizador con que se ha venido abordando la evaluación desde el siglo XVIII?

- ⇒ En términos económicos queda claro que la intención de estudios evaluativos como el de la OCDE, es el de identificar el costo-beneficio de los sistemas educativos en cuanto a su efectividad para formar capital humano e incorporar a jóvenes estudiantes a la fuerza del trabajo, que permita un crecimiento de las economías; lo que lleva a cuestionar de acuerdo con sus planteamientos y orientaciones, los estudiantes que no alcanzan el promedio establecido, ¿no están preparados para la vida y el trabajo?, ¿son sólo las habilidades para tener éxito en el mundo laboral, las que la escuela debe privilegiar?, ¿con qué criterios se determinan el tipo de aptitudes para la vida y el "mundo del mañana"?
- ⇒ El aspecto socioeducativo y cultural, es quizá el más analizado y polémico porque es indudable que para todos queda muy claro, la gran diversidad existente entre las características y condiciones sociales y educativas de estudiantes y sistemas de los países evaluados. Ante lo cual, cabe cuestionar la pertinencia de recurrir a instrumentos de medición estandarizados para situaciones y contextos diferentes y lleva a preguntar: ¿Qué tan válido es adoptar pautas igualitarias de medición en contextos socioeducativos desiguales; sin antes ajustar las características sociales y de nuestro sistema educativos?, ¿Cuáles son los intereses a los que responde la adopción de medidas universales de evaluación que no corresponden a las condiciones educativas del medio en el que se aplican?, ¿Cuál es la función y trasfondo real de la evaluación en este sentido?, En aspectos culturales, queda claro la disparidad existente entre las situaciones cotidianas a las que pruebas como las del PISA de la OCDE recurren y el escenario y realidades de los alumnos mexicanos. No se puede dejar de reparar en ¿qué tan familiarizados están los estudiantes del país con la lectura de diagramas, garantías o transacciones bancarias y comerciales?.
- ⇒ En el ámbito de lo curricular y didáctico, conviene recalcar la brecha encontrada entre el contenido y forma de la enseñanza y la evaluación en el marco internacional y la relación



que guarda con el contenido trabajado en las escuelas mexicanas y las evaluaciones realizadas a nivel nacional; lo que lleva a reparar en la tan necesaria y aún no lograda vinculación entre didáctica y evaluación y pensar en ¿qué implicaciones tiene para los sistemas de enseñanza, el contenido de evaluación de las pruebas internacionales?, y ¿qué tan válido es permitir que se evalúe lo que nos se enseña?; el gran distanciamiento existente entre sistemas de enseñanza y sistemas de evaluación ¿no estará pervirtiendo y trastocando un principio fundamental de la evaluación?

⇒ La investigación realizada deja claro que los estudios internacionales ofrecen una concepción diferente y adoptan nuevas tendencias en materia de evaluación y aprendizaje; pero en el tratamiento que se le da a los resultados obtenidos y en cómo se informa de éstos a la sociedad y sobre todo a uno de los principales actores del sistema educativo: el docente, prevalece un enfoque cuantificador y descriptivo, más que un análisis pedagógico de corte interpretativo. La manera en como se presentan y difunden los resultados, no conlleva a la función formativa y de retroalimentación de la evaluación; el saber que un tanto por ciento de los alumnos evaluados están en el nivel 1, 4 o 5, que el promedio alcanzado en matemáticas fue de 385 y que México quedó en el lugar 35 o 37, no retroalimenta la función del docente, único protagonista del verdadero cambio educativo. Lo que lleva a recapacitar en ¿cómo hacer llegar los resultados de las evaluaciones a los docentes?, ¿cómo trabajar más las evaluaciones y sus resultados a manera de retroalimentación pedagógica y no de mera calificación o cuantificación?, ¿cómo recuperar el sentido formativo de la evaluación y la necesaria vinculación entre evaluación y didáctica?

Las interrogantes y disyuntivas encontradas, reportan que en materia de evaluación del aprendizaje queda todavía mucho por hacer, el sentido pedagógico de la evaluación sigue siendo tarea pendiente y la elaboración de pruebas en México continúa con un enfoque reproductivo en materia de contenido y un tratamiento meramente técnico y arcaico; lo que demanda que los exámenes masivos que se aplican a nivel nacional, ofrezcan un



tratamiento más conceptual y cognitivo del contenido a evaluar y la incorporación de paradigmas alternativos en materia de medición y evaluación, que a nivel internacional ya se están operando.

Es importante que las pruebas mexicanas, dejen de ser instrumentos de tortura y control, centrados en la mera retención de conceptos o hechos y en la realización de ejercicios mecánicos sin sentido, recuperen su función didáctica y se conviertan, como pruebas que son, en una oportunidad de demostrar lo aprendido.

## REFERENCIAS

---

- Agut, S. y Grau, R. M. (2001). "Una aproximación psicosocial al estudio de las competencias"  
*Proyecto social* No. 9. p. 13-24 Universitat Jaime I Castellón Nacional.
- Ahumada. P. (2003). *La evaluación en una concepción de aprendizaje significativo*. Chile:  
Universitarias del Paraiso.
- Atkin, M. y Black, P. (1997). "Policy perils of international  
comparison: The TIMSS Case. *Phi Delta Kappan*. Vol. 79. Issue I.
- Atkin, M. y Black, P. (1997). "Policy perils of international comparison: The TIMSS Case. *Phi Delta  
Kappan*. Vol. 79. Issue I.
- ANEP.(2004 a). *Primer Informe Nacional. PISA 2003. Uruguay*. República Oriental del Uruguay:  
ANEP [En red]. Disponible  
en:[http://www.anep.edu.uy/qerenciaqrl/qer\\_inv\\_eva/publicaciones/evaluacion/pisa%202003%  
20primer%20informe.pdf](http://www.anep.edu.uy/qerenciaqrl/qer_inv_eva/publicaciones/evaluacion/pisa%202003%20primer%20informe.pdf).
- \_\_\_\_\_(2004 b). *Resumen de resultados. PISA 2003. Uruguay*. ANEP. República Oriental del  
Uruguay:. [En red]. Disponible  
en:[http://www.anep.edu.uy/qerenciaqrl/qer\\_inv\\_eva/publicaciones/evaluacion/pisa%202003%  
20primer%20informe.pdf](http://www.anep.edu.uy/qerenciaqrl/qer_inv_eva/publicaciones/evaluacion/pisa%202003%20primer%20informe.pdf) .
- \_\_\_\_\_(2004 c). *La evaluación Internacional PISA*. Material de divulgación para Uruguay.  
Montevideo Uruguay. [En red]. Disponible en:  
[http://www.anep.edu.uy/qerenciaqrl/qer\\_inv\\_eva/publicaciones/evaluacion/La%20Evaluaci%  
F3n%20internacional%20PISA.pdf](http://www.anep.edu.uy/qerenciaqrl/qer_inv_eva/publicaciones/evaluacion/La%20Evaluaci%F3n%20internacional%20PISA.pdf).
- \_\_\_\_\_(s.f.a). *Uruguay en el proyecto PISA 2003*. [En red]. Disponible en:  
[http://www.anep.edu.uy/qerenciaqrl/qer\\_inv\\_eva/publicaciones/evaluacion/pisa/PISA2.htm](http://www.anep.edu.uy/qerenciaqrl/qer_inv_eva/publicaciones/evaluacion/pisa/PISA2.htm)  
Recuperado en: 11/01/05.
- \_\_\_\_\_(s.f.b). *Uruguay en el proyecto PISA 2003. Evaluación de Matemática*. Material informativo  
para Profesores, Directores e Inspectores. República Oriental del Uruguay. ANEP. [En red].  
Disponible en:



[http://www.anep.edu.uy/gerenciagr/ger\\_inv\\_eva/publicaciones/documentos/Matematica.pdf](http://www.anep.edu.uy/gerenciagr/ger_inv_eva/publicaciones/documentos/Matematica.pdf)

(s.f.c). *Uruguay en el proyecto PISA 2003. Evaluación de Ciencias* informativo para Profesores, Directores e Inspectores. República Oriental del Uruguay. ANEP. [En red]. Disponible en:

[http://www.anep.edu.uy/gerenciagr/ger\\_inv\\_eva/publicaciones/documentos/Ciencias.pdf](http://www.anep.edu.uy/gerenciagr/ger_inv_eva/publicaciones/documentos/Ciencias.pdf)

Recuperado en: 11/01/05.

(s.f.). *Los estudiantes uruguayos en el proyecto PISA*. República Oriental del Uruguay. ANEP.[En red]. Disponible en:

[http://www.anep.edu.uy/gerenciagr/ger\\_inv\\_eva/publicaciones/documentos/Triptico%201.pdf](http://www.anep.edu.uy/gerenciagr/ger_inv_eva/publicaciones/documentos/Triptico%201.pdf)

Arancibia, V. (1997). *Los sistemas de medición y evaluación de la calidad de la educación* Documento 2. Laboratorio Latinoamericano de Evaluación de la Calidad Publicaciones OREAL/UNESCO Santiago. [En red]. Disponible en:

[http://www.unesco.cl/medios/biblioteca/documentos/2sistemas\\_medicion\\_evaluacion.pdf](http://www.unesco.cl/medios/biblioteca/documentos/2sistemas_medicion_evaluacion.pdf).

(s.f). "Laboratorio Latinoamericano de Evaluación de Calidad de la Educación".*Revista Iberoamericana de Educación No. 10* Organización de Estados Iberoamericanos para la Educación la Ciencia y la Cultura. (OEI). [En red] Disponible en: <http://www.campus-oei.org/oeivirt/rie10a10.html> Recuperado en: 09/11/2002.

Arce, M. (2005). "Análisis de Reactivos" *Boletín CENEVAL*. 1 de enero del 2005. Año III. No.10.

Arregui, P. (2003). *Sistemas de determinación de metas de logros de aprendizaje escolar como instrumentos para mejorar la calidad, la equidad y la responsabilidad en los procesos educativos en América Latina*. [En red]. Disponible en:

<http://www.unesco.cl/promedl7/prospectivas/arregui.pdf>.

Backhoff, E. (2003a). *Desarrollo, validación e implementación de un sistema para la selección de estudiantes a la Universidad Autónoma de Baja California*. [CD-ROM]. Tesis del Doctorado Interinstitucional en Educación. Universidad Autónoma de Aguascalientes. México.

Backhoff, E. y Solano, G. (2003b). *Tercer Estudio Internacional de Matemáticas y Ciencias Naturales: (TIMSS): Resultados de México en 1995 y 2000*. Informe Técnico. [En red] Disponible en:



[http://capacitacion.ilce.edu.mx/inee/pdf/productos/tercer\\_estudio\\_internacional\\_matematicas\\_ciencias.pdf](http://capacitacion.ilce.edu.mx/inee/pdf/productos/tercer_estudio_internacional_matematicas_ciencias.pdf).

- Barba, A. (2004). "Reprueba la OCDE a México en Ciencias". *Reforma*. 5 de febrero 2004.
- Barrón, C.(2000). "La educación basada en competencias en el marco de los procesos de globalización". En: M. A Valle. (coord). *Formación en Competencias y Certificación profesional*. col. Pensamiento Universitario, Tercera época 91 CESU-UNAMN.
- Bazán, A. y Ferrer, A., (ed.). (2002). *Estrategias de evaluación y medición del comportamiento en Psicología*. México: Instituto Tecnológico de Sonora (ITS) Universidad Autónoma de Yucatán (UADY).
- Beaton, A. E. Mullis, I. V.S., Martin, M. O., Gonzalez, E. J., Kelly, D.L. y Smith, T. (1996a). *Mathematics Achievement in the Middle School Years: IEA's Third International Mathematics and Science Study TIMSS*. TIMSS International Student Center. Boston Collage. Chestnut Hill, MA, USA. [En red]. Disponible en: <http://timss.bc.edu/timss1995i/MathB.html>.
- \_\_\_\_\_ (1996b). *Science Achievement in the Middle School Years: IEA's Third International Mathematics and Science Study (TIMSS)*. International Student Center. Boston Collage. Chestnut Hill, MA, USA. [En red]. Disponible en: <http://timss.bc.edu/timss1995i/SciencB.html>.
- Benedito, G. (1975). "El problema de la medida en psicología". En: N. Braunstein, M. Pasternac y Benedito, G. *Psicología, ideología y ciencia*. México: Siglo XXI editores.
- Bertoni, A., Poggi, M. y Teobaldo, M.(1995). *Evaluación. Nuevos significados para una práctica compleja* Kapelusz, Buenos Aires.
- Bloom, B. (1971).Taxonomía de los objetivos de educación: La clasificación de las metas educacionales. Buenos Aires: El ateneo.
- Bruer, J. (1997). *Escuelas para pensar. Una ciencia del aprendizaje en el aula*. (Traducción de Pilar Pineda Herrero) México: Biblioteca del normalista. (Original publicado en 1993).
- Cabrera, F. y Espin, J. (1986). *Medición y evaluación educativa. Fundamentos teórico- prácticos*, Barcelona: Promociones y publicaciones educativas.



- Camilloni, A., Celman, S., Litwin, E. y Palou de Maté, M.C. (1998). *La evaluación de los aprendizajes en el debate didáctico contemporáneo*. Argentina: Paidós Educador.
- Casassus, J. (1997). *Estándares en educación: conceptos fundamentales*. Documento 3. [En red]. Disponible en:  
[http://www.unesco.cl/medios/biblioteca/documentos/3estandares\\_educacion\\_conceptos.pdf](http://www.unesco.cl/medios/biblioteca/documentos/3estandares_educacion_conceptos.pdf)
- Casassus, J., Froemel, J. E. y Palafox, J. K. Cusato, S.(1998). *Primer Estudio Internacional Comparativo sobre Lenguaje, Matemáticas y Factores Asociados en Tercero y Cuarto Grado*. Santiago de Chile. LLECE UNESCO [En red]. Disponible en:  
[http://www.unesco.cl/medios/biblioteca/documentos/informe\\_laboratorio\\_1\\_espanol.pdf](http://www.unesco.cl/medios/biblioteca/documentos/informe_laboratorio_1_espanol.pdf).
- \_\_\_\_\_(2000). *Primer Estudio Internacional Comparativo sobre Lenguaje, matemática y factores asociados para alumnos del tercer y cuarto grado de la educación básica*. Segundo informe Santiago de Chile: Laboratorio Latinoamericano de Evaluación de la Calidad de la Educación / UNESCO Recuperado en abril 2002 en:  
[http://www.unesco.cl/medios/biblioteca/documentos/informe\\_laboratorio\\_2\\_espanol.pdf](http://www.unesco.cl/medios/biblioteca/documentos/informe_laboratorio_2_espanol.pdf).
- Castañeda, S., Lugo, E. Pineda, L. y Romero, N. (1998). "Estado del arte de la evaluación y el fomento de desarrollo intelectual en la enseñanza de ciencias, artes y técnicas." En S. Castañeda (coord.). *Evaluación y fomento del desarrollo intelectual en la enseñanza de ciencias, artes y técnicas. Perspectiva internacional en el umbral del siglo XXI*. México: UNAM, Facultad de Psicología, CONACYT, Porrúa.
- \_\_\_\_\_(2002 a). "Diagnostividad e interpretabilidad: retos a la medición de resultados de aprendizaje." En: A. Bazán y J. Ferrer (ed). *Estrategias de evaluación y medición del comportamiento en psicología*. México: ITSON-UADY.
- \_\_\_\_\_(2002 b). "Explorando la utilidad de los análisis clásicos y Rasch en la determinación de la consistencia interna de subpruebas cortas de homogeneidad alta." En: A. Bazán A. y J. Ferrer (ed). *Estrategias de evaluación y medición del comportamiento en psicología*. México: ITSON-UADY.
- CEDEFOP. (1994). "Las competencias: el concepto y la realidad" *Formación Profesional*. Revista Europea.



- CENEVAL.(2004a). *Cómo prepararse para el Examen Nacional de Ingreso a la Educación Media Superior*. (EXANI.1). México: CENEVAL.
- \_\_\_\_\_(2004b). *¿Quiénes somos?. Estatuto*. [En red]. Disponible en: <http://portal.ceneval.edu.mx/portalceneval/docs/100/estatuto.pdf>.
- \_\_\_\_\_(s.f.) PREXANI-1. [En red]. Disponible en: <http://www.cenevalac.edu.mx/portalceneval/index.php?q=info.fichas.ficha54>) Recuperado en: 04/06/2004.
- 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.
- Cevallos D. (2001). "Mala nota en examen de calidad". México. 5 dic. 2001 *Acción por los niños* Recuperado en: <http://acciónporlosniños.org.pe/Not-Lati1.htm> en: diciembre 2001.
- Coll, C. y Martín E. (1994). "La evaluación del aprendizaje en el currículum escolar: Una perspectiva constructivista". En: C. Coll, E. Martín, T. Mauri, M. Miras, J. Onrubia, I. Sole y A. Zabala. *El constructivismo en el aula*. Barcelona, Grao Editorial 2ª. ed. revisada.
- \_\_\_\_\_(2003). *Aprender contenidos, desarrollar capacidades. Intenciones educativas y planificación de la enseñanza*. Barcelona, España: Edebé.
- Comenio, J. A. (1922/2000). *Didáctica Magna*. México Porrúa, 10º Edición.
- Comisión Metropolitana de Instituciones Públicas de Educación Media Superior. *Concurso de Selección para la Educación Media Superior 1996*. (Documento impreso, no editado).
- Contreras, L. A. y Backhoff, E. (2004). "Metodología para elaborar exámenes criterios alineados con el currículo. En: S. Castañeda (coord). *Educación, aprendizaje y cognición. Teoría en la práctica*. México: Ed. El manual moderno. Universidad de Guadalajara. Instituto de Neurociencias. UNAM, Facultad de Psicología.
- Corvalán, O. (1993). *Una propuesta curricular, la enseñanza basada en competencias*. CIDE [En red] Disponible en: [http://mail.udlap.mx/~ciedd/RIIC/BIB/Competencias\\_Chile.pdf](http://mail.udlap.mx/~ciedd/RIIC/BIB/Competencias_Chile.pdf)
- Corriente de Opinión Fundación Chile Unido. (2002). *Calidad en Educación: ¿Cómo mejorar las prácticas pedagógicas?* [En red] Disponible en: [http://www.chileunido.cl/corrientes/docs\\_corrientes/cdeop66calidadeducacil.pdf](http://www.chileunido.cl/corrientes/docs_corrientes/cdeop66calidadeducacil.pdf).



- De Vega, M. (1986). *Introducción a la Psicología Cognitiva*. México: 2ª ed. Alianza editorial mexicana.
- Delgado, M. (2001). "Trueno OCDE a México en Educación". *Reforma* México 4 de diciembre 2001.
- Delors, J. (1996). *La Educación encierra un tesoro*. Informe para la UNESCO de la Comisión Internacional sobre la Educación para el siglo XXI. UNESCO.
- Del Valle, S. (2004). "Muestra caída, educación básica" *Reforma*, Lunes 17 de mayo del 2004.
- Díaz Barriga, A. (1988). "Una polémica en relación al examen", *Perfiles Educativos*. No. 41-42, México, UNAM, p. 65-76.
- \_\_\_\_\_ (comp.). (1993). *El examen, textos para su historia y debate*. México: Centro de Estudios sobre la Universidad. CESU. Plaza y Valdés. Universidad Nacional Autónoma de México. (Primera reimpresión, 2000).
- \_\_\_\_\_ (1997). "Tesis para una teoría de la evaluación y sus derivaciones en la docencia." En: *Didáctica y Currículum. Edición corregida y aumentada*. México: Paidós. *Didáctica y currículum..* México: Paidós. (Edición corregida y aumentada).
- \_\_\_\_\_ (2003). (coord.) *La investigación curricular en México. La década de los noventa*. (La investigación educativa en México 1992-2002-5). México: COMIE, SEP CESU.
- \_\_\_\_\_ (2004). "Comentarios de la mesa redonda: Comparabilidad de los resultados de las evaluaciones. En: *INNE Jornadas de Evaluación Educativa. Memoria*. México: INNE.
- Díaz Barriga Arceo, F. y Hernández Rojas, G. (2002). "Estrategias para el aprendizaje significativo II: Comprensión y composición de textos." En: *Estrategias Docentes para un aprendizaje significativo*. (2ª edición). México: Mc. Graw Hill.
- Díaz Barriga Arceo y Lugo E. (2003). Desarrollo del currículo". En: A. Díaz Barriga. (coord.) *La investigación curricular en México. La década de los noventa*. (La investigación educativa en México 1992-2002-5) México: COMIE, SEP CESU.
- Díaz Barriga Arceo, F. (2004). "Las rúbricas: su potencial como estrategia para una enseñanza situada y una evaluación auténtica dl aprendizaje". *Perspectiva Educativa*. (Chile), No. 43,



- Ducoing P. (coord.). (2003). *Sujetos Actores y Procesos de Formación*. Tomo I. Formación para la investigación, los académicos en México, Actores y Organizaciones. (La investigación educativa en México 1992-2002-8) México: COMIE, SEP, CESU.
- Durkheim, E. (s.f.). "Los grados", en: *Historia de la educación y de las doctrinas pedagógicas*, Madrid: La piqueta, (1ª. Ed. Francesa 1938).
- Ebel, R. (1977). "La medición educacional: perspectiva histórica". En: *Fundamentos de la medición educacional*. Buenos Aires: Ed. Guadalupe.
- Eisner, E. (1999). *Usos y límites de las pruebas de desempeño*. Programa de Promoción de la Reforma Educativa en América Latina y el Caribe. (PREAL) Grupo de Análisis para el Desarrollo. (GRADE) [En red]. Disponible en: <http://www.preal.cl/GTEE/pdf/Eisner2.pdf>.
- Ericsson, F. (1989). "Métodos cualitativos de investigación sobre la enseñanza". En: M.C. Witrock, M.C. *La investigación de la enseñanza II. Métodos cualitativos y de observación*. Barcelona-Paidós. M. E. C.
- Escudero, T. (2003). "Desde los test hasta la investigación evaluativa actual. Un siglo, el XX, de intenso desarrollo de la evaluación en educación. *Revista Electrónica de Investigación y Evaluación educativa*. (RELIEVE). Vol. 9, No. 1. [En red]. Disponible en: <http://www.uv.es/RELIEVE/>.
- Esquivel, J. M. (2001). Los próximos pasos. ¿Hacia dónde y cómo avanzar en la evaluación de los aprendizajes en América Latina? [En red]. Disponible en: <http://www.campus-oei.org/calidad/grade.PDF> 2003.
- Eurydice (2002). *Las competencias clave. Un concepto en expansión dentro de la educación general obligatoria*. Eurydice. Red Europea de Información en Educación. Recuperado en febrero 2004, [En red]. Disponible en: <http://www.eurydice.org/Documents/survey5/es/FrameSet.htm>.
- Ferrer, 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*. Documento 26, Informe de investigación. Programa de Promoción de la Reforma Educativa en América Latina y el Caribe (PREAL). [En red]. Disponible en: <http://www.preal.org/docs-trabajo/Arreguietal26.pdf>.



- Ferrer, A. y Díaz, M. I. (2004). "Análisis de componentes y la competencia para la selección de habilidades complejas". En: S. Castañeda, S., (coord). *Educación, aprendizaje y cognición. Teoría en la práctica*. México: Ed. El manual moderno. Universidad de Guadalajara. Instituto de Neurociencias. UNAM, Facultad de Psicología.
- Fernández J. L. y Quiroz, R. E. (s.f.). *La evaluación por competencias en el proceso de formación integral*. Universidad de Antioquia Colombia. [En red]. Disponible en: <http://ayura.udea.edu.co/pedagogia/lecciones%20inaugurales/lecciones2001/leonidas.htm>. Recuperado en: 08/03/2004.
- Foucault, (1977). *Vigilar y Castigar*. México, Siglo XXI. Editores.
- Foulon C. (ed). (2002). *Reformulación del sistema de selección a la educación superior. Antecedentes para la discusión*. Documento de trabajo No. 334. Chile: Centro de estudios públicos. [En red]. Disponible en: [http://www.cepchile.cl/dms/lang\\_1/doc\\_3121.html](http://www.cepchile.cl/dms/lang_1/doc_3121.html).
- Frank, W. (1997). "Evaluación educacional internacional. El ejemplo de la IEA, su investigación y desafío". En: C. De Moura, y M. Carnoy, (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). [En red]. Disponible en: <http://www.iadb.org/sds/doc/Edu%2DCCastroS.pdf>.
- García Cortés, F. (1983). *Paquete de auto enseñanza de evaluación del aprovechamiento escolar*. México: UNAM Centro de Investigaciones y servicios educativos. 2ª. Ed.
- García Garrido, J. L. (1986). *Fundamentos de Educación Comparada*. Madrid: Dykinson.
- Garduño, T., Guerra, M. E., Silva, A., Rodríguez, L. E. (2001). *Competencias. Preescolar Comunitario y Primaria Comunitario*. México, CONAFE.
- Gil, G. (s.f.). *El proyecto Pisa de la OCDE*. Recuperado en: <http://www.ince.mec.es/pub/pisa.htm> Recuperado 12/1/2005.
- Gil, M. (2004). Comentarios a la mesa: "Comparabilidad de los resultados de las Evaluaciones", por Felipe Martínez Rizo. En: INNE. *Jornadas de Evaluación Educativa*. Memoria México. INEE



- \_\_\_\_\_. (2005). *México en PISA, Análisis de Resultados*. Mesa Pública de Análisis de Evaluación Educativa. Versión estenográfica. México: INNE. [En red]. Disponible en: <http://capacitacion.ilce.edu.mx/inee/mesapublipisa2003.pdf>.
- González, E. J. y Smith, T. A. (1997). *User Guide for the TIMSS International Database Primary and Middle School Years (Population 1 and Population 2) Data Collected in 1995 TIMSS*. USA : International Study Center Boston College Chestnut Hill, MA. [En red]. Disponible en: [http://timss.bc.edu/timss1995i/database/UG\\_1and2.pdf](http://timss.bc.edu/timss1995i/database/UG_1and2.pdf).
- Griffith, J. E. y Medrich, E. A. (1992). "¿Que esperan los Estados Unidos de los estudios internacionales comparados sobre educación?" *Perspectivas Revista trimestral de educación*. Vol. XXII, No. 3, 1992 (83) Santiago de Chile: OREALC UNESCO pp. 514-523
- Haladyna, T. M., Haladyna, R. Merino, C. (s.f.). *Preparación de preguntas de opción múltiple para medir el aprendizaje de los estudiantes*. Organización de Estados Iberoamericanos. OEI. [En red]. Disponible en: <http://www.campus-oei.org/revista/deloslectores/267Haladyna.PDF>  
Recuperado 7/8/2002.
- Hambleton, R. y Van der Linden, W. (1996). *Handbook of Modern Item Response Theory*. USA: Springer- Verlag New Cork Inc.
- Harmon, M., Smith, T. A., Martin, M. O., Kelly, D. L., Beaton, E. A, Mullis, I. V. S., Gonzalez, E. J., and Orpwood, G. (1997). *Performance Assessment in IEA's Third International Mathematics and Science Study (TIMSS)* Chestnut Hill, MA: Centre for the Study of Testing, Evaluation, and Education Policy, Boston College {En red}. Disponible en: [timss.bc.edu/timss1995i/TIMSSPDF/PAToc.pdf](http://timss.bc.edu/timss1995i/TIMSSPDF/PAToc.pdf).
- Herrera, C. (2004, diciembre,7). "Empeora el rendimiento escolar en México durante el foxismo": OCDE. *La Jornada*, Sociedad y Justicia., p. 42
- Hernández, R. ( 1991). "A quienes prefieren las pruebas de ensayo", *Educación. Revista de la Universidad de Costa Rica*. Vol. 15. No. 1. Universidad de Costa Rica.



- Huerta, J. J., Pérez, I. S., Castellanos, A. R. (2000). "Desarrollo curricular por competencias profesionales". *Educ. Revista de Educación Nueva Época*. No. 13, Universidad de Guadalajara. [En red]. Disponible en: <http://educacion.jalisco.gob.mx/consulta/educar/13/13Huerta.html>.
- House, E. (1992). *Evaluación, ética y poder*. Madrid. Morata.
- Horn, R., Wolf, L., Velez, E. (1992). *Sistemas de medición de evaluación educacional América Latina. Reseña temática y experiencias recientes*. [En red] Disponible en: : <http://www.unesco.cl/pdf/actyeven/ppe/boletin/artesp/27-1.pdf>.
- Hussein, M. G. (1992). "¿Qué espera Kuwait de su participación en el estudio internacional sobre la enseñanza de las matemáticas y las ciencias? *Perspectivas Revista trimestral de educación*. Vol. XXII, No. 3, 1992 (83) ) Santiago de Chile: OREALC UNESCO pp. 499-504
- IEA's (1995a). *Third International Mathematics and Science Study 1995* International Study Center. ISC. Boston College. [En red]. Disponible en: <http://timss.bc.edu/timss1995.html>
- \_\_\_\_\_ (1995b). *TIMSS Mathematics Items: Released Set for Population 1 (Third and fourth grade)* TIMSS. [En red]. Disponible en: [http://timss.bc.edu/timss1995i/TIMSS\\_SPDF/Amitems.pdf](http://timss.bc.edu/timss1995i/TIMSS_SPDF/Amitems.pdf).
- \_\_\_\_\_ (1995c). *TIMSS Science Item. Released Set for Population 1(Third and Fourth Grade)*. TIMSS. [En red]. Disponible en: <http://timss.bc.edu/timss1995i/TIMSSPDF/ASitems.pdf>
- \_\_\_\_\_ (1995d). *TIMSS Science Item. Released Set for Population 2 (Seventh and eighth grades)*, TIMSS. [En red]. Disponible en: <http://timss.bc.edu/timss1995i/TIMSSPDF/BSitems.pdf>
- \_\_\_\_\_ (1995e). *TIMSS Mathematics Items: Released Set for Population 2 (Seventh and eighth grades)* TIMSS. [En red]. Disponible en: <http://timss.bc.edu/timss1995i/TIMSSPDF/BMIitems.pdf>
- \_\_\_\_\_ (1999a). *Third International Mathematics and Science Study Repeat 1999*. ISC Boston College. [En red]. Disponible en: <http://www.iea.nl/iea/hq/index.php?id=109&type=1>
- \_\_\_\_\_ (1999b). *Technical standards for IEA studies*, [En red]. Disponible en: [http://iea.nl/Home/IEA\\_pdf](http://iea.nl/Home/IEA_pdf)



- \_\_\_\_\_(s.f.). *Estudio de Educación Cívica*, [En red]. Disponible en: [http://www2.rz.hu-berlin.de/empir\\_bf/iea\\_s.html](http://www2.rz.hu-berlin.de/empir_bf/iea_s.html). Recuperado en: 9/06/2003.
- \_\_\_\_\_(2002). *Trends i International Mathematics and Science Study. TIMSS 2003. Mains Survey. Student Questionnaire*. TIMSS. International Study Center. Boston College. Chestnut Hill MA. [En red]. Disponible en: [http://timss.bc.edu/timss2003i/PDF/T03\\_Question.pdf](http://timss.bc.edu/timss2003i/PDF/T03_Question.pdf)
- INEE (2003). *La calidad de la educación básica en México*. Primer informe anual 2003. México, INEE. [En red]. Disponible en: <http://capacitacion.ilce.edu.mx/inee/estadisticas.htm#1>
- \_\_\_\_\_(2004). *La calidad de la educación básica en México 2004* México, [En red]. Disponible en: <http://capacitacion.ilce.edu.mx/inee/publicaciones1.2.htm>.
- \_\_\_\_\_(2004). *Evaluación PISA 2003. Resumen de los primeros resultados en España*. Madrid: INCE Recuperado en febrero 2005, en: <http://www.ince.mec.es/pub/pisa2003resumenespana.pdf>
- INSAFORP. (2001). *Una metodología de diseño curricular para programas de formación profesional por competencias*. San Salvador: Sistema de Formación Profesional. [En red]. Disponible en : <http://www.insaforp.org.sv/cdi/fulltex/insamet2.pdf>
- ISEI-IVEI. (s.f.). *Evaluación internacional de Matemáticas y Ciencias. TIMSS 2003. Primer Informe de Resultados*. Gobierno Vasco: ISEI-IVEI [En red]. Disponible en: <http://www.isei-ivei.net/cast/pub/TIMSS2003euskadicast.pdf> Recuperado 12/1/2005.
- Judges V.A. (1971). "La evolución de los exámenes", En: *El examen de los exámenes*. Buenos Aires. Ed. Estrada.
- Kant. E. (1979). *Filosofía de la historia*. (8ª reimpresión 2000). México. F. C. E.
- LLECE. *Marco Conceptual*. (1997) Documento 1- UNESCO. Publicaciones. [En red]. Disponible en: <http://llece.unesco.cl/>
- LLECE (2001). *Primer Estudio Internacional Comparativo sobre Lenguaje, Matemáticas y Factores Asociados para alumnos del tercer y cuarto grado de educación básica.*, Informe técnico. Santiago de Chile: LLECE UNESCO. [En red]. Disponible en: <http://www.unesco.cl/pdf/laboratorio/palafox/total.pdf>



- \_\_\_\_\_. (2002). *Estudio Cualitativo de Escuelas con Resultados Destacables en Siete Países Latinoamericanos*. UNESCO Santiago de Chile. Recuperado. En: octubre 2002, en: [http://lece.unesco.cl/medios/pdf/espanol/cuali\\_es.pdf](http://lece.unesco.cl/medios/pdf/espanol/cuali_es.pdf)
- López, A., y Moreno, M. L. (1997a). *Resultados españoles de Matemáticas en el TIMSS*. Madrid: INCE. [En red]. Disponible en: <http://www.ince.mec.es/timss/index.htm>
- López, A., y Moreno, M. L. (1997b). *Resultados españoles de Ciencias en el TIMSS*. Madrid: INCE. [En red]. Disponible en: <http://www.ince.mec.es/timss/ci/index.htm>
- Mabry, L. (1999). "Writing to the Rubric: Lingering Effects of Traditional Standardized Testing on Direct Writing Assessment" *Phi Delta Kappan*. [En red]. Disponible en: <http://www.pdkintl.org/kappan/kmab9905.htm>,
- Martin, M O., Rust, K. Y Adams, R. (1999). *Technical Standards for IEA Studies*. Amsterdam: IEA [En red] Disponible en: [http://www.iea.nl/iea/hq/fileadmin/IEA\\_Publications/IEA\\_TS\\_2003.pdf](http://www.iea.nl/iea/hq/fileadmin/IEA_Publications/IEA_TS_2003.pdf)
- Martin, M O, Gregory, K D, Stemler, S E. (ed.). (2000a). *TIMSS 1999 Technical Report*. Volume 1 Chestnut Hill, MA: Boston Collage. [En red]. Disponible en: [http://timss.bc.edu/timss1999i/tech\\_report.html](http://timss.bc.edu/timss1999i/tech_report.html)
- Martin, M O., Mullis, I V S., Gonzales, E J., Gregory, K D., Smith, T A., Chrostowski, S J, Garder, R A. y O'Connor, K M. (2000b). *TIMSS 1999. International Science Report*. Chestnut Hill, MA: Boston College [En red] Disponible en: [http://timss.bc.edu/timss1999i/science\\_achievement\\_report.html](http://timss.bc.edu/timss1999i/science_achievement_report.html)
- Martin, M O, Mullis, I V S., Gregory, K. D., Hoyle, C. y Shen, C. (2001) *Effective Schools in Science and Mathematics* Chestnut Hill, MA: Boston College [En red]. Disponible en: [http://timss.bc.edu/timss1995i/TIMSSPDF/T95\\_EffSchool.pdf](http://timss.bc.edu/timss1995i/TIMSSPDF/T95_EffSchool.pdf)
- Martin, M O., Mullis, I V S. y Chrostowski, S J. (eds.). (2004). *TIMSS 2003. Technical Report International* Chestnut Hill, MA: Boston College [En red]. Disponible en: [http://timss.bc.edu/PDF/t03\\_download/T03TECHRPT.pdf](http://timss.bc.edu/PDF/t03_download/T03TECHRPT.pdf)
- Martin, M O., Mullis, I V S. y Chrostowski, S J (eds.). (2004). *TIMSS 2003. Technical Report International* IEA. Chestnut Hill, MA: Boston College [En red]. Disponible en: [http://timss.bc.edu/PDF/t03\\_download/T03TECHRPT.pdf](http://timss.bc.edu/PDF/t03_download/T03TECHRPT.pdf)



- Martínez, R. y Schmelks, S. (s.f.). *Aseguramiento de la calidad de las pruebas de estándares nacionales para la educación primaria de la Secretaría de Educación pública*. Ponencia presentada en el V Congreso de Investigación Educativa México. [En red] Disponible en: ed consultado en: [http://snee.sep.gob.mx/BROW-AES/Ponencia2\\_VCongreso.htm](http://snee.sep.gob.mx/BROW-AES/Ponencia2_VCongreso.htm)  
Recuperado 12/03/2004.
- Martínez, R. y Moreno R. (2002). "Integración de la teoría sustantiva, diseño de pruebas y modelos de análisis en la medición psicológica. En: A. Bazán y J. Ferrer (ed.). *Estrategias de evaluación y medición del comportamiento en psicología*. México: Instituto Tecnológico de Sonora (ITSON) y Universidad Autónoma de Yucatán (UADY).
- Martínez, F. (2001). "Evaluación educativa y pruebas estandarizadas. Elementos para enriquecer el debate". *Revista de la Educación Superior*. Vol. XXX. No. 120. Octubre-Diciembre de 2001.
- Martínez, F. (2003). *Los Resultados de las Pruebas PISA. Elementos para su interpretación. Cuaderno. No. 1.* [En red]. Disponible en: <http://capacitacion.ilce.edu.mx/inee/pdf/Cuaderno1.pdf>
- Martínez, M. (2003). "Concepto y usos de la educación comparada". En *Educación comparada. Nuevos retos, renovados desafíos*. Madrid: La muralla.
- Martínez, N. (2003). "México reprueba en ciencias y lectura" *El Universal* Martes 01 de julio de 2003
- Martínez, N. (2004, diciembre 7). "Reprueba la OCDE a México en educación" *El Universal*, Sección A., p. 20.
- Mattos, L. (1963). *Compendio de Didáctica General*. Argentina: Kapeluz.
- Mejía, F. (2002) *La experiencia de la Dirección General de Evaluación en la Educación básica Normal. 30 años de medición de logro educativo*.
- Melgar, I. (2001, octubre, 15). Reprueba México Calidad Educativa. *Reforma*, pp. Primera Plana.
- Melgar, I. (2001, octubre, 18). Reprueba la UNESCO aprendizaje. *Reforma*, pp. 11.
- Ministerio de Educación y Cultura de Chile ( 2001). "TIMSS, evidencias para un nuevo currículo". y "TIMSS, oportunidad para crecer. Situarse para mejorar los resultados" *Revista Educación* Noviembre 2001 . Edición No. 290. Gobierno de Chile. [En red]. Disponible en: <http://www.mineduc.cl/revista/anteriores/noviembre01/index.htm#r1>



- Monereo, C. coord. (1998). *Estrategias de enseñanza y aprendizaje. Formación del profesorado y aplicación en el aula*. México: SEP/cooperación española. (Biblioteca del normalista).
- Montero, E. (2003). *La teoría de respuesta a los ítems: una moderna alternativa para el análisis psicométrico de instrumentos de medición* [En red]. Disponible en: [http://www.usma.ac.pa/publicaciones/cipsu/eiliana\\_4.html](http://www.usma.ac.pa/publicaciones/cipsu/eiliana_4.html)
- Morin, E. (1999). *Los siete saberes necesarios para la evaluación del futuro* Paris, Francia: UNESCO [En red]. Disponible en: <http://www.complejidad.org/27-7sabesp.pdf>
- Moreno, A. (ed.). (2003). *La comprensión del cerebro. Hacia una nueva ciencia del aprendizaje*. México: OCDE. Santillana Aula XXI.
- Mullis, I V S, Martin, M O, Beaton, A E, Gonzales, E J, Kelly, D L, Smith, T A.. (1998). *Mathematics and Science Achievement in the Final Year of Secondary School*: Chestnut Hill, MA: Boston Collage. [En red]. Disponible en: [http://timss.bc.edu/timss1995i/TIMSSPDF/C\\_full.pdf](http://timss.bc.edu/timss1995i/TIMSSPDF/C_full.pdf).
- Mullis, I V S, Martin, M O, Gonzalez, E J, Smith, T A. y Kelly D L. (1999). *School Contexts for Learning and Instruction, IEA's Third International Mathematics and Science Study* . Chestnut Hill, MA: Boston Collage. [En red]. Disponible en: <http://timss.bc.edu/timss1995i/TIMSSPDF/SRWhole.pdf>.
- Mullis, I.V.S., Martín, M. O., González, E. J. Gregory, K.D. Garden, R.A.. O'Connor, K.M., Chrostowski, S.J. y Smith, T.A. (2000). *TIMSS 1999 International Mathematics Report*. Chestnut Hill, MA: Boston Collage. [En red]. Disponible en: [http://timss.bc.edu/timss1999i/math\\_achievement\\_report.html](http://timss.bc.edu/timss1999i/math_achievement_report.html)
- Mullis, I V S., Martin, M O., Smith, T A., Garden, R A., Gregory, K D., Gonzales, E J., Chrostowski, S J. y O Conoor, K. (2002). *Marcos Teóricos y especificaciones de evaluación de TIMSS 2003*. Madrid: Ministerio de Educación cultura y deporte. Instituto Nacional de Calidad y Evaluación. [En red]. Disponible en: <http://www.ince.mec.es/pub/marcosteoricostimss2003.pdf>
- \_\_\_\_\_ (2003). *TIMSS Assessment Frameworks and specifications 2003*. IEA. ISC Boston College (2nd Edition). [En red]. Disponible en: [http://timss.bc.edu/timss2003i/PDF/t03\\_af\\_book.pdf](http://timss.bc.edu/timss2003i/PDF/t03_af_book.pdf)
- Muñoz, J. (1997). *Introducción a la teoría de respuesta a los ítems*. Madrid: Pirámide.



- Muñiz, J. (1998). *Teoría clásica de los test*. Madrid: Pirámide.
- Nickerson, R S., Perkins, D N. y Smith, E E. (1990). *Enseñar a pensar. Aspectos de la aptitud intelectual*, Barcelona, Paidós.
- 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, . [En red]. Disponible en: <http://www.ince.mec.es/pub/pisacomp.pdf> (Original publicado en 1999).
- OCDE (2001a). *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,. [En red]. Disponible en: <http://www.ince.mec.es/pub/pisa2000assessment.pdf> (Original publicado en 2000).
- OCDE (2001b). *Conocimientos y destrezas para la vida. Primeros resultados del Proyecto PISA 2000*. Resumen de resultados. (Versión española de Guillermo Gil Escudero) /OCDE. Madrid. Ministerio de Educación, Cultura y Deporte. INCE. [En red]. Disponible en: <http://www.ince.mec.es/pub/pisa2000-int.pdf>
- OCDE. (2002a). *Conocimientos y aptitudes para la vida. Resultados de PISA 2000*. (Traducido por Claudia Estéve) México: OCDE- Santillana. Aula XXI. (Original publicado en 2001).
- OCDE. (2002b). *Muestra de reactivos empleados en la evaluación PISA 2000. Aptitudes para la lectura, matemáticas y ciencias*. México: OCDE-Santillana Aula XXI.
- OCDE/UNESCO. (2003). *Aptitudes básicas para el mundo del mañana. Otros resultados del proyecto PISA 2000*. Resumen ejecutivo. OCDE/ Instituto de Estadística de la UNESCO. [En red]. Disponible en: [http://www.memfod.edu.uy/novedades/news\\_2003/pisa%20PLUS%202000%20-%20resumen%20ejecutivo.pdf](http://www.memfod.edu.uy/novedades/news_2003/pisa%20PLUS%202000%20-%20resumen%20ejecutivo.pdf)
- OCDE. (2004a). *Aprender para el Mundo de Mañana. Resumen de Resultados. PISA 2003*-OCDE/INCE.
- OCDE. (2004b). *Marcos teóricos de PISA 2003*: (Traducido por Encarnación Belmonte) Madrid: Ministerio de Educación y Ciencia. Instituto Nacional de Evaluación y Calidad del Sistema



- Educativo. [En red]. Disponible en: <http://www.ince.mec.es/pub/marcoteoricopisa2003.pdf>  
(Original publicado en 2003).
- OECD (2000a). *Measuring Student Knowledge and Skills - The PISA 2000 Assessment, of Reading, Mathematical, and Scientific, Literacy*, OECD, Paris, France.
- OECD (2002b). *Reading for Change. Performance and Engagement Across Countries. Results from PISA 2000* OECD publications. [En red]. Disponible en: <http://www1.oecd.org/publications/e-book/9602071E.PDF>
- OECD (2002c). *Sample tasks from the PISA 2000 Assessment- Reading, Mathematical and Scientific Literacy*. OCDE [En red]. Disponible en: <http://www1.oecd.org/publications/e-book/9602011E.PDF>
- OECD. (2003a). *The Pisa 2003 Assessment Framework – Mathematics, Reading, Science and Problem Solving and Skills* [En red]. Disponible en: <http://www1.oecd.org/publications/e-book/9603051E.PDF>
- OECD. (2003b). *Learners for Life: Student Approaches to Learning: Results from PISA 2000* OECD Paris France. [En red]. Disponible en: <http://www1.oecd.org/publications/e-book/9603101E.PDF>.
- OECD. (2003c). *Student Engagement at School: A Sense of Belonging and Participation: Results from PISA 2000* . OECD Paris, France. [En red]. Disponible en: <http://www1.oecd.org/publications/e-book/9603131E.PDF> .
- OECD- UNESCO (2003d). *Literacy Skills for the World of Tomorrow: Further Results from PISA 2000* OECD/UNESCO UNESCO Instituto for Statistic UIES. [En red]. Disponible en: <http://www1.oecd.org/publications/e-book/9603071E.PDF>
- Pajares, R., Sanz, A., y Rico, L. (2004). *Aproximación a un modelo de evaluación: el proyecto PISA 2000*. Instituto Nacional de Evaluación y Calidad. INECSE. Ministerio de Educación Cultura y Deporte. [En red]. Disponible en: <http://www.ince.mec.es/pub/aproxapisa2000.pdf>
- Pérez Gómez, A. I. (1998). *La cultura escolar en la sociedad neoliberal*. Madrid: Morata.
- Perrenoud, P. (1999). *Construir competencias desde la escuela*. Chile: Dolmen Ediciones.



- Plomp, T. (1992). "Conceptualización de un marco de investigación educativa comparada," en: *Perspectivas*. Vol.. XXII, No.3, (83) Santiago de Chile: OREALC. UNESCO pp. 297- 307.
- Plomp, T. (1998). "El potencial de los estudios internacionales comparados para el control de la calidad de la educación" *Perspectivas. Revista trimestral de educación*. Vol. XXVIII, No. 1, Santiago de Chile: OREALC UNESCO pp. 50-67.
- Poder Ejecutivo Federal.(1996). *Programa de Desarrollo Educativo 1995-2000*, México: Secretaría de Educación Pública.
- Poder Ejecutivo Federal. (2002). "Decreto por el que se crea el Instituto Nacional para la Evaluación de la Educación." En: *Diario Oficial de la Federación*, 8 de agosto, México. [En red]. Disponible en: [http://www.gobernacion.gob.mx/dof/2002/agosto/dof\\_08-08-2002.pdf](http://www.gobernacion.gob.mx/dof/2002/agosto/dof_08-08-2002.pdf)
- Posada, R. (2004). "Formación superior basada en competencias. Interdisciplinariedad y trabajo autónomo del estudiante." *Revista Iberoamericana de Educación*. OEI. [En red]. Disponible en: <http://www.campus-oei.org/revista/deloslectores/648Posada.PDF>
- Quezada, J. (2004). *Resultados de las Pruebas Nacionales de Aprovechamiento en Lectura y Matemáticas aplicadas al fin del ciclo 2002-2003* México: INEE [En red]. Disponible en: [http://capacitacion.ilce.edu.mx/inee/pdf/productos/informe\\_resultados\\_2002\\_2003.pdf](http://capacitacion.ilce.edu.mx/inee/pdf/productos/informe_resultados_2002_2003.pdf)
- Raths. L D. (1993). *Cómo enseñar a pensar. Teoría y práctica*. México, Paidós.
- Ravela, P. (2002). *¿Cómo presentan sus resultados los Sistemas Nacionales de Evaluación Educativa en América Latina?* PREAL. [En red]. Disponible en: [http://www.preal.org/docs-trabajo/ravela\\_n22.pdf](http://www.preal.org/docs-trabajo/ravela_n22.pdf) en febrero 2005.
- Rey, B. (2001). *Las competencias transversales en cuestión*. (Traducción de Alejandro Madrid Zan) Chile: Escuela de Filosofía de la Universidad de ARCIS Libro electrónico, [En red] Disponible en: <http://www.philosophia.cl/biblioteca/Rey/Competencias%20transversales.pdf>
- Roe, R. (2003). "¿Qué hace competente a un psicólogo?. *Papeles del Psicólogo. Revista del colegio oficial de psicólogos*. Diciembre. No. 86. [En red]. Disponible en: <http://www.cop.es/papeles/vernumero.asp?id=1108>



- Rojas C. y Esquivel J M. (1998). *Los Sistemas de Medición del Logro Académico en Latinoamérica*. The World Bank. [En red]. Disponible en:[http://wbln0018.worldbank.org/LAC/lacinfoclient.nsf/d29684951174975c85256735007fef12/8b08040037110db2852567ae00696787/\\$FILE/25.pdf](http://wbln0018.worldbank.org/LAC/lacinfoclient.nsf/d29684951174975c85256735007fef12/8b08040037110db2852567ae00696787/$FILE/25.pdf)
- Rosas, A., Martínez, C., Olguin, G., Basulto, M.A., Vilchis, L U., y Flores, R M. (1995). "Los profesores de la delegación mexicana nos informan sobre los objetivos del Programa TIMSS". *Educación Química*. Vol. 6, No. 4 México: Facultad de química del la UNAM. pp. 218 y 219.
- Rosales, C. (1990). "Introducción" y "Objeto y funciones". En: *Evaluar es reflexionar sobre la enseñanza* Madrid Narcea. pp. 13-18 30-34. (Educación Hoy Estudios).
- San Martín, V. (s.f.) "La formación en competencias, el desafío de la educación superior en iberoamérica." *Revista Iberoamericana de Educación*. OEI. [En red]. Disponible en: <http://www.campus-oei.org/revista/deloslectores/280SanMartin.PDF> Recuperado 12/9/2004
- Sánchez M. de A.. (1999). *Desarrollo de habilidades del pensamiento. Procesos básicos del pensamiento*. México Trillas:
- SEP. (1993). *Plan y Programa de estudio Educación Básica. Primaria*. México. SEP.
- \_\_\_\_ (1996a) *Evaluación de la Educación Primaria. Primer levantamiento de datos. Primaria Tercer Grado*. [En red]. Disponible en: [http://snee.sep.gob.mx/Documentos/Evep/1996/Prueb/terc\\_96.pdf](http://snee.sep.gob.mx/Documentos/Evep/1996/Prueb/terc_96.pdf)
- \_\_\_\_ (1996b). *Evaluación del aprovechamiento escolar*. Ciclo escolar 1999-2000 Primaria Tercer Grado. Forma 1. [En red]. Disponible en: [http://snee.sep.gob.mx/Pruebas/Aprovechamiento\\_Escolar/primaria/00/aep3\\_00g.pdf](http://snee.sep.gob.mx/Pruebas/Aprovechamiento_Escolar/primaria/00/aep3_00g.pdf)
- \_\_\_\_ (s.f.). *Evaluación del aprendizaje*. Informe de resultados. Tercero y cuarto levantamiento de datos. SEP, DGEEE, DGEPE. PAREB. COMPONENTE 11. [En red]. Disponible en: <http://www.sep.gob.mx/work/appsite/dge/index.htm>. Recuperado 17/10/ 2003:
- \_\_\_\_ (1997). *Elaboración de instrumentos de medición* .Antología Tomo II y III, México: SEP



- \_\_\_\_\_ (1999). *Nuestros nuevos alumnos en secundaria. IDANIS 99. Bases y criterios para la interpretación de resultados.* [En red]. Disponible en: <http://www.sep.gob.mx/work/appsite/dge/index.htm>
- \_\_\_\_\_ (2000a). *Estrategias y acciones en el marco del Sistema Nacional de Evaluación Educativa.* [En red]. Disponible en: <http://www.sesp.gob.mx/work/appsite/dge/index.htm>
- \_\_\_\_\_ (2000b). *Informe de labores 1999-2000.* México: SEP
- \_\_\_\_\_ (2002). *La experiencia de la Dirección General de Evaluación Educativa en la educación básica y normal. 30 años de medición de logro educativo.* [En red]. Disponible en: <http://www.sep.gob.mx/work/appsite/dge/index.htm> Recuperado 5 diciembre 2002
- \_\_\_\_\_ (s.f.). *Fichas técnicas de los programas nacionales de evaluación.* [En red]. Disponible en: [http://capacitación.ilce.edu.mx/inee/c\\_taller\\_1](http://capacitación.ilce.edu.mx/inee/c_taller_1) Recuperado: 9/5/2004.
- San Martín, V. (s.f.). "La formación en competencias: el desafío de la educación superior en iberoamericana". *Revista Iberoamericana de Educación*, OEI.
- Sepúlveda, L. (2001). *El Concepto de Competencias Laborales en Educación. Notas para un Ejercicio Crítico.* [En red]. Disponible en: <http://www.reduc.cl/reduc/sepulveda.pdf>
- Schmidt, W. H. (1992). "Análisis del currículo en el marco del TIMSS: descripción longitudinal de la enseñanza de las disciplinas" *Perspectivas Revista trimestral de educación comparada*. Vol. XXII, No. 3, 1992 (83) Santiago de Chile: OREALC UNESCO pp. 351-359.
- Schwartzman, S. (2001). *Las pruebas de medición del desempeño escolar en América Latina: Dificultades e inferencias inválidas* Tomado de: *El futuro de la educación en América Latina y el Caribe*. Santiago de Chile: UNESCO-SANTIAGO, pp.57-59. [En red]. Disponible en: <http://www.unesco.cl/promed17/elfuturo.doc>
- Simons, H. (1999). "Naturaleza y evolución de la investigación sobre evaluación" y "Las consecuencias políticas de la teoría de la evaluación: un examen más detenido". En: *Evaluación democrática en instituciones escolares*, Madrid: Morata. pp. 25- 133.
- Stevens (1959). *La medición y el hombre.* Suplemento del Seminario de Problemas científicos y filosóficos No. 19. México: UNAM.



- Stufflebeam, D. y Shinkfield, A. J. (1987). *Evaluación Sistemática. Guía teórica y práctica*. Barcelona; México: Paidós (Tr. Carlos Losilla).
- Thorndike, R. y Hagen, E. (1978). *Tests y técnicas de medición en psicología y educación: Elaboración, diseños, investigación, aplicación*. México: Trillas.
- Tiana A. (2000). *Cooperación internacional en evaluación de la educación en América Latina y el Caribe. Análisis de la situación y propuestas de actuación*. Washington, D.C. Banco Interamericano de Desarrollo. BID Serie de Informes Técnicos del Departamento de Desarrollo Sostenible. [En red]. Disponible en: <http://www.iadb.org/sds/doc/EDU%2D115S.pdf>
- Tirado, F. (1991). "Efectos de la escolaridad en las habilidades de reflexión" *Revista Latinoamericana de estudios educativos*. México, Vol. XXI, No. 3.
- Torney-Purta, J. y Orozco, L. (s.f.) *Ciudadanía y educación cívica en 28 Países: Resumen de los resultados del estudio comparado de la IEA*. E.U.A. Universidad de Maryland del College Park. [En red]. Disponible en: [http://www.wam.umd.edu/~iea/exec\\_summ/ExecutiveSummSpanish.htm](http://www.wam.umd.edu/~iea/exec_summ/ExecutiveSummSpanish.htm) Recuperado 4/07/2003.
- Tristán, A. (1998). *Análisis de Rasch para todos. Una guía simplificada para evaluadores educativos*, México: CENEVAL.
- Tristán, A. (2004). "Sistema para calificación de pruebas referidas a criterio y definición de estándares". En: Castañeda, S., (coord.). *Educación, aprendizaje y cognición. Teoría en la práctica*. México: Ed. El manual moderno. Universidad de Guadalajara. Instituto de Neurociencias. UNAM, Facultad de Psicología.
- Tyler, R. (1973). "Cómo evaluar la eficacia de las actividades de aprendizaje, en *Principios Básicos del Currículo*. Argentina: Troquel.
- 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. [En red]. Disponible en: <http://www.unesco.cl/promedi7/balance.pdf>



- UNESCO/OREALC (1981-2000). *Proyecto Principal de Educación en América Latina y el Caribe. Sus objetivos, características y modalidades de acción*. Documento electrónico [En red]. Disponible en: <http://www.unesco.cl/03prelac.htm>
- Velásquez, V. M. (1996). "La evaluación como recurso para elevar la calidad de la educación en México" *Revista Iberoamericana de Educación Organización de Estados Iberoamericanos para la educación, la ciencia y la cultura*. [En red]. Disponible en: <http://www.campus-oei.org/oeivirt/rie10a08.htm> .
- Weinstein y Mayer ( 1986). *Caracterización de las estrategias de aprendizaje*. En. Witrock, M. y Baker, E. (1998). *Test y Cognición. Investigación cognitiva y mejora de las pruebas psicológicas*. España. Paidós.
- Vidal, R., Díaz, M. A. y Loyola del Río, J. (2003). *El Proyecto PISA. Su aplicación en México*. Recuperado en octubre 2004, en: <http://capacitacion.ilce.edu.mx/inee/pdf/Cuaderno9.pdf> .
- Vidal, R.; Díaz, M A. y Loyola del Río, J G. ( 2004). *Resultados de las pruebas PISA 2000 y 2003 en México. Habilidades para la vida en estudiantes de 15 años*. INNE. Recuperado en febrero 2005 <http://capacitacion.ilce.edu.mx/inee/pdf/PISA/InformePisaMexico2003.pdf>.
- VV. A. A. (2001). *Establecimiento de estándares para la educación primaria*. OEI. [En red]. Disponible en: <http://www.campus-oei.org/estándares/centrometo.htm>.
- Weikart, D. P. y Barriao, J. (1992). "El proyecto de la IEA sobre educación preescolar: estudios preliminares en Portugal y en China" *Perspectivas Revista trimestral de educación*. Vol. XXII, No. 3, 1992 (83) Santiago de Chile: OREALC UNESCO, p.506-512.
- Wiley, D. Y Wolfe R:G. (1992). "Problemas que plantea la concepción del Tercer Estudio Internacional de la IEA sobre Matemáticas y Ciencias". *Perspectivas Revista trimestral de educación*. Vol. XXII, No. 3, 1992 (83) ) Santiago de Chile OREALC UNESCO pp. 319-339
- Wilson, D. ( 2003). "The future of comparative and international education in a globalised World." En: *International Review of education*. Special Issue: Comparative Education: Continuing Traditions, New Challenges, and New Paradigms. Volume 49 Issues 1 and 2. Netherlands: UNESCO Institute for Education and Kluwer Academia Publishers. pp. 17 -18.



- Wittrock, M. y Baker, E. (1998). *Test y cognición. Investigación cognitiva y mejora de las pruebas psicológicas*. España. Paidós.
- Wolf, L. (1999). *Imágenes de Enseñanza: El proyecto de grabaciones del TIMSS*. BID. [En Red].  
Disponible en: <http://www.iadb.org/sds/doc/25EduTecn.pdf>.
- Woolfolk, A.. (1996). "Evaluación y Calificación de la clase." En: *Psicología Educativa México*: Prentice Hall.