

Los Resultados de las Pruebas PISA

Elementos para su Interpretación

Felipe Martínez Rizo

CUADERNO No. 1



Instituto Nacional para la
Evaluación de la Educación

COLECCIÓN CUADERNOS
DE INVESTIGACIÓN

ISSN 1665-9457

Los Resultados de las Pruebas PISA

Elementos para su Interpretación

Felipe Martínez Rizo*

CUADERNO No. 1

*Director General del Instituto Nacional para la Evaluación de la Educación

MÉXICO, JULIO, 2003

CONTENIDO

Introducción	3
Características de las evaluaciones de PISA	4
Primera aproximación a los resultados	5
Interpretación de los resultados	6
Multidimensionalidad de las pruebas	7
Traslapes de las puntuaciones	10
Límites de las interpretaciones en términos de <i>rankings</i>	11
De la interpretación a la explicación	14
Explicaciones por factores del entorno y la escuela	18
Resultados dentro de los países y equidad	19
Comparaciones con otros países	20
En resumen	21
Conclusiones	22
Bibliografía	24

INTRODUCCIÓN

La Organización para la Cooperación y el Desarrollo Económico (OCDE) agrupó en sus inicios a los países más industrializados del mundo, en particular Estados Unidos, varias naciones de Europa Occidental y Japón. Posteriormente su membresía se ha ampliado, incorporando a países con niveles intermedios de desarrollo, como México.

El trabajo de la OCDE incluye la recopilación de estadísticas y la realización de estudios sobre temas relacionados con el desarrollo económico. Los temas educativos han sido atendidos con especial cuidado, por la importancia de la educación para la preparación de una fuerza laboral calificada y de una ciudadanía preparada para la democracia.

Por ello, la OCDE ha desarrollado un sistema internacional de indicadores educativos, del que se deriva la conocida publicación anual llamada *Education at a Glance*. Desde los inicios de esos esfuerzos podía contarse con datos razonablemente confiables sobre el número de escuelas, maestros y alumnos, la cobertura de los sistemas y su eficiencia terminal, el gasto educativo y otros aspectos cubiertos por las estadísticas educativas nacionales. En cambio, la ausencia de instrumentos equiparables de evaluación de los aprendizajes impedía que se tuviera información comparable sobre el nivel de conocimientos y habilidades alcanzado por los alumnos.

De la misma manera, como complemento de su sistema de indicadores, la OCDE decidió poner en marcha un programa internacional de evaluación de aprendizajes, que se conoce ahora como PISA, por sus siglas en inglés: *Programme for International Student Assessment* (Programa para la Evaluación Internacional de Estudiantes) en el que México participa, como miembro de la Organización.

La primera aplicación de las pruebas PISA realizó en el año 2000, en 32 países: 28 de la OCDE, incluyendo a México, y 4 más: Rusia, Letonia, Liechtenstein y Brasil. Los resultados se difundieron a fines de 2001. En 2003 se aplicó la segunda ronda de pruebas, cuyos resultados se darán a conocer a fines de 2004. Once países que no participaron en el estudio de 2000 se sumaron al programa, aplicando en 2002 las pruebas utilizadas dos años antes en los 32 países iniciales. Los nuevos fueron Albania, Argentina, Bulgaria, Chile, Indonesia, Israel, Hong Kong, Macedonia, Perú, Rumania y Tailandia. Por razones técnicas, los resultados de Holanda y Rumania no se incluyeron en los reportes, por lo que los países de los que se tienen resultados son, finalmente, 41.

La aplicación complementaria se conoce como PISA PLUS; sus resultados son los que la OCDE dio a conocer el 1 de julio de 2003. Como las pruebas de la ronda PISA PLUS fueron

las mismas de PISA 2000, los resultados permiten comparar, con las salvedades que luego se comentarán, a los 41 países que participaron en una u otra aplicación.

Dada la importancia de la educación para el desarrollo nacional, los resultados de estas evaluaciones son de gran interés para los maestros, los padres de familia, las autoridades educativas y, en general, para toda la sociedad mexicana.

Para contribuir a su difusión y a la adecuada interpretación de sus resultados, de manera que resulten provechosos para el mejoramiento del Sistema Educativo Nacional, el Instituto Nacional para la Evaluación de la Educación difunde este resumen, con una primera aproximación al tema.

CARACTERÍSTICAS DE LAS EVALUACIONES DE PISA

Población

Las pruebas se aplican a jóvenes de 15 años de edad que estén inscritos en cualquier grado y modalidad de la enseñanza secundaria o media superior.

Los sistemas educativos difieren en aspectos como cobertura del nivel preescolar, edad de inicio y fin de la enseñanza obligatoria, duración de la enseñanza común a todos los alumnos o el inicio de variantes educativas de orientación especializada y carácter más o menos selectivo. Buscando la mayor comparabilidad, la población objetivo de las pruebas PISA es la formada por todos los jóvenes de una edad que coincide, en muchos lugares, con el fin de la escolaridad obligatoria; de esta forma es posible apreciar la medida en que una sociedad prepara a la siguiente generación de sus ciudadanos.

Muestra

Las pruebas PISA se aplican a muestras representativas, de un tamaño que va de 4 mil 500 a 10 mil sujetos en cada país. Este tamaño de muestra permite hacer inferencias para el país en conjunto, pero no para regiones o entidades en su interior; para ello es necesario aplicar las pruebas a una muestra mucho mayor. Esto es lo que se hizo en México en el estudio de PISA 2003, que consideró cerca de 40 mil alumnos.

Contenidos

PISA es un programa permanente, que evaluará cada tres años los conocimientos y habilidades de los sustentantes en comprensión lectora, matemáticas y ciencias naturales, destacando uno de estos tres campos en cada aplicación. PISA 2000, al igual que PISA PLUS, enfatizó el área de lectoescritura; PISA 2003 puso el acento en matemáticas; y PISA

2006 lo pondrá en ciencias.

Las pruebas PISA no se limitan a los contenidos curriculares; incluyen conocimientos avanzados y habilidades complejas, como se requiere para una vida adulta bien integrada en las sociedades contemporáneas.

Instrumentos y procedimientos

Se trata de pruebas estandarizadas de papel y lápiz, que combinan reactivos de respuesta estructurada y abierta. Las pruebas son diseñadas y aplicadas, y sus resultados analizados, con los más altos estándares de calidad, por un equipo internacional. Cada país traduce las pruebas a su lengua nacional y la traducción es verificada y validada por la OCDE.

Los alumnos requieren alrededor de dos horas para responder la prueba; adicionalmente se distribuyen cuestionarios para obtener información sobre la escuela y el contexto, a los directores y a los propios alumnos.

La OCDE selecciona aleatoriamente las escuelas en que se aplican las pruebas, con escuelas de reemplazo para los casos en que no fuera posible la aplicación en alguna de las originales, lo que no puede rebasar un porcentaje estipulado. La aplicación es organizada por cada país y supervisada por la OCDE, al igual que la calificación de las respuestas y el análisis de los resultados.

Primera aproximación a los resultados

En orden alfabético de países, se presenta el promedio y la desviación estandar de los puntajes alcanzados por los alumnos en lectoescritura, matemáticas y ciencias. Las escalas tienen una media de 500 y una desviación estandar de 100

PAÍSES	LECTURA		MATEMÁTICAS		CIENCIAS	
	media	desv.	media	desv.	media	desv.
01 ALBANIA	349	99	381	107	376	94
02 ALEMANIA	484	111	490	103	487	102
03 ARGENTINA	418	109	388	120	396	109
04 AUSTRALIA	528	102	533	90	528	94
05 AUSTRIA	507	93	515	92	519	91
06 BÉLGICA	507	107	520	106	496	111
07 BRASIL	396	86	334	97	375	90
08 BULGARIA	430	102	430	110	448	96
09 CANADÁ	534	95	533	85	529	89
10 CHILE	410	90	384	94	415	95
11 COREA	525	70	547	84	552	81
12 DINAMARCA	497	98	514	87	481	103
13 ESPAÑA	493	85	476	91	491	95
14 ESTADOS UNIDOS	504	105	493	98	499	101
15 FINLANDIA	546	89	536	80	538	86
16 FRANCIA	505	92	517	89	500	102
17 GRECIA	474	97	447	108	461	97
18 HONG KONG	525	84	560	94	541	85
19 HUNGRÍA	480	94	488	98	496	103
20 INDONESIA	371	72	367	85	393	75
21 IRLANDA	527	94	503	84	513	92
22 ITALIA	487	91	457	90	478	98
23 ISLANDIA	507	92	514	85	496	88
24 ISRAEL	452	109	433	131	434	125
25 JAPÓN	522	86	557	87	550	90
26 LETONIA	458	102	463	103	460	98
27 LIECHTENSTEIN	483	96	514	96	476	94
28 LUXEMBURGO	441	100	446	93	443	96
29 MACEDONIA	373	94	381	98	401	83
30 MÉXICO	422	86	387	83	422	77
31 NORUEGA	505	104	499	92	500	96
32 NUEVA ZELANDA	529	108	537	99	528	101
33 PERÚ	327	96	292	108	333	90
34 POLONIA	479	100	470	103	483	97
35 PORTUGAL	470	97	454	91	459	89
36 REINO UNIDO	523	100	529	92	532	98
37 REPÚBLICA CHECA	492	96	498	96	511	94
38 RUSIA	462	92	478	104	460	99
39 SUECIA	516	92	510	93	512	93
40 SUIZA	494	102	529	100	496	100
41 TAILANDIA	431	77	432	83	436	77

INTERPRETACIÓN DE LOS RESULTADOS

Una manera simple de dar sentido a los resultados anteriores es ordenar los países según sus resultados y no alfabéticamente. Esto es sencillo y, a primera vista, parece que ofrece una visión clara e indiscutible de la calidad de los sistemas educativos de los países que participaron en las pruebas. En realidad, aunque las pruebas PISA sean buenas y sus resultados confiables, un análisis tan elemental es engañoso, por varias razones.

Desde un punto de vista técnico, debemos considerar por lo menos dos: la naturaleza multidimensional de las pruebas y el considerable traslape de los resultados obtenidos por los alumnos de muchos países.

Otras razones para tener cuidado en la interpretación son las que se refieren a la cobertura de la muestra respecto a la población objetivo y a la influencia diferencial del contexto socioeconómico y la escuela sobre los resultados. La abundante y valiosa información que se obtiene mediante los cuestionarios que se aplican junto con las pruebas PISA a los alumnos, maestros y directivos escolares, permite, en efecto, análisis mucho más finos de los resultados. Estos aspectos se analizarán posteriormente; por ahora centremos la atención en las dos aspectos técnicos mencionados antes, como razones para tener cuidado con las interpretaciones simples: la multidimensionalidad de las pruebas y el traslape de los resultados.

MULTIDIMENSIONALIDAD DE LAS PRUEBAS

Las pruebas PISA comprenden, como se ha dicho, las áreas de lectura, matemáticas y ciencias. Aunque los resultados de cada país en las tres áreas de la prueba se sitúan, por lo general, en un rango cercano, la coincidencia no es perfecta, por lo que los puntajes promedio de cada área dan lugar a una clasificación diferente.

Pero la multidimensionalidad de las pruebas llega más lejos. Dentro de una misma área del conocimiento, las pruebas de PISA comprenden aspectos varios, cada uno de los cuales puede dar lugar también a una clasificación diferente.

Como se ha explicado, las pruebas PISA 2000 y PLUS cubrieron con más profundidad el área de comprensión lectora. Además de la escala general de comprensión lectora, las pruebas permiten construir tres escalas particulares que se refieren, respectivamente, a la habilidad de localizar información (lo más sencillo); interpretar textos (nivel intermedio de complejidad); y reflexionar sobre el contenido de la lectura y valorarlo (lo más complejo).

Las tablas siguientes presentan las diversas clasificaciones que pueden hacerse de los países participantes, según sus resultados en las tres áreas de conocimiento que evaluaron las pruebas o en las dimensiones de la prueba de comprensión lectora.

Es obvio que el lugar que ocupa un país en la tabla de los participantes en las pruebas PISA no es idéntico en todas las escalas. En la escala de lectura, México, por ejemplo, ocupa el lugar 34 y Argentina el 35; en ciencias, México mantiene el lugar 34 y Argentina cae al 37, pero en matemáticas la situación se invierte, con Argentina en el 34 y México en el 35. Si se revisan las subescalas de lectura, México mantiene el lugar 34 en la escala de interpretación de textos, baja al 35 en la de localización y alcanza el 31 en la de reflexión, mientras que Argentina lo supera en localización (33) y queda por debajo en las otras dos: 36 en interpretación y 35 en reflexión.

Clasificación de países en las tres áreas de PISA

Lugar	Lectura		Matemáticas		Ciencias	
	País	Media	País	Media	País	Media
01	Finlandia	546	Hong Kong	560	Corea	552
02	Canadá	534	Japón	557	Japón	550
03	N. Zelanda	529	Corea	547	Hong Kong	541
04	Australia	528	N. Zelanda	537	Finlandia	538
05	Irlanda	527	Finlandia	536	Reino Unido	532
06	Hong Kong	525	Australia	533	Canadá	529
07	Corea	525	Canadá	533	N. Zelanda	528
08	Reino Unido	523	Suiza	529	Australia	528
09	Japón	522	Reino Unido	529	Austria	519
10	Suecia	516	Bélgica	520	Irlanda	513
11	Austria	507	Francia	517	Suecia	512
12	Bélgica	507	Austria	515	Rep. Checa	511
13	Islandia	507	Dinamarca	514	Francia	500
14	Noruega	505	Islandia	514	Noruega	500
15	Francia	505	Liechtenstein	514	USA	499
16	USA	504	Suecia	510	Hungría	496
17	Dinamarca	497	Irlanda	503	Islandia	496
18	Suiza	494	Noruega	499	Bélgica	496
19	España	493	Rep. Checa	498	Suiza	496
20	Rep. Checa	492	USA	493	España	491
21	Italia	487	Alemania	490	Alemania	487
22	Alemania	484	Hungría	488	Polonia	483
23	Liechtenstein	483	Rusia	478	Dinamarca	481
24	Hungría	480	España	476	Italia	478
25	Polonia	479	Polonia	470	Liechtenstein	476
26	Grecia	474	Letonia	463	Grecia	461
27	Portugal	470	Italia	457	Rusia	460
28	Rusia	462	Portugal	454	Letonia	460
29	Letonia	458	Grecia	447	Portugal	459
30	Israel	452	Luxemburgo	446	Bulgaria	448
31	Luxemburgo	441	Israel	433	Luxemburgo	443
32	Tailandia	431	Tailandia	432	Tailandia	436
33	Bulgaria	430	Bulgaria	430	Israel	434
34	México	422	Argentina	388	México	422
35	Argentina	418	México	387	Chile	415
36	Chile	410	Chile	384	Macedonia	401
37	Brasil	396	Albania	381	Argentina	396
38	Macedonia	373	Macedonia	381	Indonesia	393
39	Indonesia	371	Indonesia	367	Albania	376
40	Albania	349	Brasil	334	Brasil	375
41	Perú	327	Perú	292	Perú	333

Clasificación de países en escalas de comprensión lectora

	Escala general		Localización		Interpretación		Reflexión	
01	Finlandia	546	Finlandia	556	Finlandia	555	Canadá	542
02	Canadá	534	Australia	536	Canadá	532	R. Unido	539
03	N. Zelanda	529	N. Zelanda	535	Australia	527	Hong Kong	538
04	Australia	528	Corea	530	N. Zelanda	526	Finlandia	533
05	Irlanda	527	Canadá	530	Irlanda	526	Irlanda	533
06	Hong Kong	525	Japón	526	Corea	525	Japón	530
07	Corea	525	Irlanda	524	Hong Kong	522	N. Zelanda	529
08	R. Unido	523	R. Unido	523	Suecia	522	Australia	526
09	Japón	522	Hong Kong	522	Japón	518	Corea	526
10	Suecia	516	Suecia	516	Islandia	514	Austria	512
11	Austria	507	Francia	515	R. Unido	514	Suecia	510
12	Bélgica	507	Bélgica	515	Bélgica	512	USA	507
13	Islandia	507	Noruega	505	Austria	508	España	506
14	Noruega	505	Austria	502	Francia	506	Noruega	506
15	Francia	505	Islandia	500	Noruega	505	Islandia	501
16	USA	504	USA	499	USA	505	Dinamarca	500
17	Dinamarca	497	Dinamarca	498	Rep. Checa	500	Bélgica	497
18	Suiza	494	Suiza	498	Suiza	496	Francia	496
19	España	493	Liechtenstein	492	Dinamarca	494	Grecia	495
20	Rep. Checa	492	Italia	488	España	491	Suiza	488
21	Italia	487	España	483	Italia	489	Rep. Checa	485
22	Alemania	484	Alemania	483	Alemania	488	Italia	483
23	Liechtenstein	483	Rep. Checa	481	Liechtenstein	484	Hungría	481
24	Hungría	480	Hungría	478	Polonia	482	Portugal	480
25	Polonia	479	Polonia	475	Hungría	480	Alemania	478
26	Grecia	474	Portugal	455	Grecia	475	Polonia	477
27	Portugal	470	Letonia	451	Portugal	473	Liechtenstein	468
28	Rusia	462	Rusia	451	Rusia	468	Israel	467
29	Letonia	458	Grecia	450	Letonia	459	Letonia	458
30	Israel	452	Luxemburgo	433	Israel	458	Rusia	455
31	Luxemburgo	441	Israel	431	Luxemburgo	446	México	446
32	Tailandia	431	Bulgaria	422	Tailandia	439	Luxemburgo	442
33	Bulgaria	430	Argentina	407	Bulgaria	434	Tailandia	439
34	México	422	Tailandia	406	México	419	Bulgaria	431
35	Argentina	418	México	402	Chile	419	Argentina	430
36	Chile	410	Chile	383	Argentina	415	Brasil	417
37	Brasil	396	Brasil	365	Brasil	400	Chile	412
38	Macedonia	373	Macedonia	362	Macedonia	381	Indonesia	378
39	Indonesia	371	Indonesia	350	Indonesia	375	Macedonia	360
40	Albania	349	Albania	336	Albania	352	Albania	350
41	Perú	327	Perú	289	Perú	342	Perú	323

TRASLAPES DE LAS PUNTUACIONES

Es importante tener presente que los resultados de las pruebas PISA no son el promedio de los puntajes de todos los sujetos de la población objetivo –en este caso los jóvenes de 15 años de cierto país– sino sólo de los obtenidos por los integrantes de una muestra. Por ello, inevitablemente, existe un margen de error. Un buen diseño muestral permite estimar con precisión ese margen de error, con una probabilidad determinada, lo que se hace en el caso de las pruebas PISA.

Cuando la distancia que separa los promedios obtenidos por los alumnos de dos países es menor al margen de error que se maneja con cierto grado de probabilidad, no se puede afirmar sin más que el resultado de quienes tienen el puntaje superior sea efectivamente mejor que el de los otros.

En esos casos, los países en cuestión deberán considerarse técnicamente empatados: es posible que el nivel del país con el puntaje superior, sea mejor que el del que tiene un puntaje menor, pero estrictamente hablando también es posible que ocurra todo lo contrario, o que el nivel de los dos países sea estrictamente igual.

En el caso de la aplicación de PISA 2000, el margen de error de los resultados obtenidos en la escala de competencias de lectura, hace que la clasificación resultante de los países pudiera variar de manera importante. En las dos columnas numéricas de la tabla siguiente se presenta el lugar más alto y el más bajo que podrían ocupar los países.

La tabla muestra que Finlandia, que ocupa el primer lugar, y los tres últimos países en el orden (Luxemburgo, México y Brasil) tienen puntajes tales, que la posibilidad de que el verdadero lugar que ocupen sea diferente al que se les asigna puede descartarse con 95 por ciento de confianza. En todos los demás casos no ocurre así: Canadá, que se ubica en segundo lugar, podría también ser tercero o cuarto, situándose detrás de Nueva Zelanda y Australia, en lugar de precederlos, pero Nueva Zelanda podría caer hasta el octavo lugar y Australia hasta el noveno, según la misma lógica. En el caso más llamativo, los Estados Unidos, que se ubican en el lugar 16°, podrían situarse realmente entre el 10° y el 20°.

Rango de las posiciones que podrían ocupar los países de la aplicación PISA 2000, con 95 por ciento de confianza

País y puntaje		Posición del país	
		La más alta	La más baja
Finlandia	546	1	1
Canadá	534	2	4
Nueva Zelanda	529	2	8
Australia	528	2	9
Irlanda	527	3	9
Corea	525	4	9
Reino Unido	523	5	9
Japón	522	3	10
Suecia	516	9	11
Austria	507	11	16
Bélgica	507	11	16
Islandia	507	11	15
Noruega	505	11	16
Francia	505	11	16
Estados Unidos	504	10	20
Dinamarca	497	16	19
Suiza	494	16	21
España	493	17	21
República Checa	492	17	21
Italia	487	19	24
Alemania	484	21	25
Liechtenstein	483	20	26
Hungría	480	21	26
Polonia	479	21	27
Grecia	474	23	28
Portugal	470	24	28
Rusia	462	27	29
Letonia	458	27	29
Luxemburgo	441	30	30
México	422	31	31
Brasil	396	32	32

LÍMITES DE LAS INTERPRETACIONES EN TÉRMINOS DE RANKINGS

El carácter multidimensional de las pruebas de aprendizaje de PISA, junto con los traslapes de las puntuaciones, hacen claramente inadecuadas las interpretaciones de los resultados en términos del lugar que ocupen los países en un ordenamiento de los puntajes obtenidos. Cuando se difundieron los resultados de PISA 2000, algunos resumieron una versión derrotista de los de México en los siguientes términos:

México, penúltimo lugar del mundo

Si se tiene en cuenta que en PISA 2000 sólo participaron 32 países de los más de 200 que hay en el mundo, es obvio el endeble sustento de tal interpretación. Los resultados de México en PISA 2000 podrían haberse interpretado también en forma diametralmente opuesta, triunfalista con afirmaciones como estas:

México, primer lugar de América Latina Aplasta México a Brasil

Estas formas alternas de expresar los resultados de México suenan absurdas, pero en sentido estricto son soportadas por los datos, al menos en una medida similar a la que puede aducirse en apoyo de la versión derrotista. Ambas versiones tienen un soporte muy endeble y no reflejan la complejidad de la realidad. Ninguna puede servir para que se tomen decisiones positivas para el mejoramiento de la educación. Los resultados de PISA PLUS podrían leerse también en formas extremas, como sigue:

Cae México del lugar 31 al 34 México por encima de Argentina y Chile

El sustento de estas nuevas versiones no es mejor que el de las anteriores. Para llegar a interpretaciones más adecuadas de los resultados de pruebas como las de PISA habrá que emplear acercamientos diferentes; en las páginas siguientes se presentan algunos de ellos.

Proporción de buenos y malos lectores

Los resultados de las pruebas de PISA permiten situar a cada sustentante en un nivel de competencia determinado, según una escala de cinco niveles en cada área y subárea. Los niveles de la escala de lectura, de inferior a superior, significan que el alumno puede:

Nivel

- 1 Localizar una información puntual, captar la idea principal de un texto...
- 2 Localizar información más compleja pero explícita, hacer inferencias simples...
- 3 Integrar información dispersa, captar relaciones entre partes diferentes...
- 4 Localizar información implícita, captar matices finos, evaluar el texto...
- 5 Dominio sofisticado de la lectura, evaluación crítica, manejo de hipótesis...

Una forma de aproximarse a los resultados de PISA es comparando las proporciones de sustentantes que se sitúen en cada nivel de competencia de una escala. Esto es lo que se intenta a continuación, utilizando la escala combinada de lectura, dado que esa fue el área principal evaluada en las aplicaciones de PISA 2000 y PISA PLUS.

Para simplificar la interpretación, agrupamos los niveles de competencia de la manera siguiente: consideramos buenos lectores a los sustentantes que se ubiquen en los niveles 5 y 4 de la escala; lectores regulares a los que estén en los niveles 3 y 2; y malos lectores a los que se sitúen en el nivel inferior de competencia, el 1, o por debajo de él.

Si comenzamos observando la proporción de lectores buenos, regulares y malos que hay entre los alumnos de 15 años de edad del sistema educativo mexicano, veremos que sólo poco menos del 7% (6.9) pueden definirse como buenos lectores, en tanto que 44.2% caen en la categoría de malos lectores, y el 49.1 restante en el nivel intermedio.

La cifra mexicana no es, desde luego, halagadora. Sin embargo la comparación con otros países ofrece referentes interesantes para darle sentido. Observamos así que en Argentina, pese a que el país ocupa el lugar inferior siguiente en la clasificación de PISA, hay más buenos lectores que en México. La explicación radica en el hecho de que también hay más lectores muy deficientes: 22.6% se ubican por debajo del nivel 1, contra 16.1% en México.

Los países que ocupan los siguientes lugares en la lista de PISA, desde el 36 (Chile) hasta el 41 (Perú) tienen cada vez menos buenos lectores y cada vez más lectores deficientes. Sólo 5.3% de los primeros frente a 48.2% de los segundos en Chile; 3.7% vs. 55.8% en Brasil; y 1.1% vs. 79.6% en Perú.

En el otro extremo de la tabla, la proporción de buenos lectores sube a medida que se asciende en los lugares: poco más de 10% en Bulgaria y Luxemburgo; alrededor de 20% en Israel, Letonia, Rusia, Portugal y Grecia; cerca de 25% en Italia y España; 28% en Alemania; algo más de 30% en Estados Unidos y Francia; más del 35% en Suecia y Corea; alrededor del 40% en Japón, el Reino Unido, Hong Kong, Irlanda y Australia; casi 45% en Nueva Zelanda y Canadá; y 50.1% en Finlandia.

El hecho de que en el país mejor ubicado en la escala de lectura, solamente la mitad de los jóvenes alcance el nivel de buen lector, muestra la dificultad de llegar a ese nivel de competencia, y obliga a matizar el severo juicio que los resultados de México y otros países provocan en una primera lectura.

Distribución de los alumnos por niveles de dominio en porcentaje

Lugar	País	Buenos lectores Niveles 4 y 5	Lectores regulares Niveles 2 y 3	Malos lectores Niveles 1 y menos
34	México	6.9	49.1	44.2
35	Argentina	10.3	45.8	43.9
36	Chile	5.3	46.6	48.2
37	Brasil	3.7	40.6	55.8
38	Macedonia	1.9	35.5	62.6
39	Indonesia	0.4	30.9	68.7
40	Albania	1.4	28.3	70.3
41	Perú	1.1	19.4	79.6
33	Bulgaria	11.2	48.5	40.3
32	Tailandia	5.3	57.6	37.0
31	Luxemburgo	12.9	52.1	35.1
30	Israel	18.8	48.1	33.2
29	Letonia	17.9	51.5	30.6
28	Rusia	17.4	56.1	27.5
27	Portugal	21.0	52.8	26.3
26	Grecia	21.7	54.0	24.4
22	Alemania	28.2	49.1	22.6
21	Italia	24.8	56.2	18.9
19	España	25.3	58.5	16.3
16	USA	33.7	48.4	17.9
15	Francia	32.2	52.6	15.2
10	Suecia	36.8	50.7	12.6
9	Japón	38.7	51.3	10.0
8	Reino Unido	40.0	47.1	12.8
7	Corea	36.8	57.4	5.7
6	Hong Kong	40.8	50.2	9.1
5	Irlanda	41.3	47.6	11.0
4	Australia	42.9	44.7	12.4
3	N. Zelanda	44.5	41.8	13.7
2	Canadá	44.5	46.0	9.6
1	Finlandia	50.1	43.0	6.9

DE LA INTERPRETACIÓN A LA EXPLICACIÓN

Sobre la interpretación

Toda evaluación consiste finalmente en un juicio que contrasta dos elementos: un dato empírico que refleja, con más o menos precisión, la situación de una realidad; y un referente con el que se compara el dato empírico. La calidad de una evaluación, en consecuencia,

depende en parte de la calidad de los datos empíricos, pero también de la pertinencia del referente seleccionado para la comparación. Para que los datos empíricos –los resultados mismos de las pruebas– cobren sentido, es indispensable que se escojan puntos de referencia adecuados.

En temas complejos, multidimensionales, como los educativos, los referentes deberán ser varios, ya que ninguno será adecuado considerado en forma aislada. Por ello, no es irrelevante compararse con los países más desarrollados como referentes lejanos, en cierta forma ideales; pero es necesario complementar esa comparación con otras, por ejemplo con países similares, con la propia situación en el pasado y con las metas que se hayan fijado para el futuro.

Sobre la explicación

La evaluación de la calidad educativa es un medio, no un fin: cobra sentido en la medida en que contribuye a que dicha calidad mejore. Por lo mismo, además de interpretar adecuadamente los resultados de evaluación es necesario ir más allá: buscar las causas de las situaciones detectadas. Para diseñar e instrumentar programas de mejoramiento educativo es indispensable, en efecto, identificar los factores que inciden en la calidad y entender la forma en que lo hacen. De otra manera, evaluar es una tarea inútil, que puede llevar a la frustración.

Los cuestionarios de contexto que se aplican con las pruebas PISA ofrecen una rica gama de informaciones, que pueden dar lugar a acercamientos explicativos. Esto deberá hacerse en análisis posteriores. Por el momento nos limitaremos a una primera aproximación sobre los dos grandes rubros en que suelen agruparse los factores que inciden en la calidad educativa: los factores del entorno social y los de la escuela misma.

El cuadro siguiente presenta la relación de países participantes (con excepción de Liechtenstein y Francia, excluidos por razones técnicas) ordenados según la varianza total de los resultados obtenidos por sus alumnos: el país con mayor igualdad es Corea y el más desigual Alemania. Las dos columnas restantes indican, en porcentaje, la proporción de las diferencias atribuible a factores escolares (varianza entre escuelas) y la atribuible a factores extraescolares (varianza total menos entre escuelas).

Un cuadro más ofrece información sobre dos indicadores importantes de los dos grupos básicos de factores del rendimiento escolar: el ingreso per cápita (expresado en dólares controlados según la paridad del poder adquisitivo, ppa ó ppp) es un indicador grueso de la riqueza o pobreza de un país; y el gasto educativo por alumno lo es de la mayor o menor abundancia de recursos de un sistema educativo. Las gráficas subsecuentes muestran la relación de ambos indicadores con el desempeño de los alumnos en las pruebas PISA.

Varianza total de los resultados. Porcentaje de varianza entre escuelas y otra

Países	Varianza total	% de varianza entre escuelas	% de varianza restante
Corea	4,833	37.4	62.6
Indonesia	5,184	42.7	57.3
Tailandia	5,929	30.9	69.1
Hong Kong	7,056	48.3	51.7
España	7,181	20.7	79.3
Japón	7,358	45.4	54.6
México	7,370	53.4	46.6
Brasil	7,427	43.1	56.9
Finlandia	7,994	12.3	87.7
Chile	8,100	56.7	43.3
Italia	8,356	54.0	46.0
Rusia	8,466	37.1	62.9
Suecia	8,495	9.7	90.3
Islandia	8,529	7.6	92.4
Austria	8,649	60.0	40.0
Irlanda	8,755	17.8	82.2
Hungría	8,810	67.2	32.8
Macedonia	8,836	44.9	55.1
Canadá	8,955	17.6	82.4
Perú	9,216	58.0	42.0
Promedio	9,277	35.2	64.8
República Checa	9,278	53.4	46.6
Grecia	9,436	50.4	49.6
Portugal	9,436	36.8	63.2
Dinamarca	9,614	18.6	81.4
Albania	9,801	40.9	59.1
Polonia	9,958	63.2	36.8
Luxemburgo	10,088	30.8	69.2
Inglaterra	10,098	21.4	78.6
Australia	10,357	18.8	81.2
Bulgaria	10,404	55.4	44.6
Suiza	10,408	43.4	56.6
Letonia	10,435	31.2	68.8
Noruega	10,743	10.9	89.1
Estados Unidos	10,979	29.6	70.4
Bélgica	11,455	59.9	40.1
Nueva Zelanda	11,701	16.2	83.8
Argentina	11,881	51.2	48.8
Israel	11,881	42.9	57.1
Alemania	12,368	59.8	40.2

Desempeño educativo, ingreso per cápita y gasto educativo

Países	Lugar y puntaje en la escala de lectura		PIB per cápita en dlls. controlando por ppa	Gasto educativo total por alumno
Finlandia	01	546	25,357	47,854
Canadá	02	534	28,130	59,808
Nueva Zelanda	03	529	20,372	---
Australia	04	528	26,325	55,987
Irlanda	05	527	28,285	34,329
Hong Kong	06	525	25,153	---
Corea	07	525	15,186	30,246
Reino Unido	08	523	24,964	46,175
Japón	09	522	26,011	54,737
Suecia	10	516	26,161	54,845
Austria	11	507	28,070	77,027
Bélgica	12	507	26,392	49,489
Islandia	13	507	28,143	---
Noruega	14	505	36,242	63,599
Francia	15	505	25,090	55,086
USA	16	504	34,602	72,119
Dinamarca	17	497	28,755	65,244
Suiza	18	494	29,617	66,214
España	19	493	20,195	41,267
República Checa	20	492	13,806	22,606
Italia	21	487	25,095	58,868
Alemania	22	484	26,139	44,800
Liechtenstein	23	483	---	---
Hungría	24	480	12,204	21,997
Polonia	25	479	9,547	18,586
Grecia	26	474	15,885	24,671
Portugal	27	470	16,780	41,166
Rusia	28	462	8,377	---
Letonia	29	458	7,045	---
Israel	30	452	20,131	---
Luxemburgo	31	441	48,239	---
Tailandia	32	431	6,402	---
Bulgaria	33	430	5,710	---
México	34	422	9,117	12,189
Argentina	35	418	12,377	18,839
Chile	36	410	9,417	17,820
Brasil	37	396	7,625	10,269
Macedonia	38	373	5,086	---
Indonesia	39	371	3,043	1,164
Albania	40	349	3,506	---
Perú	41	327	4,799	3,479

EXPLICACIONES POR FACTORES DEL ENTORNO Y LA ESCUELA

PISA analiza los factores que inciden en la calidad de la educación; coincidiendo con hallazgos previos de la investigación, en todos los países se aprecia, en diverso grado:

- La importancia de los factores socioeconómicos de los alumnos en su rendimiento;
 - La importancia no despreciable de los factores de la escuela;
 - La confluencia de ambos grupos de factores.
-
- ◆ México tiene un desarrollo inferior al de la mayoría de países participantes en PISA; los jóvenes mexicanos tienen, en promedio, condiciones de vida menos favorables al aprendizaje que los de otros países. Por el peso de los factores socioeconómicos no sorprende que los resultados del país sean inferiores a los de los más desarrollados. El que México tenga resultados mejores o comparables a los de Argentina, Brasil, Chile, Perú y otros países de desarrollo bajo o medio se explica por las mismas razones.
 - ◆ Hay una relación clara entre nivel de desarrollo general y resultados educativos: la correlación entre el PIB per capita, ajustado según el poder adquisitivo, y los puntajes obtenidos en PISA, es de 0.53; y entre el gasto por alumno y los puntajes de 0.47. La influencia del nivel socioeconómico de los alumnos sobre sus resultados escolares es clara y lineal, con excepciones que muestran que no se trata de algo inexorable.
 - ◆ La tendencia, en efecto, no es perfecta: algunos países obtienen resultados mejores a los que podrían esperarse por su nivel económico, como Finlandia, Japón y Corea; otros, en cambio, obtienen resultados inferiores a los esperables por su nivel de gasto por alumno. En esta situación se encuentran países ricos como Estados Unidos o, marcadamente, Luxemburgo, y otros de nivel de desarrollo menor, como todos los de América Latina, incluyendo a México y, señaladamente, a Perú.
 - ◆ Los resultados de México se explican, en parte, por su menor desarrollo económico, pero incluso así, nuestro sistema educativo podría obtener mejores resultados: el PIB per capita de Polonia es igual al de México; sus resultados son mucho mejores.
 - ◆ Los factores escolares, influenciados por las políticas educativas, también pesan. En este sentido también hay diferencias importantes: algunos países nórdicos europeos (Noruega, Suecia, Finlandia, Islandia) muestran pocas diferencias entre sus escuelas, aunque en el primero las diferencias entre alumnos sean altas; en otros las diferencias entre escuelas son fuertes; es el caso de Polonia y Hungría, y también de México.
 - ◆ Las diferencias fuertes entre escuelas que se observan en algunos países pueden deberse a razones étnicas (Bélgica); curriculares, (Alemania); o sociales (Bulgaria).
 - ◆ México presenta un grado de desigualdad de los resultados bastante bajo, lo que posiblemente refleje la homogeneidad que produce la elevada deserción que ocurre antes del fin de la secundaria, que alcanza sólo poco más de la mitad de los alumnos, cifra

muy inferior a la de todos los países desarrollados. La proporción de las diferencias entre escuelas es relativamente alta, posiblemente debidas sobre todo a razones sociales.

RESULTADOS DENTRO DE LOS PAÍSES Y EQUIDAD

Las diferencias de los resultados entre los alumnos de un mismo país es mayor que las que existen entre unos países y otros.

- ◆ En todos los países hay alumnos de alto y bajo desempeño, pero la proporción de ambos es muy diferente; en general un resultado global mejor se asocia, obviamente, a proporciones mayores de alumnos con resultados altos, y un promedio inferior con más alumnos con resultados bajos.
- ◆ En el interior de cada país las diferencias pueden ser más o menos grandes: en Corea, con resultados globales muy buenos, la proporción de alumnos con resultados muy altos no es tan elevada, pero muy pocos obtienen resultados bajos; en Estados Unidos la desigualdad es mucho mayor: la proporción de alumnos en el nivel más alto es elevada, pero también hay muchos en el nivel más bajo.

El que algunos países consigan un alto nivel promedio en los resultados de sus alumnos, y a la vez una brecha pequeña entre los mejores y los peores... sugiere que el alcanzar altos niveles de desempeño no implica necesariamente una gran desigualdad. De manera similar, el hecho de que la fuerza de la relación entre el contexto socioeconómico y los resultados de aprendizaje varíe considerablemente entre países, y que en algunos las escuelas logren moderar tal relación... ilustra que los bajos resultados de los alumnos de extracción social más modesta no son inevitables. Las escuelas y los tomadores de decisiones pueden hacer algo para mejorar esos bajos resultados. (PISA, 2002, 202-234)

- ◆ Los resultados de PISA permiten afirmar que es posible alcanzar un nivel más alto y homogéneo de efectividad en las escuelas. En otras palabras, eficiencia y equidad no necesariamente se oponen, y pueden perseguirse conjuntamente.
- ◆ Los resultados de PISA muestran también que altos resultados van regularmente asociados a niveles muy reducidos de rezago escolar y viceversa: los países en los que proporciones elevadas de alumnos se encuentran en grados inferiores al que debería esperarse por su edad, tienen también resultados bajos.
- ◆ El rezago escolar es resultado de un ingreso tardío a la escuela o, más frecuentemente, de la reprobación. Esta última no parece contribuir a elevar los resultados escolares. El que los países de resultados más altos, como Finlandia, tengan prácticamente a todos sus jóvenes de 15 años en el grado en que teóricamente deberían estar, muestra que la reprobación efectivamente es nula en esos sistemas educativos, como lo establecen sus políticas.

- ◆ El que aún en esos países haya cierta proporción de alumnos en el nivel inferior de las escalas de aprovechamiento, o incluso por debajo de él, muestra que en esos países la promoción de un grado al siguiente no depende de que se alcancen los niveles mínimos de aprendizaje establecidos en el currículo. Las concepciones pedagógicas que se manejan superan la disyuntiva aprobación/reprobación.

COMPARACIONES CON OTROS PAÍSES

Además de tener en cuenta la naturaleza multidimensional de las evaluaciones de PISA, y los traslapes que se presentan en las puntuaciones resultantes, hay otros aspectos que deben tenerse en cuenta al hacer comparaciones entre países, que tienen que ver con las características demográficas, socioeconómicas y culturales que los hacen más o menos comparables.

Sabemos que los países de mayor nivel de ingreso *per cápita* se caracterizan también, por lo general, por tener poblaciones con proporciones de niños y jóvenes menores que las que hay en los países menos desarrollados. Esto representa una carga mayor sobre los presupuestos de los segundos, que son además muy inferiores a los de los primeros. El resultado de un menor numerador y un mayor denominador es obvio: un gasto educativo por alumno muy inferior en los países de menor desarrollo.

Por otra parte, en los países pobres la proporción de jóvenes que siguen en la escuela a los 15 años dista mucho de ser cercana al cien por ciento, como ocurre en los países desarrollados.

Como PISA define su población objetivo como la de 15 años que sigue inscrita en la escuela, los resultados que obtienen en promedio los alumnos de un país de menor desarrollo estarán sobreestimados, en relación con los que obtienen los de un país más desarrollado, en la medida en que la proporción de jóvenes que siguen en la escuela sea menor, ya que es obvio que, en promedio, los resultados de los jóvenes no evaluados por haber abandonado ya el sistema escolar serán inferiores a los de quienes siguen en él.

Es el caso, por ejemplo, de Argentina y Chile en comparación con México, Brasil y Perú: en los dos países del cono sur la proporción de jóvenes de 15 años en la escuela es muy superior a la de los otros tres países latinoamericanos que participaron en PISA; por ello el que los puntajes de México sean ligeramente mejores que los de Argentina y Chile, no sustenta una afirmación simple de superioridad de nuestro sistema educativo.

La proporción de jóvenes en la escuela está relacionada, desde luego, con los niveles de pobreza y marginalidad. Esto hace, a su vez, que la mayor deserción temprana que se da en México, Brasil y Perú, no deba imputarse sin más a deficiencias de sus respectivos sistemas educativos.

Si a la pobreza añadimos los problemas que plantea la diversidad étnica y lingüística que se da en países con importantes poblaciones indígenas como México y Perú, más la elemental cuestión de las dimensiones del sistema educativo, resulta claro que los resultados de países grandes y étnicamente complejos, como México o Brasil, no pueden compararse sin más con los de países menos grandes y más homogéneos social y culturalmente.

Es claro que México, con más de 2 millones de alumnos en cada cohorte, de los cuales un buen número pertenecen a grupos étnicos cuya lengua materna no es el español, no puede compararse sin muchas salvedades con un país chico, rico y culturalmente homogéneo como Finlandia, con menos de setenta mil jóvenes por cohorte.

Resultados de México y de otros países en contexto

	México	Argentina	Brasil	Chile	Perú	España	USA	Finlandia
Puntaje en:								
Lectura gral.	422	418	396	410	327	493	504	546
Localización	402	407	365	383	289	483	499	556
Interpretación	419	415	400	419	342	491	505	555
Reflexión	446	430	417	412	323	506	507	542
Matemáticas	387	388	334	384	292	476	493	536
Ciencias	422	396	375	415	333	491	499	538
PIB per cápita	9,117	12,377	7,625	9,417	4,799	20,195	34,602	25,357
I. Gini ingreso	.519	---	.591	.575	.462	.325	.408	.256
Gasto educ./alumno	12,189	18,893	10,269	17,820	3,479	41,267	72,119	47,854
Pobl. 15 años (miles)	2,127.5	662.0	3,464.3	263.9	546.6	462.1	3,876.0	66.6
Pobl. 15 en escuela	1,098.6	505.4	1,841.8	230.5	358.8	451.7	3,836.0	66.6
% pobl. 15 en escuela	51.6	76.3	53.2	87.3	65.6	97.7	99.0	100.0
Cobertura PISA	0.45	0.77	0.69	0.82	0.50	0.86	0.81	0.94
Cobertura neta sec.	57%	76%	68%	72%	61%	91%	87%	95%
Cobertura bruta sup.	20%	48%	15%	38%	29%	58%	72%	84%
Varianza total	7,370	11,881	7,427	8,100	9,216	7,181	10,979	7,994
% varianza entre esc.	53.4	51.2	43.1	56.7	58.0	20.7	29.6	12.3

EN RESUMEN:

◆ PISA no es una evaluación de la efectividad de las escuelas de un país, sino de los niveles de aprendizaje de sus jóvenes de 15 años, resultado tanto de la escuela como de las influencias familiares y sociales. Los resultados de PISA, por consiguiente, pueden verse como una evaluación de la calidad educativa de una sociedad en conjunto, y no sólo de su sistema escolar:

Si los puntajes de un país son significativamente más altos que los de otro, no se puede inferir automáticamente que las escuelas del primero sean más efectivas que las del segundo, puesto que el aprendizaje comienza mucho antes que la escolaridad y ocurre en una amplia gama de lugares, dentro y fuera de la escuela; lo que puede concluirse legítima-

mente es que en el primer país el impacto acumulado de las experiencias de aprendizaje, desde la primera infancia hasta los 15 años, y tanto en la escuela como en el hogar, ha producido resultados más adecuados en los terrenos que evalúa PISA. (PISA, 2002, 55-73)

- ◆ Los resultados de México en PISA no deberían sorprendernos: se sitúan en el rango de lo esperable, debido al peso de los factores socioeconómicos y a los recursos con que pueden contar las escuelas. El desempeño promedio de nuestros alumnos no es similar al de los países más desarrollados, pero tampoco inferior al de los que tienen un nivel de desarrollo comparable al mexicano.
- ◆ Lo anterior no debe leerse en términos conformistas: la experiencia internacional muestra que es posible alcanzar niveles de aprendizaje mejores, aún con nuestro nivel de desarrollo y nuestros recursos. Eso no será fácil e implicará un gran esfuerzo, pero es posible si se adoptan políticas educativas apropiadas y se les apoya en forma consistente a lo largo del tiempo.
- ◆ Debe subrayarse que el cambio y la mejora de la educación implican años de esfuerzo y no pueden esperarse en lapsos cortos; el esfuerzo debe ser continuado.

Los resultados de pisa no deben abrumarnos; tampoco pueden dejarnos impasibles: deben ser un acicate para emprender o reforzar acciones de mejoramiento.

CONCLUSIONES

- ◆ Un buen análisis de los resultados implica entender el sentido de los puntajes, según las características de los instrumentos y las muestras; las semejanzas y diferencias de los países participantes respecto a México; las diferencias dentro de nuestro país; los factores explicativos de la situación, tanto del entorno como de la escuela.
- ◆ El mejoramiento educativo no se logra en poco tiempo, ni con acciones apresuradas de emergencia. La peor forma en que un maestro puede reaccionar ante un bajo resultado es ponerse a enseñar para la prueba, tratando de mejorar artificialmente los resultados de sus alumnos, preocupándose antes que nada por su imagen, al tiempo que descuida aspectos esenciales de su tarea, que posiblemente empeoren. Lo que no debemos hacer es intentar un esfuerzo que busque incrementar espectacularmente y en poco tiempo los resultados nacionales en próximas pruebas; tal propósito estaría condenado de antemano al fracaso, y probablemente traería consecuencias negativas.
- ◆ Necesitamos mejorar las evaluaciones nacionales:
 - Aplicando instrumentos de calidad similar a los internacionales, y relacionándonos con ellos, de manera que se puedan hacer comparaciones de sus resultados.
 - Manejando muestras mayores, que permitan hacer análisis regionales, estatales y de subconjuntos de escuelas de ciertas características.

- Combinando evaluaciones basadas en instrumentos de tipo objetivo y aplicación masiva, con metodologías cualitativas que aporten otros elementos.
- Utilizándolas de modo que sirvan a maestros y escuelas para mejorar su quehacer, al tiempo que sirvan para la rendición de cuentas a la sociedad.

◆ Debemos participar en evaluaciones internacionales:

- No para autoflagelarnos porque estamos más bajos que otros países.
- Tampoco para jactarnos en caso de estar mejor cuando ocurra.
- Sí para tener referentes externos, cuidando su pertinencia y comparabilidad.
- Sí para aprovechar las mejores tendencias internacionales en el campo de la calidad educativa y su evaluación.
- Sí para validar nuestras propias evaluaciones y mejorarlas.
- Como resultado de lo anterior, deberíamos lograr una participación más completa del país en las evaluaciones internacionales, sacando el máximo provecho de ellas e incluso contribuyendo a mejorarlas.
- Para que sean útiles, debemos dar continuidad a las evaluaciones. De los resultados de una sola evaluación, como PISA 2000, no podemos concluir si avanzamos o retrocedemos; lo sabremos sólo si damos continuidad tanto a las evaluaciones nacionales como a la participación de México en las internacionales.

BIBLIOGRAFÍA

Edmonds, Ronald (1979). "Effective Schools for the Urban Poor". *Educational Leadership* No. 37: 15-27.

Hamilton, Laura S., Brian M. Stecher y Stephen P. Klein Eds. (2002) *Making Sense of Test-Based Accountability in Education*. Santa Mónica CA, Rand Education.

OECD (2001). Programme for International Student Assessment. *Knowledge and Skills for Life. First Results From Pisa 2000*. Paris, France.

_____*Reading for Change. Performance and Engagement Across Countries. Results from Pisa 2000*. Paris, France.

_____(2002). Muestra de reactivos empleados en la evaluación PISA 2000. Aptitudes para la lectura, matemáticas y ciencias. Aula XXI. Santillana. México.

_____(2002). Conocimientos y Habilidades para la Vida. Primeros Resultados de Pisa 2000. Traducción realizada por la Dirección General de Evaluación para consulta interna, del resumen disponible en: www.pisa.oecd.org