

avance y
PERSPECTIVA

ÓRGANO DE DIFUSIÓN DEL
CENTRO DE INVESTIGACIÓN Y DE ESTUDIOS AVANZADOS DEL I. P. N.



Evaluación



Maestría

El alumno interesado en estudiar la maestría en Bioquímica, debe cumplir con los siguientes requisitos:

Estudios a nivel licenciatura terminados en alguna de las carreras de ciencias de la salud o biológicas.

Promedio mínimo de 8 en la licenciatura.

Examen profesional ya realizado o en proceso, en su defecto presentar constancia de ser candidato a obtener el título a través de la maestría.

Presentar por duplicado la documentación solicitada en la página del departamento.



Doctorado

Los candidatos al programa de doctorado son aquellos maestros en ciencias titulados en la especialidad de bioquímica o en áreas afines (favor de comunicarse con la coordinación académica del departamento para recibir mayor información acerca del proceso de admisión al doctorado).

Líneas de investigación

Canalopatías de músculo esquelético

Dr. Guillermo Ávila Flores
gavila@mail.cinvestav.mx

Bioquímica médica

Dr. J. Víctor Calderón Salinas
jcalder@mail.cinvestav.mx

Dinámica de lípidos

Dr. Jorge Cerbón Solórzano
jcerbon@mail.cinvestav.mx

Sistemas modelo y propiedades fisicoquímicas de membranas Biológicas

Dra. Marta S. Fernández Pacheco
msfernandez@mail.cinvestav.mx

Sistemas fotosintéticos bacterianos

Dr. Carlos Gómez Lojero
cgomez@mail.cinvestav.mx

Regulación del calcio intracelular

Dr. Agustín Guerrero Hernández
aguerrera@mail.cinvestav.mx

Bioquímica celular y molecular de patógenos intracelulares

Dr. Ricardo Mondragón Flores
rmflores@mail.cinvestav.mx

Mecanismos de transducción de señales hormonales

Dr. J. Alberto Olivares Reyes
jolivare@mail.cinvestav.mx

Biología molecular y bioquímica del desarrollo

Dr. Luis M. Salgado Rodríguez
lsalgado@mail.cinvestav.mx

Metabolismo del pre-mRNA

Dr. Jesús Valdés Flores
jvaldes@mail.cinvestav.mx

Inscripciones

La inscripción a los cursos propedéuticos está abierta durante el primer semestre del año. Los cursos propedéuticos inician en Agosto. La inscripción al doctorado está abierta todo el año.

La entrega de documentos y solicitud de ingreso se realiza directamente en las oficinas del departamento de Bioquímica, CINVESTAV, ubicadas en Av. I.P.N. 2508, Col. San Pedro Zacatenco, México, D.F.

Los planes de estudio se pueden consultar en la página del departamento.

Cualquier duda o aclaración, puede llamar al (52) (55) 5061-3800, Ext. 5213, con la secretaria Martha Montes, o bien, al teléfono 50-61-39-50 con el coordinador académico.

<http://www.biochem.cinvestav.mx>

<http://www.cinvestav.mx>





Centro de Investigación
y de Estudios Avanzados del IPN
CINVESTAV



La obra que se reproduce en este número es del ceramista y escultor José Antonio Gurtubay Arenas, el fundidor de las piezas es el maestro Enrique Hernández Caro y la fotógrafa es Denise Escamilla, a todos les agradecemos su colaboración.

Portada

"María" bronce a la cera perdida

Avance y Perspectiva, órgano de difusión del Centro de Investigación y de Estudios Avanzados del IPN, CINVESTAV, es una publicación trimestral. El número correspondiente a enero-marzo de 2005, volumen 24, Núm. 1, se terminó de imprimir en febrero de 2005. El tiraje consta de 5,000 ejemplares. *Editor responsable:* Susana Quintanilla. Oficinas: Av. IPN No. 2508 esquina calzada Ticomán, apartado postal 14-740, 07000, México, D.F. Certificados de litud del título del No. 1001 otorgados por la Comisión Calificadora de Publicaciones y Revistas Ilustradas de la Secretaría de Gobernación. Reserva de Título No. 577-85 otorgado por la Dirección General del Derecho de Autor de la Secretaría de Educación Pública. Publicación periódica: Registro No. 01603-89, características 220221122, otorgado por el Servicio Postal Mexicano. *Negativos, impresión y encuadernación:* Gama sucesores, S.A. de C.V., ingenieros civiles No. 94, delegación Iztapalapa, C.P. 09420, México, D.F. *Avance y Perspectiva* publica artículos de divulgación y notas sobre avances científicos y tecnológicos. Los artículos firmados son responsabilidad de los autores. Las instrucciones para los autores que deseen enviar contribuciones para su publicación aparecen en la última página de esta revista. Se autoriza la reproducción parcial o total del material publicado en *Avance y Perspectiva*, siempre que se cite la fuente. *Avance y Perspectiva* se distribuye en forma gratuita a los miembros de la comunidad del CINVESTAV y a las instituciones de educación superior. Suscripción personal por un año: \$150.00

CINVESTAV

Rosalinda Contreras Theurel
Directora General

José Mustre de León
Secretario Académico

Luis Alfonso Torres Gómez
Secretario de Planeación

Enrique Campesino Romeo
Secretario Administrativo

AVANCE Y PERSPECTIVA

Susana Quintanilla
Directora editorial

María de Ibarrola
Coordinación de este número

Ricardo Salas § Frontespizio
Begoña Sáinz
Diseño y cuidado de la edición

Verónica Arellano y Ma. Eugenia Ávila
Apoyo editorial

Josefina Miranda López
María Gabriela Reyna López
Captura

CONSEJO EDITORIAL

J. Víctor Calderón Salinas
Bioquímica

Luis Capurro Filograsso
Unidad Mérida

Marcelino Cerejido
Fisiología

María de Ibarrola Nicolás
Investigaciones Educativas

Eugenio Frixione
Biología Celular

Jesús González
Unidad Querétaro

Luis Herrera Estrella
Unidad Irapuato

Gabriel López Castro
Física

Luis Moreno Armella
Matemática Educativa

Ángeles Paz Sandoval
Química



Centro de Investigación
y de Estudios Avanzados del IPN
CINVESTAV

avance@mail.cinvestav.mx

T/F 57 47 70 76

www.cinvestav.mx/publicaciones

SUMARIO

AVANCE Y PERSPECTIVA	
Presentación editorial <i>María de Ibarrola</i>	5
Evaluación de investigadores. Diálogo entre disciplinas e instituciones <i>Marta de Ibarrola</i>	9
El Seminario de Evaluación de la Evaluación <i>Arturo Escobosa, Eugenio Frixione, Luis Gorostiza, Walid Kuri, Michael Porter, Enrique Ramírez de Arellano</i>	19
Evaluación académica: empate de dos imposibilidades <i>Bogdan Mielnik</i>	23
¿Medir para conocer o para regular?: evaluación y acreditación de la educación superior en México <i>Sylvie Didou Aupetit</i>	31
La internacionalización de la evaluación de los aprendizajes en la educación básica <i>Emilia Ferreira</i>	37
Efectos de las evaluaciones estandarizadas en los sistemas educativos <i>Antonia Candela</i>	45
La evaluación de profesores de educación básica en México: la Carrera Magisterial <i>Ruth Mercado</i>	57
Criterios de la evaluación educativa: bases y perspectivas <i>Andrés Sánchez Moguel</i>	71
NOTICIAS DEL CINVESTAV	78
LIBROS Y REVISTAS	
La ciencia como proyecto de vida <i>Antonio Gómez Nashiki</i>	85



PRESENTACIÓN EDITORIAL

MARÍA DE IBARROLA

EN 1982 SE INSTALÓ EN EL CINVESTAV LA COMISIÓN EXTRAORDINARIA DE Reclasificación y Becas de Exclusividad. Promovida por el Dr. Manuel Ortega, director de nuestra institución de 1978 a 1982, esta Comisión logró instaurar un mecanismo de evaluación del personal académico sustentado en los juicios de los “pares” –como debe de ser en toda evaluación de académicos. La evaluación tenía cuatro características innovadoras: era interna a la institución (a diferencia de las evaluaciones extra institucionales e indirectas de los “colegios invisibles” en las que habían confiado directores y jefes de departamento hasta entonces); abarcaba en un mismo procedimiento a todas las áreas de conocimiento que se cultivan en el Centro; fijaba como su objeto la productividad alcanzada en un periodo determinado, y sancionaba el resultado con un estímulo económico adicional al salario.

La importancia y la necesidad de evaluar los resultados individuales alcanzados por los investigadores con fines de reordenar la clasificación y promoción del personal académico no se cuestionaron; la introducción de un estímulo económico adicional fue bienvenida en su momento. Pocos años después el Sistema Nacional de Investigadores coincidiría con este modelo, pero a través de una evaluación metainstitucional. Al paso del tiempo todas las instituciones científicas y educativas del país incorporaron modelos semejantes para la asignación diferenciada y selectiva de mayores ingresos no salariales.

Todavía recuerdo el impacto que me provocó una de las propuestas iniciales sobre la “productividad” que sería objeto de esa evaluación de los pares. La hizo, con absoluta seguridad, en el seno de una junta de jefes de departamento convocada con ese fin, uno de nuestros investigadores de más prestigio académico. “Es muy sencillo, declaró: artículos en revistas arbitradas y de prestigio internacional: 5 puntos; artículos en revistas nacionales, 1 punto; tesis terminadas de doctorado: 20 puntos; tesis terminadas de maestría: 5 puntos.” “¿Dónde queda entonces todo lo que hacemos en el DIE”, fue mi pregunta angustiada y silenciosa, seguramente compartida por varios más de los presentes.

La Dra. María de Ibarrola es investigadora titular del Departamento de Investigaciones Educativas del CINVESTAV. Correo electrónico: ibarrola@cinvestav.mx

A lo largo de este número de *AyP* observamos que a veinticinco años de distancia las políticas científicas y educativas del país le dan absoluta prioridad a la evaluación de la productividad o el desempeño de los individuos y del supuesto impacto inmediato de su trabajo (o más bien de aquello que se puede medir y contar) para decidir sobre el financiamiento y la conducción de toda la actividad académica y educativa del país. Se evalúan todas las funciones: la investigación, la docencia, la difusión, la vinculación; se extrapolan criterios e indicadores de una función a otra o de un nivel a otro sin la debida atención a su pertinencia; se evalúan no sólo a los investigadores sino a los profesores de todos los niveles. Intervienen en la evaluación organismos públicos e instituciones privadas. Se crean instituciones especializadas, se aplican instrumentos estandarizados de alcance internacional. Las evaluaciones no sólo repercuten en la política nacional e institucional, sino que ocupan una parte importante del tiempo de las instituciones y del interés y la emotividad de individuos. A pesar de las críticas y cuestionamientos a su validez, encauzan en buena medida las actividades y los proyectos de investigadores y docentes.

Los artículos que constituyen este número de *AyP* documentan y analizan una parte significativa de los programas de evaluación ahora vigentes en México. También denuncian los extremos y las distorsiones a las que han llegado. Los investigadores y docentes afectados —abrumados por la “monstruosa” penetración de la evaluación en nuestras actividades cotidianas y en nuestros proyectos a corto y largo plazo— ya no sólo debatimos si los productos de nuestro trabajo están bien identificados y medidos. Empezamos a cuestionar a fondo las

estructuras y los mecanismos que se han ido instalando: a cuestionar el quién evalúa y a quién se evalúa, el cómo, el cuándo, el por qué y el para qué. No se trata de eliminar la evaluación, sino de analizar y evaluar cuidadosamente sus efectos positivos y negativos, a corto y a largo plazo, tomar nota de las advertencias sobre cómo interpretar los resultados y, sobre todo, encontrar alternativas al rendimiento de cuentas que ciertamente debemos a nuestras instituciones y a la sociedad que nos subsidia.

PRODUCTION DATA AND
QUALITY CONTROL





EVALUACIÓN DE INVESTIGADORES. DIÁLOGO ENTRE DISCIPLINAS E INSTITUCIONES

MARÍA DE IBARROLA

EN MARZO DE 2004 EL DEPARTAMENTO DE INVESTIGACIONES EDUCATIVAS y el Seminario Interdepartamental de Evaluación de la Evaluación del CINVESTAV, convocaron a un simposio orientado a compartir entre colegas de diferentes disciplinas e instituciones el apasionante debate que suscita la evaluación de los investigadores en nuestro país, tanto en lo referente a las propuestas de las instituciones de educación superior como a la implementada por el Sistema Nacional de Investigadores desde 1984.

Participaron la Dra. Rocío Grediaga, del Área de Sociología de las Universidades, de la UAM-Atzacapotzalco; el Dr. Bogdan Mielnik, del Área de Ciencias Exactas del CINVESTAV; el Dr. Eduardo Ibarra Colado, del Área de Estudios Organizacionales de la UAM-Iztapalapa; el Dr. Carlos Beyer, del Área de Ciencias Biológicas del CINVESTAV; la Dra. Susana García Salord, del Instituto de Matemáticas Aplicadas y Sistemas de la UNAM; el Dr. Arturo Escobosa, del Área de Ciencias de la Ingeniería del CINVESTAV; la Dra. Hebe Vessuri, del Departamento de Estudios de la Ciencia del Instituto Venezolano de Investigación Científica y la Dra. María de Ibarrola, del Área de Ciencias Sociales del CINVESTAV.

Al evento asistieron medio centenar de investigadores de casa y alrededor de veinte más de otras instituciones de educación superior del Distrito Federal que se arriesgaron a recorrer la distancia citadina y acudir al Auditorio José Ádem de la Unidad Zacatenco del CINVESTAV.

El debate estuvo orientado a identificar modelos diferentes de evaluación de académicos tanto en México como en otros países y a analizar los criterios que se aplican de manera más generalizada: la manera como se constituyen y priorizan los índices de productividad, los requisitos que se establecen para la promoción y los niveles del escalafón que se fijan como parte del recorrido promocional. También se propuso que discutiéramos el significado de la evaluación en el fortalecimiento (o no) de la vocación de investigador, los costos y la eficiencia de los procesos de evaluación y todo lo que no

Coordinación y relatoría del evento: María de Ibarrola; organización: María de Ibarrola y Eugenio Frixione.

se toma en cuenta cuando se evalúa a los investigadores. En cierta medida las dos horas del debate se organizaron conforme a esas preguntas iniciales, pero no necesariamente se respetó un orden secuencial ni se trataron los diferentes temas con la misma intensidad.

MODELOS GENERALES O MODELOS PARTICULARES DE EVALUACIÓN

En la experiencia de los ponentes –y de los participantes que realizaron múltiples aportaciones y comentarios a lo largo de la sesión– resulta muy importante que existan distintos modelos de evaluación. Una especificidad que debiera tomarse en cuenta –para todos obvia– son las diferencias entre disciplinas. Dada la mayoría de investigadores de las “ciencias duras” que poblaba el auditorio, la discusión al respecto se ejemplificó mediante las diferencias en el tipo de citas y en el factor de impacto que registra el *Science Citation Index*. Sin dudar de la validez del criterio mismo para la evaluación de un investigador, los matemáticos fueron enfáticos en que han sido reiteradamente subvaluados, situación reconocida incluso en los más recientes informes sobre productividad científica reportados por la Academia Mexicana de Ciencias. Los pocos investigadores de las “ciencias blandas” presentes en la reunión (entre ellos cuatro ponentes) insistieron en dejar constancia de que en su caso no sólo no se puede hablar de subvaloración de las publicaciones de estas disciplinas, sino que el criterio mismo es inoperante debido a la ausencia de registros en el *Social Science Citation Index* de la mayor parte de los diversos medios por los que se difunde la investigación en ciencias sociales: libros, capítulos de libros, publicaciones orientadas no nada más a los pares, sino a los distintos tomadores de decisiones en materia social. A pesar de ello, el CINVESTAV, por ejemplo, no acepta eliminar o sustituir el criterio

de las citas para la evaluación del investigador de ciencias sociales, mientras que el área de humanidades y ciencias de la conducta del SNI sí lo hace.

La importancia de reconocer diferencias institucionales en los modelos de evaluación no parecía tan obvia a los miembros del CINVESTAV. Sin embargo, los invitados fundamentaron las diferencias en las actividades académicas que se derivan de ellas. La mayor parte de las instituciones de educación superior del país deben atender todas las funciones asignadas: en particular la docencia de licenciatura tiene un enorme peso en la programación del tiempo de trabajo de los académicos de esas instituciones. Las actividades de docencia en estos niveles y la formación y atención de los jóvenes estudiantes no se pueden reducir fácilmente a cursos impartidos o tesis dirigidas, puesto que exigen mucho más a los profesores responsables. También la difusión de la cultura conduce a diferenciar prioridades en el trabajo de los académicos en las distintas instituciones, y aunque ha sido generalmente desvalorizada, poco a poco se reconoce su relevancia y su enorme grado de dificultad. La importancia que alcanza recientemente la “vinculación” académica –caracterizada ahora como transferencia y coproducción de conocimiento entre centros de investigación y centros de producción de bienes y servicios– apenas empieza a estar contemplada en los distintos reglamentos existentes e involucra dificultades especiales para la evaluación del trabajo de los investigadores, por un lado, la confidencialidad exigida en los convenios que impide la publicación de los resultados; por otro, la complejidad de la administración de proyectos, en ocasiones sumamente costosos, y cuya gestión no es reconocida como “actividad académica”.

En los hechos, los modelos difieren en cuanto al referente de la evaluación: si sólo se toma en cuenta una función académica, la investigación

o la docencia, o se reconocen todas; si sólo se reconoce el trabajo realizado en la institución de adscripción o toda la productividad del académico en el amplio espacio de la educación superior pública en el país e incluso fuera del mismo. Otra diferencia radica en la periodicidad de las evaluaciones: semestral, anual, bianual o trianual. Incluso se denunciaron casos en que se realizan evaluaciones trimestrales. Los modelos difieren también en cuanto a la conexión con otros sistemas de evaluación. Al respecto, resaltó el hecho de que los académicos pueden estar sometidos a múltiples evaluaciones diferentes sin conexión entre sí.

Otra diferencia sustancial es la manera como se conforman las comisiones evaluadoras: si son elegidas por los evaluados o impuestas por las autoridades, si se conforman por internos, por externos o de manera mixta. Al respecto se cuestionó el concepto de “pares”, que por lo general rige la elección de los miembros de las comisiones evaluadoras de los académicos—intra o inter institucionales—en el país. Un dato interesante es que efectivamente los evaluadores directos son académicos o investigadores igual que los evaluados. Quiénes son nuestros iguales o “pares” queda sujeto sin embargo a discusión. Un ilustre matemático de nuestra institución cuestionó severamente si los investigadores de otras áreas que conforman la Copei pueden considerarse “pares” de un matemático.

Las evaluaciones difieren según la manera como se decide la naturaleza y el peso de los indicadores que serán empleados. En ese punto el diálogo fue largo y nutrido alrededor de temas que parecen haberse tratado hasta la saciedad y que se desmenuzan en reglamentos de evaluación tan detallados como los de la Universidad Autónoma Metropolitana o el propio CINVESTAV, o tan genéricos como los de la UNAM y el Sistema Nacional de Investigadores. En realidad no se llegó a conclusiones respecto de cuál modelo resulta mejor.

Reconocidas las múltiples y posibles diferencias que afectan el trabajo académico sujeto a los procesos de evaluación, y la composición y procedimientos de las comisiones evaluadoras, no se analizó, sin embargo, el tema de cómo lograr equivalencias que reconozcan méritos semejantes entre investigadores de distintas instituciones o disciplinas, entre diferentes productos o resultados del trabajo y entre los criterios aplicados.

Pero en cada caso, según nos informan los investigadores que se han dedicado al estudio de estos procesos de evaluación en otras instituciones, los diferentes “espacios de aplicación” rompen esquemas en los hechos, cambian criterios y generan tensiones que se dirimen de acuerdo con las relaciones de fuerza que hay en cada comisión evaluadora o incluso en cada dependencia. Un aspecto fundamental en este último caso es el grado de conocimiento sobre las reglas del juego que tienen los afectados, la suficiencia de información sobre los criterios y mecanismos de evaluación y sobre los juicios—y sus fundamentos—que emiten las comisiones. Las quejas respecto de la insuficiencia e incluso ausencia de comunicación adecuada sobre estos rubros fueron prácticamente unánimes.

SOBRE LOS ÍNDICES Y CRITERIOS DE PRODUCTIVIDAD

Uno de los grandes riesgos—y rasgos—de la evaluación de los académicos es el de evaluar aquello que se puede cuantificar y medir: el número de publicaciones y las citas registradas de las mismas, y el conjunto cada vez mayor de requisitos que deben reunir quienes otorgan constancia y registro de los productos del trabajo académico: si las revistas están indexadas, cuál es el factor de impacto de las citas; si los programas de docencia están en el padrón de excelencia; si las editoriales que publican son “reconocidas”. Son prácticamente inexistentes las comisiones evaluadoras que no

se basan en la cuantificación de estos productos sino en un análisis directo de los evaluadores sobre la calidad de los mismos, muy difícil de realizar por cierto. La posibilidad de que el evaluado elija y someta a los evaluadores sólo la mejor expresión de sus trabajos se pierde ante el peso que alcanza una determinada expectativa –no muy precisa– de lo que debe haber producido en el periodo. Prácticamente todos los ponentes y el público participante coincidieron al ilustrar los efectos perversos de esos indicadores; abundaron en ejemplos del valor excesivo asignado a ciertos productos en detrimento de otros, la manera en que la cantidad ha ido en detrimento de la calidad; se constituyen en obstáculos burocráticos a las promociones; favorecen sistemas de castas e incluso pervierten vocaciones. Algunos de estos efectos se caracterizaron y ejemplificaron: el desaliento a emprender problemas desafiantes o polémicos que requieren trabajo amplio, paciente y de largo plazo, porque no se van a encontrar resultados –parciales– publicables en corto plazo. Se denunciaron comportamientos de “autismo académico”: la escasa dedicación a leer, a escuchar a los demás, a dialogar, la falta de tiempo para madurar ideas, proyectos y trabajos de investigación y en el caso de las ciencias sociales, el silenciamiento de la reflexión crítica, porque los “ensayos” en los que el investigador plasma su conciencia social no son “productos de investigación”.

COSTOS DE LA EVALUACIÓN

Uno de los temas originales que se trataron en la reunión fue el relativo a los elevados costos y la posible ineficiencia de los continuos procesos de evaluación a los que estamos ahora sometidos. En realidad no hay datos pero la simple idea de calcular el costo del tiempo que dedican los investigadores en lo personal a presentar sus informes y del que dedican las

comisiones de evaluadores a evaluarlos podría dar una idea de los excesos en los que se ha incurrido.

La eficiencia de los procesos de evaluación, es decir, la relación entre el esfuerzo y el costo de los mismos frente a los resultados, en realidad no se conoce bien. Los resultados se manejan en términos cuantitativos: cuántos investigadores hay de tal o cual nivel, qué cantidad de productos de x o z tipo registran, en qué instituciones o zonas geográficas del país se localizan. Existe la tendencia a colocar los parámetros de comparación en la cúspide de la productividad: los premios Nobel, los premios nacionales o internacionales, los mejores investigadores constituyen el referente de la productividad de los demás, y no la norma estadística; de hecho no hay análisis estadísticos públicos que digan cuál es el promedio de productividad de los investigadores mexicanos por disciplina o por edad, por institución, experiencia laboral y antigüedad en el oficio. No se conoce el impacto del trabajo de investigación más allá de su reconocimiento y uso por otros académicos (ese es el sentido de las citas) y no hay criterios claros para analizar la eficiencia de la investigación en las políticas públicas, la solución de problemas nacionales y el impacto en otros sectores sociales. Más bien predominó el entendido de que este impacto difícilmente es directo y menos aún es inmediato. Se reiteró que los tiempos y objetivos de la investigación son diferentes que los de las políticas y se plantearon de nuevo los tradicionales ejemplos que registra la historia de la ciencia sobre el tiempo que tardaron algunos descubrimientos trascendentales de teoría básica en encontrar aplicaciones concretas. Hubo menos participaciones en el sentido de intentar identificar indicadores de los diversos tipos de impacto, que en el caso de las ciencias sociales, sin embargo, pueden ser frecuentes.

Una preocupación fue la consecución de resultados diferentes a los declarados –denominados

“perversos” en las ciencias sociales— que se estarían desprendiendo de la influencia tan poderosa que los instrumentos más técnicos de la evaluación (las listas de indicadores, su ponderación) está teniendo sobre el comportamiento de los investigadores y la reorientación de sus actividades en función de los productos a los que se da mayor puntuación y valoración. La mayor parte de los asistentes coincidió en que el perfil de la comunidad académica mexicana se ha ido moldeando por influencia del peso de los procesos e indicadores mismos de evaluación. En el extremo, una revisión a fondo de las políticas de evaluación nos obligaría a profundizar en problemas mucho más complejos, como el papel que las corrientes internacionales tienen en una posible división internacional del trabajo de investigación, que asigne a México cómo debe ser su investigación, de qué tipo y de qué alcance. Surgió también en la discusión el problema de la conformación de mercados internacionales de conocimiento y la privatización de este último, en clara oposición a la función pública —literalmente abierta a todos— y a favor del desarrollo del país reconocida históricamente para la investigación y el desarrollo científico apoyados por el Estado. Se incluyó en esta problemática una crítica a la orientación de las políticas recientes de apoyo a la investigación que instrumenta ya el Conacyt, de asignar fondos a entidades empresariales privadas.

EL MODELO “MEXICANO” DE EVALUACIÓN DE LOS INVESTIGADORES

Una duda que afectó a los ponentes y asistentes fue si la discusión versaba sobre los modelos posibles de evaluación de los investigadores o sobre el modelo específico de evaluación que se ha ido imponiendo en México. La tendencia fue darle peso a la experiencia de los académicos mexicanos en estos asuntos, aunque

salpicada de ejemplos puntuales de otros países aportados por ponentes y asistentes.

Los modelos de evaluación de investigadores se identifican a partir de la respuesta que se dé a los siguientes tópicos y de la interacción entre las respuestas mismas: qué se evalúa, quién evalúa, cómo, cuándo, por qué y para qué. Cada uno de ellos admite diferentes respuestas. De ahí que lo importante debería ser reconstruir modelos integrales de trabajo académico e identificar el lugar que ocupa la evaluación en el proceso completo, lugar que define también al modelo.

El “modelo mexicano” de evaluación de investigadores se puede caracterizar de esa manera. Cabe señalar que se trata de las tendencias dominantes que tipifican el modelo, aunque durante la discusión no se dieron ejemplos de otras formas de evaluar que se apliquen en el país.

Qué se evalúa. Se evalúa aquello que se puede medir y contabilizar a partir de “indicadores”, esto es, expresiones contables y medibles de productos concluidos de la actividad académica: de investigación, de docencia, de difusión, de vinculación, de preferencia previamente evaluados por otros. Se analizan los registros de esas evaluaciones tal y como se expresan en documentos que deben quedar físicamente accesibles a los miembros de las comisiones: los artículos publicados en revistas con arbitraje y “prestigio internacional”; las citas a esos trabajos registradas por el *Science Citation Index*; los grados otorgados por cierto tipo de instituciones; las actas de los alumnos graduados; las constancias de los cursos impartidos; las constancias de la participación en los congresos, etcétera. El extremo es tal que no basta la relación que haga el investigador de su productividad. Si no hay un documento comprobatorio externo, el expediente no pasa los filtros burocráticos y el investigador queda fuera del proceso y reprobado *por default*.

Quién evalúa: Evalúan los “pares”, pero en el contexto de reglamentos que distorsionan la evaluación hacia criterios burocráticos y hacia los criterios que se pueden medir y contabilizar más fácilmente. Evalúan comisiones centrales, nacionales o institucionales, y esa evaluación afecta toda la trayectoria y toda la productividad del investigador en un periodo determinado, lo que conlleva además a la aplicación de formatos únicos y vías únicas de acceso al formato para todos los investigadores.

Cómo evalúan: Se evalúa de manera indirecta en el fondo se trata de una contabilidad y clasificación de los documentos probatorios. Son raros los casos en que los evaluadores analizan de manera directa el trabajo presentado.

Cuándo evalúan: Las evaluaciones suceden en periodos fijados por criterios de conveniencia para las instituciones evaluadoras. El SNI, por ejemplo, evaluó durante mucho tiempo el periodo comprendido de junio a mayo, causando graves confusiones respecto de la vigencia de los trabajos que el investigador somete a evaluación. La periodicidad de las evaluaciones puede no reconocer los tiempos y las dinámicas propias de los objetos de investigación o de los métodos requeridos por diferentes disciplinas. Se propician así discusiones entre investigadores de distintas disciplinas sobre la frecuencia posible para la entrega de resultados originales: los químicos, cada tres meses; los matemáticos, cada dos o tres años. Las metodologías cualitativas de las ciencias sociales exigen seguimiento de procesos largos y no sólo el análisis de aplicación puntual de algún cuestionario, test, encuesta, por lo que puede ser mucho el tiempo para alcanzar resultados. Salen perdiendo los proyectos de largo aliento que no tengan capacidad de producir resultados publicables intermedios conforme a los mismos criterios de un producto acabado. El modelo

implica, además, someter al mismo investigador a diferentes evaluaciones en periodos relativamente cercanos. Esas evaluaciones difieren en cuanto a los criterios burocráticos, los formatos, las fechas de aplicación y los efectos de las evaluaciones.

Por qué se evalúa: El fundamento de la evaluación de investigadores mexicanos se sustenta en una noción confusa y vaga de baja calidad de la investigación en el país, de recursos mal utilizados, de preparación inadecuada, falta de responsabilidad de los investigadores y poca productividad de los mismos.

Para qué se evalúa: Se evalúa sobre todo para asignar recursos escasos y para *legitimar* la asignación diferenciada de sueldos, estímulos, plazas académicas, complementos de sueldos y proyectos personales e institucionales en las diferencias entre académicos en cuanto a su productividad.

Este es el factor que determina el modelo: al no haber recursos suficientes todos los demás parámetros quedarán supeditados. Los recursos, por supuesto, se calificaron como extremadamente escasos. Sin embargo, también se puso sobre la mesa el hecho de que en términos absolutos las cifras asignadas por investigador a las instituciones académicas son muy superiores a las de los otros niveles educativos, razón por la cual los gobiernos y las sociedades no dudan en calificarlas de muy onerosas y poco rentables, dado el escaso impacto inmediato y directo de los resultados de la investigación en el desarrollo económico de los países.

LA DETERMINACIÓN HISTÓRICA DEL MODELO MEXICANO

El origen de los programas de evaluación de académicos en México fue minuciosamente analizado para identificar los rasgos que impone a los modelos vigentes: desde el inicio la finalidad

principal no fue crear un “modelo” de evaluación para alimentar la posibilidad de mejorar la investigación en el país, sino un mecanismo para redistribuir recursos muy escasos ante la necesidad impostergable de incrementar los ingresos de los investigadores. Los mecanismos y criterios que se reglamentaron para atender inicialmente sólo a este reducido grupo se empezaron a ampliar para atender a los profesores de carrera de las universidades e incluso a los profesores de asignatura, técnicos académicos y auxiliares de investigación, con los inevitables sesgos y vicios de origen.

La sobrepolitización de las instituciones académicas en las décadas de los setenta y ochenta y el abuso de los recursos públicos en países como México también fueron situaciones –se reconoció durante el evento– que justificaron mayormente la introducción de las múltiples políticas de evaluación que ahora aplica el Estado mexicano a las instituciones académicas para asignar cualquier recurso: evaluaciones de alumnos para el ingreso y el egreso; evaluaciones de profesores y de investigadores para complementar sus ingresos o para otorgar recursos para proyectos de investigación; evaluación de programas institucionales e incluso de instituciones completas. En el país ha habido una tendencia histórica diferencial de darle peso a distintas dimensiones en la puesta en operación de las políticas de impulso a la investigación y no a su integralidad: en alguna época se dio prioridad a la creación de instituciones; en otra, a la formación de recursos humanos. Estamos ahora en la era de la evaluación de la productividad de los investigadores. Sin embargo, los resultados de estas evaluaciones tienden a replicar el modelo cuantitativo, a usar como referentes indicadores no cuestionados, a dejar de lado las consideraciones sobre circunstancias e historias de las instituciones dedicadas a la investigación, o de las diferentes disciplinas académicas.

LOS EFECTOS DE LAS EVALUACIONES REALIZADAS

No es fácil llegar a conclusiones sobre los efectos que ha tenido la evaluación de los investigadores. En principio, todo este enorme esfuerzo ha permitido un mayor conocimiento del número de científicos en México y de su productividad. Se cuenta ahora con ciertos elementos que permiten la comparación entre instituciones y disciplinas en distintas zonas geográficas del país, a la vez que se están dando importantes esfuerzos por elaborar cada vez mejores indicadores y parámetros y por ser transparentes en las evaluaciones. Se han generado líneas de investigación y grupos de académicos dedicados a lograr mejores enfoques de conocimiento, investigación y evaluación de la evaluación.

Uno de los ponentes señaló que la exigencia de publicar a nivel nacional e internacional, impulsada fuertemente por los reglamentos de evaluación, tuvo un impacto positivo muy importante para la ciencia en América Latina. Obligó a pasar de una “cultura oral” –casi una cultura de poetas– en la que participaban incluso los investigadores, a una cultura escrita, que permitió discriminar positivamente a la actividad fundamental de la investigación en los medios académicos. Los procesos de evaluación han estimulado una mayor visibilidad internacional de la ciencia de nuestros países y han contribuido a generar una cultura de responsabilidad en el uso de los recursos públicos que se han asignado a la ciencia.

Un efecto muy discutido fue el peso que tienen las evaluaciones sobre los ingresos de los académicos y el grado de estabilidad de los montos derivados de esas evaluaciones. Caben en cada uno de esos rubros múltiples variantes, según las instituciones, pero se destacó el hecho de que para los académicos mexicanos, los estímulos derivados de las evaluaciones llegan a significar más del 50% de los ingresos; estímulos

y "becas" que pueden ser suspendidos por el simple mecanismo de la evaluación y que no tienen repercusiones en ninguna de las prestaciones contractuales.

Otro efecto de las evaluaciones responde claramente según los ponentes de las áreas de ciencias sociales al identificado por un notable sociólogo norteamericano, Robert K. Merton, quien lo denominó "Efecto Mateo", en alusión al apóstol que —según los Evangelios— más recibe: "al que más tiene más se le dará," situación que se expresa cada vez con mayores exigencias, como es la disposición más reciente del Conacyt, de otorgar recursos para proyectos de investigación sólo a los investigadores nacionales de nivel II por lo menos.

Este rejuego entre las evaluaciones y la asignación de recursos suplementarios obliga a los investigadores a argumentar la pertinencia social de la investigación, aunque sea distante e indirecta; su capacidad de generar —en el largo plazo— respuesta a los problemas que tenemos como países, sean disciplinarios o sociales, o por lo menos aprender a negociar y a argumentar ante otros sectores sociales cuál es la pertinencia, la importancia y la calidad de su trabajo, superando el grave riesgo de quedar atados a la coyuntura y al corto plazo.

Para algunos de los presentes en la reunión pensar en explicar el valor de la ciencia, por ejemplo la física pura, a los congresistas que aprueban los presupuestos nacionales parecía una afrenta. Sin embargo, la nueva situación política del país y el papel del Congreso en la aprobación del presupuesto, pero en particular el reconocimiento de que los recursos públicos son generados por toda la sociedad por la vía de los impuestos, obliga —sin duda— a las instituciones de investigación a rendir cuentas a la sociedad que los financia. Las ciencias sociales tienen mucho mayor experiencia al respecto que las ciencias duras.

SE DEBE O NO EVALUAR A LOS INVESTIGADORES

El consenso entre ponentes y asistentes fue que la evaluación es parte integrante de todo proceso racional. Implica comparar los logros o resultados alcanzados con los objetivos planeados, pero también requiere valorar la validez y la pertinencia de los objetivos mismos. En el caso de México no está claro cuáles son los referentes: ¿El desarrollo nacional? ¿Qué sentido del desarrollo? ¿El bienestar de la población? ¿Las mejores condiciones de vida de los mexicanos? ¿El avance internacional de la ciencia? ¿El lugar de la tecnología en el impulso a la productividad? ¿De las grandes empresas? ¿De las pequeñas y medianas empresas? Evaluar implica entonces tomar en consideración la eficacia y eficiencia de los medios puestos al servicio de los objetivos con la finalidad de alcanzar resultados acordes; implica también la capacidad de identificar de manera válida tanto los objetivos, como los medios y los resultados. Pero no es sólo un proceso de insumo-producto de un sistema cerrado. La actividad académica constituye un sistema abierto a múltiples intervenciones, a escala nacional e internacional, afectado por un amplio conjunto de circunstancias y de historias que determinan, delimitan o por lo menos propician la definición y la identificación de cada uno de esos elementos.

El problema ha sido la manera tan particular de exigir cuentas del quehacer científico en nuestros países, el exceso de control burocrático y la *monstruosidad* que han alcanzado los múltiples ejercicios de evaluación a los que nos obligan ahora diferentes organismos. Lo enormemente costoso e improductivo de revisar y resolver páginas y páginas de formularios, que en alguna época nos inundaron de papel, y que ahora resuelve —aparentemente— la computadora, creando un nuevo tipo de agresión a los investigadores que no encuentran en

los formatos preestablecidos y en la terquedad de la máquina la posibilidad de dar cuenta de lo que efectivamente hacen.

La evaluación de los investigadores mexicanos no ha permitido todavía una adecuada evaluación del papel de la ciencia en México. No se han desprendido de la evaluación de los científicos mejoras en los presupuestos asignados, en las instituciones, no hay plazas nuevas para graduados, no hay repatriaciones. El lugar actual de la evaluación en el proceso de producción científica en el país fue calificado como excesivo, irritante y aparentemente inútil.

EL PAPEL DE LA "COMUNIDAD CIENTÍFICA"

Otra situación claramente identificada y puesta sobre la mesa de discusión por los investigadores de las áreas sociales, fue que la idea precisa de cómo evaluar la investigación no la dio el gobierno mexicano sino que la propuso la "comunidad científica", en particular un cierto liderazgo de la misma, con la intención de sacar a los académicos del famoso "modelo salarial de educación superior" que en la década de los ochenta, y como respuesta a las continuas huelgas declaradas por los sindicatos de las instituciones de educación superior de todo el país, fijó por igual los salarios a todas las instituciones de educación superior. La finalidad era asignar a los investigadores ingresos adicionales, a cambio de la certificación de su mayor preparación y la evaluación de su productividad científica. La propuesta también tenía ingredientes atractivos para los empleadores públicos: se trataba de un estímulo, no de un incremento salarial, sin repercusiones sobre prestaciones presentes o futuras y que podría ser suspendido según el decreto de creación del Sistema, o por ausencia de recursos o por incumplimiento del investigador.

Resolver entonces los problemas derivados de la aplicación actual de la evaluación de

los investigadores es fundamentalmente un asunto de la propia comunidad y de cada uno de los grupos que la constituyen. El cuestionamiento es a nosotros mismos como investigadores y académicos y va por dos dimensiones:

a) Hasta qué punto somos nosotros quienes estamos violentando nuestro comportamiento académico, nuestras actividades profesionales, nuestras posiciones valorales y éticas con respecto al sentido y significado de la ciencia y nuestras vocaciones para apegarnos a lo que reconocen —y premian con dinero— los instrumentos de evaluación. Hasta qué punto, por las retribuciones implicadas y los incrementos en nuestros ingresos, aceptamos personalmente caer en todos los vicios que denunciábamos derivados de las propuestas ahora vigentes: dejamos de trabajar proyectos de largo plazo, partimos un resultado de investigación en el mayor número de publicaciones posibles, descuidamos los ensayos en los que podemos manifestar posiciones de crítica social, descuidamos el trabajo colectivo, el diálogo con los colegas o con los alumnos, etcétera.

b) La segunda dimensión es hasta qué punto podemos oponernos al modelo, salirnos del juego y oponernos activamente a los mecanismos vigentes de evaluación, con el riesgo, por supuesto, de quedarnos sin los ingresos adicionales que nos reporta el sometimiento y aceptación de las reglas vigentes.



EL SEMINARIO DE EVALUACIÓN DE LA EVALUACIÓN

ARTURO ESCOBOSA
EUGENIO FRIXIONE
LUIS GOROSTIZA
WALID KURI
MICHAEL PORTER
ENRIQUE RAMÍREZ DE
ARELLANO

ESTE NÚMERO DE *AVANCE Y PERSPECTIVA* SE AÑADE A CUANDO MENOS veinte anteriores en los que fueron publicados artículos sobre el tema de la evaluación del rendimiento en investigación y docencia a nivel superior, desde aquel anuncio de “un acuerdo inicial sobre cómo evaluar a los profesores adjuntos de los Departamentos de Física, Química y los del Área Biológica” del CINVESTAV en 1981,¹ al cual sucedió meses después la noticia de la instalación de la primera Comisión de Promoción y Becas de Exclusividad.² Tal frecuencia de contribuciones acerca del asunto refleja que éste ha permanecido como motivo de reflexión y discusión interna por parte de nuestro personal académico. Es más, en los últimos cinco años las manifestaciones públicas y privadas en acuerdo o divergencia con el sistema vigente han alcanzado niveles preocupantes. En cualquier caso, no existe la certeza de que el mecanismo de evaluación implantado hace dos décadas, y hasta ahora tan sólo acrecentado sin introducir en él verdaderas modificaciones estructurales, cuente hoy con un respaldo mayoritario del profesorado para legitimar la continuidad de su aplicación.

Las críticas y razones para la inconformidad son múltiples y variadas, si bien casi todas ellas se refieren a la discordancia fundamental de aplicar un mismo procedimiento de evaluación general para todos los sectores de una estructura científica esencialmente diferenciada no sólo por áreas, sino también por disciplinas e incluso por especialidades dentro de una misma disciplina. Estas diferencias derivan, entre otras características particulares, de los tipos de productos esperables, de sus ritmos de producción y difusión, de las clases de impacto o utilidad que puedan tener, de los números de coautores que intervienen en su elaboración, y de los índices de repercusión bibliográfica que alcanzan en análisis publicados periódicamente por organismos extranjeros.

El Dr. Arturo Escobosa es investigador del Departamento de Ingeniería Eléctrica del CINVESTAV. Correo electrónico: escobosa@cinvestav.mx

El Dr. Eugenio Frixione es investigador del Departamento de Biología Celular y del Departamento de Fisiología, Biofísica y Neurociencias del CINVESTAV. Correo electrónico: frixione@cel.cinvestav.mx

El Dr. Walid Kuri es investigador del Departamento de Biología Celular del CINVESTAV. Correo electrónico: walid@cel.cinvestav.mx. Los doctores Luis Gorostiza, Michael Porter y Enrique Ramírez de Arellano son investigadores del Departamento de Matemáticas del CINVESTAV. Correos electrónicos: lgorosti@math.cinvestav.mx, myke@math.cinvestav.mx, eramirez@math.cinvestav.mx

Es, sin embargo, indispensable que una institución dedicada a la investigación y a la formación de investigadores de primer nivel cuente con un mecanismo de evaluación justo, claro y ágil para clasificar a su personal académico. Porque una evaluación bien ponderada, además de recompensar el esfuerzo de los investigadores en función de sus resultados, señala a las nuevas generaciones una trayectoria a seguir en la carrera profesional. Ambos factores se complementan para la construcción y preservación de una estructura académica sana. Por el contrario, un método de evaluación inapropiado, que no reconoce distintas clases de mérito en la investigación, puede llegar a distorsionar la actividad científica misma, como se ha señalado que sucede cada día más con la búsqueda frenética de puntos a través de cualquier medio posible. El problema de retribuir con justicia el trabajo intelectual es pues no sólo difícil, sino crucial para la consolidación y el desarrollo futuro del CINVESTAV.

En reconocimiento de esta realidad, y de que únicamente los propios investigadores pueden diseñar un mecanismo apropiado para la evaluación de su trabajo, un grupo de colegas pertenecientes a las diferentes áreas académicas convino por iniciativa y decisión propias realizar un seminario mensual con el propósito de analizar el problema y examinar posibles soluciones.

Las sesiones se han celebrado con toda regularidad a partir de junio de 2003, con asistencia de entre ocho y diez participantes cada vez.³ Aparte de una discusión progresiva en torno a las principales dificultades del proceso, el Seminario de Evaluación de la Evaluación ha realizado actividades orientadas a recabar información actualizada y relevante al respecto. Entre éstas cabe mencionar una sesión de intercambio de puntos de vista con el Dr. José Mustre de León —Secretario Académico del CINVESTAV—, la compilación de una bibliografía clasificada en la que se enlistan ya más de cien

fuentes de diversos tipos, la celebración de una mesa redonda para el diálogo entre varias instituciones nacionales de educación superior —descrita por la doctora María de Ibarrola en otras páginas de este mismo número—, y una consulta internacional acerca de procedimientos y criterios de evaluación de investigadores, tal como se practican hoy en distintas instituciones científicas en el extranjero.

La consulta mencionada se realizó mediante un breve cuestionario en inglés, distribuido por correo electrónico entre colegas de diversos países, el cual fue respondido por la misma vía. Una de las observaciones más significativas de este ejercicio muestra que en la mayoría de las instituciones, y sobre todo en aquellas de mayor prestigio, la jerarquía académica consta de sólo tres niveles o categorías, a los que se accede principalmente por consideraciones acerca de la calidad más que la cantidad de investigación. En ningún caso se reportó el uso de un sistema numérico de puntuación para la evaluación. Además, en general las respuestas reflejaron que los procedimientos de evaluación son flexibles y se adaptan al contexto específico de cada disciplina.

En conjunto, el trabajo del Seminario ha logrado hasta ahora conciliar parcialmente algunas diferencias entre los puntos de vista habituales de las distintas áreas académicas. Así, por ejemplo, hay consenso en que debiera reducirse el número de las categorías de investigador en la estructura académica del CINVESTAV, lo cual permitiría distinguir mejor entre los respectivos perfiles, además de disminuir la frecuencia de solicitudes por individuo y por lo tanto aligerar la tarea de los colegas que integran las comisiones evaluadoras. Sin embargo, queda aún por definir cómo pudiera instrumentarse el proceso de reducción en número de categorías.

Hay también un acuerdo mayoritario en que debieran existir al menos dos, o quizá tres, ins-

NOTAS

¹ Nota anónima (1981), "Se reclasificó y promovió a 40 profesores", *Avance y Perspectiva* 2: 7.

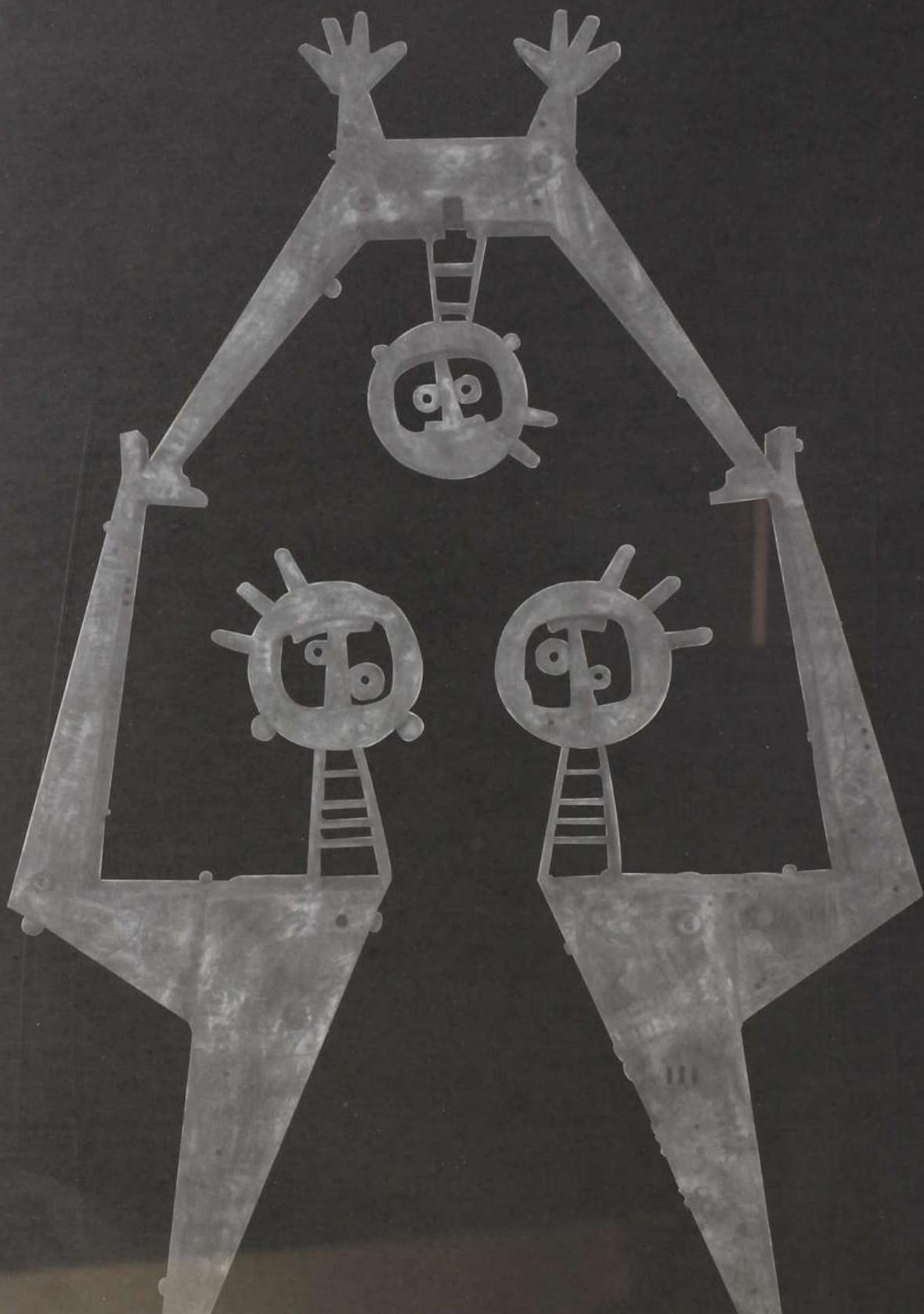
² Nota anónima (1982), "Se instaló la Comisión de Promoción y Becas de Exclusividad", *Avance y Perspectiva* 11: 23.

³ Entre los asistentes más asiduos están Carlos Beyer, Marcelino Cerejido, Arturo Escobosa, Eugenio Frixione, Luis Gorostiza, Carlos Hoyo, Walid Kuri, Bogdan Mielnik, Guillermo Morales, Miguel Ángel Pérez Angón, Michael Porter y Enrique Ramírez de Arellano.

tancias en serie en el mecanismo de evaluación. La primera de ellas pudiera encargarse de revisar el contenido de las solicitudes y verificar que reúnan los requisitos para cada categoría, eliminando así gran parte del trabajo a la siguiente instancia, que se concentraría principalmente en dictaminar sobre los casos difíciles y las promociones para acceder a las altas categorías. En esta segunda instancia participarían especialistas en el área o la disciplina correspondientes a los solicitantes, tanto pertenecientes al propio CINVESTAV como adscritos a otras instituciones. Una posible tercera instancia, que quizá pudiera ser el Consejo Académico Consultivo o un comité representativo del mismo, se ocuparía de vigilar un adecuado balance entre las promociones autorizadas a las diferentes áreas que componen a la institución, pero no tendría relación directa con los solicitantes salvo en casos excepcionales. No obstante estas convergencias generales, persisten algunas discrepancias en cuanto a las funciones específicas y la integración de cada una de las instancias evaluadoras.

No será fácil diseñar un modelo de evaluación de investigadores alternativo al que existe ahora en el CINVESTAV, que rescate las ventajas del modelo vigente y sea a la vez equiparable con los sistemas que califican a nuestros pares y colegas en universidades nacionales y extranjeras. Lo que no puede dudarse es que, de frente al nuevo tiempo que vivimos, será todavía más difícil robustecer verdaderamente la institución sin una evaluación correcta y legítima de nuestro personal académico. Y esto dependerá, más que de las autoridades administrativas, de un diálogo constructivo entre colegas de diferentes áreas del conocimiento.





EVALUACIÓN ACADÉMICA: EMPATE DE DOS IMPOSIBILIDADES

BOGDAN MIELNIK

RESUMEN

SE CONSIDERA LA NECESIDAD DE REFORMULAR LAS REGLAS DE LA EVALUACIÓN pero también se advierte de riesgos de las reformas precipitadas.

1. INTRODUCCIÓN

Todo indica que revisar las reglas de la evaluación académica es como abrir la caja de Pandora. La caja debe abrirse, y no debido a una curiosidad vana. El problema fundamental es que el sistema actual ni puede continuar sin cambios ni tampoco cambiarse a la ligera. Para motivar ambas partes de este diagnóstico trataré de hacer una lista preliminar de todas las virtudes y desventajas de este sistema. Empezaré por el lado oscuro.

2. LOS ESCLAVOS DEL MARKETING

*¿A qué se debe tan increíble facilidad
de multiplicar los reportes burocráticos?
A que los burócratas, en general, no efectúan el
trabajo burocrático; lo ceden a sus víctimas.*

En lo que sigue me voy a concentrar en el esquema de la evaluación de la Copei del CINVESTAV¹, pero discusiones análogas pueden ocurrir en otras instituciones. El centro de la polémica es el sistema de evaluación numérica, es decir, el de “evaluar sin leer” sobre la base del número de trabajos, citas y estudiantes graduados. El problema es que el sistema basado en tan simple *marketing*, con los criterios unificados para todos, se vuelve visiblemente obsoleto.

El obstáculo principal son las discrepancias de los “índices de productividad” entre los departamentos y dentro de cada uno de éstos. Como resultado, la evaluación por medio de indicadores numéricos comunes es bastante injusta: lo que algunos investigadores cumplen

El Dr. Bogdan Mielnik es investigador titular del Departamento de Física del CINVESTAV. Correo electrónico: bogdan@cinvestav.mx

de una manera casi automática, otros lo pagan con inseguridad, estrés y daños de salud. Destaca la crítica situación de los matemáticos: sus trabajos en algunas áreas esperan 2 años o incluso 3 antes de ser aceptados y otro tanto antes de tener repercusión. Cualquier signo objetivo de su calidad científica aparece con un frustrante retraso.

Además de estas injusticias, el sistema numérico induce una visible superficialidad del esfuerzo científico. La prisa de reportar los mejores números desalienta el método fundamental de la ciencia, el esfuerzo paciente de acumular el conocimiento. Forzada por los proyectos burocráticos, la pseudo-productividad intimida los temas arriesgados en los que el tiempo y los resultados de la investigación no se pueden prever de antemano. Y por desgracia, estos temas precisamente son el elemento crucial del progreso científico mientras no lo son los temas con resultados previsibles, ni los proyectos con un "calendario de actividades".

Aumenta el peligro de los trabajos publicados apresuradamente para satisfacer los números requeridos que podrían ser exitosos si estuvieran mejor escritos y más pacientemente elaborados. Abundan a la vez los trabajos irrelevantes que en condiciones normales no deberían ver la luz y que rara vez (o nunca) serán leídos. Existirán en las revistas como archivos muertos, o quizás servirán para que las futuras generaciones reciclen papel.

Crece también las amenazas sobre la seriedad de los grados académicos que otorgamos. El problema de hecho ya es endémico en toda la cadena de dependencias administrativas. La enseñanza en los niveles de secundaria y preparatoria está bajo la presión de reportar éxitos a cualquier precio. Como resultado nos llegan estudiantes de nuevo ingreso etiquetados como "buenos" pero con conocimientos deficientes. A su vez, los centros de educación superior también están bajo presión de aumentar

los números utilizando cualquier medio para lograrlo. El problema no se resolverá gracias a los "doctorados directos", quizá buenos para los éxitos numéricos, pero a costa de la autenticidad. Los coordinadores académicos notan que, a pesar de todo, los procesos doctorales están bastante atrasados pero esto es resultado de la presión, la prisa y de la práctica de aceptar, a toda costa, aun a los candidatos muy mal preparados. Los (cortos) programas de maestría no bastan para compensarlo. Además, las calificaciones de algunos profesores sufren de un exceso de optimismo, por miedo de crear imágenes negativas. Los problemas no son exclusivos de nuestro entorno, últimamente las cosas van muy mal en muchos centros mundiales. Durante mi reciente visita a Varsovia, Polonia, recibimos un mensaje alarmante bajo el título "Publish and Perish", acerca de la degradación de los estudios doctorales en varios países de Europa (¡dándonos unos momentos de satisfacción barata!). Mientras tanto, "la cabalgata del *marketing*" avanza, causando la inevitable *infantilización* de los estudios. Muy pronto los alumnos de las secundarias empezarán a graduar a los bebés en las guarderías infantiles.

El futuro cercano no parece más prometedor. En la cadena de dependencias en la cual cada una presiona a sus subordinadas para reportar más éxitos, ninguna está en condiciones de regresar a la realidad. Así, el CINVESTAV está bajo presión de la SEP para reportar mayor número de publicaciones y de graduados y, a su vez, tiene que presionar a los investigadores. Cualquier intento por escapar del esquema puede causar la derrota institucional. La cabalgata se asemeja a las procesiones penitenciarias de la Edad Media: cada uno imparte latigazos y a su vez los recibe; el proceso no se puede detener sin arriesgar el desastre de la cadena.

A esto se agrega el afán de exigir reportes burocráticos. En ciertos periodos del año las

demandas de reportar causan una verdadera fiebre ejecutiva. El problema se agrava debido a que la burocracia en muchos niveles no efectúa el trabajo burocrático. Todo el sistema está incapacitado para procesar la información acumulada (con frecuencia se piden de nuevo los datos que ya se tienen). En tales ocasiones es muy difícil dedicarse a la investigación; la ciencia se convierte en *El amor en los tiempos del cólera*. ¿Cómo regresar a la realidad? La respuesta no es nada fácil.

3. LA DANZA DE LAS RELACIONES PÚBLICAS

¿Por qué fallan las revoluciones?

¿Por suponer que las soluciones son simples!

¡Y desgraciadamente no lo son! En nuestro seminario sobre la evaluación crecen las demandas de abandonar los criterios numéricos y pasar a la evaluación por expertos, tal y como lo practican muchos centros europeos. De hecho, la *evaluación por pares* ya se convirtió en una frase clave. Sin embargo, la solución no parece ni fácil ni viable. En primer lugar, con todas las críticas presentadas aquí, el sistema actual ciertamente no se reduce a los puros aspectos negativos.

El "reglamento café" de 2002¹ y el blanco de 1996 cumplen el propósito principal de proteger las ciencias básicas (que no son necesariamente teóricas, ¡la buena ciencia experimental es igualmente básica!).

Lo importante también es que el "sistema Copei" logra evitar problemas sociales muy serios, creando a la vez una especie de "paraíso laboral" en donde cada uno se promueve según sus logros medibles y cada promoción tiene consecuencias presupuestales. La valoración basada en los "criterios objetivos" no puede sustituirse por las evaluaciones realizadas por los *referees*, "por pares" y/o por cartas de recomendación. Debido a la situación difícil en algunas áreas

(por ejemplo matemáticas), las opiniones de los expertos inevitablemente van a servir como un método auxiliar, pero no deberían convertirse en el método dominante. Las razones son tanto técnicas como fundamentales.

La razón técnica es el alto número de niveles y promociones que maneja el Centro. Al efectuar siempre la evaluación a fondo (incluyendo el refereo de los trabajos), el CINVESTAV se sumergiría en el paroxismo de las evaluaciones a tiempo completo. La razón fundamental es que las "promociones por pares", o por cartas de recomendación, son por definición procesos con un elemento subjetivo; pueden ser sensitivos a las modas pasajeras e incluso a los cambios de los grupos y personajes dominantes en el ambiente científico. Al menos, no es obvio que el método asegure a los jóvenes la oportunidad de defender sus logros con base en criterios bien definidos.

Ya que "el demonio está en los detalles" trataré de imaginar cómo precisamente trabajaría tal método y cuáles son los peligros.

Las demandas de acudir a los *expertos* suponen que aquellos deben de ser *independientes*. Esto, de hecho, parece indispensable para reducir el papel de los grupos de influencia o de las buenas o malas relaciones públicas en el ámbito local. Sin embargo, la propuesta no explica de dónde exactamente la Copei va a sacar tales expertos. ¿Empezarán los departamentos del CINVESTAV a bombardear a los colegas de otros centros mundiales solicitando que se ocupen de solucionar nuestros problemas de evaluación? Por desgracia, los expertos de hoy están muy sobrecargados, casi enfermos por la prisa. Aun las solicitudes de refereo de las revistas mundiales quedan con frecuencia sin respuesta, o esperan un largo rato (en realidad, el refereo en las revistas ya no está en las de los expertos, sino en manos de los usuarios; es decir, de los autores que en alguna ocasión publicaron algo). Dadas las circunstancias,

la hipótesis de que “los pares” se volcarán a resolver nuestros enredos de promoción ¡parece un tanto surrealista! A veces, si un investigador tiene un jefe influyente y si éste escribe a sus amigos en algunas instituciones mundiales pidiéndoles su opinión, algunos contestarán, pero esto es precisamente el escenario que se debe evitar (¿qué pasará pues con un investigador sin un jefe suficientemente poderoso?). Además, ¿qué tal si un departamento tuviera varios candidatos para ser promovidos: ¿se enviarían todos los currículos al mismo experto internacional? ¿Imagínense simplemente la reacción del experto! En cambio, al consultar a varios, los candidatos frustrados inmediatamente objetarán que los expertos no fueron iguales.

Así, parece inevitable que los intentos de involucrar a “los pares” acabarían con los expertos del entorno local. El obstáculo aquí es la baja saturación del terreno por los especialistas en múltiples áreas. ¿Cuántos expertos tenemos en el CINVESTAV por ejemplo, en teoría de cuerdas? ¿Cuántos en óptica cuántica? ¿Cuántos más en solitones? Inevitablemente, el proyecto conduciría a las comisiones interdisciplinarias de unos cuantos colegas (siempre los mismos), algunos del CINVESTAV, otros de la UNAM, quizás algunos del ININ; algunos de fuera del Distrito Federal, formando un grupo de “alta nomenclatura”, tremendamente abrumado pero concentrando un notable poder interinstitucional. Se abrirían los espacios para la evaluación discrecional por una “dinocracia omnipotente”. Al paso del tiempo, sería inevitable que los miembros de aquellas comisiones se consulten unos a otros en privado acerca de los casos concretos de la evaluación. Tales arreglos abren la posibilidad de varios mecanismos degenerativos. Así, como la siguiente fase del proceso sociológico, no es imposible imaginar un paraíso de relaciones públicas en el cual el ascenso de los jóvenes surge como un intercambio de cortesías entre “influyentes”. Quizá tales consultas se llevarían

a cabo de una manera tradicional y humana, durante alguna comida (¡muy bueno en la industria restaurantera, pero no tanto para la evaluación académica!).

Los antagonismos entre los pares serían más peligrosos aún. Las guerras entre las deidades fueron conocidas en las mitologías egipcia y griega y siempre las víctimas fueron los fieles. El mismo fenómeno destaca en las ciencias políticas, económicas y sociales. Menos en las ciencias naturales y exactas, pero aun aquí los conflictos entre “los grandes expertos” causan que cada uno de ellos menosprecie los logros del otro y de sus alumnos. Las batallas en los mercados de prestigio (los mundiales y los locales) ocurren con tanta frecuencia que sería inútil negarlo: más vale admitir que somos humanos y tratar de sacar conclusiones.

Aun idealizando al extremo, si la evaluación “por pares”, recomendaciones, etcétera, fuese más o menos creíble, en nuestra realidad actual de todas maneras se desatarían profundos antagonismos, grillas, reproches, acusaciones de arbitrariedad o de subjetivismo. La institución, inevitablemente, quedaría sumergida en dolorosos conflictos internos, mucho más intensos que los generados por el actual sistema de la Copei.

Por otro lado, debido a la dificultad de evaluar algunos casos y disciplinas, parece inevitable que las opiniones de expertos van a ser necesarias de una forma u otra. ¿Qué hacer entonces?

4. REFORMAR CON CUIDADO

Propondría proceder con cautela: *¡primum non nocere!* Tenemos sólo un CINVESTAV. Recomendaría reformas cuidadosas, paso a paso, sin fascinación por las revoluciones pintorescas. Me parece que el camino correcto, en términos generales, sería reducir la importancia del número de publicaciones, prestando a la vez mucha

más atención a los logros cualitativos. Estos logros deberían ser visibles no sólo para los “bien informados” o los “expertos competentes”, sino detectables también por un sistema de *criterios cualitativos* cuidadosamente formulados. Así, por ejemplo, si el trabajo de un investigador está discutido en un artículo de un experto mundial, esto nos debe importar mucho más que cualquier carta de recomendación. Esto no excluye las opiniones de los pares o los referencios como método complementario, cuando sea preciso salir de las reglas demasiado rígidas. De hecho, ambos métodos deberían ejercer el control mutuo: la presencia de los expertos impediría la aplicación ciega de los criterios numéricos, mientras que la existencia de las reglas escritas impediría la evaluación arbitraria por parte de los “pares”.

Para facilitar el trabajo no me parece adecuado dividir a la Copei en comisiones departamentales, o aun menores. En particular, la idea de crear los comités evaluadores en cada departamento (¿las Copei minúsculas?) para efectuar la primera etapa de la evaluación, no me parece afortunada (tal método tendría que perturbar el trabajo normal, introducir el estrés y profundizar antagonismos dentro de los departamentos). Sin embargo, la Copei podría trabajar al estilo del SNI, en comisiones interdisciplinarias más amplias. Por ejemplo, “matemáticas, física y control automático” o “química y farmacología”, y otras. Los contactos y discusiones en un amplio gremio interdisciplinario parecen un factor importante para conservar el equilibrio y el sentido común.

Como reformas estructurales recomendaría:

- Reducir la cantidad de niveles. En este momento los departamentos del CINVESTAV tienen entre 9 y 12; me parece que con 6 sería más que suficiente.
- Fortalecer los gremios de la Copei para asegurar la participación de expertos capaces

de evaluar cada área, especialmente en caso de surgir dudas sobre la aplicación de criterios cualitativos. De hecho, no parece acertado que los trabajos de un matemático o físico teórico sean evaluados solamente por dos cansados colegas, uno de física experimental, otro de química, o al revés.

- Calibrar los criterios numéricos de promoción y/o de la beca de desempeño según las normas mundiales en cada área.

Aunque modestos, estos puntos no son fáciles y pueden ser la clave para asegurar una evaluación cuidadosa. En el nivel de las reglas concretas valdría la pena buscar más mecanismos que permitan reconocer y premiar la calidad del trabajo académico. En particular, me parece indispensable:

1. Revisar con cuidado los criterios cualitativos. Reducir el número de los criterios requeridos para cada promoción pero, a la vez, subir el nivel de cada uno para asegurar que cada criterio corresponda a un verdadero éxito que cuenta en el entorno académico. Quitar algunas prescripciones de baja seriedad que reconocen un criterio cualitativo como el equivalente de un trabajo publicado³, y otras que equiparan la dirección de una tesis de maestría con tres criterios cualitativos⁴. ¡El criterio cualitativo debe de ser mucho más que esto! Para fomentar la calidad, los criterios deberían estar bien calibrados, y no siempre lo están. En particular, las instrucciones detalladas para aplicar los criterios necesitan una revisión muy cuidadosa².

2. Aceptar que para promoverse, un científico no necesita ser un “milusos” que presente una increíble vajilla de virtudes (recibe premios, participa en comités organizadores, pertenece a los *editorial boards*, es un *invited speaker* in the *international conferences with payed travel*

expenses, es autor de artículos de revisión, de libros, recibe jugosas becas, pero como dice la canción, todo esto se vale "solamente una vez". Así, por ejemplo, después de publicar una monografía exitosa, el autor debe cambiar de actividad y buscar éxitos diferentes; de otra manera no alcanzará el conjunto completo de siete criterios para promoverse (¡qué suerte que no debe ganar también una competición deportiva!). Las reglamentaciones de este tipo son increíblemente infantiles; lo que premian es un amplio activismo y no los resultados científicos. Por el contrario, me parecen mucho más importantes "las chimeneas" que "las vajillas del éxito". Hay que permitir que cada investigador haga lo que puede hacer bien y me parece muy natural la promoción mediante un solo éxito si éste es bastante notable. La regla de "continuar lo que uno puede" es muy clara en los currículos de los científicos más destacados del mundo (muy pocos de ellos fueron milusos y no debemos exigirlo a nuestros investigadores).

3. Sería útil formular más criterios cualitativos, difíciles de cumplir pero adecuados para varias áreas y especialidades. La tarea de formular tales criterios en 1996 y en 2002 dejó muchas opciones abiertas para reconocer las aportaciones teóricas⁵ y experimentales⁶. En particular, sería razonable ofrecer un mayor reconocimiento a los productos de enseñanza y divulgación de alta calidad. Un buen libro de texto o una exitosa monografía de divulgación publicada en una casa editorial de prestigio no es una cosa trivial y debería representar un fuerte criterio cualitativo.

4. Sin embargo, no me parece prudente abrir nuevamente los caminos de promoción por razones extra científicas (ejecutivo-administrativas u otras). Tales promociones fueron cerradas en 1996, pues se notó que facilitaban un ascenso a los más altos niveles de la jerarquía científica casi sin investigación científica *sensu*

stricto. El fenómeno fue elemental: después de llegar a cierto nivel el investigador empezaba a fungir, dirigir y organizar a otros. Después de un rato recibía un premio; la escalada de puestos, premios y otros reconocimientos lo llevaba arriba hasta el grado 3F, o incluso a 3G o 3H (G=genio, H=héroe) prácticamente sin resultados científicos. El efecto fue recientemente descrito por M. Cereijido⁷, así que podemos hablar del "Efecto Cereijido" (brevemente EC). Nótese, que el EC no se limita a las jerarquías ejecutivas, es un mecanismo universal y se alimenta muy bien con todos los criterios cualitativos por actividades extracientíficas. Así, por ejemplo, el texto de este artículo no pertenece a las publicaciones científicas (ni de investigación, ni educativas) y entonces, no califica para un premio. Sin embargo, si ustedes insisten, entro de inmediato a la convocatoria; ¡Que me den un premio por revelar que no me deben dar premio! Regresando a nuestra realidad cotidiana, me parece que nos acercamos al momento en el que Secretaría Académica debe dar pasos firmes para modernizar todo el sistema. Estas fueron sólo algunas ideas libres que quizá puedan servir a la comunidad.

NOTAS

¹"Bases para la clasificación, promoción y otorgamiento de estímulos al personal académico del Cinvestav", Copei 2002.

²Me limitaré a discutir aquí el criterio cualitativo 4.4. El criterio se ofrece por: "Haber publicado artículos de investigación que hayan sido citados de manera extraordinaria, usando como guía el nivel sobresaliente de citas de la especialidad correspondiente al artículo (ver por ejemplo el *Scientometric Datafile*"). La instrucción detallada (Reglamento Copei 2002, Apéndice 3, p. 38) explica: "El criterio se ha cumplido cuando uno de los artículos tiene del orden de 10 veces o más el promedio de citas sobresaliente de la especialidad o con cuatro artículos, cada uno con el número de citas igual o mayor al nivel sobresaliente o con dos artículos, cada uno con el triple o más de citas que el nivel sobresaliente." La regla sufre una alarmante "lógica invertida". El criterio se otorga por un trabajo con número de citas que sobrepasa 10 veces el "sobresaliente" en la clasificación del *Scientometric Datafile* (SD), lo que efectivamente no es tan fácil de lograr. ¡El mismo criterio, sin embargo, se cumple con un menor número de citas (6 veces sobresaliente) para 2 trabajos y con menos citas aún (4 veces sobresaliente) por 4 artículos! Esta última aplicación del criterio abarca ya situaciones muy comunes,

castigando un logro verdadero en comparación con los resultados bastante corrientes. En física matemática se tiene "sobresaliente" = 7 citas por trabajo y es bastante difícil tener 70 citas en total para un sólo trabajo. Mientras tanto, tomando en cuenta que el término "sobresaliente" en el SD se refiere al número de citas por trabajo durante 2 años, no implica nada especial tener 28 citas para 4 trabajos. Puede ocurrir que un investigador tenga un trabajo excelente que coleccionó 60 citas, pero estará aún por debajo de 4.4, mientras que otro acumuló $4 \times 7 = 28$ citas con 4 trabajos durante algunos años. Un resultado mediocre que será premiado, mientras que el auténtico éxito, 60 citas por una obra, estará todavía por debajo del umbral. Sugeriría recalibrar el criterio en la dirección opuesta para premiar los picos del éxito en vez de las distribuciones planas.

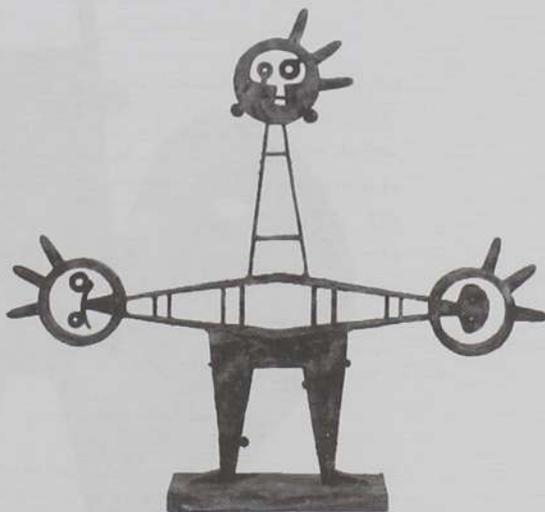
³Ver e.g. "Investigador 3-A", punto 1, Copei 2002, p.18.

⁴Ver e.g. "Investigador 3-C", punto 2, Copei 2002, p.19.

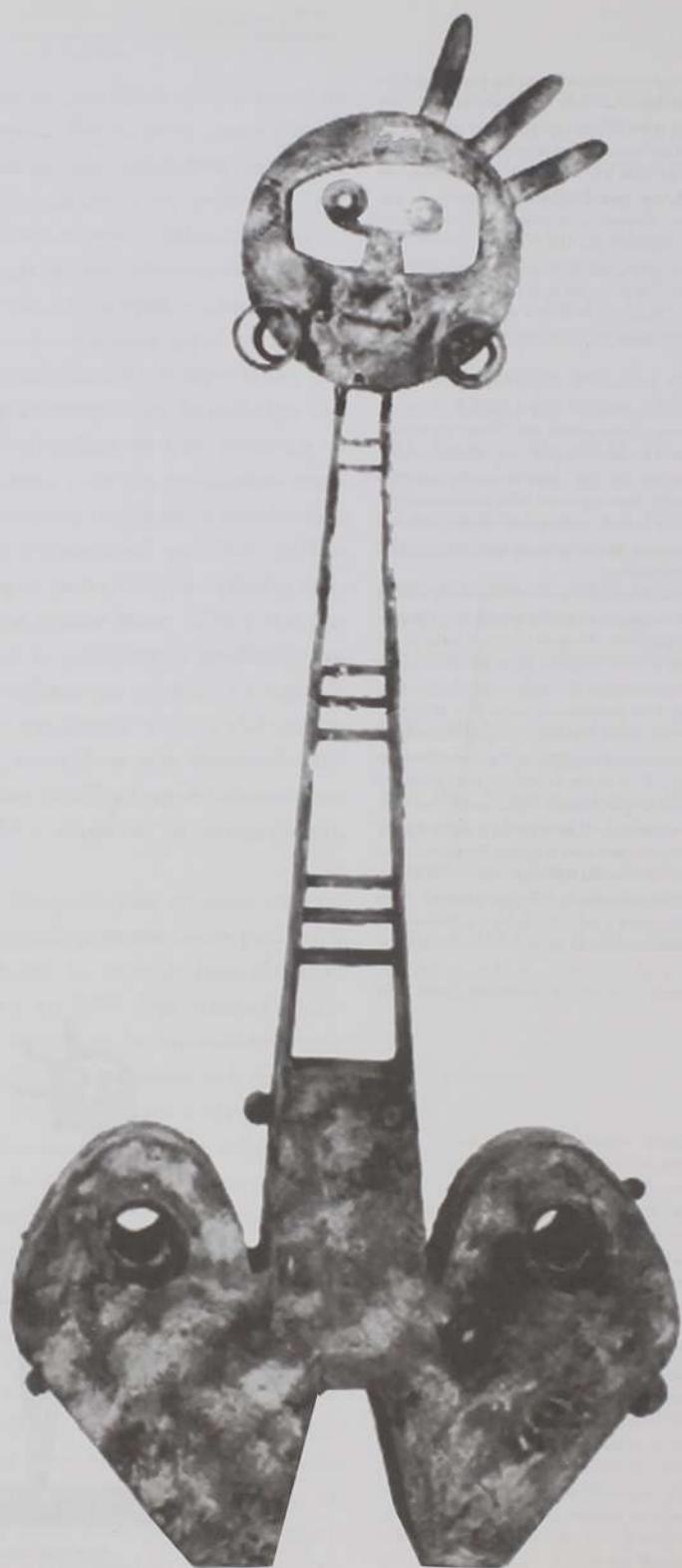
⁵Uno de los criterios muy obvios podría ser: "Tener un trabajo en una revista de muy bajo factor de impacto pero muy altamente citado" (tener cuidado que la barra sea alta, por ejemplo, un trabajo en la *Revista Mexicana de Física* con 40 citas, ya es un éxito internacional!). Otro ejemplo: "Tener un trabajo que sobreviva citado durante muchos años" (otra vez vale cuidar que tenga muchas citas en total; esto sería un criterio para los investigadores maduros).

⁶Para los colegas de los laboratorios experimentales que publican artículos con gran número de coautores es indispensable basar la evaluación en algunos principios genéricos. No en el simple cálculo de publicaciones y citas, tampoco en la falta temporal de las publicaciones y citas, pues todo esto no permite distinguir sus logros individuales. (Las comisiones académicas evalúan a las personas concretas y no al laboratorio como tal). El método óptimo hasta hoy día es un problema abierto, pero parece que deberá basarse en las contribuciones e ideas personales importantes para el laboratorio. Con el fin de evaluar tales éxitos, la Copei debería tener acceso al *ranking* de méritos y prestigio de los individuos y/o subgrupos en la comunidad del laboratorio y de las contribuciones consideradas de valor especial para todo el gremio. Para acreditarlo, los documentos de las discusiones internas deberían tener valor decisivo, con o sin publicaciones multitudinarias de la "persona moral". Sería también preciso prestar más atención a las actividades y publicaciones independientes, especialmente si enriquecen conocimientos de la institución de adscripción.

⁷M. Cerejido, *Recomendaciones para la evaluación*, junio 2004 (no publicado).



"LAS TRES GRACIAS" bronce a la cera perdida



¿MEDIR PARA CONOCER O PARA REGULAR?: EVALUACIÓN Y ACREDITACIÓN DE LA EDUCACIÓN SUPERIOR EN MÉXICO

SYLVIE DIDOU AUPETIT

EN MÉXICO, LA DÉCADA DE LOS 90 HA SIDO LA DE LA EVALUACIÓN Y EL aseguramiento de calidad en educación superior, tanto en la óptica de las políticas públicas como en la de la investigación. Es larga la lista de artículos y libros sobre la contribución de esas medidas a la reforma, (o incluso a la mutación del sistema de educación superior) sobre sus impactos en los establecimientos y sobre sus repercusiones en las prácticas académicas. También son numerosas las reflexiones sobre su papel como instrumento de gestión, principalmente en materia de asignación financiera y de redistribución del prestigio, sea institucional o personal. En todas esas perspectivas, la evaluación y la acreditación han sido abordadas en tanto vectores de cambios e interpretados desde diversos esquemas de valoración.

¿Qué se puede añadir a toda esa literatura, sea gris o referenciada, a veinte años de que iniciara el proceso? No sólo parece que todo ha sido dicho sobre esos temas. Parece también que, en los últimos años, el tópico despertó menos interés. Si bien algunos investigadores siguen cuestionando la forma como opera la evaluación, se está ampliando un asentimiento sobre el principio de su necesidad. En consecuencia, recientemente la atención de los especialistas se ha desplazado hacia otros objetos: las competencias, el mercado, el bien público y las nuevas tecnologías.

Más allá de señalar esa pérdida de relevancia discursiva —la cual no implica un arrinconamiento estratégico de la política— ¿qué se puede todavía aportar al debate? Recientemente, diversos grupos de investigadores han explorado un tema que remite indirectamente a la evaluación y, a la par, enriquece su lectura, el de la medición. De allí estas breves (y primeras) consideraciones sobre la evaluación y la acreditación, como soportes de una medición que aporta luces y crea sombras.

La Dra. Sylvie Didou Aupetit es investigadora titular del Departamento de Investigaciones Educativas del CINVESTAV. Es titular de la cátedra UNESCO-CINVESTAV sobre «Aseguramiento de calidad y nuevos proveedores de educación superior» y actualmente en el marco de un año sabático, es investigadora invitada en el Instituto de Investigación y Desarrollo, París Francia. Correo electrónico: didou@cinvestav.mx

MEDIR PARA EVALUAR, EVALUAR PARA MEDIR: DOS NOCIONES VINCULADAS

Hace dos décadas las autoridades empezaron a instalar dispositivos orientados a corroborar la calidad de la educación superior. Con base en un sistema de inducción financiera amarrado a la posibilidad de tener acceso a fondos de apoyo adicionales a los presupuestarios, crearon programas que sirvieron para canalizar diferencialmente recursos a las universidades públicas en función de la coherencia de sus proyectos de desarrollo, pero también a los investigadores y docentes, según su productividad. En paralelo, el gobierno apoyó la formación de asociaciones civiles encargadas de acreditar la calidad de los programas de estudio o de los procesos de gestión, así como de certificar las competencias de los actores, sean académicos, estudiantes o egresados.

Inicialmente, la apuesta por procesos evaluativos de diversa naturaleza fue estrechamente ligada a las elecciones de apertura (económica primero, y luego social y política) del país. Para legitimar su adopción, en una fase inicial, las autoridades, secundadas por parte de los investigadores, se refirieron al Tratado de Libre Comercio de América del Norte (TLCAN) y a las disposiciones sobre la movilidad de las personas de negocios —categoría que incluía a los profesionistas. Al poco tiempo, sin embargo, quedó claro que la integración macro regional, tantas veces citada, operaba más como pretexto que como motivo. En un sistema de educación superior cuya articulación era deficiente en lo relativo a la estructuración de la oferta de formación, al desarrollo de las capacidades de investigación, a sinergias internas, a costos y a cobertura, la evaluación y la acreditación fungieron como soluciones técnicas, diseñadas para incidir en la organización de las instituciones y en la del sector público de educación superior. Rápidamente dejaron de ser legitimadas en referencia a la coyuntura internacional para serlo en un contexto

nacional de conducción, gestión y pilotaje que, en su enunciación misma pero también en sus instrumentos, expresaba una voluntad de modernización, vía la regulación y la diferenciación de los apoyos.

Por lo tanto, evaluación y acreditación impulsaron esencialmente el reordenamiento del sistema de educación superior. Sirvieron para exigir responsabilidad (*accountability*) a las instituciones y a los individuos más que para alimentar una reflexión en profundidad. Los conocimientos producidos vía la producción de indicadores legitimaron modos de intervención inéditos en la educación superior: auspiciaron la transformación de actores ubicados hasta fechas recientes en lugares periféricos en agentes centrales, cuya visibilidad y relevancia estribaban en la producción de arbitrajes razonados, con base en mediciones, las cuales a su vez justificaban nuevos arreglos, en lo que concernía a la ocupación de posiciones jerárquicas por las instituciones y los individuos en el campo de la educación y en el interior de las comunidades allí asentadas. Para ello, indujeron también comportamientos diferentes en los académicos, entre la participación de facto y el rechazo argumentado.

MEDIR PARA EVALUAR LOS DESEMPEÑOS Y LOS CONOCIMIENTOS DE LOS INDIVIDUOS

Los dispositivos complementarios de evaluación y de acreditación, aunados progresivamente unos a otros, constituyen en México uno de los sistemas más complejos y diferenciados de aseguramiento de calidad en América Latina: abarca actualmente individuos, programas, procesos de investigación, de docencia e incluso de administración. Lo animan agencias gubernamentales, asociaciones civiles, agencias y organismos internacionales, especializados en la medición, vía la supervisión de calidad institucional o disciplinaria y la certificación de las competencias o

de los conocimientos individuales. Convergen en su operación agentes públicos y privados, revelando por cierto esas actividades, en una forma ejemplar, la hibridación e interpenetración de lo público y de lo privado, mediante la delegación a intermediarios de funciones que tradicionalmente eran inscritas en el ámbito de la acción pública así como la rápida (y poco documentada) emergencia de un sector mercantil de otorgamiento de constancias de calidad.

En 1984, la evaluación arrancó en forma suave, con la inauguración del Sistema Nacional de Investigadores (SNI), abocado a elevar el ingreso, vía compensaciones no salariales, de los investigadores más activos de tiempo completo. Su productividad era medida con base en los artículos, capítulos y libros publicados, en el número de estudiantes dirigidos y/o titulados, en el de cursos impartidos y el de congresos en los cuales se presentaban ponencias. Al paso del tiempo se mundializaron los circuitos de producción y de consumo de la investigación, los criterios de publicación, sobre todo, fueron reinterpretados en función del idioma (con predominancia absoluta del inglés en todas las disciplinas, de las ciencias duras a las humanidades), el tipo de casa editorial y la participación en redes internacionales. Reforzado prontamente por dispositivos institucionales convergentes, el SNI funcionó durante la mayor parte de su historia, asignando puntos a productos específicos en detrimento de un análisis cualitativo de un recorrido evolutivo. Con menos recursos y quizá menos visibilidad, medidas similares fueron destinadas a los docentes. En una reflexión sobre la medición como base de la evaluación, la decisión política de no revalorización directa del salario no plantea tantos problemas como la elección técnica relativa a la valoración de una trayectoria intelectual en función de productos precalificados. En efecto, ¿qué es lo que se mide por esta vía: la relevancia de los aportes individuales o la estandarización forzosa de las pautas de producción a modelos

preestablecidos? En otras palabras, ¿cuáles son las incidencias de un dispositivo que aprecia la adecuación a un patrón de exigencias y deprecia, penalizando al individuo en sus ingresos y en su prestigio, cualquier variación ante ésta?, ¿no se estará hipotecando el futuro, al elegir inhibir la posibilidad de recorridos atípicos y premiar el conformismo?

Numerosos trabajos sobre los sistemas de incentivos indican que, más allá de lo pretendido, esos sirven para resolver conjuntos de problemas, de características organizacionales, más que para mejorar la calidad de la producción intelectual de los individuos, en una contradicción flagrante entre lo dicho y lo hecho. Permiten imponer reglas legítimas de profesionalización, reconociendo, de manera diferencial, compromisos laborales ejercidos también en forma diversa. No olvidemos que en México fueron instalados durante la década perdida de los 80 cuando el lema era: "Si hacen como que me pagan, hago como que trabajo". En ese contexto, por el lado positivo, apostar por un sistema de evaluación basado en mediciones contribuyó fuertemente a redefinir las exigencias mínimas de la profesión, planteando las bases de una carrera académica y justificó la identificación de nuevas elites así como los sesgos en los ingresos y en las trayectorias, dentro de un mercado académico al cual no se accedía (y todavía no se accede) en una forma que, en sí misma, otorga legitimidad profesional y social a sus integrantes.

Por el lado negativo, los sistemas de incentivos midieron a partir de la duda y de la sospecha, que el ingreso a la profesión dependía en muchos lugares de criterios laxos de selección. Todavía hoy, como regla general, un académico sólo obtiene legitimidad, ante sus pares y sus patrocinadores, si es parte de los sistemas de incentivos, los cuales comprueban, más allá de la pertenencia a la profesión, la inclusión a un clan de elites, nunca asegurada del todo, y recurrentemente reconfirmada. Así, la evaluación

de los individuos no implicó fundamentalmente mejorar las condiciones para la producción intelectual sino imponer orden en medio de la anarquía. A quince años de haber arrancado, si bien ha sido útil para designar grupos modelo, es insuficiente para alimentar una reflexión innovadora sobre el estado de la ciencia en el país y sobre las condiciones de trabajo de las comunidades científicas. Hoy, agotada su función de delimitación de nuevas elites intelectuales y de regulación, crea una cortina de humo más de lo que otorga un significado al quehacer académico en un país como México.

En paralelo a los procesos de evaluación de los académicos, definitorios en las oportunidades de llevar a cabo estrategias de desarrollo profesional, se miden también los conocimientos adquiridos por los estudiantes, los cuales, según las instituciones, están obligados u optan voluntariamente por presentar los exámenes, diseñados y aplicados por el Centro Nacional de Evaluación A.C. (Ceneval) en diversos momentos del ciclo escolar. Pese a las críticas de que han sido objeto los cuestionarios de opción múltiple, el número de sustentantes de esos exámenes aumentó constantemente, debido a decisiones institucionales y a elecciones personales abocadas a mejorar las condiciones de ingreso a los mercados de trabajo. Gracias a ello, el Ceneval dispone ahora de información de primera mano sobre los conocimientos efectivos de los alumnos, *versus* los que teóricamente deberían dominar, según su institución, su condición de género, su origen social e incluso étnico. A pesar de que existen todavía algunas informaciones reservadas, principalmente sobre la distribución de los resultados obtenidos por los alumnos en función de sus establecimientos de adscripción, dentro de una política reciente de información pública, el Ceneval contribuyó a formalizar, en cada una de las disciplinas que cubre, la definición de competencias mínimas exigibles, propuestas por comisiones de pares.

MEDIR PARA ACREDITAR LA CALIDAD DE LOS PROCESOS INSTITUCIONALES

A la par de esos esfuerzos de medición de desempeños para evaluar la calidad de los procesos de investigación y de enseñanza, a través respectivamente de la producción y de los aprendizajes de los individuos, otros organismos dieron cuenta de la calidad de los planes de estudio, de la licenciatura hasta el doctorado, mediante diversas fórmulas de evaluación o de acreditación, fundamentadas en combinaciones variadas de evaluación experta y verificación de cumplimiento de indicadores. Independientemente de la pertinencia de los esquemas y criterios seleccionados, su significado más relevante estriba en las transformaciones inducidas en el perfil de la oferta de formación, en las elecciones estudiantiles de carrera, en la organización de las instituciones y en la estructuración del sistema de educación superior. En los establecimientos, la necesidad de cumplir con los requisitos de responsabilidad, de eficiencia y de transparencia implicó la revaloración, como departamentos estratégicos, de las instancias de planeación y de información institucional; en el interior del segundo, *"la montée en puissance"* de actores externos, asociaciones profesionales y agencias de acreditación, hizo estallar la interacción binaria, tradicionalmente establecida entre el gobierno y las instituciones. En conjunto, la producción de indicadores de calidad permitió a una parte de los usuarios documentar sus decisiones de ingreso, en función de las ventajas diferenciales brindadas por las carreras acreditadas o insertas en padrones de excelencia, aun cuando todavía una mayoría de ingresantes lo hace conforme a una lógica de proximidad, más que con una lógica informada de apuesta a la excelencia.

Por ende, es preciso aceptar que el dispositivo de monitoreo de los desempeños, pese a estar legitimado en dos nociones, la de servicio

público (y, en consecuencia, de gasto público) y la de defensa del consumidor, ha tenido como resultado principal impulsar un nuevo modelo articulado de gestión, de distribución de los flujos de recursos humanos y de reorganización de la profesión docente. Por ese motivo, no se ha reflexionado tanto sobre la medida en sí, sus aportes y sus disonancias sino sobre sus impactos.

CONSIDERACIONES FINALES

La falta de reflexión acerca de los indicadores, de sus implicaciones en la autodefinición de las elecciones de desarrollo, individual, institucional y sistémico y de las representaciones implícitas de calidad que vehiculan, no contribuye a explicitar la heterogeneidad de los proyectos y de los objetivos: por esa razón, la medida funge más como pretexto político que como procedimiento científico, más como instrumento de administración que como herramienta para producir conocimientos. Remite esencialmente a una voluntad de poner en marcha mecanismos de conducción flexibles, adecuados para el Estado en términos de variabilidad de sus compromisos de financiamiento a largo plazo.

Sin embargo, auspició cambios estructurales. Hoy día, los organismos de medida y de evaluación representan organismos *brokers*, ejercen una fuerte influencia sobre el sistema de educación superior, desde una posición de exterioridad y a partir de un *ethos* de neutralidad. Desencadenan procesos complejos de ayuda a la decisión y dinámicas de recomposición. Los usos de sus medidas son no obstante problemáticos, entre confidencialidad y uso público, entre transparencia y opacidad, entre comercialización y regulación. Alimentan dos lógicas de conocimiento de la educación superior, una articulada sobre la investigación y el conocimiento y la otra sobre la planificación mediante la presentación de indicadores. Mucho se ganaría con definir y separar una y otra escala de referencia

y mucho también con reflexionar sobre lo que se mide y lo que se evalúa, tomando en cuenta las diferencias y las contradicciones entre uno y otro proceso. Faltaría ubicar la evaluación y la acreditación, con sus respectivas herramientas, en su justa perspectiva, más allá de las convicciones, abordando la cuestión del aseguramiento de calidad, a partir de nuevos interrogantes:

¿Cuáles son los cambios así documentados, cuáles son los ocultados? ¿Permite el sistema actual de mediciones asir los cambios emergentes en el escenario educativo, los cuales implican tanto una transformación del papel asumido por los actores tradicionales como el surgimiento de proveedores transnacionales o comerciales?

¿Cuáles son los usos de la medición en términos políticos (en cuanto a asignación de fondos, legitimación y estandarización), intelectuales (reforzamiento del comparatismo en la investigación educativa) y económicos (competencias de los recursos humanos móviles)?

¿Cuáles son, en los circuitos internacionales y nacionales de toma de decisión, el lugar y el poder de los organismos encargados de dar fe de la calidad de los sistemas y de las instituciones así como de las competencias y los conocimientos de las personas?

¿Cuál es el costo directo de la medición y de la evaluación, cuando los servicios proporcionados por esta vía están consumidos sobre una base comercial de intercambio? ¿Cuál es su costo indirecto cuando la medición sirve para filtrar el acceso a los financiamientos públicos y por tanto incide en la relación de la institución con la sociedad?

¿Cómo funcionan los mercados de la evaluación?



LA INTERNACIONALIZACIÓN DE LA EVALUACIÓN DE LOS APRENDIZAJES EN LA EDUCACIÓN BÁSICA

EMILIA FERREIRO

LA EVALUACIÓN ES INTRÍNSECA AL FUNCIONAMIENTO DE UN SISTEMA educativo, el cual tiene la responsabilidad de garantizar que sus egresados tengan los conocimientos y las habilidades requeridas por el programa de formación. Siempre ha habido evaluación en el contexto educativo, a todos los niveles. Pero lo novedoso es la *internacionalización o globalización de las evaluaciones*. También son novedosas las *agencias evaluadoras*. La Organización para la Cooperación y el Desarrollo Económico (OCDE), cuyos orígenes tenían muy poco que ver con la educación, ha plegado a sus miembros —y a varios países no miembros— a una evaluación internacional, que ha tenido gran impacto, la cual trataré más adelante. Las autoridades locales (las secretarías o ministerios de educación de los diferentes países) hacen evaluaciones que están en gran medida “modeladas”, o cuando menos influenciadas, por los parámetros internacionales. América Latina no es una excepción.

A nivel de la educación superior hace años que resultan visibles los efectos benéficos, tanto como los desfavorables, de los múltiples tipos de evaluación que se aplican (evaluaciones de instituciones, de programas, de individuos, etcétera) y las variadas instituciones o grupos que funcionan como evaluadores. No es el propósito de este artículo hablar de eso. Por lo que toca a la educación básica la evaluación tiene, en México y en toda América Latina, mucho menos tradición que la evaluación universitaria. De eso trataremos en este artículo.

EVALUACIÓN DE APRENDIZAJES

La Conferencia Mundial realizada en Jomtien, Tailandia, en 1990, marcó un parteaguas en varios sentidos. Primero, porque el Banco Mundial firmó junto a UNESCO un compromiso expresado en la declaración de “Educación para Todos” y en metas a ser conseguidas antes del 2000; segundo, porque a partir de allí el enfoque tradicionalmente centrado en la enseñanza cambió hacia un enfoque en los aprendizajes.

Desde la reunión de Jomtien el Banco Mundial impulsa fuertemente reformas, acciones y proyectos educativos concretos destinados a modificar

la gestión de los sistemas educativos y a garantizar “productos medibles”, o sea, aprendizajes sujetos al escrutinio de pruebas cuyos resultados son expresados por números.

En 1990, en la región latinoamericana, “sólo Chile contaba con un sistema público e institucionalizado que recogía en forma recurrente y sistemática mediciones del rendimiento académico. Diez años después, (...) todos los sistemas educativos cuentan con sistemas institucionales sofisticados de este tipo” (Casassus, 2001)¹.

Actualmente, el interés está puesto en la adquisición de *competencias* más que en conocimientos, y en ese sentido las reformas curriculares tienden en todas partes, incluido México, a formularse en términos de competencias, sin que haya una definición clara de qué se entiende por ello, más allá de un conjunto —que se supone articulado— de conocimientos, habilidades y destrezas capaz de ser transferido a contextos diferentes de aquellos en los cuales se produjo el aprendizaje.²

¿Por qué este cambio de énfasis de programas centrados en contenidos a programas centrados en competencias? Esto tiene que ver con la rápida expansión del conocimiento que permiten los medios electrónicos y la acelerada multiplicación de artículos científicos o técnicos en circulación. Se estima que “la riqueza global de conocimiento acumulado se duplica actualmente cada cinco años” (Brunner, 2001)³. Las revistas científicas han pasado de unas 10,000 en 1900 a más de 100,000 al final del siglo xx.⁴ Es evidente que ningún sistema educativo —menos aún a nivel de la educación básica— puede ofrecer a sus alumnos un conocimiento actualizado. “Aprender a aprender” se ha convertido entonces en un *leitmotiv*.

Esta sobreproducción de conocimientos ha llevado a las grandes agencias (particularmente el Banco Mundial y la OCDE) a la convicción de que todos los conocimientos tienen fecha de caducidad, como productos del supermercado. Puede ser cierto del conocimiento tecnológico en ciertas áreas; no es cierto del conocimiento científico.

Darwin, Durkheim, Einstein y tantos otros continúan alimentando las discusiones científicas. Los grandes constructores de esquemas interpretativos de amplio alcance (que llamamos, en sentido estricto, “teorías científicas”)⁵ no son muchos y tienden a perdurar. Incluso pensadores anteriores al calendario cristiano tienen vigencia: Aristóteles es el punto de partida ineludible para hablar de lógica o retórica y Euclides sigue siendo referencia obligada para hablar de geometría. Entonces, no todo tiene fecha de caducidad, a pesar de la tendencia hacia el conocimiento “inmediatamente aplicable y rápidamente perimible” propiciado por las grandes agencias financiadoras.

Curiosamente, no es una reflexión epistemológica sino práctica la que produce una extraña coincidencia: buscan las competencias básicas, transferibles a distintos dominios, quienes apoyan una formación de base sólida, que garantice aprendizajes futuros, y también las buscan quienes están convencidos del conocimiento perimible y quieren garantizar que la gente siga aprendiendo (el *slogan* de “aprendizaje a lo largo de la vida” está incorporado a todos los discursos de las grandes agencias, lo cual no es ajeno a la inestabilidad del mercado laboral). En muchos discursos de estos últimos se trata al aprendizaje como si fuera “chips” que se cargan y descargan a voluntad, y como los seres humanos no venimos equipados con una tecla *delete*, algunos incluso hablan de la necesidad de entrenar *forgetting abilities*.

¿A QUIÉN SE EVALÚA?

Mientras los programas educativos enfatizan con progresivo énfasis las actividades grupales (discusión para la preparación de actividades y la toma de conciencia de los logros obtenidos) y subrayan el interés de obtener productos grupales (como preparar el periódico mural o una página web), la evaluación sigue siendo individual. Se evalúan los logros de aprendizaje de sujetos individuales. No se evalúa la oferta educativa como tal. Hay en esto

una profunda injusticia. Es como evaluar el efecto de cierta ingesta alimenticia, sin garantizar que dicha ingesta haya tenido lugar.

Evaluar la oferta educativa, o sea, evaluar las situaciones efectivas que favorecerían la obtención de una determinada competencia, supone, en principio, emitir dudas sobre la calidad de la intervención educativa, y nadie quiere hincarle el diente a ese problema, que se aborda de manera sumamente tangencial. Resultado: tenemos evaluaciones individuales hechas con toda independencia de la evaluación de las condiciones que permitirían acceder a dichas competencias. Esto es así, sin que ello involucre un pronunciamiento con respecto al contenido mismo de las evaluaciones, sobre cuya pertinencia y calidad es incluso posible coincidir, como veremos enseguida.

El aprendizaje es una relación tripartita entre un cierto contenido “objeto de aprendizaje”, un aprendiz y un adulto encargado de organizar las condiciones didácticas específicas para hacer posible ese aprendizaje. No corresponde decir *adulto encargado de transmitir el conocimiento* porque no hay ninguna “transmisión” directa. A lo sumo, el adulto puede “presentar” el contenido a ser aprendido, pero el aprendizaje es apropiación y la apropiación supone una actividad cognitiva muy diferente de la pura recepción. Los seres humanos no somos “portadores de conocimientos” como podemos ser portadores de determinados virus, sin saberlo y a pesar de nuestra voluntad.

Siendo una relación tripartita, la responsabilidad del aprendizaje no puede ser exclusivamente del alumno. Sin embargo, a partir de los resultados de aprendizaje se pueden emitir dudas acerca de la calidad de la oferta educativa a menos de suponer, por ejemplo, que todos los niños de determinado país, con resultados muy pobres, son igualmente “pobres de espíritu” (como se decía antes), infradotados (como se dijo en algún momento) o con débiles capacidades de aprendizaje, como se diría ahora. Es evidente que ningún país puede permitirse el desliz de afirmar que bajos

resultados de aprendizaje reflejan falta de capacidad de aprendizaje de su población infantil, porque eso supondría hipotecar su propio futuro (además de no ser “políticamente correcto” en época de atención preferencial a las minorías, del tipo que ellas sean).

Los únicos candidatos, pues, a ser señalados como responsables de los bajos resultados son, en primer lugar, la institución escolar en tanto institución encargada de producir aprendizajes y la familia —o los grupos sociales de pertenencia— caracterizados eventualmente como “deficitarios”. La visión empresarial vigente tiende a atribuir responsabilidad a la institución escolar, en lugar de diluirla en agentes mal identificados sobre los cuales no es posible incidir de manera directa. La institución escolar requiere asignación de recursos financieros y éstos pueden variar según los resultados que sus unidades exhiban.

¿QUIÉN EVALÚA?

Por supuesto, los sistemas educativos han practicado siempre la evaluación del aprendizaje de los alumnos. A menudo, con pruebas cambiantes de un periodo presidencial al otro y con resultados conservados bajo llave, para “insumo interno”.

Los cambios actuales son radicales: una agencia internacional organiza evaluaciones de tal manera que una misma prueba es administrada a alumnos de un número importante de países y los resultados se publican en la página web de dichas agencias, involucrando a la prensa escrita en la difusión de los resultados. Los países se ven entonces como participantes en una especie de *ranking* donde aparecen como mejores o peores que sus vecinos en términos de puntajes globales, o de puntajes desglosados: el sector público con respecto al privado, escuelas urbanas con respecto a rurales y muchas otras relaciones de este tipo.

La OCDE se ha convertido en una de estas grandes agencias de evaluación. La razón por la cual dicha agencia —lejana de la educación en sus

finés originales— aparece encabezando estudios en este campo ni siquiera requiere justificación: “Como respuesta a la necesidad de evidencias del desempeño estudiantil comparable internacionalmente, la OCDE introdujo el programa PISA (Programme for International Student Assessment)”.

PISA se aplica a jóvenes de 15 años, al final de la escolaridad obligatoria, cualquiera que sea el tipo de escuela secundaria al que asistan. “PISA se concentra en temas que los jóvenes de 15 años necesitarán en el futuro y busca evaluar lo que pueden hacer con lo que han aprendido”. Los contenidos corresponden a las áreas de lectura, matemáticas y ciencias.

Se trata de una prueba que se aplica en los mismos países cada tres años, con un enfoque dominante en cada aplicación: en 2000 el énfasis estuvo puesto en lectura; en 2003, en matemáticas; en 2006 estará puesto en ciencias; en 2009 nuevamente lectura, y así en adelante. El énfasis en cierta área se expresa en el hecho de que las dos terceras partes del tiempo de evaluación (2 horas) se dedican a esa área.

En la aplicación del año 2000 se involucró a 28 países de la OCDE—entre ellos México— y a otros 4 países, entre ellos Brasil. En 2002 se sumaron otros 13 países a la misma prueba, incluyendo a Chile y Perú.

En 2003 se sumaron otros 3 países, siendo Argentina uno de ellos. Por lo tanto, tenemos un total de 48 países (entre ellos 5 latinoamericanos) que participaron en lo que oficialmente se conoce como “PISA 2000”.

PISA no utiliza como término clave el de “competencias” sino el de “aptitudes”: “El término *aptitudes* se emplea para englobar los conocimientos y las competencias”. “La aptitud se mide conforme a un patrón continuo, no como algo con lo que un individuo cuenta o no”⁴.

A través de un largo proceso de consultas, traducciones y reformulaciones se construyeron situaciones de prueba que apuntan a evaluar la “aptitud” para el dominio en cuestión o las aptitudes que se consideran involucradas. Hay aquí un evidente

desliz en el modo de expresión. Se intenta evaluar o valorar una cierta aptitud a través de determinadas respuestas. Como se les pone un número a las distintas categorías de respuestas, y se realizan operaciones estadísticas con dichos números, se habla entonces de medición⁸.

No están disponibles todos los ítems utilizados para evaluar “la aptitud para lectura” (unos 140), pero los ejemplos disponibles son bastante ilustrativos como para imaginar el resto. En sus propias palabras “la aptitud para lectura se define en PISA como la capacidad de comprender, emplear y reflexionar sobre textos escritos”. Incluye textos en prosa de tipo narrativo, expositivo y argumentativo. Pero también exige interpretar y comparar gráficas y diagramas, lo cual es muy novedoso en este tipo de pruebas.

Los estudiantes deben comprender la información presentada y responder a algunas preguntas o tareas particularmente exigentes. Por ejemplo, se presenta un texto de una página a cuatro columnas, con diagramación propia de revista contemporánea (un texto rodeado por tres recuadros informativos, con diferente color de fondo, todo sin ilustraciones). El título del texto es “Armas científicas de la policía” y trata sobre el uso del análisis genético en la investigación criminológica. Los recuadros agregan información de tipo divulgación sobre ADN, la “credencial de identificación genética” y temas similares. La tarea del estudiante es leer con atención el texto y responder preguntas de comprensión acerca de su propósito y de la información específica presentada. Algunas respuestas requieren elegir entre cuatro opciones, bastante próximas entre sí, cuya diferencia requiere hacer una comparación de la información presentada en el texto principal con la que se encuentra en alguno de los recuadros.

Una lectura comparativa y crítica no es la que se suele practicar en el común de las escuelas. No es extraño que los países de América Latina se hayan ubicado en los últimos lugares de esta evaluación. México ocupó el último lugar en el primer grupo de países que participó en PISA 2000, pero luego Brasil se

agregó, ocupando el último lugar. Finalmente Perú desplazó a Brasil de esta última posición. En los tres países mencionados los resultados de PISA tuvieron repercusiones políticas. En México, el Instituto Nacional de Evaluación de la Educación (INEE) publicó varios documentos para ayudar a leer esos resultados y matizar el impacto de las gráficas comparativas (análisis disponibles en su página electrónica). En Brasil, la comisión de educación de la Cámara de Diputados organizó varias sesiones de debate alrededor del tema del retraso en lectura. Finalmente, Perú fue declarado en estado de "emergencia educativa" al darse a conocer los resultados que le conciernen, como si se tratara de una catástrofe producida por inundaciones o movimientos telúricos.

No es mi propósito discutir esos resultados, sino dar una idea del impacto que tienen estas evaluaciones. La pregunta de difícil respuesta es: ¿A qué se deben esos magros resultados? Vayamos al otro extremo de la escala. El gran ganador de este certamen comparativo fue Finlandia. Y los finlandeses respondieron a la inquietud de los otros países con expresiones muy honestas: "La razón más importante reside en los profesores de educación básica y en su capacitación. En Finlandia, la profesión docente es muy valorada. Son muchos los postulantes, podemos escoger a los mejores, seleccionamos apenas al 10% de los postulantes. En todas las encuestas los profesores de educación básica aparecen como teniendo el mismo prestigio social que los profesores universitarios, los médicos y los abogados". ¿Cuándo podremos decir algo así en América Latina?

OTRAS EVALUACIONES INTERNACIONALES

PISA evalúa expresamente comportamientos deseables, al nivel de un lector competente del mundo contemporáneo. No se preocupa por saber si esas competencias se han enseñado o no. Más bien, trata de incidir, a través de los resultados de evaluación, sobre los países involucrados, para que esas

competencias se enseñen. Como tal, la evaluación internacional tiene un efecto de rebote muy fuerte sobre el currículo y la organización de la enseñanza.

Otra evaluación con fuerte repercusión fue la que realizó UNESCO-ORELAC, a través del Laboratorio Latinoamericano de Evaluación de la Calidad de la Educación (LLECE), en 1997, en once países de la región. Dos diferencias muy importantes con respecto a PISA: se evaluaron a niños de 3° y 4° grado de educación primaria y en este caso las pruebas se basaron en los programas curriculares de cada país, logrando un consenso entre los países después de dos años de negociaciones para construir los ítems de una prueba única¹⁰. Contrariamente al caso de PISA, no se han hecho públicos los ítems tal cual sino categorías genéricas. Por ejemplo, en el caso de lenguaje, se evaluó la identificación de tipos de textos; el reconocimiento de la función de un texto; la distinción entre emisor y destinatario de un texto; la distinción entre una secuencia temporal y causal explícita; y la identificación de concordancias gramaticales y categorías tales como sustantivo, adjetivo, verbo, pero no sabemos a través de cuáles reactivos. La evaluación incluía también una parte de redacción, que finalmente no fue integrada en el informe porque no hubo criterios homogéneos en los distintos países para atribuir puntajes. Por lo tanto, los resultados publicados corresponden a pruebas de lectura.

La sorpresa de estos resultados fue Cuba. Los niños cubanos sobrepasan a todos los otros de América Latina tanto en lectura como en matemáticas y tanto en tercero como en cuarto grado. Las diferencias son tan grandes que no hay manera de formar un grupo con Cuba y otros países. Cuba constituye un grupo por sí mismo. Probablemente estos resultados reflejen más de 30 años de priorización de la educación en los hechos y no solamente a nivel declarativo, así como los resultados a largo plazo de la superación del analfabetismo y el incremento general del nivel educativo de la población.

En todo caso, es de lamentar que Cuba no haya sido incluida en PISA 2000. De todos modos,

el contraste Finlandia-Cuba sirve para mostrar cuán difícil es saber a qué hay que atribuir los buenos resultados en estas pruebas.

OBSERVACIONES FINALES

Los resultados de aprendizaje no tienen un patrón de medida único. Si cambian las pruebas, cambian los resultados. No hay instrumentos consensuados sino instrumentos que se inventan con motivo de cada evaluación, ya sea porque han cambiado los programas de estudio o porque se pone énfasis en nuevas competencias que se estiman necesarias para el futuro (también, por supuesto, porque las pruebas anteriores fueron fuertemente criticadas por algunos sectores sociales con "peso específico").

En Latinoamérica no hay experiencia acumulada sobre las consecuencias del bilingüismo de buena parte de nuestra población autóctona en los resultados de aprendizaje. Las evaluaciones se hacen en español para todos los niños. El informe 2003 del INEE tiene datos reveladores como el siguiente: al comparar los resultados de pruebas nacionales en lectura (alumnos de 6° grado) se constata que "sólo 2.5 % de los alumnos que egresan de las escuelas de educación indígena logran niveles altos de lectura, mientras el 14% de los egresados de escuelas urbanas públicas lo alcanzan" (señalemos, de paso, que ambos porcentajes son preocupantes)¹¹. En el análisis de los resultados del LLECE también se señala la incidencia de las lenguas indígenas en los resultados de pruebas de lectura en Bolivia y México (Perú y Ecuador no participaron en esta evaluación).

En razón del alto impacto social de la difusión de los resultados de evaluación, "la prueba pasa a convertirse en el foco de orientación de los procesos instruccionales en la escuela", dice Juan Casassus, un prominente miembro del LLECE, experto en los problemas de evaluación. Este autor señala, en el artículo ya citado, que "en la medida que se crea una cultura evaluativa, la prueba pasa

a tener un peso mayor que el currículo en la cotidianidad del proceso pedagógico en el aula. En ese momento, la prueba tiende a transformarse, en los hechos, en el currículo real". Si sumamos a esto que las reformas educativas tienden cada vez más a expresar los objetivos en términos de competencias y no de conocimientos, el resultado, según este autor, es que "los currícula, en la práctica, se vacían de contenido". Esta tendencia se acentúa con la periodicidad de administración de las pruebas. Entre las pruebas internacionales, las regionales y las nacionales es difícil encontrar un año calendario en el cual alguno de los niveles de la educación básica no sea sometido a escrutinio valorativo. Preguntarse a qué intervalos pueden administrarse exámenes de evaluación sin interferir en el desarrollo normal de los procesos de aprendizaje no es, pues, una banalidad ya que la evaluación, como lo estamos viendo, no es algo que se hace "por fuera" del proceso educativo sino que constituye una fuerte intervención en el proceso mismo.

Cuando se constatan resultados preocupantes, no es evidente lo que se debe hacer para cambiar la situación. Es difícil comparar resultados de aprendizaje, en sí mismos, entre países que difieren en todos los parámetros posibles de comparación, como ocurre con México cuando se lo compara con los otros países de la OCDE (Canadá, Estados Unidos, Finlandia, Suecia, Noruega, Alemania, Japón, etcétera.). De nada sirve declarar al país en "emergencia educativa". ¿Qué es lo que hay que corregir para estar a la altura de los otros? ¿El salario de los docentes y su preparación? ¿El nivel cultural general de la población? ¿Aumentar las horas de asistencia a clase? ¿Poner computadoras esperando que, mágicamente, este recurso técnico suplante al docente mal preparado y peor pagado? ¿Reducir el número de alumnos por clase? ¿Publicitar los casos extremos, o sea, una escuela que funciona a pesar de un entorno desfavorable, esperando que el ejemplo cunda? ¿Lograr que el sindicato docente incorpore la superación académica como

una de sus demandas? ¿Entregar libros y renovarlos periódicamente? ¿Publicitar el *ranking* de escuelas según el rendimiento de sus alumnos? ¿Hacer investigación didáctica para detectar los procedimientos que garantizan mejores experiencias de aprendizaje?

Lo que seguramente no hay que hacer es aplicar ciegamente las recetas de los organismos internacionales, ésos que actúan como si todo lo supieran de antemano, que ignoran las diferencias culturales, que sugieren escuelas mixtas en Afganistán y aseguran que los 300 días efectivos de clase declarados van a coincidir, en México, con un calendario ideal que hace abstracción de las festividades locales y de los "puentes" nacionales. La evaluación incide en la cultura escolar, pero la cultura escolar resiste. Los mecanismos de simulación también forman parte de la resistencia al cambio. Finalmente, de eso se trata: una institución —la escuela— que ayudó a forjar una nación porque tenía una utopía y la supo perseguir, y que ahora ha perdido el prestigio social y ha perdido los ideales que la vieron surgir debe cambiar, es evidente. ¿Serán las evaluaciones internacionales el agente de cambio? ¿Los docentes se sentirán convocados o excluidos? ¿De dónde vendrá la chispa que los encienda nuevamente y que los ponga a funcionar cambiando y siendo agentes de cambio, pero sabiendo cuál es la dirección y el sentido del cambio?

El *para qué* y el *por qué* de la educación están en crisis porque nuestros países latinoamericanos también lo están, sufriendo una globalización opresora y enajenante. La educación es quizás el lugar de manifestación de esta crisis.

NOTAS

¹ Artículo disponible en <http://llece.unesco.cl/documentosdigitales/>

² Por ejemplo, el nuevo Programa de Preescolar de la SEP 2004 señala que: "Una competencia es un conjunto de capacidades que incluye conocimientos, actitudes, habilidades y destrezas que una persona logra mediante procesos de aprendizaje y que se manifiestan en su desempeño en situaciones y contextos diversos", p.22.

³ Brunner, José Joaquín, "Globalización, educación, revolución tecnológica". *Perspectivas*, vol. XXXI, n°2, 139-156.

⁴ Este incremento en las revistas científicas no tiene sólo que ver con un incremento del conocimiento publicable, sino también con los

requisitos de evaluación a los que estamos sometidos los científicos.

⁵ La palabra "teoría" ha sido banalizada en la tradición anglófona, al menos en ciencias sociales, donde se aplica a micro-teorizaciones de un campo restringido del conocimiento.

⁶ Todas las citas referidas al programa PISA corresponden a la publicación en español "Conocimientos y aptitudes para la vida – Resultados de PISA 2000", co-edición OCDE-Santillana, 2002.

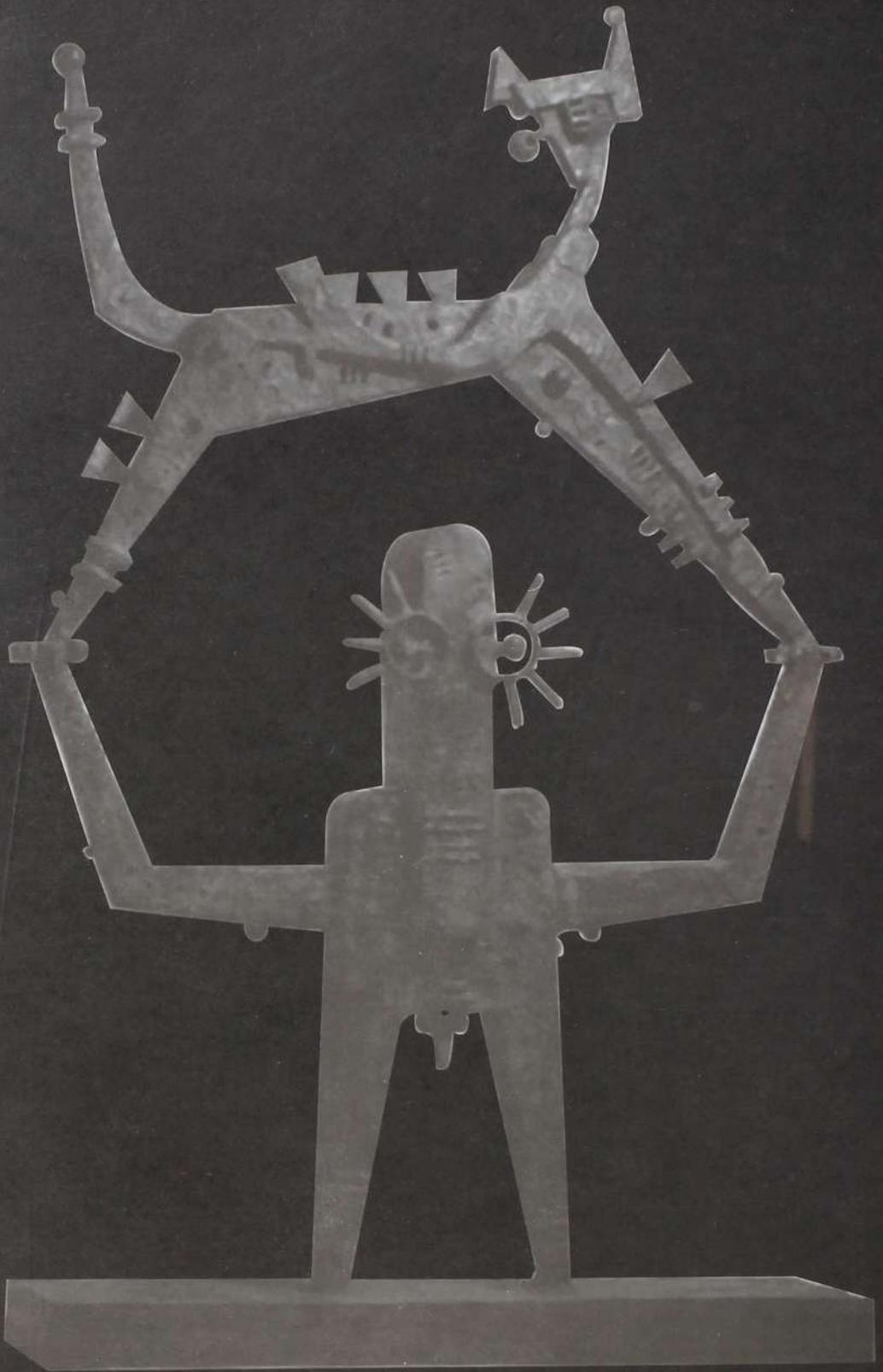
⁷ La variabilidad en las denominaciones, y sus fluctuaciones al pasar del inglés a las distintas lenguas, muestra la ausencia de teorización sobre ese "constructo mental" (competencia, aptitud) al cual se hace vagamente referencia.

⁸ "Inteligencia es lo que mide mi test" decía a principios del siglo XX Lewis Terman, el exitoso adaptador de la primera escala de medición de la inteligencia, construida por Binet y Simon. Como puede apreciarse, seguimos en la misma longitud de onda, un siglo después.

⁹ Entrevista con Jouni Ensio Valijarvi publicada en el periódico *Clarín*, Buenos Aires, 28 septiembre 2003, y reproducida en varios sitios Internet.

¹⁰ Ver en la página web de LLECE el informe con el título "Primer estudio internacional comparativo sobre lenguaje, matemática y factores asociados en tercero y cuarto grado".

¹¹ Instituto Nacional para la Evaluación de la Educación. *Panorama educativo de México – Indicadores del sistema educativo nacional 2003*.



EFECTOS DE LAS EVALUACIONES ESTANDARIZADAS EN LOS SISTEMAS EDUCATIVOS

ANTONIA CANDELA

1. INTRODUCCIÓN

UNA DE LAS PRINCIPALES POLÍTICAS IMPULSADA POR LA ACTUAL Administración educativa federal (la Secretaría de Educación Pública) consiste en promover una cultura de la evaluación como uno de los mecanismos para mejorar la calidad de la educación en México, una forma de garantizar la calidad de la educación, y un requisito para tener la información que permita tomar las decisiones que posibiliten mejorar la educación en todos sus niveles (INEE, 2003). Para ello se ha creado el Instituto Nacional de Evaluación Educativa (INEE) que tiene la tarea de diseñar esta política nacional, realizar pruebas en todo el país, básicamente en educación media y básica, y coordinar los esfuerzos de todas las instituciones que realicen evaluaciones. Aunque éstas siempre han existido en el sistema educativo, la novedad (siguiendo ciertas políticas globalizadoras) es hacer evaluaciones llamadas estandarizadas diseñadas desde fuera del sistema educativo para tener resultados nacionales, e incluso internacionales, que permitan hacer comparaciones sobre la calidad de la educación entre diferentes contextos regionales y socioeconómicos de un país, así como entre varios países y darles una amplia difusión.

Varias son entonces las razones por las que éste se convierte en un tema de gran importancia. En primer lugar, porque se pretende que esta cultura se convierta en una práctica permanente en todos los niveles del sistema educativo, desde preescolar hasta posgrado, evaluando a los alumnos, maestros, escuelas y currículos. En segundo lugar, porque se pretende que la evaluación sea el insumo fundamental para mejorar la calidad de la educación. En tercer lugar, porque la aplicación continua de pruebas y la importancia que se le otorga a la divulgación de los resultados en los medios masivos de comunicación y a la comparación entre escuelas, alumnos, sistemas educativos y países, la anuncian como un factor que muy probablemente va a tener un fuerte y permanente impacto en las prácticas cotidianas de nuestro sistema educativo.

A continuación reflexionaré acerca de las evaluaciones estandarizadas analizando algunos efectos que han tenido sobre el sistema educativo

La Dra. Antonia Candela es investigadora titular del Departamento de Investigaciones Educativas del CINVESTAV. Correo electrónico: acandela@cinvestav.mx

en países en los que éstas se han aplicado ya por varios años, sobre todo aquellos efectos que pueden afectar el buen desenvolvimiento del sistema. El propósito de estas reflexiones es aportar elementos a los actores del sistema educativo para interpretar los resultados de dichas evaluaciones, ubicando en contexto lo que éstos indican. También me interesa analizar cuáles podrían ser algunas de las medidas que permitirían complementar la información sobre los procesos educativos y así contrarrestar algunos de los efectos negativos que estas evaluaciones pudieran tener en el sistema educativo de nuestro país, en donde esta política está empezando a implementarse. Asimismo, revisaré el tipo de información que aportan estos modelos de evaluación comparándola con la información educativa que es necesaria para tomar medidas que mejoren la calidad de la educación. Por tanto, se plantean algunas de las características que debieran tener las evaluaciones para ser un instrumento útil para la toma de decisiones en manos de los propios actores educativos.

2. ALGUNOS POSIBLES EFECTOS DE LAS EVALUACIONES ESTANDARIZADAS A LARGO PLAZO

Una de las formas de prever los posibles efectos de una política a largo plazo es analizar el impacto que ha tenido su aplicación en los países en los que dicha política se ha sostenido por varios años, haciendo las correcciones pertinentes según las diferencias de contexto y de circunstancias. Por esta razón describiremos algunos de los efectos de las políticas de evaluación aplicadas nacionalmente por más de 10 años en Chile, Brasil y Estados Unidos. Los primeros dos países tienen condiciones similares a las de México: México 34° lugar, Brasil 35° lugar, Chile 36° lugar en las pruebas internacionales del Programme for International Student Assessment (PISA), aplicadas por la Organización para la

Cooperación y el Desarrollo Económico (OCDE), mientras que Estados Unidos tiene mejores condiciones: el 16° lugar en los resultados de las mismas pruebas. Asumimos y aclaramos que cuando se quiere evaluar un sistema nacional, las evaluaciones estandarizadas son necesarias pero, como más adelante analizaremos, es imprescindible combinarlas con otro tipo de evaluaciones para tratar de evitar sus efectos deformantes.

a) Efecto de aumento de la marginación

En esta sección el análisis se centrará en lo que podríamos llamar los “efectos sociales de las pruebas estandarizadas sobre la educación”, poniendo atención en algunos riesgos que es conveniente evitar.

En Chile, después de 15 años de aplicación de pruebas nacionales estandarizadas a través del Sistema de Medición de la Calidad de la Educación (Simce), que incluye pruebas de castellano y matemáticas a los cuartos y octavos básicos en los años pares e impares respectivamente, los resultados muestran que hay un mejoramiento de entre 10 y 20% de los conocimientos que adquieren los alumnos, pero aumenta la distancia de los resultados entre las escuelas privadas de los sectores de más altos ingresos y las escuelas públicas a las que asisten las capas bajas de la sociedad (Mella, 1999). Este mejoramiento de la calidad, sin embargo, no ocurre siempre: en Brasil los resultados nacionales han descendido después de 10 años de aplicación de pruebas de matemáticas a nivel nacional en educación básica (De Freitas, 2004). Por otro lado, aun cuando los resultados mejoren, como ocurre en Chile, el incremento en la distancia entre los resultados de las escuelas privadas y las públicas para sectores pobres cuestiona la relación que se ha pretendido establecer entre equidad y educación.

Uno de los mecanismos que produce este efecto de aumento de la desigualdad es el que

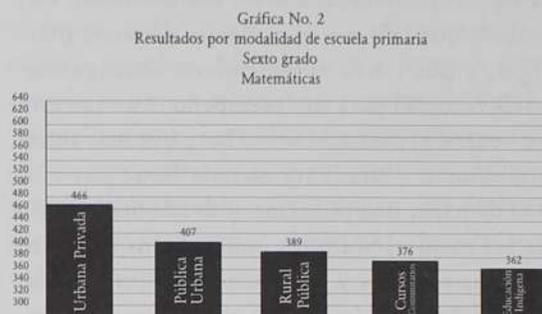
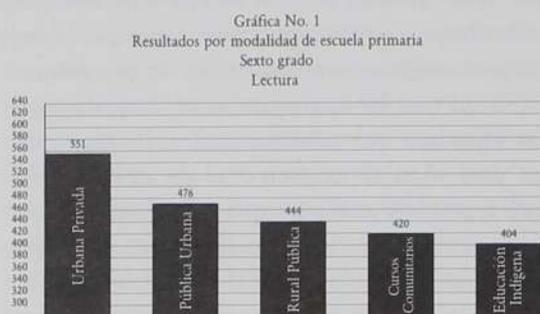
se ha llamado el “Efecto Mateo”. Éste actúa de la siguiente manera: a nivel institucional, una escuela que sale mejor evaluada adquiere prestigio, y por tanto posiblemente obtenga mayores recursos para su desempeño. Aunque esto no ocurra, el prestigio social hace que más alumnos acudan a ella. Esta escuela tiene posibilidad de hacer una selección más adecuada tanto de los mejores alumnos como de los maestros más capacitados y por tanto aumentarán sus posibilidades objetivas de obtener mejores resultados en evaluaciones posteriores. En educación es bien conocido el efecto de la autoestima y de la valoración social como uno de los mecanismos más efectivos para optimizar el desempeño educativo, tanto de alumnos como de maestros. No es casual que en Finlandia, que tiene el primer lugar en las pruebas de PISA, los docentes de educación básica sean reconocidos al mismo nivel que los profesionistas del más alto prestigio social. Por lo tanto, el acrecentamiento del prestigio social actúa también como un factor subjetivo que contribuye a mejorar el desempeño educativo en las escuelas, alumnos y maestros que salen bien evaluados. Por otro lado, el prestigio de las escuelas que obtiene peores resultados disminuye y con ello posiblemente disminuya la autovaloración de sus actores y el estímulo por el trabajo. El efecto continuo de las pruebas contribuye a una selección en la que los peores alumnos o al menos los alumnos que están en peores condiciones para sacar buenos resultados en el tipo de pruebas estandarizadas, se va concentrando en las “peores escuelas” (Pérez Mora, 2004).

A nivel individual el mecanismo de marginación, como ha ocurrido no sólo en Chile sino en Brasil y sobre todo en Estados Unidos, actúa de la siguiente manera. Un alumno que sale mal evaluado es etiquetado como un mal alumno lo que también actúa en su autoestima y en lo que el maestro espera de él (Efecto Pigmalion), de modo que es muy pro-

bable que disminuya su desempeño educativo. Pero también hay un efecto objetivo que contribuye a irlo marginando. Para que la escuela obtenga mejores resultados, un mecanismo habitual que ha aparecido es el de pedirles a los alumnos que salieron mal evaluados que no presenten el examen o el de obligarlos a que repitan el año para que así obtengan mejores resultados (Valenzuela, 2002). Investigaciones realizadas en EU muestran que la reprobación aumenta hasta en un 50% la posibilidad de que un alumno sea expulsado de la escuela (Intercultural Development Research Association, 2001). Así, las evaluaciones externas representan el riesgo de contribuir primero a aumentar la reprobación y después a aumentar la expulsión de los alumnos que obtienen peores resultados en las pruebas estandarizadas.

Si bien las evaluaciones estandarizadas son una necesidad cuando se quiere analizar todo un sistema educativo porque diseñan parámetros (ítems) que buscan comparar el desempeño educativo en diferentes medios, es conveniente considerar las limitaciones que esta comparación tiene y los efectos secundarios que puede tener.

Al construir ítems estandarizados para las pruebas que se aplican a los actores educativos, necesariamente se homologan las demandas desde unas características lingüísticas y desde unas prácticas sociales y culturales determinadas que ponen en desventaja a los sectores que no pertenecen culturalmente a ellas. La mayor parte de estos ítems son diseñados desde parámetros de la cultura occidental de los países a los que pertenece la OCDE. En un país multicultural como el nuestro, los grupos indígenas y muchos de los sectores bajos de la sociedad, como los campesinos pobres, están muy distantes lingüística y culturalmente de estos referentes y por tanto hay que asumir que muy probablemente parten de una posición de desventaja para enfrentarse a las características de pruebas como éstas.



Gráficas 1 y 2: Fuente: Instituto Nacional para la Evaluación de la Educación, *La Calidad de la Educación Básica en México*, México, INEE, 2003, p. 56.

Los resultados de las primeras pruebas nacionales de lectura y matemáticas aplicadas por el INEE en 2003 a una muestra amplia de alumnos de 6° grado de primaria de escuelas urbanas públicas y privadas, rurales, comunitarias e indígenas, muestran esta tendencia.

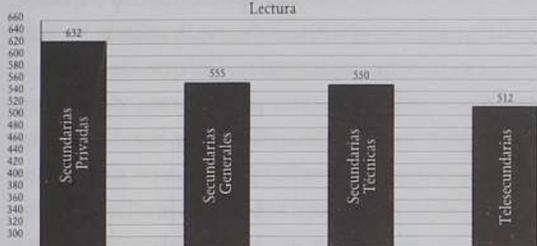
Por tanto, son los sectores indígenas y de comunidades rurales pobres (que de entrada son los más marginados educativamente en nuestro país y cuyo contexto social y familiar no desarrolla las competencias que la escuela valora, como la lectura, o que tienen más dificultades para entender el contenido de las preguntas en español y para responder a exámenes largos por escrito) los que tienen más probabilidades de salir mal evaluados en las pruebas estandarizadas. Esto no es debido a su incompetencia sino a que tienen una lengua y una cultura que no es tomada en cuenta como referente para las pruebas. En países como EU los grupos étnicos distintos de los dominantes, como los latinos y los sectores más pobres de la sociedad, son los que obtienen los peores resultados en estas pruebas (Valenzuela, 2003). Lo anterior nos indica las limitaciones de la pretendida objetividad de las pruebas y la importancia de hacer un esfuerzo por construir parámetros contextualizados a las condiciones sociales y culturales, al menos del país, y por

relativizar los resultados de las pruebas estandarizadas a través de incluir, al menos, el factor socioeconómico en las comparaciones (Zorrilla, 2004; Himel, 2004).

Aparte de los factores ya analizados de desventaja por el origen lingüístico y cultural de los alumnos, existen otros elementos que deben ser tomados en cuenta para interpretar los resultados de las evaluaciones estandarizadas, sobre todo cuando éstas ponen el acento en la evaluación de los aprendizajes de los alumnos. La investigación educativa ha puesto en evidencia desde hace mucho tiempo que el aprendizaje depende del contexto social en el que ocurre (Cicourel, 1974; Bruner, 1984; Coll, 1986); sin embargo, las pruebas estandarizadas tienden a descontextualizar el aprendizaje y evalúan con los mismos parámetros a un alumno cuya escuela tiene las mejores condiciones de enseñanza (buenos maestros, instalaciones apropiadas, etcétera) y a un alumno que estudia en una escuela que no tiene ni los medios materiales ni tal vez los recursos de enseñanza adecuados.

Si la evaluación estandarizada se centra en los aprendizajes logrados por los alumnos, sin tomar en cuenta lo que se les ha enseñado y las condiciones en las que se imparte esta enseñanza, se pone en desventaja a los sectores

Gráfica No. 3
Resultados por modalidad de escuela secundaria
Tercer grado
Lectura



Gráfica No. 4
Resultados por modalidad de escuela secundaria
Tercer grado
Matemáticas



Gráficas 3 y 4: Fuente: Instituto Nacional para la Evaluación de la Educación, *La Calidad de la Educación Básica en México*, México, INEE, 2003. p. 57

que, por diversos factores, trabajan en peores condiciones de enseñanza. De esta manera, se corre el riesgo de culpar a los individuos de los resultados de las pruebas, sin tomar en cuenta la calidad de la enseñanza y los contextos de desigualdad de la oferta educativa que son responsabilidad no de los individuos sino del sistema educativo en su conjunto. Esto es, se culpa a los individuos de que el sistema les proporcione una escuela con menos recursos y condiciones que a otros. Esta situación nuevamente pone en desventaja a los sectores más desfavorecidos económica y socialmente, que son los que frecuentemente tienen apoyos educativos más deficientes.

De esta manera, la aplicación de pruebas estandarizadas o iguales a los desiguales y la comparación entre ellos con base en estas pruebas (a través de lo que se denomina "ranqueo") es un mecanismo que ha conducido tanto, en países del primer mundo por ejemplo EU, como en países con condiciones muy semejantes a las nuestras, como Chile y Brasil, al aumento de la desigualdad por escuela y por alumno. Estos sistemas permanentes de evaluaciones estandarizadas tienen el gran riesgo de actuar en dirección contraria a la equidad educativa al aumentar la distancia entre los sectores que tienen mejores resultados y los que tienen peores

resultados en las pruebas aplicadas (Martinić, 2003; Valenzuela, 2002).

b) Efectos sobre el currículo y sobre los enfoques de enseñanza

En esta sección se revisarán los posibles efectos de las pruebas estandarizadas sobre el contenido educativo. Uno de los efectos ampliamente reconocidos de la aplicación de cualquier sistema de evaluación, sobre todo cuando éste es diseñado desde fuera de la práctica misma, es el del impacto sobre la práctica, que en este caso es sobre lo que se enseña y la manera como se enseña. Por ejemplo en educación superior y en investigación científica es claro el impacto que han tenido los sistemas de evaluación como el SNI al orientar el quehacer investigativo hacia lo que otorga mejores calificaciones. Cuando estos resultados tienen un impacto sobre los recursos que recibe el docente o investigador, y cuando sus resultados tienen una difusión social amplia que conduce al fenómeno de etiquetar a los evaluados e influir en el prestigio social, estos sistemas tienen el efecto de ir modificando las prácticas eliminando aquello que no se evalúa positivamente.

En el caso de la educación, la presión a los docentes para que sus alumnos y ellos mismos salgan bien evaluados tiene el efecto de

Puntajes de los alumnos en Lectura y Matemáticas, según contexto sociocultural de la escuela, 2001

Nivel sociocultural	Lectura	Matemáticas
Muy desfavorable	431.7	445
Desfavorable	435	449.1
Medio	464.7	471.1
Favorable	489.7	487.1
Muy favorable	542.5	519.4

Fuente: Tabaré Fernández, *Contextualización sociocultural de escuelas de la muestra de Estándares 1998-2002, 2003*, en INEE, *La calidad de la educación básica en México, 2003*.

que tiendan a “enseñar para el examen”. Por tanto los contenidos que no son planteados en las pruebas, muchos de los cuales tienen que ver con el contexto local de los alumnos o con prácticas, van siendo dejados de lado. Asimismo, las prácticas que, aunque sean necesarias para la comprensión, no son tomadas en cuenta en las pruebas, pueden tender a no ser consideradas por los maestros si ellos priorizan el obtener mejores resultados en las pruebas a la importancia de que sus alumnos comprendan bien el contenido. El riesgo es que los criterios sobre cómo enseñar y qué enseñar ya no se basen en la necesidad de una mejor formación ni en las necesidades locales de cada sector educativo, sino en la transmisión de los recursos que se requieren para pasar bien las pruebas. Por tanto, un problema importante es que el currículo real en la práctica vaya siendo el contenido y los enfoques que tienen implícitos las pruebas estandarizadas. Cuando las pruebas están diseñadas desde fuera del sistema educativo y no están vinculadas al currículo real y a las condiciones de enseñanza, o sea, al contexto local y a lo que se enseña y cómo se enseña a los alumnos, la aplicación continua de estas pruebas va desplazando al currículo y sustituyéndolo por los contenidos de las evaluaciones.

La construcción de los ítems y de las pruebas se realiza necesariamente desde una cierta concepción

educativa, desde enfoques sociales y culturales de lo que es deseable que la educación promueva y favorezca, pero estas concepciones generalmente no se plantean explícitamente y quedan ocultas en los criterios de evaluación. Es fundamental conocer los criterios que determinan la construcción de los parámetros de evaluación para poder identificar los modelos educativos que estas pruebas promueven y por tanto permitir la reflexión sobre lo que implica tener éxito o fracaso en las pruebas. La información sobre los criterios técnicos, los modelos y enfoques educativos implícitos en las pruebas es indispensable para permitir el debate académico y de política educativa necesario en nuestro país, así como para evitar que se vayan imponiendo los modelos educativos que están ocultos en las pruebas estandarizadas (Manzi, 2004) y se marginen los currículos nacionales o locales y las políticas compensatorias para apoyar a los más atrasados.

Si el sistema nacional de evaluación se basara sólo en este tipo de pruebas su efecto a largo plazo corre el riesgo de ir marginado el currículo nacional por parámetros internacionales que pueden no ser necesariamente los más convenientes para nuestro país. Asimismo, se ponen en riesgo políticas educativas como la equidad de la educación, las políticas compensatorias, las políticas interculturales que están tratando de que las escuelas y nuestro sistema educativo tomen en cuenta la cultura de origen de los alumnos y la valoren como un medio para educar en el reconocimiento a la diversidad y el fortalecimiento de la identidad de los distintos grupos culturales en México.

3. ALTERNATIVAS Y EVALUACIONES COMPLEMENTARIAS

Las pruebas estandarizadas que hasta ahora hemos analizado se basan en cortes diacrónicos del sistema que son como una fotografía que nos aporta cifras agregadas que permiten conocer

el estado del sistema educativo comparando los resultados de sus diferentes estratos y niveles. Sin embargo, estos datos no nos dan información sobre las causas por las que se obtienen ciertos resultados ni nos permiten analizar las características de los procesos educativos que ocurren en las aulas, sin los cuales es difícil entender cuáles son los problemas educativos que llevan a ciertos resultados y saber qué decisiones y medidas tomar para contribuir a resolver estos problemas y, en consecuencia, a mejorar la calidad educativa.

Existen otros tipos de pruebas que son complementarias a las antes descritas y que aportan información relevante para la toma de decisiones dirigidas a mejorar la calidad educativa como son los procesos asociados al aprendizaje vinculados a la cultura escolar y a las prácticas educativas.

a) Evaluaciones de valor agregado

Estudios numerosos de perspectiva etnográfica así como análisis de seguimiento de cohortes de alumnos buscan analizar el efecto de la escuela sobre los alumnos. Se ha encontrado que tanto en EU (Arancibia, 1992) como en América Latina (Schieffelin y Farrel, 1982) y en estudios comparativos entre países ricos y pobres (Heyneman y Loxley, 1982) la diferencia entre cómo entran a la escuela y cómo salen es mucho mayor en alumnos que provienen de sectores pobres. Es decir, estos estudios denominados de "valor agregado" permiten analizar los conocimientos que les aporta la escuela a los niños, cuánto aprendieron en relación a lo que sabían y no en relación a otros niños que tal vez entraron a la escuela en distintas condiciones. De esta manera se compara lo comparable, pues se compara a los alumnos consigo mismos y con otros alumnos y escuelas de las mismas condiciones socioculturales y económicas que las de ellos.

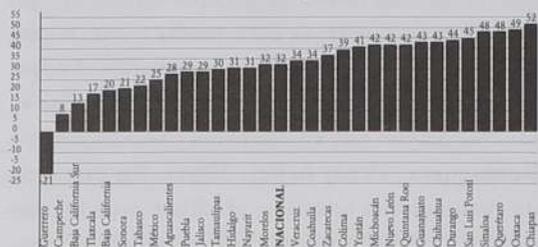
Estas son, por tanto, evaluaciones contextualizadas que controlan los factores socioeconómicos

de entrada que tan fuerte impacto tienen sobre los resultados comparativos a nivel nacional. En este tipo de evaluaciones se encuentra que las escuelas de los sectores pobres son mucho más efectivas que las de clase media y clase media alta pues el cambio en conocimientos y competencias de los alumnos de estas escuelas es mucho mayor que en los otros casos (Mizala y Romaguera, 1998; Bellei, 2001). En este tipo de evaluaciones se refleja el esfuerzo de los docentes y de los alumnos por superar sus condiciones de marginación y se puede notar el reconocimiento y el importante valor social que tiene la escuela para sectores y países pobres como los nuestros (Pérez Mora, 2004).

b) Evaluaciones cualitativas

A diferencia de las evaluaciones que tienden a dar números, y por tanto a medir los resultados de la educación, las evaluaciones cualitativas recuperan el sentido profundo de evaluar como el proceso que nos permite emitir un juicio de valor sobre un asunto determinado y no necesariamente un número. Así, es importante reconocer que evaluar no implica necesariamente medir y que las mediciones son sólo un tipo de evaluación pero no la única ni necesariamente la más importante para darnos información sobre los procesos educativos. A la evaluación cualitativa pertenecen las que realizan de manera continua los docentes acerca de sus alumnos. Como muestra la investigación educativa, los maestros no sólo hacen pruebas para comprobar a veces midiendo y otras considerando con criterios cualitativos el avance en el manejo de los contenidos que tienen sus alumnos, sino que en el trabajo cotidiano van formando un expediente de cada uno de ellos con el conocimiento que van teniendo de sus potencialidades cognitivas, intelectuales y afectivas así como de las posibilidades o limitaciones que les aporta su contexto familiar y sociocultural (Luna, 1994).

Gráfica No. 23
Diferencia de promedios en lectura entre las escuelas urbanas y rurales



Fuente: Instituto Nacional para la Evaluación de la Educación. *La Calidad de la Educación Básica en México*, México, INEE, 2003. p. 79.

Es importante recuperar y tomar en cuenta estas evaluaciones en los análisis nacionales ya que son las que mejores posibilidades tienen de considerar todos los factores que intervienen en el desempeño de los alumnos.

También hay evaluaciones cualitativas de tipo etnográfico que pueden ser realizadas por investigadores externos pero que requieren de estancias en el contexto a evaluar, observaciones y registros detallados sobre todo lo que ocurre en el aula o en la escuela y que requieren de análisis cualitativos para tratar de comprender cuáles son los procesos que ocurren en esta situación, para describir las condiciones en las que se realiza la educación y los factores que influyen en los posibles resultados educativos. Estas evaluaciones son costosas en tiempo y en recursos económicos y humanos ya que la descripción de los procesos requiere comprender la lógica de los actores educativos, que frecuentemente es diferente de la perspectiva del investigador. Por esta razón no se pueden hacer evaluaciones cualitativas de todo el sistema sino que hay que hacer estudios de casos representativos. Para esto es interesante tomar en cuenta las evaluaciones estandarizadas o cuantitativas como un punto de partida para decidir qué es lo que conviene analizar más detalladamente. Por ejemplo, el informe del INEE sobre las pruebas

estandarizadas aplicadas nacionalmente aporta una gráfica de la diferencia de promedios en lectura entre las escuelas urbanas y rurales en todos los estados de la República. En esta gráfica, que se muestra a continuación, las escuelas urbanas tienen mejores resultados que las rurales en todos los estados excepto en Guerrero y Campeche. Este es un dato que no puede ser explicado más que si se realiza algún tipo de estudio cualitativo para analizar qué es lo que pasa en esos estados a diferencia de los demás que aporta un resultado así.

c) Evaluaciones con base en un proyecto educativo de cambio y por zona escolar

En algunos lugares, como en el estado de Campiñas en Brasil (De Freitas, 2004), se está probando un tipo de evaluación en la que de manera voluntaria los docentes y las escuelas de una región con condiciones semejantes de nivel cultural y socioeconómico deciden plantearse un proyecto educativo de cambio para mejorar la calidad de su enseñanza. En este caso los propios docentes son los que deciden cuáles son los ítems que les interesa que sean evaluados de acuerdo al enfoque que tienen y a los propósitos del proyecto educativo con el que pretenden mejorar su trabajo. Los estándares máximos se establecen respecto a la mejor escuela de la zona. Los investigadores aportan su conocimiento para diseñar las pruebas y los sistemas de evaluación con las mejores técnicas de acuerdo con los propósitos del proyecto. De esta manera los resultados vienen a ser significativos para los propios actores educativos ya que responden a las necesidades concretas de información que ellos tienen y son base para la reflexión colectiva y para la toma de decisiones sobre las medidas a elegir para superar la calidad de sus procesos educativos. Así, la evaluación continua tiene un sentido más dinámico pues está vinculada a los cambios permanentes que se van implementando con base

en los resultados de las evaluaciones. Por otro lado, este tipo de proyectos compara a las escuelas equiparables en cuanto a contexto sociocultural y permite socializar la experiencia de las escuelas que, teniendo las mismas condiciones socioeconómicas y de trabajo educativo, obtienen mejores resultados.

4. LEGITIMACIÓN NECESARIA PARA EL PROCESO DE EVALUACIÓN

Uno de los requisitos fundamentales para que una evaluación registre un efecto sobre la práctica educativa de modo que contribuya a mejorar la calidad de la educación es que la evaluación tenga credibilidad, esto es, que los actores educativos le otorguen un sentido positivo como contribución para mejorar su tarea educativa. Para esto, es necesario que la evaluación tenga legitimidad en varios sentidos. Debe contar con legitimidad técnica y académica ya que debe ser realizada con el conocimiento más avanzado de los recursos de medición, de análisis estadísticos sofisticados y de evaluaciones de diferente tipo que tienen gran complejidad técnica. Pero para que la evaluación se vincule al trabajo cotidiano, para que pueda aplicarse y para que sus resultados sean retomados en la práctica, es necesario que tenga legitimidad social y política. Una evaluación sin legitimidad política corre el riesgo de sólo servir para aportar datos que no tienen, en el mejor de los casos, ninguna repercusión en las prácticas educativas, y en el peor, que tienen efectos secundarios que pueden dañar al sistema empeorando la educación a través de mecanismos de marginación y estrechamiento del currículo real que ya antes analizamos.

Si la necesidad de legitimación política y social es necesaria para la aplicación de cualquier sistema de evaluación que pretenda retroalimentar y tener un efecto en el mejoramiento de las prácticas educativas, esto es indispensable en

las evaluaciones que buscan analizar los procesos que ocurren en la escuela. No se pueden evaluar procesos sin contar con la participación de los maestros, los alumnos, los padres y las autoridades escolares. Esto es, los actores educativos deben tener un compromiso de cambio, pero para que la evaluación contribuya a mejorar la calidad educativa es imprescindible también que las autoridades relacionen la evaluación con una política de mejoramiento y de financiamiento a los sectores que obtienen los peores resultados.

5. ALGUNAS CONCLUSIONES

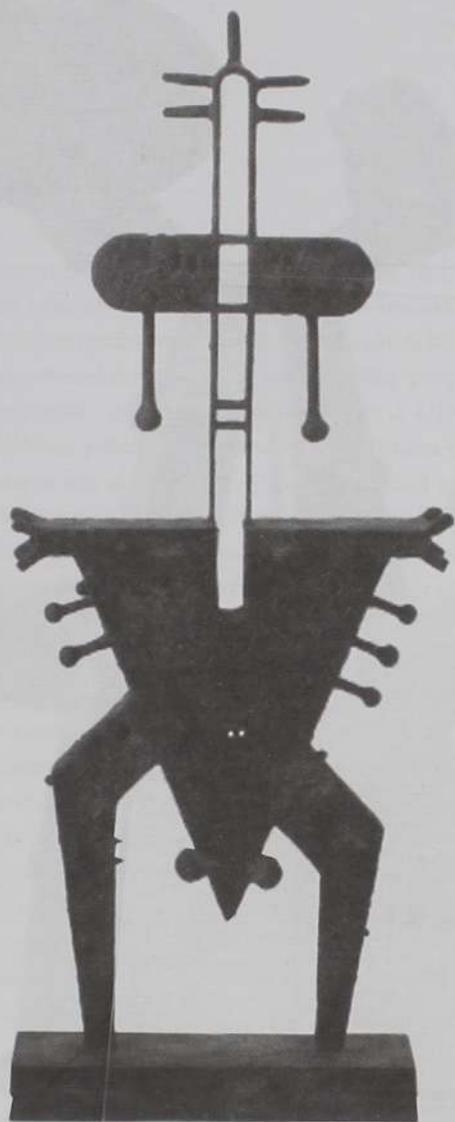
Cuando se pretende conocer el estado del sistema educativo en su conjunto son necesarias las pruebas estandarizadas que proporcionan datos cuantitativos que permiten hacer comparaciones entre diferentes sectores del sistema. Sin embargo, estas pruebas son muy costosas y como hemos visto tienen riesgos de producir efectos secundarios no deseados sobre el sistema. A la larga pueden causar más problemas a la educación que los que pueden resolver. Para tratar de evitar efectos como el aumento de la marginación y las modificaciones curriculares para terminar enseñando sólo para el examen, es necesario que los actores educativos comprendan los alcances y las limitaciones de esta información, esto es, que la puedan interpretar. Para ello es necesario, por un lado, que al mismo tiempo que se da la información sobre los resultados de las pruebas, se informe acerca de cómo se hicieron los ítems o indicadores y sobre cómo se aplicó la prueba. No se pueden interpretar linealmente los resultados de las pruebas estandarizadas ya que como hemos visto un mal resultado no necesariamente refleja una mala educación. Es importante informar de estos resultados con las precauciones adecuadas para evitar descalificaciones, desprestigio social y los efectos de desmoralización y disminución de la autoestima.

Por otro lado es de gran importancia que, más que evaluaciones continuas del mismo tipo a todo el sistema sin que se hayan realizado cambios en él, se implementen formas alternativas de evaluación con enfoques diversos, como las descritas u otras, que aporten información complementaria que permita comprender las causas de los resultados de las pruebas estandarizadas y que den elementos para conocer las características de los procesos que ocurren en las aulas y las escuelas. Sólo con información sobre los procesos educativos en las aulas y sobre las condiciones de enseñanza se pueden tener los elementos informativos que posibiliten reconocer dónde están los problemas y así buscar los recursos que permitan superarlos y mejorar la calidad de la educación. También tiene gran relevancia que las pruebas estandarizadas se adecuen lo más posible al currículo nacional, como ya está intentando hacerlo el INEE. Al mismo tiempo que se realizan evaluaciones del sistema, es necesario descentralizar otras evaluaciones para que éstas sean acordes a las condiciones locales de cada región, para que tomen en cuenta los factores socioeconómicos y culturales tan diversos que hay en nuestro país. De esta manera se permite que las evaluaciones se relacionen con políticas locales de cambio en las que se comprometan los actores educativos, mientras que las autoridades aporten los apoyos necesarios.

Las evaluaciones por sí mismas no mejoran la calidad, por lo que es importante tomar las medidas para que se comprenda que la cultura de la evaluación sólo es un medio que tiene riesgos como los de aumentar la inequidad y estrechar el contenido de la educación si no se vincula con una cultura de la calidad. Para esto, es necesario articular los resultados de distintos tipos de evaluación con el compromiso, la participación y los recursos de todos los actores educativos y con los apoyos materiales y sociales necesarios por parte de las autoridades, en torno a proyectos locales de cambio.

REFERENCIAS

- Arancibia, V., 1992, *Efectividad escolar. Un análisis comparado*, Santiago, CEP. Documentos de Trabajo No. 174
- Bellei, C., 2001, *¿Ha tenido impacto la reforma educativa chilena?*, Santiago, Ministerio de Educación (m.s.)
- Brunet, J., 1984, *Acción, pensamiento y lenguaje*. Comp. J.L. Linaza. Madrid: Alianza Psicología.
- Cicourel, A., 1974, "Some basic theoretical issues in the assessment of the child's performance in testing and classroom settings", en *Language Use and School Performance*. USA: Academic Press, 300-351.
- Coll, C., 1984, "Estructura grupal, interacción entre alumnos y aprendizaje escolar" en *Infancia y Aprendizaje*, 27/28, pags. 119-138.
- De Freitas, L. C., 2004, "Evaluación institucional y evaluación de sistemas: hacia un abordaje integrado", en *XIV Congreso Mundial de Ciencias de la Educación, Educadores para una Nueva Cultura*, Santiago, p. 434
- Heyneman, S.P. y W. A. Loxey. 1982, "Influences on academic achievement across high and low income countries: a re-analysis of Iea Date, en *Sociology of Education*, vol. 55, pp. 1-32
- Himmel, E., 2004, "Modalidades de informes de resultados de mediciones nacionales y su impacto en el uso" en *XIV Congreso Mundial de Ciencias de la Educación, Educadores para una Nueva Cultura*, Santiago, p. 435
- Instituto Nacional para la Evaluación de la Educación, 2003, *Política Nacional de Evaluación*, México, INEE.
- Instituto Nacional para la Evaluación de la Educación, 2003, *La calidad de la educación básica en México*, México, INEE.
- Instituto Nacional para la Evaluación de la Educación, 2003, *¿Qué hace a una escuela una buena escuela?*, México, Fundación Este País.
- Intercultural Development Research Association, 2001, *Retention in grade: policy brief update*, San Antonio, Texas.
- Luna, M.E., 1994, "Los maestros y la construcción del expediente cotidiano", *Investigación en la Escuela* No. 22, pp. 105-114.
- Manzi, J., 2004, "El rol de las evaluaciones estandarizadas en la elevación de los resultados educacionales en Chile," en *XIV Congreso Mundial de Ciencias de la Educación, Educadores para una Nueva Cultura*, Santiago, p. 436.
- Martín, S., s/f, "Las representaciones de la desigualdad en la cultura escolar" (mimeo).
- Mella, O., 1999, *Equidad y reforma educacional en Chile*, Santiago, Documento interno.
- Mizala, A. y P. Romaguera, 1998, *Desempeño escolar y elección de colegios: la experiencia chilena*, Santiago, U. de Chile-CEA, Documento de Trabajo No. 36.
- Pérez Mora, R., 2004, "Estandarización vs. equidad en la evaluación del desempeño de los profesores. Un ejercicio en torno a los cuestionarios para evaluación de la docencia", en *XIV Congreso Mundial de Ciencias de la Educación, Educadores para una Nueva Cultura*, Santiago, p. 437.
- Schieffelin, E. y J. Farrel, 1982, *Eight years of their lives: through schooling to the labor market in Chile*, Ottawa, IDRC.
- Valenzuela, A., 2002, "High-Stakes-Testing and U.S.- Mexican Youth in Texas: The Case for Multiple Compensatory Criteria in Assessment," en *Harvard Journal of Hispanic Policy*, vol. 14, pp. 97-116.
- Valenzuela, A., 2003, "Accountability and the Privatization Agenda" en: Valenzuela, A. *Leaving Children Behind: Why Texas-Style Accountability Fails*. Latino Youth, Albany, N.Y.: State University of New York Press.
- Zorrilla, M., 2004, "La evaluación del aprendizaje a gran escala en México. Una mirada a contraluz" en *XIV Congreso Mundial de Ciencias de la Educación, Educadores para una Nueva Cultura*, Santiago, p. 440, University of New York Press.



"EL GUERRO" *bronce a la cera perdida*



LA EVALUACIÓN DE PROFESORES DE EDUCACIÓN BÁSICA EN MÉXICO: LA CARRERA MAGISTERIAL ¹

RUTH MERCADO

INTRODUCCIÓN

EN EL PRESENTE TRABAJO SE DESCRIBEN ALGUNOS PROGRAMAS QUE EN el ámbito educativo mexicano implican incentivos y evaluación del desempeño docente en educación básica. Asimismo se identifican algunos debates en torno a dichos programas. En el país funcionan diversos programas que tienden a elevar la calidad de la educación básica y tratan de atender a población escolar que de otro modo quedaría sin el servicio. Algunos de esos programas implican evaluación e incentivos al desempeño docente.

Sin pretender hacer una lista exhaustiva, pueden mencionarse el programa dirigido a la educación primaria para migrantes, así como el de escuelas adscritas al desarrollo del proyecto escolar que opera en algunos estados de la República; de igual modo, funciona el programa de escuelas de calidad donde el incentivo es por unidad escolar, el cual ya ha tenido una evaluación externa (Bracho; 2001). Hay también otros programas que cuentan con financiamiento internacional que se mencionan en otra parte del documento y que no fueron tratados aquí porque estaban por desaparecer o no se referían a la educación básica.

La investigación para elaborar este documento se basó en búsqueda documental, así como en la red internacional (Internet) y en entrevistas a funcionarios responsables de las áreas involucradas en la operación y evaluación del programa de Carrera Magisterial (CM) presentado en la primera parte del texto. En la segunda parte del documento se presentan de manera más breve los Programas Compensatorios (PC) que se desarrollaron a partir del Programa para Abatir el Rezago Educativo iniciados en los años noventa.

LOS INCENTIVOS Y LA EVALUACIÓN DEL TRABAJO DOCENTE COMO FACTOR DEL MEJORAMIENTO DE LA CALIDAD DE LA EDUCACIÓN

La evaluación educativa es un tema que ha sido extensamente tratado en la literatura internacional y en la que se produce en la región

La Dra. Ruth Mercado es investigadora titular del Departamento de Investigaciones Educativas del CINVESTAV.
Correo electrónico: rmercado@cinvestav.mx

de América Latina y el Caribe. El interés de los sistemas nacionales se ha desplazado de las preocupaciones por lograr mayor cobertura de la demanda, que por otro lado, en algunos casos está cerca de ser satisfecha, para ubicarse en el terreno de la calidad de la educación básica. Muestra de ello es la comparación entre las metas sólo asociadas con la ampliación de la cobertura que se observaba en las conferencias de Ministros de Educación y la UNESCO en las décadas de los sesenta y los setenta con la Conferencia de Quito o la Declaración Mundial sobre "Educación para Todos" de Jomtien. Como señala Toranzos: "Existe un consenso creciente acerca de que es necesario, no sólo que todos los niños asistan a una escuela, sino además, que esa escuela incorpore efectivamente los conocimientos y competencias necesarios para desempeñarse y participar en la sociedad en la que viven" (Toranzos, 1996:1).

Y como lo declara el artículo 4º de la Declaración Mundial sobre "Educación para Todos" de Jomtien, cuando dice: "Es necesario determinar niveles aceptables de adquisición de conocimientos mediante el aprendizaje en los planes de educación y aplicar sistemas mejorados de evaluación de los resultados" (UNESCO, 1994).

Es así como el concepto de calidad puede constituirse, como lo propone Toranzos (1996:2) en la "idea aglutinadora de los consensos sociales y políticos indispensables para el desarrollo de nuestros sistemas educativos en el presente siglo."

Entre los estudiosos del tema, la noción de calidad viene asociada con la idea de evaluación ya que en la búsqueda de la calidad se vuelve indispensable saber hasta qué punto se ha alcanzado la calidad deseada y qué se debe mejorar para lograrlo². Es así que durante las décadas de los ochenta y los noventa se han desarrollado tanto mecanismos institucionales como centros de evaluación de los sistemas educativos

en busca de mejorar la calidad de los mismos. Se ha hecho en "países como Francia, Suecia, Noruega, España, Argentina o Chile, así como planes sistemáticos de evaluación en el Reino Unido, Holanda, Francia, Argentina, Chile, República Dominicana o México." (Tiana, 1996). De manera creciente en la pasada década, organismos internacionales han promovido programas para desarrollar políticas de evaluación educativa, entre ellos la Organización para la Cooperación y el Desarrollo Económico (OCDE), la UNESCO, la Unión Europea y la Organización de Estados Americanos. Dado lo anterior, puede hablarse, como sugiere Tiana de: "La extensión de un interés creciente por la evaluación de los sistemas educativos, que ha producido como efecto, una rápida evolución de la evaluación entendida como disciplina científica y como práctica profesional" (House, 1993; Tiana, 1996).

Por otro lado, también es necesario reconocer que los consensos no son aún muy claros en cuanto al qué y cómo evaluar. Es importante tomar en cuenta a ese respecto las reservas de algunos estudiosos cuando identifican el problema de los criterios con los cuales se mide y el de los métodos para medir.

El tema de los incentivos a las buenas prácticas escolares viene, a su vez, asociado a los esfuerzos que hacen los países por lograr la calidad de los servicios de educación básica. El desarrollo de programas que contemplen estos incentivos, ya sea individuales o por institución, es aún incipiente y todavía más el estudio de su impacto. Uribe (1999), nos presenta un panorama al respecto. La autora examina «modalidades de incentivos utilizadas alrededor del mundo bajo el marco conceptual del Ciclo del Recurso Humano (Fonbrum, C. J., Tichy, N. H., Devana, M. A., 1984).

Uno de los problemas que destaca Uribe es que "pocos países han establecido criterios profesionales para identificar y premiar las

mejores prácticas docentes." De igual manera, "pocos han establecido políticas de rendición de cuentas para hacer responsables a las escuelas y los maestros del aprendizaje de los estudiantes." En los países de la región se observa que más que incentivos al desempeño docente, lo que abunda son factores que lo desincentivan como los bajos salarios y las precarias condiciones en que los maestros realizan su trabajo.

Sin embargo, Uribe presenta un interesante inventario de "factores que se encontraron que operan como incentivos económicos y no-económicos para los maestros en un estudio llevado a cabo en Haití, Liberia, Somalia y la República Árabe de Yemen." (Kemerer y Thiagarajan, 1989).

En ese estudio Uribe señala cómo los autores destacan tres elementos para el diseño de incentivos a los docentes: consistencia o que las metas estén claramente establecidas, suficiencia o la condición de ser suficiente para las metas propuestas, justicia o que el esquema sea considerado justo por los actores a quienes se dirige.

Por otra parte, la autora describe diversas modalidades de incentivos docentes; dentro de los esquemas de incentivos económicos están: el incremento salarial generalizado, por un lado, y establecimiento de esquemas de pago diferenciado por otro, éste último dividido en 6 modalidades. Una de éstas es la del escalafón docente, que por sus características es la que tiene lugar en México con la Carrera Magisterial que establece, precisamente, un escalafón horizontal para la promoción de los mejores maestros, como se describirá en este documento.

ALGUNAS CARACTERÍSTICAS DEL SISTEMA DE EDUCACIÓN BÁSICA EN MÉXICO

Desde el final de la Revolución y la promulgación de la Constitución de 1917 cuyo artículo tercero consagra el derecho de todos los mexicanos a

la educación y la obligación del Estado a proporcionarla, éste ha promovido diferentes programas y reformas que pretenden, inicialmente, mejorar la cobertura de la educación básica y después garantizar su calidad. Sin embargo, dichos esfuerzos se han enfrentado con la realidad de la desigualdad social que prevalece en el país y que ha dificultado tanto universalizar la educación básica (aunque en este renglón hay avances significativos) como conseguir la equidad y la calidad de los servicios educativos proporcionados. Tan grande como es el tamaño del sistema de educación básica mexicano, que cuenta con aproximadamente 1,000,000 de maestros y 23.8 millones de alumnos (SEP, 2001 a y b), son los esfuerzos que realiza el Estado para cubrir la demanda y apoyar el servicio educativo, por ejemplo, con la distribución de aproximadamente 200 millones de libros gratuitos en este año.

Actualmente, se reconocen severas inequidades entre los estados y dentro de ellos y, preponderantemente, entre la educación proporcionada en medios urbanos y medios rurales (Instituto Nacional de Geografía Estadística e Informática, 1990; SEP, 2000). Asimismo, las condiciones en que trabajan gran cantidad de escuelas rurales es reconocida en términos de su aislamiento geográfico y las carencias de las comunidades en que se encuentran (Mercado, 2000; Ezpeleta y Weiss, 2000) Además, del 30 al 40 % de las escuelas primarias en el país son multigrado (CONAFE, 1996) sobre todo en el medio rural y en algunos estados llegan al 70% (Weiss y Ezpeleta, 2000:277).

Por otra parte, hay aún problemas de eficiencia y de rezago del sistema, como por ejemplo, un 58.11% de eficiencia terminal como media nacional con el agravante de que en algunos estados sólo un promedio de tres de cada diez niños terminan la primaria y en zonas rurales e indígenas lo hace uno de cada diez (Consejo Nacional de Fomento Educativo, 1994).

Frente a las inequidades descritas en la educación básica, el Estado mexicano, además de emprender acciones regulares y de las reformas de los noventa para mejorar y apoyar estos servicios educativos, ha suscrito acuerdos internacionales como los de la Conferencia de Jomtien y aquéllos donde se expresan estos acuerdos en sus implicaciones para la región de América Latina (CEPAL-UNESCO, 1997). De igual forma, ha puesto en marcha un programa nacional de incentivos docentes y ha promovido los programas compensatorios. A continuación se presentan los programas con los que en México se pretende elevar la calidad de la educación básica y en éstos confluyen los incentivos y la evaluación del trabajo docente. Dichos programas se llevan a cabo por la Secretaría de Educación Pública (SEP) con la participación de los estados de la federación. El más importante de ellos por su cobertura nacional, es el de Carrera Magisterial; otros son los programas compensatorios que cuentan con apoyo financiero internacional.

CARRERA MAGISTERIAL

Este es el programa de incentivos y de evaluación relacionado con la calidad del trabajo educativo de mayor impacto en el país por la cobertura que tiene y los recursos con los que opera. "Carrera Magisterial es un sistema de promoción horizontal en el que los docentes participan de forma voluntaria e individual y tienen la posibilidad de incorporarse o promoverse si cumplen con los requisitos y se evalúan conforme a lo indicado en los lineamientos normativos" (SEP, 1998:4). El objetivo general del programa consiste en "elevar la calidad de la educación nacional a través del reconocimiento e impulso a la profesionalización del magisterio y del mejoramiento de las condiciones de vida y labores de los docentes" (SEP, 1998:10).

Como objetivos específicos el programa pretende:

- "Valorar la actividad docente fortaleciendo el aprecio por la función social del profesor.
- Motivar a los profesores para que logren un mejor aprovechamiento en sus alumnos.
- Promover el arraigo profesional y laboral de los docentes.
- Reconocer y estimular a los profesores que prestan sus servicios en escuelas ubicadas en comunidades de bajo desarrollo y escasa atención educativa, así como a los que trabajan con alumnos que requieren mayor atención.
- Reforzar el interés por la actualización, capacitación y superación profesionales del magisterio así como la acreditación de cursos de mejoramiento académico" (SEP, 1998:11).

ANTECEDENTES DEL PROGRAMA

El programa de Carrera Magisterial tuvo su origen en una propuesta del Sindicato Nacional de Trabajadores de la Educación (SNTE) emitida en su congreso extraordinario de febrero de 1990. La propuesta se integró posteriormente por el gobierno federal, con el acuerdo de los estados, al Acuerdo Nacional de Modernización de la Educación Básica de mayo de 1992 (SEP, 1992). Dicho acuerdo dio lugar a las reformas que desde entonces se han implementado en México en la educación básica y normal. En ese sentido, las reformas en cuyo contexto se inició Carrera Magisterial se dieron en el marco de los procesos de mejoramiento y apoyo que se desarrollaron en los años noventa para la educación básica y la formación de maestros en los países de la región. Incluso en esa época se promulgó la Ley General de Educación que establece la obligatoriedad de la evaluación nacional del servicio para la SEP (SEP, 1993).

Posteriormente, se constituyó la Comisión Nacional Mixta SEP-SNTE y las comisiones mixtas

estatales. La Comisión Nacional es la máxima autoridad del programa de CM y quien emite la normatividad correspondiente y vigila desde entonces la operación del mismo.

Las autoridades educativas de la Comisión Mixta SEP-SNTE consideran un logro que la Carrera Magisterial sea gobernada a partir de una comisión en la que se encuentra el SNTE como contraparte (EDCMSEP-SNTE; 2002)³. Ello implica contar con el acuerdo y la colaboración sindical en políticas educativas conducentes a mejorar la calidad de la educación básica en el país y su acuerdo para implementar los incentivos y la evaluación del desempeño docente. Como se sabe, el SNTE es una de las mayores organizaciones sindicales de América Latina ya que agrupa aproximadamente a un millón de maestros y su influencia en el éxito de las políticas educativas nacionales es importante.

FUNCIONAMIENTO, COBERTURA Y RECURSOS DEL PROGRAMA

En la operación del programa de CM intervienen varios organismos tanto federales como estatales. Entre los federales interviene la Comisión Mixta SEP-SNTE para carrera magisterial que es la autoridad máxima del programa constituida por autoridades de la SEP y sus contrapartes del SNTE. Otros organismos federales involucrados en CM son la Dirección General de Evaluación de la SEP. Y el Programa Nacional de Actualización de Profesores de Educación Básica en Servicio, conocido como PRONAP. En los estados operan las Comisiones Paritarias Estatales SEP-SNTE, las representaciones estatales de la SEP y algunos otros órganos educativos estatales.

El programa de CM fue creado para funcionar como estímulo al desempeño de los mejores profesores. Es decir, que quien ingresa al mismo recibe un estímulo económico de acuerdo a la categoría del programa a la que acceda (son cinco categorías, de la A a la E).

El estímulo económico en CM incluye la evaluación del desempeño del personal. Es decir, que para ingresar al programa o promoverse en el mismo se pasa por un proceso de evaluación de diferentes factores de desempeño. Pueden entrar a CM tanto profesores de grupo (de las diferentes modalidades educativas en el nivel básico), como directivos y personal técnico de la educación básica, que constituyen las tres vertientes del programa.

En el proceso de evaluación se contabiliza el puntaje que obtenga el personal en seis factores que tienen peso diferente y en total deben dar 100 puntos. Dichos factores de evaluación y su respectivo valor son los siguientes:

Antigüedad 10, grado académico 15, preparación profesional 28, actualización y superación profesional 17, desempeño profesional 10 y aprovechamiento escolar 20.

Respecto a los cuatro últimos factores, en la evaluación se procede de la manera siguiente:

- La Preparación profesional se evalúa con un examen que diseña y aplica la SEP a través de la Dirección General de Evaluación. En este rubro se evalúa "el conocimiento del maestro sobre los contenidos del grado o de la materia que imparte, de los enfoques metodológicos, de los materiales de apoyo. Esto sirve para diseñar los programas de actualización en los cursos estatales con los que también se realiza parte de la evaluación del programa, la que corresponde a la Actualización y superación profesional." (EDCM-SNTE, 2002).

Además, el examen "diseñado y aplicado por la Dirección General de Evaluación de la SEP se elabora a partir de un Banco Nacional de Reactivos conformado por la misma SEP que se actualiza con base en talleres que se llevan a cabo en todas las entidades del país" (Canales *et al.*, 2002).

- La evaluación del factor de Acreditación de cursos de actualización y superación del magisterio, "consiste en la acreditación de cursos de actualización, capacitación y superación profesional,

para fortalecer los conocimientos en general y los relativos a aspectos pedagógicos y didácticos en particular [...]. Para el diseño de dichos cursos la autoridad educativa tomará en cuenta los resultados de los factores de Preparación profesional y Aprovechamiento escolar, de modo que se subsanen las deficiencias detectadas por medio de la evaluación de estos factores. Este factor considera cursos Nacionales y Estatales” (SEP, 1998: 29).

-Para el factor de desempeño profesional se evalúa “el conjunto de acciones cotidianas que realizan los docentes en el desempeño de sus funciones. Los docentes son evaluados en cuatro aspectos: planeación de la enseñanza, desarrollo del proceso enseñanza-aprendizaje, participación en el funcionamiento de la escuela y participación en la interacción escuela-comunidad.” La evaluación de cada profesor la realiza el Órgano de Evaluación Escolar que en cada escuela se constituye con todos los profesores y un representante sindical y que es presidido por el director de la escuela (EDCMSEP-SNTE, 2002).

- En el factor de aprovechamiento escolar “se evalúan, por medio de un examen, los aprendizajes que los alumnos de cada maestro participante han obtenido en su grado o asignatura.” El instrumento de evaluación es elaborado por la Dirección General de Evaluación de la SEP “y aplicado por ésta y la autoridad educativa estatal. El procesamiento de los resultados obtenidos en la evaluación será responsabilidad de la SEP. Los reactivos que conformen los instrumentos de evaluación serán elaborados a partir de los contenidos de los planes y programas de estudio y de los Libros de Texto Gratuitos” (SEP, 1998:31). Finalmente, es la Comisión Mixta SEP-SNTE, tanto nacional como estatal en cada caso, la que dictamina qué docentes serán incorporados o promovidos en CM.

Como puede apreciarse, en el funcionamiento de CM confluyen varias líneas de acción mediante las cuales se incentiva y se evalúa

el trabajo docente, a la vez que se actualiza a los profesores de educación básica en el país. Mediante el ingreso y la promoción en CM, los profesores acceden a incentivos económicos como “el primero que es del 34 % en números redondos del incremento sobre la plaza inicial y llega hasta el 256 % en el nivel más alto de CM.” En total, para el año 2000-2001 “se invirtieron 24,000 millones de pesos” o 2,400 millones de dólares (EDCMSEP-SNTE).

En total participan en CM entre 600,000 y 700,000 maestros⁴ de los cuales la mayoría son profesores de grupo.⁵ Por otro lado, en el año 2000-2001 se examinaron a 6, 000, 000 de alumnos de educación básica para evaluar el factor de Aprovechamiento escolar.

ALGUNOS DEBATES EN TORNO A CM

Algunos sectores del magisterio, entre ellos los disidentes de la organización sindical o Coordinadora Nacional de Trabajadores de la Educación (CNTE), han cuestionado al programa por su tendencia a las evaluaciones individualistas y por promover la competitividad eficientista que, en su opinión, denota una visión empresarial sobre la educación. Se cuestiona desde esa perspectiva, que el fomentar esos valores entre el magisterio propicie una mejor calidad de la educación.

Se dice también, desde posiciones críticas al programa, que éste ha representado para el gobierno un medio para desatender las demandas de incremento salarial que corresponderían al magisterio en general, dada la pérdida creciente del poder adquisitivo de los salarios del magisterio nacional.

Además, parece haber algunos efectos perversos del programa, entre otros, el que muchos maestros abandonen el medio rural para estar más cerca de los medios urbanos donde es más fácil acceder a los cursos de actualización lo cual es un factor de evaluación importante para CM. Otro efecto no deseado es que

la evaluación al interior de los planteles por el Órgano de Evaluación Escolar, donde participan los docentes para evaluar a sus compañeros de CM, origina tensiones y conflictos que afectan negativamente la vida diaria del plantel y por tanto, su quehacer educativo. Existen también otros debates en los que no sólo participan los grupos disidentes y estudiosos de la educación, sino los propios directivos del programa, además de otros agentes interesados. Entre esos debates están los que a continuación se señalan:

- El impacto del programa en los resultados del quehacer educativo

Este debate coloca en el centro una de las cuestiones que en realidad es de las principales en el marco mismo de CM. Se trata de la aspiración del programa por repercutir en una mejor calidad de la educación; para muchos de los involucrados esa mejor calidad debe expresarse en los productos finales del quehacer escolar. Sin embargo, no hay consenso sobre si se ha logrado la mejoría buscada. Para algunos, el impacto de CM ha sido positivo tanto para promover una cultura de la evaluación entre el magisterio como para propiciar su actualización permanente, como para esperar que su práctica educativa haya mejorado con este programa (EDCMSEP-SNTE, 2002). En cambio, otros opinan que no ha habido un impacto importante en el resultado final del trabajo escolar para los sectores más desfavorecidos (EDDGE-SEP, 2002). La Dirección General de Evaluación de la SEP reporta, por ejemplo, que: "La preparación profesional de los maestros y su compromiso encauzado en CM son elementos que están actuando con una gran influencia en el aprovechamiento escolar de los alumnos; no obstante, debe señalarse que el programa ha incorporado más maestros de escuelas urbanas que de rurales e indígenas" (Velázquez, 1996).

Sin embargo, en todos los casos se reconoce que no hay datos confiables para evaluar con precisión las diferencias que hicieron posible identificar, en los exámenes de los alumnos, una mejoría de la calidad de la educación proporcionada por las escuelas que tienen maestros en CM. Incluso hay maestros de la disidencia magisterial que opinan que los resultados de las evaluaciones a los alumnos de docentes que están en CM, comparados con los que no lo están, resultan prácticamente iguales. "Por lo tanto, nos están pagando un salario diferenciado por hacer el mismo trabajo con semejantes resultados" (EMSSIX).

- La CM y la cultura de la evaluación

La CM ha sido un programa que ya se ha instalado en la vida profesional de los maestros de educación básica del país. Las acciones que deben emprender los maestros para ingresar y promoverse en el programa son parte de su vida cotidiana. En opinión de algunos directivos esto ha implicado el incipiente cultivo de una "cultura de la evaluación" entre los docentes. Sin embargo, otros agentes opinan que no es eso lo que ha ocurrido, sino que se ha propiciado una cultura de la simulación. Esto se dice porque, por ejemplo, en la evaluación del Desempeño docente se ha encontrado que hay una tendencia que consiste en que los compañeros otorguen siempre la máxima calificación a quien se está evaluando para CM (EDCMSEP-SNTE, 2002, EDDGE-SEP, 2002). Más allá de ese ejemplo, las dudas sobre las posibles simulaciones y a veces hasta la filtración de las pruebas o la omisión de algunos exámenes que benefician a algunos maestros cercanos a grupos regionales sindicales ponen en cuestión al programa en cuanto a su transparencia y a su real incidencia en una mejor calidad de la educación básica.

Por otra parte, en relación con la evaluación también hay diferentes opiniones sobre

si ésta debe ser del conocimiento público de acuerdo con la tendencia actual de los sistemas gubernamentales hacia la llamada "rendición de cuentas." Algunos se manifiestan porque los resultados de las evaluaciones tanto a nivel del sistema escolar como de las propias escuelas y del trabajo de los maestros sean públicas. En ese aspecto, hay quienes proponen que los resultados individuales de las escuelas y de los maestros no sean utilizados para sancionar, sino para apoyar a los que tienen resultados deficientes y para retroalimentar al sistema de manera que puedan tomarse medidas compensatorias y correctivas (Schmelkes, 2002:9).

Por otro lado, existen inquietudes entre los involucrados en CM y otros interesados en los resultados de este programa porque se realice una evaluación del mismo en tanto que éste ya ha funcionado por diez años y se cuenta con información para evaluar cuáles han sido sus aportaciones y sus límites e insuficiencias (EGCPRONAP, 2002; EDGE-SEP, 2002; EDCMSEP-SNTE, 2002).⁶

- La CM y la actualización permanente de los maestros

En este aspecto, hay apreciaciones que valoran positivamente el hecho de que la participación en CM implica forzosamente tomar y aprobar los cursos de actualización para los maestros. Desde otras perspectivas, sin embargo, se aprecia que esto ha sido un obstáculo para que los maestros se interesen por el valor propio de la actualización y sólo acudan a ella de manera forzada y por el beneficio económico que representa la CM. Con ello se desvirtúa el propósito principal de la actualización que es propiciar una cultura de la capacitación profesional permanente del magisterio alejada del credencialismo. Incluso para algunos la actualización profesional debería ser un proceso independiente de CM (EDPRONAP, 2002).

- Diferencias sobre el aspecto más importante a evaluar

Hay también diferencias sobre el aspecto más relevante para ser evaluado en CM; para algunos es el factor de Desempeño profesional y para otros el de Aprovechamiento escolar. Es decir, hay puntos de vista donde destaca el factor de Desempeño profesional como aquél en el que se expresa el trabajo cotidiano de los maestros, el que incluye las tareas más importantes que éste implica, ya sea dentro o fuera del aula. Además este factor incluye la autoevaluación de los docentes más la evaluación externa que es la que le otorgan sus colegas y el director, o en el caso de los maestros multigrado, la supervisión escolar. Así fue señalado, por quien en los inicios de CM era el coordinador de la representación sindical de la Comisión Nacional Mixta SEP-SNTE de CM (Callejas, 1992).

En cambio, para otros actores el factor más importante a evaluar es el que resulte de los exámenes de aprovechamiento escolar, como lo argumentaba el actual Director General de Evaluación de la SEP. Desde su perspectiva, que es compartida por otros (Órnelas, 2000:180; Schmelkes, 2001:1) es éste el factor donde se expresan los resultados finales del trabajo docente y escolar, en el aprovechamiento de los alumnos quienes son los usuarios últimos del servicio educativo; sería ahí donde se expresaría la calidad del servicio proporcionado por la escuela.

Sin embargo, persiste la discusión respecto de los instrumentos de evaluación del Aprovechamiento escolar mediante tradicionales exámenes "objetivos" que presentan limitaciones por todos conocidas respecto de las áreas de conocimiento que pueden explorar. De hecho, por esas limitaciones, en CM no se evalúa ese factor en los dos primeros grados de la primaria, ni en preescolar (EDDGE-SEP, 2002). A la vez se cuestionan "las fuentes de los reactivos, las formas de aplicación o los procedimientos de evaluación" (Canales, et al., 2002).

Como un comentario final sobre CM vale la pena referirse al acceso que éste ha brindado respecto al conocimiento social de sus resultados. Esto ha sido un reclamo que se ha hecho al programa desde distintos sectores interesados en la educación. Lo que informa la SEP sobre este tema es que: "Con respecto a los resultados del factor Aprovechamiento escolar del programa de CM, desde 1997 éstos se entregan de manera agregada a las autoridades educativas de las entidades y, a partir de 1999, se detallan por escuela; algunos estados distribuyeron a los planteles sus resultados desde ese año. En agosto y septiembre del 2000, la SEP proporcionó a cada entidad federativa la información del periodo 1995-1999 detallada por docente, en formatos diseñados para ese propósito, solicitándoles que los distribuyeran y los utilizaran para mejorar lo alcanzado, 28 entidades federativas han iniciado ya este procedimiento" (SEP, 2000:679).

En el mismo documento informa la SEP que también se imprimió el folleto: "Distribución de los planteles públicos de educación primaria y secundaria según el nivel de aciertos de sus alumnos en los exámenes de carrera magisterial" para todos los docentes de las escuelas evaluadas, así como para instituciones y personas del medio educativo e investigadores del campo.

Como puede verse, son algunos los logros y muchas las materias que quedan pendientes en torno a CM. Como programa que implica estímulos y evaluación del desempeño docente que ya se ha consolidado en el país requiere de una revisión "a fondo", como muchos interesados y aun responsables del mismo han señalado.

OTROS PROGRAMAS DE INCENTIVOS Y EVALUACIÓN DEL TRABAJO DOCENTE: LOS PROGRAMAS COMPENSATORIOS

Otros programas existentes en México para elevar la calidad de la educación y por tanto abatir

los rezagos educativos identificables en las regiones más pobres son los Programas Compensatorios (PC) del Consejo Nacional de Fomento Educativo (CONAFE)⁷ y que también implican incentivos y evaluación a los docentes. Los PC operan en todas las entidades de la República Mexicana en aquellas regiones donde se presenta mayor rezago educativo entendiendo este término como lo describen Torres y Tenti (2000:186). Y como señala la SEP (citada en García, 2002:14) estos programas intentan "equilibrar las desigualdades, nivelar en relación con un parámetro o resarcir una carencia o un daño" (SEP 2000:186).

El desarrollo de los PC se promovió a partir del interés de los países de la región expresado en el espíritu de Jomtien (UNESCO, 1994) "por abatir las desigualdades y los rezagos educativos y en el marco de las tendencias en la década de los noventa" de brindar a los niños más pobres la permanencia y la promoción en la escuela, así como de asegurar la calidad de los servicios educativos que se proporcionan a estos segmentos de la población (García, 2002: 34; Órnelas, 2001: 152).

Además, la implementación de los PC se inscribe en el contexto de la descentralización del sector educativo, que se impulsó en la década de los noventa, tanto en nuestro país como en otros de la región. En ese sentido, los estados de la República "definen las necesidades y propuestas de acción, así como la definición de las regiones y municipios con referencia a los indicadores de nivel central" (SEP-CONAFE, 1994, citado en García, 2002:37).

Fue a partir de 1991 que la SEP inició los PC para ser operados por el CONAFE. Entre los objetivos generales de los programas está el de implementar acciones de compensación para la situación de las escuelas que tienen mayores desventajas en el sistema. Además, se espera, como resultado de esos programas, que los estudiantes obtengan mejores resultados de aprovechamiento y que las escuelas objeto de ellos

lleguen a tener los índices de rendimiento de las escuelas urbanas (García, 2002:36; CONAFE, 1996).

Por otro lado, en México los PC se sustentan en el Artículo 34 de la Ley General de Educación donde se les define como las «acciones adicionales tendientes a disminuir las disparidades, avanzar en relación con parámetros establecidos y subsanar las carencias más importantes del sector educativo (Diario Oficial, 1993; García, 2002:37).

Los PC que han funcionado en México cuentan con el financiamiento por parte de la federación y de los estados; además, son apoyados por el Banco Mundial y el Banco Interamericano de Desarrollo. Aunque antes de la intervención de estos organismos «las ideas sobre los PC ya estaban sembradas en el ambiente académico y organizaciones civiles los promovían en otras latitudes», como lo señala Órnelas (2001: 168), refiriéndose a los datos y reflexiones sobre el tema que había escrito Latapí a ese respecto (Latapí, 1973).

Los principales, aunque no los únicos, PC que han funcionado en México para la educación básica son:

-Programa para Abatir el Rezago Educativo (PARE), previsto para desarrollarse durante cinco años funcionó entre 1991 y 1997. Desde entonces y hasta el año 2000 funcionó con recursos del PAREB.

-Programa para Abatir el rezago en Educación Básica (PAREB) que sustituyó al PARE.

-Programa Integral para Abatir el Rezago Educativo (PIARE). Está dirigido a educación inicial, preescolar y primaria y además, educación para los adultos de las comunidades beneficiadas. En 1997 el PAREB se incorporó al PIARE (SEP, 1999:45). Otros programas con financiamiento internacional que no se han descrito en este texto son el Programa para el Desarrollo de la Educación Inicial (PRODEI) y el Programa de Apoyo a Escuelas en Desventaja (PAED).

Los PC implican un incentivo económico para los maestros, el cual tiene el propósito de «arraigarlos» en las comunidades donde se encuentran las escuelas y de estimular el trabajo docente que se realiza en condiciones difíciles.

Por otra parte, los PC constan en general de cinco componentes que, a su vez, son los objetos de evaluación de dichos programas y que incluyen: capacitación a docentes y directivos, incentivos, material didáctico, infraestructura y equipamiento y fortalecimiento institucional.

Según informa la SEP (2001) los PC funcionan en 31 entidades en «47,878 escuelas primarias rurales, (o 95% de las 50,636 existentes, incluidas el total de escuelas indígenas) y para 4,336,250 alumnos» (SEP, 2002).

De los PC que han funcionado hasta ahora en el país sólo ha sido objeto de evaluación por parte de la SEP y de instituciones especializadas externas a ésta el PARE en 1996 (Velázquez, 1996; Muñoz, 1996; Ezpeleta y Weiss, 2000). En esos estudios, se ha encontrado, en general, que los componentes no lograron cumplir los objetivos que se propuso el Programa. No pudieron asociarse los resultados con un mayor aprovechamiento de los alumnos y cambiar algunas de las prácticas no deseables que permanecen tanto en la administración como en la gestión escolar y en las prácticas pedagógicas. Los estudios mencionados abundan en las causas por las que ocurrió esto y hacen recomendaciones importantes para subsanar dichas limitaciones del Programa, como la atención más específica, con materiales especiales y capacitación, referida al trabajo pedagógico, particularmente en lo que toca a la enseñanza multigrado.⁸ Se hacen también otras recomendaciones como por ejemplo, que: «Los programas compensatorios deben escapar de la estructura formal del sistema, trabajar en paralelo y al margen de los intereses sindicales. Ahí, junto con las prácticas burocráticas heredadas del corporativismo, reside

la fuente de la precariedad institucional y de la perversión de los programas que persiguen los mejores fines" (Órnelas, 2001: 182).

En una de las evaluaciones mencionadas se destaca también el papel que tiene cursar el preescolar para que los niños tengan mayor éxito en la primaria (Velázquez, 1996).

En México los programas compensatorios donde están implicados los incentivos y la evaluación del trabajo docente y que se desarrollaron a partir de los noventa, no han sido evaluados, al menos, de manera que esas evaluaciones sean socialmente conocidas, excepto por las evaluaciones al PARE ya mencionadas, de los demás no hay información conocida sobre su operación y el desempeño de sus componentes.

Por otro lado, parece ser que se ha intentado hacer ajustes a los programas posteriores al PARE atendiendo a los problemas encontrados en las evaluaciones que se le realizaron a éste. Según informaciones provenientes del CONAFE se prepara una evaluación del PIERE que está por realizarse a partir de 2002 (García, 2002).

Asimismo, la Dirección General de Evaluación de la SEP informa que como resultado de su línea de acción "Evaluación en el Aula" y: "Dentro del ámbito del PARE se publicó el folleto *Prácticas de evaluación en el aula* para el docente de educación primaria con un tiraje de 95 mil ejemplares el cual apoyaría la aplicación de estrategias para que los docentes de educación primaria desarrollaran y fortalecieran los procesos de evaluación más adecuados en el salón de clases" (SEP, 2000:657).

Finalmente, en términos más globales, la SEP reporta diversas acciones tendientes a consolidar un ya iniciado programa llamado Estudio de Evaluación de Educación Primaria que ha implicado evaluaciones anuales desde el inicio del PARE hasta el 2000 y el trabajo continuo hacia las entidades federativas para proseguir el trabajo ya desarrollado con ellas durante la

década de los noventa para el Programa de Instalación y Fortalecimiento de las Áreas Estatales de Evaluación. El trabajo en este campo – según la SEP– "ha consistido en el establecimiento de un área específica en cada entidad, con una estructura mínima pero con la capacidad suficiente para desarrollar las acciones de evaluación previstas tanto por los programas compensatorios como por el Sistema Nacional de Evaluación Educativa" (SEP, 2000:257).

Si bien entre el programa de CM y los PC se cubre una población importante y la que más rezagos presenta en materia de educación básica a la vez que se incentiva y evalúa el trabajo docente, faltan todavía, según los agentes involucrados y especialistas en el tema, más y mejores acciones tendientes tanto a la evaluación y presentación de resultados como a medidas correctivas sobre lo que no ha funcionado en dichos programas. Esperamos que algo de ello se cumpla con la consolidación y mayor desarrollo de los sistemas de evaluación nacionales sobre la calidad de la educación que se presta en México.

NOTAS

¹ Este artículo se basa en el trabajo que presenté en: *Profesores en América Latina: carrera, incentivos y evaluación de profesores de educación básica en México*. Conference «Teachers performance in Latin America: time for new priorities». The interamerican development BID-UNESCO, Brasilia, Brasil, 2002.

² Desde luego, las ideas sobre calidad de la educación y su evaluación, en su producción y circulación social, están preñadas de debates que no pueden describirse en este texto. Entre ellos se cuentan los relativos a los valores implicados en la evaluación y a las implicaciones políticas de la misma. (Tiana, 1996).

³ La clave corresponde a la entrevista. Las que se usan en este documento son: EDCMSEP-SNTE: Entrevista Director Comisión Mixta SEP-SNTE, por parte de la SEP; EDDG-SEP: Entrevista Director de la Dirección General de Evaluación de la SEP; EDPRONAR: Entrevista Directora del Programa Nacional de Actualización de Profesores de Educación Básica en Servicio; EMSISX: Entrevista maestro de la Sección Sindical IX que representa a los maestros de primaria del D.F. tradicionalmente disidente del SNTE.

⁴ Este es un número aproximado declarado por algunos de los directores de la SEP entrevistados, pero con la salvedad de que no pueden precisar exactamente cuántos maestros están en el programa.

⁵ Actualmente, existe una presión del magisterio para que todos los docentes de educación básica ingresen a CM por lo menos en la primera categoría.

⁶ Este es un tema que se espera que pueda ser abordado por el recientemente instaurado (2002) Instituto Nacional de Evaluación Educativa.

⁷ Esta institución opera en todo el país, promueve y desarrolla, bajo el marco normativo de la SEP programas de apoyo pedagógico y material para las zonas rurales que presentan los mayores índices de rezago educativo en la República. Es la encargada del programa de Cursos Comunitarios que es una alternativa de educación primaria y preescolar para zonas rurales aisladas que no cuentan con el servicio regular para estos niveles educativos. Cursos Comunitarios funcionan en México desde la década de los setenta.

⁸ En México se ha recomendado en un estudio sobre escuelas multigrado [Fuenlabrada y Weiss (coords.) Candela, Ezpeleta, Fuenlabrada, Kalman y Mercado 1997] que para mejorar el trabajo docente en esa modalidad se recurra al modelo de cursos comunitarios de CONAFE que ha mostrado en diferentes evaluaciones tener resultados comparables o aun superiores a escuelas de educación indígena o rurales regulares (SEP, 2000:656).

REFERENCIAS

- Bracho, Teresa, 2001, *Diseño de una política educativa: El programa Escuelas de Calidad. Evaluación externa*. México, Centro de Investigación y Docencia Económicas.
- Callejas Arroyo, Juan N., 1992 *Carrera Magisterial, problemática y evaluación*, Canales, Alejandro; María De Ibarrola; Pablo Latapí; Felipe Martínez; Javier Mendoza; Carlos Muñoz; Elsie Rockwell; Roberto Rodríguez y Lorenza Villa, 2002, "Observatorio Ciudadano", *La Jornada*, agosto 13 de 1999, www.observatorio.org/comunicados/comun14.html.
- CEPAL-UNESCO, 1997, *Educación y conocimiento, eje de la transformación productiva con equidad*.
- Consejo Nacional de Fomento Educativo, 1994, *Asesoría pedagógica, Documento del Docente*, México.
- Consejo Nacional de Fomento Educativo, 1996, *Diario oficial*, 1993.
- Ezpeleta y Weiss, 2000, *Cambiar la escuela rural. Evaluación cualitativa del Programa para Abatir el Rezago Educativo*, México, Departamento de Investigaciones Educativas del Centro de Investigación y de Estudios Avanzados del IPN
- Fonbrum, C.J., Tichy, N.H., Devana, M.A., 1984, *Strategic Human Resource Management*, New York, Wiley.
- Fuenlabrada, I; E. Weiss (coords.) A. Candela; J. Ezpeleta; I. Fuenlabrada; J. Kalman; R. Mercado, 1997, *Las prácticas escolares y docentes en las escuelas multigrado de la educación primaria. Lineamientos para un nuevo modelo*, reporte elaborado para el CONAFE, DIE-CINVESTAV, México.
- García O., Mónica, 2002, *El papel de la experiencia docente en los cursos de capacitación para maestros multigrado*, tesis maestría, México, Departamento de Investigaciones Educativas del Centro de Investigación y de Estudios Avanzados del IPN.
- House, E. R., 1993, *Professional Evaluation*, Social Impact and Political Consequences London and New Delhi: Newbury Park, SAGE.
- Instituto Nacional de Geografía Estadística e Informática, 1990, *Censo Nacional de Población*.
- Kemerer y Thiagarajan, 1989, *Teacher Incentive Systems Final Report*, Policy Research Initiative, USAID.
- Latapí, Pablo, 1973, *Mitos y realidades de la educación mexicana. 1971-1972. Una opinión independiente*, Centro de Estudios Educativos.
- Muñoz I., Carlos, 1996, *Principales resultados y recomendaciones de la evaluación del PARE*, México, Centro de Estudios Educativos.
- Mercado, Ruth, 2000, *El trabajo docente en el medio rural*, México, Secretaría de Educación Pública.
- Ornelas, Carlos, 2001, "Equidad: educación comunitaria y programas compensatorios," en Carlos Ornelas (comp.) *Investigación y política educativas: Ensayos en honor de Pablo Latapí*, México, Santillana, aula XXI, 135-185.
- Schmelkes, Sylvia, 2001, *Evaluación de la educación básica*, México, Departamento de Investigaciones Educativas del Centro de Investigación y de Estudios Avanzados del IPN
- Schmelkes, 2002, *La difícil relación entre la evaluación educativa y la calidad de la educación*, www.sep.gob.mx
- Secretaría de Educación Pública, 1992, *Acuerdo Nacional para la Modernización de la Educación Básica*.
- Secretaría de Educación Pública, 1993, *Ley General de Educación*.
- Secretaría de Educación Pública, 1998, *Lineamientos Generales de Carrera Magisterial*
- Secretaría de Educación Pública-Consejo Nacional de Fomento Educativo, 1994.
- Secretaría de Educación Pública, 1999, *Perfil de la Educación en México*.
- Secretaría de Educación Pública, 2000, *Memoria del Quehacer Educativo 1995-2000*.
- Secretaría de Educación Pública, 2001a, *Informe de Labores 1999-2000*.
- Secretaría de Educación Pública, 2001b, *Programa Nacional de Educación 2001-2006*.
- Secretaría de Educación Pública, 2002, www.sep.gob.mx
- Tiana, Alejandro, 1996, "La evaluación de los sistemas educativos" en *Revista Iberoamericana de educación*, Evaluación de la calidad de la educación, Organización de Estados Iberoamericanos N° 10, www.campus.oei.org/oeivirt/rie10a02.htm
- Toranzos, Lilia, 1996, "Evaluación y Calidad", en *Revista Iberoamericana de educación*, Evaluación de la calidad de la educación, Organización de Estados Iberoamericanos N° 10 www.campus-oei.org/oeivirt/rie10a03.htm
- Torres, Rosa María y Emilio Tenti Fanfani (2000), "Políticas Educativas y Equidad en México: la experiencia de la Educación Comunitaria, la Telesecundaria y los Programas Compensatorios". En *Equidad y Calidad en la Educación Básica. La experiencia del CONAFE y la Telesecundaria en México*, CONAFE, México.
- UNESCO, 1994 *Conferencia Mundial sobre Educación para Todos*, Jomtien, Tailandia.
- Uribe, Claudia, 1999, *Políticas e incentivos que contribuyen al mejoramiento del desempeño y motivación de los maestros*.
- Velázquez, Víctor, M., 1996, "La evaluación como recurso para elevar la calidad de la educación en México," en *Revista Iberoamericana de educación*. Evaluación de la calidad de la educación, Organización de Estados Iberoamericanos N° 10 www.campus.oei.org/oeivirt/rie10a08.htm
- Weiss, E., J. Ezpeleta, 2000, "Cambiar la escuela rural", en Ezpeleta, J y E. Weiss, *Evaluación cualitativa del Programa para Abatir el Rezago Educativo*, México, Departamento de Investigaciones Educativas del Centro de Investigación y de Estudios Avanzados del IPN.

CRITERIOS DE EVALUACION PEDAGOGICA BASES Y PERSPECTIVAS

La obra de arte es un producto cultural que surge de la actividad creativa del ser humano. Su función principal es la de expresar sentimientos, ideas y valores. En el ámbito educativo, el arte sirve como un medio para desarrollar la sensibilidad, la imaginación y la capacidad de análisis crítico. A través del estudio de las obras de arte, los estudiantes pueden comprender mejor la cultura y la historia de su sociedad y de otras culturas. Además, el arte puede ser utilizado como una herramienta para fomentar la creatividad y la innovación en el aprendizaje.



"MINOTAURO" *bronce a la cera perdida*



CRITERIOS DE EVALUACIÓN EDUCATIVA: BASES Y PERSPECTIVAS

ANDRÉS SÁNCHEZ MOGUEL

RESUMEN

LA EVALUACIÓN EDUCATIVA SE BASA EN UNA SERIE DE SUPUESTOS QUE no siempre son evidentes para los observadores externos. Incluso, en ocasiones, una mala comprensión de lo que implica la evaluación por parte de quienes la llevan a cabo conduce al empleo de prácticas evaluativas incorrectas. Los criterios de aplicación comienzan desde el momento en que se plantean las preguntas “¿qué evaluar?”, “¿para qué evaluar?”, “¿quién lleva a cabo la evaluación?” y “¿cómo hacerla?”. Tras tomar estas decisiones, el uso de criterios de evaluación se multiplica: deben tomarse en cuenta los criterios técnicos que garanticen una evaluación imparcial y justa, los criterios cualitativos que garanticen la validez y la pertinencia del proceso y los criterios sociales que garanticen la utilidad y el desarrollo del sistema de evaluación. El trabajo en extenso explica algunos de estos criterios, sus interrelaciones y la necesidad de definir comunitariamente los faltantes.

La evaluación del aprendizaje es una práctica que ha ido cobrando relevancia en la reflexión de los estudiosos educativos de nuestro país en los últimos 30 años. En 1979, García Cortés señalaba el problema de la “realización paupérrima de estudios de evaluación educativa en nuestro medio”. El estado de conocimiento sobre la evaluación del aprendizaje del II Congreso Nacional de Investigación Educativa (Martínez y colaboradores., 1993) señalaba que eran muy pocos los trabajos realizados a lo largo de la década de 1982 a 1992 sobre este aspecto de la evaluación educativa (se encontraron 81 estudios, tesis y artículos únicamente, en revisión exhaustiva). En el III Congreso Nacional de Investigación Educativa (1995) se observaba aún la dispersión de la temática, ya que los trabajos presentados eran en general técnicos y referentes a problemas relativamente aislados (Sánchez Moguel, 1997). El Programa Nacional de Educación 2001-2006 (2001) ya reconoce que “la evaluación permanente y sistemática (...) es un importante instrumento de gestión (...) para llevar a cabo acciones de mejora, así como para rendir

cuentas a la sociedad". Actualmente, existen instituciones de creación reciente (el Instituto para la Evaluación de la Educación es un ejemplo) cuyas preocupaciones giran en torno a desarrollar la evaluación de la educación así como la reflexión de la misma. En el reciente informe de seguimiento de la Educación para Todos en el Mundo, de la UNESCO 2004 se señala, por ejemplo, que "los niveles de aprovechamiento escolar son muy bajos en los países con ingresos escasos o medios, así como entre los grupos desfavorecidos de algunas naciones industrializadas", al analizar los resultados de 127 países.

Este *boom* de la evaluación de lo educativo contrasta con la falta de interés en lo técnico-teórico del tema, y el consiguiente poco espacio dedicado al análisis de estos aspectos en las instituciones formadoras de maestros, lo cual ha llevado a la existencia, en algunos casos, de prácticas evaluativas del logro escolar en el interior del aula (¡y en ocasiones, también encontradas en evaluaciones de más amplio alcance!) con las siguientes características:

- a) Falta de reflexión sobre las razones por las cuales se evalúa, dando prioridad al cumplimiento administrativo sobre la utilidad real de la información.
- b) Procedimientos e instrumentos de evaluación poco planeados y mal estructurados.
- c) Poco análisis de lo obtenido en las evaluaciones, priorizando los intereses crediticios ("aprobé o no", "15 alumnos reprobaron", etcétera) sobre los logros académicos ("hasta qué punto se ha aprendido este contenido", "ya hay un conocimiento generalizado de esta materia o no", etcétera).
- d) Una serie de factores que distorsionan la medición de lo que los estudiantes realmente saben, como ligar la conducta a la calificación, las altas posibilidades de fraude o las pruebas que privilegian lo memorístico sobre lo reflexivo.

Por todo ello, es necesario regresar a los fundamentos, a las razones por las cuales se hace la evaluación educativa. El objetivo del presente trabajo es analizar algunas cuestiones de fondo de la teoría de la evaluación, en su relación específica con la evaluación del aprendizaje, con la intención de mostrar a los partidarios de las posturas extremistas que las prácticas evaluativas en el aula, en la escuela y en el sistema educativo no tienen por qué ser ciegas recetas que inventó un técnico, y que debe haber una reflexión seria sobre la manera de abordar una tarea de evaluación y de extraer el significado de cada dato obtenido.

La evaluación educativa es una estrategia de recolección de información sobre los diferentes momentos, actores y auxiliares del proceso enseñanza-aprendizaje, y de reflexión y juicio de valor con respecto a esta información. Si bien es indispensable que cada profesor haga evaluaciones particulares y a profundidad de estos elementos en su espacio de trabajo, es también necesario contar con perspectivas más generales de la labor académica que sirvan de monitor a la comunidad general de las escuelas. Por tanto, la evaluación de grandes cantidades de estudiantes a partir de instrumentos estandarizados, la evaluación reflexiva entre los maestros y la evaluación institucional, se presentan como prácticas útiles en el acopio de datos globales sobre la situación escolar. Múltiples circunstancias han hecho que en algunas ocasiones la evaluación general se lleve a cabo con una perspectiva de conteo y control que recaba datos con instrumentos que no han sido depurados, asigna calificativos por simple tradición numérica¹, y genera listados llenos de cifras a los que no se da un uso en el perfeccionamiento de los esquemas educativos.

La intención de quienes realizan el proceso de evaluación y quienes lo promueven es decisiva en el énfasis que se va a dar a unos u otros elementos del sistema de evaluación que se genere. Así, ante el único interés de cumplir

de la manera más eficiente posible con una exigencia administrativa o estatutaria, los evaluadores prefieren hacer instrumentos que estén listos para ser aplicados en muy poco tiempo, que sean lo más económicos posible en dinero, esfuerzo y tiempo, así como análisis de los datos resultantes de las aplicaciones que sean rápidos y no evidencien las carencias de sus instrumentos, aunque estos análisis no tengan uso fuera de los expedientes. En cambio, ante el interés de obtener información útil en el análisis de la situación educativa, las preferencias son por desarrollar instrumentos con un nivel suficiente de validez, confiabilidad y pertinencia, que lleven a obtener datos adecuados para el propósito de conocer los resultados escolares, permitiéndose hacer mayores inversiones de recursos que en otros modelos.

Varios puntos de decisión se plantean en este proceso: ¿quién debe decidir las áreas a evaluar, los contenidos de las áreas a evaluar y los métodos de evaluación?, ¿quién debe crear los instrumentos de evaluación?, ¿qué características deben tener los ítems de los instrumentos y/o métodos de evaluación?, ¿en qué momento se pueden considerar adecuados los instrumentos y/o métodos de evaluación?, ¿qué criterios deben tomarse en cuenta para analizar la información obtenida?, ¿qué segmentos y agrupaciones de la información obtenida son más útiles, y a quién? La pertinencia y utilidad del proceso dependen de que se den respuestas razonables a estas preguntas, en los momentos oportunos, perfilando un sistema.

Existe una serie de cuestiones respecto de la forma de abordar el trabajo de la evaluación masiva que ha sido resuelta con base en criterios poco claros, e incluso sin considerar siquiera los problemas teóricos y de aplicación, planteando las decisiones en términos meramente técnicos y/o administrativos. García Cortés (1979) explicó hace tiempo la gran importancia que tiene determinar, para cada caso

específico, las respuestas a ¿para qué evaluar? y ¿qué evaluar? Responder estas dos preguntas señala criterios que generalmente son de gran ayuda para tomar decisiones sobre la manera de operar un programa de evaluación.

Es conocido el hecho de que un instrumento y/o método de medición (desde una regla hasta un fotocolorímetro) que va a ser utilizado en repetidas ocasiones, y para sacar conclusiones al hacer comparaciones, debe cumplir ciertos criterios de confiabilidad y validez, así como ser pertinente. Dado que los sistemas de evaluación sistemáticos utilizan instrumentos o métodos de medición en muchas ocasiones, permítaseme hacer un repaso de los conceptos de confiabilidad, validez y pertinencia que, por otra parte, muchas veces son mal utilizados.

La confiabilidad se refiere a la estabilidad del instrumento a través del tiempo y de las muestras. Sabemos que las condiciones y cualidades de los actores educativos son dinámicas, así que esta primera definición no parece sernos muy útil en la escuela. Una segunda aproximación refiere que la medida confiable es aquella que se encuentra libre de error. Sin embargo, aunque esto parece ser suficientemente exacto en las ciencias naturales (por ejemplo, en la medición del contenido de sodio en un compuesto), en las ciencias sociales es muy ingenuo pensar en alcanzar la exactitud (puede incluso plantearse la duda de la posibilidad o la necesidad de ella a nivel filosófico). Una definición que nos parece más viable para la tarea que nos ocupa es la de considerar semejante a lo que es semejante y diferente a lo que lo es, lo cual acerca la noción cuantitativa de confiabilidad a la noción cualitativa de imparcialidad (Fernández Ríos, 1994) con un toque de lógica difusa.

Los Estándares para la Evaluación Educativa y Psicológica por Medio de Pruebas (APA, 1985) señalan que "la validez es la consideración más importante en la evaluación por medio

de pruebas. El concepto se refiere a la pertinencia, significación y utilidad de las inferencias específicas que se hagan de los puntajes de una prueba." Es muy difundida la definición básica de validez en instrumentos de evaluación que indica que éstos son válidos cuando miden lo que pretenden medir (Magnusson, 1975). Sin embargo, el concepto de validez aparentemente tan simple se encuentra en el centro de una polémica que aún se lleva a cabo. Gray (1997), haciendo una pequeña revisión, señala que:

...en 1949 Cronbach declaró que la definición de validez como "la extensión con que una prueba mide lo que pretende medir" era comúnmente aceptada, aunque él prefería una ligera modificación: una prueba es válida en el grado en que sabemos qué mide o predice. Cureton (1951) provee una definición similar: La cuestión esencial de la validez en las pruebas es qué tan bien realizan la tarea para la cual se les esté usando. La validez es definida entonces en términos de la correlación entre los puntajes de una prueba y los "verdaderos" puntajes del criterio. La perdurable definición de Anastasi (usada desde 1954 hasta 1997), "la validez es qué mide una prueba y qué tan bien lo hace", es también citada ampliamente.

Gray señala también que aunque Cronbach trató de evitar redefinir el término a partir de 1949, en 1971 hizo un comentario que reavivó la controversia: 'validación es el proceso de examinar la precisión de una predicción o inferencia específica hecha a partir de los puntajes de una prueba', o bien, como señalan otros autores "la validez se refiere no a las puntuaciones o datos en sí mismos, sino a las inferencias que se hagan a partir de ellos bajo determinadas circunstancias" (Silva y Martorell, 1991); "lo que se valida no es el instrumento, sino la interpretación de los datos obtenidos por medio de un procedimiento especificado" (Aragón, 1990); "la validez depende de la "adecuación y pertinencia de inferencias y acciones" basadas en

los resultados de la evaluación" (Messick, 1989, en Linn y Baker, 1996).

Finalmente, es importante señalar que, aunque muchos autores (Rudner, 1993; Niemi, 1996; Aragón, op. cit.; Tourón, 1989; Burns, 1996; GAO, 1991) reportan al menos tres tipos "clásicos" de validez, actualmente existe una tendencia a considerar un tipo único (cf. Gray, op. cit.; Silva y Martorell, op. cit., quienes incluso sugieren que el concepto de confiabilidad también es mucho más cercano al de validez de lo que se ha pensado), de la cual, eso sí, se obtienen distintos tipos de evidencias: "Se ha sugerido que la validez de constructo abarca tanto a la validez de criterio como a la de contenido. Sheperd anotó que la validez de constructo incluye los requisitos teóricos y empíricos de la validez de contenido y de criterio. Anastasi (1986) coincide en que la validez de constructo subsume los requisitos de la validez de contenido y de criterio" (Stapleton, 1997).

En resumen, debemos considerar como una cualidad primordial de las pruebas la posibilidad de extraer de manera correcta y verdadera el significado de sus puntajes. Dado que esto no sólo depende de la prueba sino de las circunstancias de aplicación y los objetivos de la misma, diferentes aspectos de esta cualidad pueden ser considerados. Aunque esto puede parecer sencillo cuando los instrumentos de medición son muy cercanos a la realidad física, la tarea se complejiza conforme el objeto de evaluación se vuelve abstracto o difícil de observar directamente. Tal es el caso de la evaluación del aprendizaje.

Existen tres puntos de especial importancia en cuanto a la pertinencia de un procedimiento de evaluación: 1) Que el tipo de información arrojada sea realmente un indicador útil sobre los conocimientos y/o habilidades de la población; 2) Que existan criterios fundamentados para interpretar los datos obtenidos en la examinación de grandes grupos; 3) Que la información obtenida llegue a los

destinatarios que pueden darle utilidad, es decir, los profesores, planificadores académicos de la escuela y los propios estudiantes. A continuación, se desarrollan estos puntos:

1) Existe una discusión importante con respecto a los instrumentos de evaluación que se utilizan en educación. En realidad, el origen de la discusión está en el pseudoproblema de lo cuantitativo *vs.* lo cualitativo. Algunos autores como Díaz Barriga (1982) han señalado que la evaluación no debe hacer uso de la tecnología de medición generada por la psicometría pues “se minimiza tanto el proceso mismo de la evaluación del aprendizaje como la noción de aprendizaje y la de docencia.”; otros plantean problemas técnicos en el uso de ciertos tipos de evaluación “objetiva”, por ejemplo que sólo se mide lo que el alumno memoriza, o la posibilidad de acertar por azar (Fermín, 1971). Finalmente, otros autores, reconociendo los problemas de “el hiato indudable entre la medida y lo que pretendemos medir” y “el uso de la medición en la evaluación educativa” (Tourón, 1989) confían en el uso del método científico para la valoración escolar y generan estrategias cada vez más refinadas para salvar los problemas mencionados (cf. Tourón, 1989; Tirado y Serrano, 1989; Rodríguez y García, 1982). Nosotros consideramos importante rescatar nociones de cada uno de estos planteamientos que equilibren una práctica evaluativa eficaz, eficiente y útil. Así, creemos que el diseño de un instrumento de evaluación debe estar firmemente enraizado en una reflexión del para qué y el qué evaluar, de tal modo que si una técnica y el qué evaluar se revelan incompatibles, debe ser la técnica la que cambie.

2) Una problemática común entre los que atacan el problema de la evaluación desde un punto de vista social y/o filosófico, que en cambio es poco tocado por quienes tienen el punto de vista únicamente técnico, es el criterio de pase/

reprobación en el caso de evaluaciones con fines de acreditación, o el criterio de “aceptabilidad/inaceptabilidad” en caso de evaluaciones para toma de decisiones. Sabemos que existen en este sentido juicios “por criterio” y juicios “por norma”. En el primero, se establece de antemano el mínimo aceptable, que depende de una discusión teórica de lo que se va a evaluar, y en el segundo se juzga cada caso individual con base en la cercanía o lejanía que tenga con la media de ejecución (por ejemplo, a partir del número de desviaciones estándar), y el sentido de esta distancia (positivo o negativo). En papel, estos criterios pueden parecer fáciles de aplicar, pero en la práctica vale la pena reflexionar profundamente en los motivos y las consecuencias de permitir, por ejemplo, que sean acreditados estudiantes de medicina con calificaciones apenas pasables, además de relativas (dado que ante el examen de una escuela podrían sacar altas calificaciones y ante el de otra podrían sacar calificaciones bajas). En efecto, no hay una estandarización en la dificultad que deben tener este tipo de pruebas, ni normas o consejos de uso generalizado para establecer los criterios. Por todo ello, el conjunto de la sociedad escolar debe dedicar tiempo a la reflexión de este problema, aterrizándolo en programas concretos en que se trabaje y tomando decisiones con respecto a los criterios a emplear en ellos.

La relatividad llegó a la física —una de las ciencias más duras y clásicas— hace unos ochenta años; tal vez ya es tiempo de que llegue a la educación: no existen criterios ni fórmulas universales para llevar a cabo las tareas evaluativas, ni deben existir. Cada sociedad escolar debe definir los suyos propios. “Las interpretaciones válidas del significado y la verdad son hechas por gente que comparte decisiones y las consecuencias de las decisiones”, escribe Steinar Kvale a propósito del conocimiento (según traducción inédita de César Carrascosa). Estos términos, llevados a la evaluación educativa

implican el compromiso y la reflexión de todos los participantes de la educación.

3) El último problema que planteamos para reflexionar en cuanto a la pertinencia de la evaluación es el de decidir la manera de presentar la información y el análisis realizados con base en la aplicación de los métodos e instrumentos, así como los modos e instancias de distribución de estos datos. Consideramos útil discutir de antemano estos elementos, y evaluar la certeza de nuestras decisiones tras cada experiencia de divulgación, mejorando cada vez las estrategias de difusión con base en las observaciones que se hagan. También consideramos útil consignar el proceso de búsqueda de las mejores estrategias en escritos que puedan ser de utilidad a otros en su práctica evaluativa y recuperar la experiencia para facilitar el camino a otros evaluadores.

CONCLUSIONES

Con base en lo reflexionado, concluyo que la evaluación de lo educativo es una tarea fundamental, por su función de retroalimentación del sistema y sus subsistemas. La época de decidir entre una evaluación cuantitativa y una cualitativa ya pasó, aunque es cierto que subsiste el problema técnico de que algunos tipos de evaluación, por su naturaleza, tienden a pertenecer mayormente a uno de estos dos campos. La evaluación de lo educativo debe ser llevada a cabo, en primera instancia, por la comunidad involucrada en el proceso educativo. Debe haber participación de los actores educativos en las diferentes fases de la evaluación, principalmente en las de fundamento (cuando se establecen los criterios con base en valores reconocidos por el grupo) y en las de retroalimentación propiamente dicha. Una cultura de evaluación no significa una época de terror, de premios y castigos basados en procesos desconocidos que asignan números bajo reglas cabalísticas oscu-

rísimas: esa es la cultura de la zanahoria y el palo para hacer andar al "motor ecológico". En una cultura de evaluación hay un interés de los participantes del proceso educativo por conocer el desempeño personal y grupal para analizar lo alcanzado y dirigir esfuerzos con conocimiento de causas que aumenten las probabilidades de éxito. Asimismo, hay un esfuerzo sostenido por revisar y mejorar constantemente los medios por los que se obtiene la información que sirve de base para los análisis.

NOTA

¹ Nos referimos especialmente a la conocida "escala de cero a diez", en que 6 o más significa "aprobado", es decir, adecuado, y 5 o menos significa "reprobado", es decir, inadecuado. Del mismo modo podemos hablar del sistema NA-MB.

REFERENCIAS

- American Psychological Association. 1985. *Standards for Educational and Psychological Testing*, Washington, D.C., APA.
- Aragón, Laura. 1990. *Elaboración de un instrumento de evaluación conductual, con validez de contenido y de tratamiento, para niños disléxicos*. Tesis de Grado. ENEP-Iztacala, UNAM.
- Burns, William. 1996. Chapter 5: Content Validity, Face Validity and Quantitative Face Validity. En Barrett, Richard (ed.), *Fair employment strategies in human resource management*. Connecticut, Quorum Books.
- Carrascosa, César (trad.) s.d. Traducción inédita de "Psicología Posmoderna: ¿Una Contradicción de Términos?" de Steinar Kvale.
- Díaz Barriga, Ángel. 1982. «Tesis para una teoría de la Evaluación y sus Derivaciones en la Docencia». En *Perfiles Educativos*, No. 15, enero-marzo, pp. 16-37.
- Fermin, M. 1971. *La Evaluación, los Exámenes, las Calificaciones*. Madrid, Kapelusz.
- Fernández Ríos, L.F. 1994. *Manual de Psicología Preventiva: Teoría y Práctica*. Madrid, Siglo XXI editores. Capítulo 7.
- General Accounting Office. 1991. *Designing Evaluations*. United States General Accounting Office.
- García Cortés, Fernando. 1979. "La Evaluación en la Educación". En *Perfiles Educativos*, No. 3, enero-marzo, pp. 37-43.
- Gray, B.T. 1997. «Controversies Regarding the Nature of Score Validity: Still Crazy After All These Years» Presentado en la reunión anual de la Southwest Educational Research Association, Austin, enero.
- Linn, R.L. y Baker, E.L. 1996. *Assessing the validity of the National Assessment of Educational Progress: NAEP technical review panel white paper*. U.S. Department of Education.
- Magnusson, D. 1975. *Teoría de los Tests*. México, Trillas.
- Martínez, F.F.; Fuentes Trejo, G.; Cepeda Hinojosa, B.; Burgos Fajardo, R. 1993. *Estado de Conocimiento 8: Evaluación del Aprendizaje*. México, Comité Organizador del Segundo Congreso Nacional de Investigación Educativa/Sindicato Nacional de Trabajadores de la Educación.

- Niemi, D. 1996. *Instructional influences on content area explanations and representational knowledge: evidence for the construct validity of measures of principled understanding*. Estados Unidos de América, National Center for Research on Evaluation, Standards, and Student Testing.
- Rodríguez Cruz, H.M. y García González, E. 1982. *Evaluación en el Aula*. México, Trillas.
- Rudner, L.M. 1993. *Test Evaluation*. ERIC /AE. (URL: <http://136.242.172.58/intass.htm>).
- Sánchez Moguel, Andrés. 1997. Evaluación de la Educación. Introducción. En *Currículum, Evaluación y Planeación Educativa*. Ángel Díaz Barriga (Coord.), México, Comie-CESU-ENEP Iztacala.
- Secretaría de Educación Pública. 2001. *Programa Nacional de Educación 2001-2006*. México, SEP.
- Silva, F. y Martorell, C. 1991. Evaluación Conductual y Evaluación Tradicional: la Cuestión Psicométrica. En: Caballo, V. E. (Ed.), *Manual de técnicas de terapia y modificación de conducta*. Madrid, Siglo XXI.
- Stapleton, C.D. 1997. "Basic Concepts in Exploratory Factor Analysis (EFA) as a Tool to Evaluate Score Validity: A Right-Brained Approach". Presentado en la reunión anual de la Southwest Educational Research Association, Austin, Enero.
- Tirado Segura, Felipe y Serrano Carrillo, Víctor. 1989. "En torno a la calidad de la educación pública y privada en México", en *Ciencia y Desarrollo*, vol. XV, núm. 85, marzo-abril.
- Tourón, J. 1989. "La Validación de Constructo: su Aplicación al CEED (Cuestionario para la Evaluación de la Eficacia Docente)". En *Bordón*, V. 41 (3-4).



EL CINVESTAV EXTIENDE SU MÁS CÁLIDA FELICITACIÓN A SUS INVESTIGADORES, ESTUDIANTES Y AUXILIARES de investigación por los logros y distinciones académicas alcanzadas en 2004. Durante este año se graduaron 340 estudiantes de Maestría y 165 de Doctorado. A estas cifras las respalda la calidad de los programas de posgrado pues, actualmente, de los 32 programas de Competencia Internacional, de acuerdo al Padrón Nacional de Posgrado del Conacyt, 12 son del CINVESTAV

Premio Canifarma 2003 en el área de Investigación Básica, por el trabajo "La Neocarzinostatina induce apoptosis e inhibe la proliferación celular, en forma dependiente de P53 en células derivadas de cáncer cervical", otorgado por la Cámara Nacional de la Industria Farmacéutica
Dr. Patricio Gariglio Vidal

Primer Premio en la Reunión Anual 2004 de la Western Pharmacology Society para estudiantes mexicanos, por el trabajo "Pharmacokinetic and Pharmacodynamic Characterization of Acemetacin a produg of Indomethacin"
M. en C. Miguel Gil Flores (estudiante)
Dr. Gilberto Castañeda Hernández, Director de Tesis

Primer Lugar en el área de Investigación Publicada Clínica del XIV Premio Nacional de Investigación de la Fundación GlaxoSmithKline, por el trabajo "Evaluation of opioid peptide and muscarinic receptors in human epileptogenic neocortex: An autoradiography study"
Dra. Luisa L. Rocha Arrieta y Colaboradores

Presentación del libro "Breve Historia de la Computación y sus Pioneros"
Dr. Carlos A. Coello Coello

Director del Centro Latinoamericano de Física (CLAF)
Dr. Feliciano Sánchez Sinencio

Vicepresidente de la Asociación de México de Control Automático para el Bienio 2004 - 2005
Dr. Eduardo Aranda Bricaire

Publicación del Artículo "Mangroves enhance the biomas of coral reef fish communities in the Caribbean" en la Revista Nature
Dr. J. Ernesto Arias González

Premio "Young Scientist Award 2003" a la mejor tesis doctoral, en el Cinvestav, que otorga el Dr. Heinrich Nöth, Presidente de la Academia de Ciencias de Bavaria
Dr. Juan Carlos Gálvez Ruiz (estudiante)
Dra. Angelina Flores Parra, Directora de Tesis

Premio Weizmann-Kahn a la mejor tesis doctoral en Investigación Tecnológica 2003,

por el trabajo "Desarrollo de una prótesis con cuatro grados de libertad activos para reemplazo por arriba de codo"
Dr. Apolo Zeus Escudero Uribe (estudiante)
Dr. Lorenzo Leija Salas, Director de Tesis

Premio a la Investigación 2003 en el Área de Ciencias Biológicas y de la Salud otorgado por la Universidad Autónoma Metropolitana, por el trabajo "Feasibility of Fishmeal Replacement by Shrimp Head Silage Protein Hydrolysate in Nile Tilapia (Oreochromis L) Diets"
Dr. Miguel A. Olvera Novoa

Seleccionado para recibir financiamiento por el programa "U.S. - México Higher Education Partnership for Environmental Sciences and Toxicology" que forma parte de la iniciativa "Alianza para la Prosperidad", auspiciada por las Presidencias de México y de los Estados Unidos del programa Enlaces (TIES - Training Internships, Exchanges and Scholarships), de la Agencia para el Desarrollo Internacional de los Estados Unidos
Dr. Mariano E. Cebrián García

Miembro integrante del Comité Administrativo de la Electron Device Society (EDS) del IEEE en calidad de Miembro del Comité de Actividades Educativas y Presidenta de la Región Latinoamericana de la EDS
Dra. Magali Estrada del Cueto

Miembro del Comité Científico del Noveno Congreso Mundial de Farmacología Clínica y Terapéutica, a realizarse en Quebec, Canadá, en el año 2008
Dr. Gilberto Castañeda Hernández

Premio CARPERMOR en la categoría "A", por el trabajo "Single Multiplex Polymerase Chain Reaction to Detect Diverse Loci Associated with Diarrheagenic Escherichia coli", otorgado por el Grupo Diagnóstico PROA
Dra. Teresa Estrada García

Primer Lugar en la Categoría "B" del Premio CARPERMOR 2004, por el trabajo "Infection with ureC, cagA and vacA Helicobacter pylori strains is associated with metaplasia progression in patients with chronic gastritis", otorgado por la Asociación Mexicana de Bioquímica Clínica, A.C.
Dr. Patricio Gariglio Vidal

Premio Iberoamericano a la Excelencia Educativa 2004, otorgado por el Consejo Iberoamericano en Honor a la Calidad Educativa
Dr. Tonatiuh Matos Chasin

Premio Weizmann 2003 a la mejor tesis doctoral, otorgado por la Academia Mexicana de Ciencias
Dr. José Gabriel Merino Hernández (estudiante)
Dr. Alberto Vela Amieva, Director de Tesis

Miembros regulares de la Academia Mexicana de Ciencias
Dr. José Victor Calderón Salinas
Dr. Jesús Contreras Nuño
Dr. Rubén Gerardo Contreras Patiño
Dr. Francisco Cordero Osorio
Dra. Rosa Ma. Del Ángel Núñez de Cáceres
Dr. Felipe de Jesús González Bravo
Dr. Fernando Antonio Hitt Espinosa
Dr. Alexander Loukianov Georgievich
Dr. Omar Gustavo Miranda Romagnoli
Dr. Luis Enrique Moreno Armella
Dr. Alberto Sánchez Hernández
Dra. Matilde Mineko Shibayama Salas
Dr. Yuriry Shkvarko
Dr. Sergio Armando Tomás Velázquez

Beca de Disertación Doctoral en Ciencias Biomédicas "Hugo Aréchiga Urtuzuástegui", otorgada por el Colegio de Sinaloa
M. en C. Vicente Adrián Canzales Román (estudiante)
Dr. Fernando Navarro García, Director de Tesis

Miembro del Comité Científico Internacional del 5th International Fructan Symposium, realizado en Cuba en Diciembre de 2004
Dra. Mercedes G. López Pérez

Miembro de la Mesa Directiva del Congreso Internacional de Triquinosis, realizado en San Diego, CA, en Agosto de 2004
Dra. Guadalupe Ortega Pierres

Boron in the Americas Award for Distinguished Achievements in Boron Science
Dra. Rosalinda Contreras Theurel

Premio y Medalla Alzheimer 2004, por el trabajo "Conformational changes and truncation of tau protein during tangle evolution in Alzheimer's disease"
Dr. Francisco García Sierra

Miembro del Comité Organizador del International Congress on Mathematics Education (ICME-11), a realizarse en la Cd. de México en el año 2008
Dra. Olimpia Figueras Mourut de Montpellier

Miembro del Consejo Científico de la Comisión Federal para la Protección contra Riesgos Sanitarios (COFEPRIS) de la Secretaría de Salud
Dr. Mariano E. Cebrián García

Miembro del Comité de Directores del Consorcio de la Cuenca del Pacífico para las Ciencias del Ambiente y la Salud
Dr. Mariano E. Cebrián García

Ad hoc Specialist (site visitor) por el Consejo de Acreditación de la Association for Assessment and Accreditation of Laboratory Animal Care International
Lic. Jorge Fernández Hernández (Auxiliar de Investigación)

Premio al Especialista en Animales de Laboratorio 2004 otorgado por la International Council for Laboratory Animal Science (ICLAS)
MVZ. Ricardo Gaxiola Centeno (Auxiliar de Investigación)

Best Presentation Award del Graduate Student Workshop, celebrado durante la Genetic and Evolutionary Computation Conference (GECCO'2004)
M. en C. Efrén Mezura Montes (estudiante)
Dr. Carlos A. Coello Coello, Director de Tesis

Miembro del Consejo Editorial de la Revista "Archives of Medical Research"
Dra. Isaura Meza Gómez Palacio

Primer Premio en Investigación Ficológica Aplicada, XVIII International Seaweed Editor Asociado de la revista "International Journal of Dynamics of Continuous, Discrete and Impulsive Systems, Part B: Applications and Algorithms"
Dr. Edgar Nelson Sánchez Camperos

Reconocimiento por el trabajo en el campo de la Matemática Educativa en Latinoamérica, otorgado por la Sociedad Cubana de Matemática y Computación
Dra. Rosa María Farfán Márquez

Premio Estatal de la Juventud del Gobierno del Estado de Yucatán, en el área de Actividades Académicas (Traectoria académica ejemplar elaboración de investigaciones o estudios científicos, artículos académicos, distinciones recibidas, etc).
M. en C. Omar de la Peña Seaman (estudiante)

Premio Bianual otorgado por la Sociedad de Dispositivos Electrónicos del IEEE, por el trabajo "The integral function methods: a new method to determine the nonlinear harmonic distortion" presentado en el 18TH International Symposium on Microelectronics Technology and Devices (SBMicro 2003)
Ing. Miguel Angel Alemán Arce (estudiante)
Dr. Antonio Cerdeira Aluzarra, Director de Tesis

Ingreso como Miembro Correspondiente de la Academia para el Avance de la Ciencia, la Tecnología y las Humanidades de Puebla
Dr. Eusebio Juaristi y Cosío

Accésit en la categoría de Divulgación del Premio "Bruno Mascanzoni" al libro IMP, por el trabajo "El chapopote y las chapopoterías del Golfo de México"
M. en C. María Concepción Rosano Hernández (estudiante)
Dra. Beatriz Xoconostle Cázares, Directora de Tesis

Premio Maestro Leonardo Oliva 2004, a nivel tesis de doctorado, otorgado por la Cámara Nacional de Comercio de Guadalajara, por el trabajo "Desarrollo de un marcador "SCAR" para dos especies de hongos micorrízicos arbusculares establecidas en cultivo monoxénico"
Dr. Juan Florencio Gómez Leyva (estudiante)
Dr. Juan Pablo Martínez Soriano, Director de Tesis

Premio AgroBIO México 2004 a la mejor Tesis de Licenciatura en Biotecnología Agrícola, por el trabajo "Caracterización molecular del hongo Fusarium oxysporum patógeno del Agave tequilana Weber var. Azul y la implementación de un método para su detección oportuna"
Biól. Sara Mendoza Chagolla (Tesisista)
Dr. Juan Pablo Martínez Soriano, Director de Tesis

Premio AgroBIO México 2004 a la mejor Tesis de Doctorado en el área de Biotecnología Agrícola, por el trabajo "Transducción de señales extra-celulares asociadas al proceso micoparasítico de Trichoderma: papel de las MAP cinasas"
Dr. Artemio Mendoza Mendoza (estudiante)
Dr. Alfredo Herrera Estrella, Director de Tesis

"La Mujer 2004" para representar a la mujer mexicana en la categoría 'Salud', otorgado por la revista Glamour en español y MasterCard
Dra. Esther Orozco Orozco

Primer Lugar en el Simposio "La Investigación y el Desarrollo Tecnológico en Querétaro 2004", otorgado por el Gobierno del Estado de Querétaro a través de su Consejo de Ciencia y Tecnología, por el trabajo "Influencia de las condiciones de

crecimiento en películas delgadas de p-CdTe en la eficiencia de celdas solares del tipo V/n-CdO/p-CdTe"
Dra. Rebeca Castanedo Pérez, Dr. Sergio Joaquín Jiménez Sandoval, Dr. Gerardo Torres Delgado, Dr. Orlando Zelaya Ángel, M. en C. José Santos Cruz (estudiante) y Lic. Joaquín Márquez Marín (Auxiliar de Investigación)

Premio Interciencia - 2004 en Ciencias de la Vida, otorgado por el Gobierno de Canadá, en colaboración con la Asociación Francophone pour le Savoir
Dra. Mayra de la Torre Martínez

Premio de Investigación 2004, otorgado por la Academia Mexicana de Ciencias Área de Ciencias Naturales
Dr. Jean-Philippe Vielle Calzada
Área de Ciencias Exactas
Dr. Héctor Hugo García Compeán

Apoyo de la Fundación Miguel Alemán, A.C., para el proyecto "Clonaciones de genes y caracterización de proteínas con afinidad por el receptor sensor del calcio normal y una variante activa"
Dra. Guadalupe Reyes Cruz

Apoyo de la Fundación Miguel Alemán, A.C., para el proyecto "Fisiología y fisiopatología molecular de los canales de calcio neuronales sensibles al voltaje transmembranal"
Dr. Diego Ricardo Félix Grijalva

Primer Lugar en el Simposio "La Investigación y el Desarrollo Tecnológico en Querétaro 2004" dentro del área Estudios y Proyectos para el Desarrollo Sustentable en Querétaro, otorgado por el Gobierno del Estado de Querétaro a través de su Consejo de Ciencia y Tecnología, por el trabajo "La detoxificación de las aflatoxinas, lograda mediante el proceso tradicional de nixtamalización, es reversible"
Dr. Gerónimo Arámbula Villa y M. en C. Jesús Abraham Méndez Albores (estudiante)

Tercer Lugar en el Simposio "La Investigación y el Desarrollo Tecnológico en Querétaro 2004", dentro del área Avance Tecnológico de la Industria, otorgado por el Gobierno del Estado de Querétaro a través del Consejo de Ciencia y Tecnología, por el trabajo "Optical approach to the source of color by interference on austenitic stainless steels"
Dr. Alejandro Manzano Ramírez, Dr. Juan Francisco Pérez Robles y M. en C. Luis Dante Vázquez Santoyo (estudiante)

Distinción en el Concurso de Proyectos de investigación terminados, convocado por el Consejo Estatal de Ciencia y Tecnología del Estado de Coahuila de Zaragoza, por el proyecto "Modelo Dinámico para el Control del Enfriamiento Secundario en la Colada Continua de Planchones Delgados de Acero"
Dr. A. Humberto Castillejos Escobar, Dr. Francisco Acosta González y M. en C. Emilio Camporeddo Saucedo (estudiante)

Distinción en el Concurso de Proyectos de investigación terminados, convocado por el Consejo Estatal de Ciencia y Tecnología del Estado de Coahuila de Zaragoza, por el proyecto "Fabricación de Compósitos Al/SiC por infiltración de preformas de SiC en apoyo para combatir la contaminación ambiental"
M. en C. Mario Rodríguez Reyes (estudiante)

Primer Lugar en el Área de Proyectos Multidisciplinarios Categoría "AAA" del Certamen Nacional Juvenil de Ciencia y Tecnología 2004, por el trabajo de tesis "La proteína VP8 de la superficie de rotavirus abre las uniones estrechas y facilita el paso de fármacos a través de los epitelios"
QFB Porfirio Nava Domínguez (estudiante)
Dra. Lorenza González-Mariscal y Muriel, Directora de Tesis

Miembro Correspondiente Extranjero de la Academia de Ciencias Farmacéuticas de Chile
Dr. Pedro Joseph-Nathan

Nombramiento de Investigadora Emérita
CINVESTAV
Dra. Isaura Meza Gómez-Palacio

Premios Arturo Rosenblueth 2003 a las mejores tesis doctorales del CINVESTAV Área de Ciencias Exactas y Naturales
Dr. Carlos Enrique Valencia Oleta (estudiante)
Dr. Rafael Heraclio Villarreal Rodríguez, Director de Tesis

Área de Ciencias Biológicas y de la Salud
Dr. Mario Enrique Rueda Hernández (estudiante)
Dr. Omar Domingo Defeo Gorospe, Director de Tesis

Área de Tecnología y Ciencias de la Ingeniería
Dr. Emilio Barocio Espejo (estudiante)
Dr. Arturo Román Messina, Director de Tesis

Premio Internacional de la Investigación sobre Cultura Escrita 2002 del Instituto de Educación de la UNESCO, por el libro "Saber lo que es la Letra" editado por la UNESCO, Siglo XXI Editores y la SEP
Dra. Judith Kalman Landman

Miembro de la Academia de Ciencias del Tercer Mundo (TWAS)
Dr. Luis Herrera Estrella

Mejor Poster de la Conferencia COASTFISH 2004 Pesquerías Costeras en Latino América y el Caribe, por el trabajo "Descripción de la captura del recurso mero a través de dos tipos de artes de pesca, en la plataforma yucateca y su posible impacto sobre poblaciones explotadas"
M. en C. Armin Tuz Sulub (estudiante)
Dr. Thierry Brulé Demarest, Director de Tesis

Nombramiento de Senior Member de la IEEE
Dr. Carlos A. Coello Coello

Miembros del grupo de 300 líderes mexicanos 2004
Dr. Luis Herrera Estrella,
Dra. Rosalinda Contreras Theurel y
Dr. Octavio Paredes López

Miembro del Comité Editorial de la Revista Engineering Optimization
Dr. Carlos A. Coello Coello

Premio Nacional del Mérito Tecnológico "TECNOS 2004", mejor publicación tecnológica, otorgado por la Subsecretaría de la Industria, Comercio y Desarrollo Tecnológico, Secretaría de Desarrollo Económico del Gobierno del Estado de Nuevo León
M. en C. Emilio Camporeddo Saucedo (estudiante), Dr. A. Humberto Castillejos Escobar y Dr. Francisco Acosta González

Publicación de su artículo número 100 en Matemáticas
Dr. Onésimo Hernández-Lerma

Miembro del Comité de Asuntos Latinoamericanos de la International Brain Research Organization (IBRO)
Dr. Alonso Fernández Guasti

VISITA

POR INICIATIVA DE LA EMBAJADA DE LOS ESTADOS Unidos en México el CINVESTAV recibió al Dr. Arden L. Bement Jr., presidente interino de la National Science Foundation y director del National Institute of Standards and Technology (NIST), y al Dr. Stephen Carpenter, director de la oficina de asuntos internacionales del NIST.

La Dra. Rosalinda Contreras, directores de Unidad y jefes de Departamento, expusieron a los visitantes la estructura académica y líneas de investigación del Centro. Posteriormente, visitaron los laboratorios del grupo de nanoestructuras semiconductoras de los doctores Isaac Hernández y Miguel García en el Departamento de Física. Allí fue presentado el desarrollo del proyecto de fabricación y caracterización de pozos y puntos cuánticos ultradelgados para la emisión en el espectro visible.

También visitaron el laboratorio de la Dra. Carmen Sánchez, donde se expusieron los avances del proyecto sobre el desarrollo de terapia inmunogénica en el tratamiento de cáncer renal con presencia de metástasis.





NOMBRAMIENTOS RECIENTES EN EL
CINVESTAV

Dr. Iván Ortega Blake

Director de la Unidad Mérida

Dr. Agustín Guerrero Hernández

Jefe del Departamento de

Bioquímica

Dr. Rafael Castro Linares

Coordinador Académico de la Sección de

Mecatrónica

Dr. Juan Carlos Martínez García

Coordinador Académico del Departamento de

Control Automático



LIBROS Y REVISTAS

LA CIENCIA COMO PROYECTO DE VIDA

ANTONIO GÓMEZ NASHIKI

LA TRAYECTORIA DE JUAN MANUEL GUTIÉRREZ-Vázquez, a lo largo de 55 años dedicados a la educación es uno de los mejores ejemplos de lo que es ser un científico y de lo que la ciencia representa cuando se adopta como proyecto de vida.

A través de una exposición clara y apoyado en una narrativa que tiene como característica principal atrapar al lector y obligarlo a recorrer el texto hasta el final, Juan Manuel Gutiérrez-Vázquez escribe tanto para los versados en los temas educativos, como para los lectores nóveles en busca de nuevos referentes, con rigor y calidad.

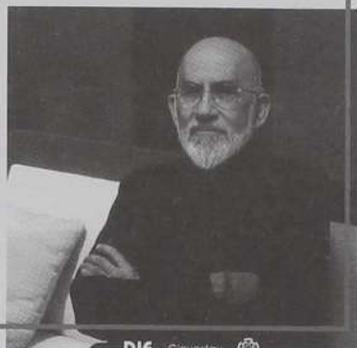
La selección de textos educativos contenidos en este libro, se basa en una filosofía particular y en una convicción producto de una larga trayectoria académica como maestro, investigador y coordinador de múltiples instituciones y proyectos educativos.

El libro está dividido en cinco capítulos precedidos de una entrevista, que a manera de introducción, muestra varias facetas de la vida personal, intelectual y profesional del autor. Esta sección constituye un acercamiento íntimo a acontecimientos centrales que explican en buena parte la actitud y la inclinación hacia la ciencia fomentada desde su infancia. Un aspecto relevante de esta entrevista lo conforman las fotografías que el autor proporcionó de su archivo personal que permiten un acercamiento a sus amigos y familiares y que dan cuenta de los momentos importantes de su vida.

Con paso lento y agitadamente

Ensayos educativos

Juan Manuel
Gutiérrez-Vázquez



A lo largo de estas páginas se suceden historias entrelazadas de trayectorias de sujetos; proyectos, la fundación de instituciones y posiciones ante la vida, en el marco de un esfuerzo educativo de larga duración y que a la fecha sigue vigente.

El capítulo uno, “La enseñanza de las ciencias” es una muestra de las principales tendencias en la enseñanza de la ciencia, los debates y perspectivas predominantes en la década de los ochenta; los retos que afronta quien se dedica a la impartición de la ciencia y la tecnología, así como las particularidades que asume el quehacer científico en un país como el nuestro.

“Vivir para aprender”, el segundo capítulo, es un largo ensayo que muestra la posición del autor en

torno a los problemas de la ciencia, sus fines y objetivos, así como la concepción y dilemas que afronta quien se dedica a esta actividad. Dos son los puntos centrales de este apartado, uno tiene que ver con el desarrollo de la actitud científica que debemos mantener en todo momento, y el otro se refiere a lo que el autor denomina como el *currículum de vida o muerte*, que es, sin duda, una reflexión producto de la amplia experiencia de Gutiérrez-Vázquez en donde se muestran con claridad los elementos que deben estar incluidos en la elaboración de un currículo que excede con mucho los límites de la educación denominada formal, siendo básicamente una propuesta para pensar en un currículo para la vida.

El tercer apartado, "Reformas curriculares", reúne tres ensayos cuya importancia radica en otorgar al currículo el lugar central que ocupa en el desarrollo de las reformas educativas: su concepción, valoración, trascendencia, seguimiento y evaluación, a través de las experiencias educativas en donde el autor participó activamente: la reforma de la enseñanza de las ciencias naturales en la escuela primaria en 1976; la reforma curricular en Belice; y su propuesta de crear una universidad intercultural indígena para Michoacán.

En el capítulo cuatro, "El libro de texto", se aborda un aspecto central en el trabajo educativo y académico de la obra del autor, pues de los diversos proyectos en los que ha participado, el de los libros de texto guarda un significado especial, pues por una parte es una historia íntimamente ligada a la gestación del Departamento de Investigaciones Educativas, DIE (institución que surgió con la misión de producir evidencias de investigación para fundamentar las futuras políticas educativas nacionales, y elaborar los Libros de Texto Gratuitos de Matemáticas y de Ciencias Naturales). Y por otra, la apuesta del autor está enfocada en la promoción y perfeccionamiento de libros y materiales educativos en beneficio de alumnos y maestros.

El capítulo final, "Materiales educativos", inclu-

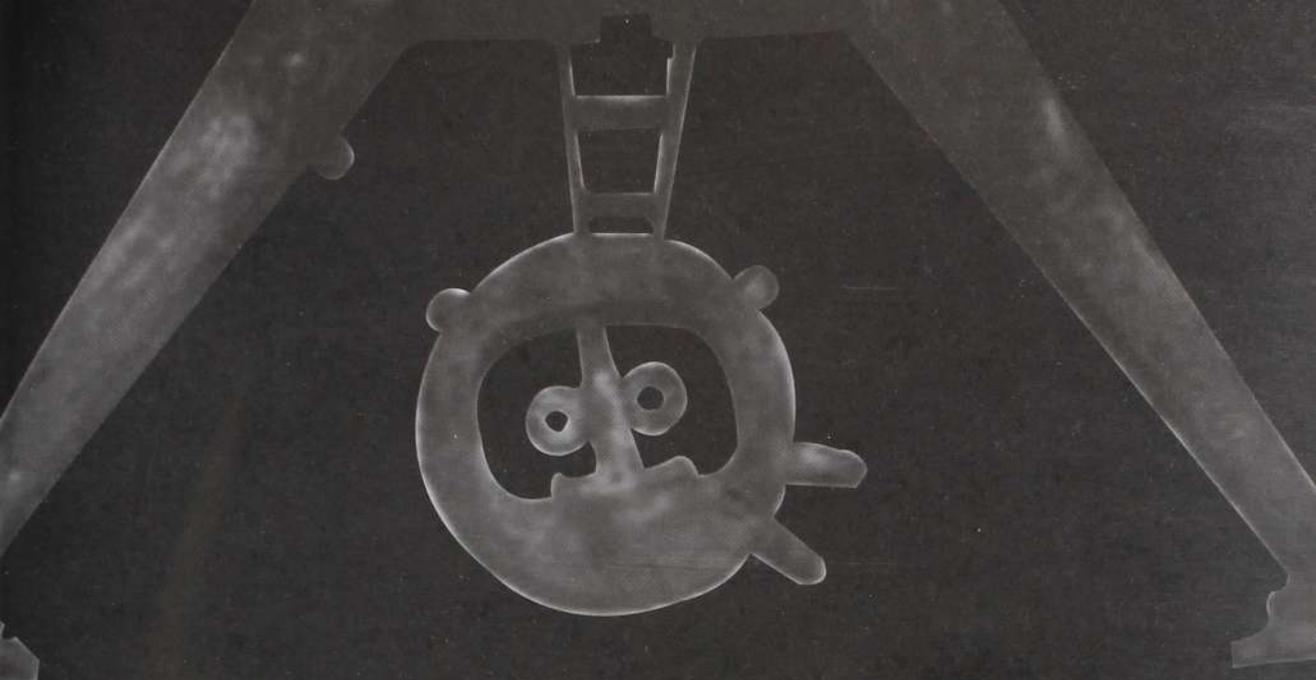
ye trabajos de corte didáctico que muestran aspectos útiles para estructurar los materiales educativos, así como recomendaciones para los estudiantes a la hora de afrontar la enseñanza de los diversos contenidos académicos.

Con paso lento y agitadamente. Ensayos educativos, es un libro donde la preocupación central está apostada en el desarrollo de una actitud científica, reflexiva y comprensiva del entorno, de la promoción de la ciencia como algo cercano y útil, el cual muestra contra las creencias generales, que la actividad científica no es una acción individual, ni el producto del trabajo científico aislado. Todo lo contrario, como toda actividad humana, la ciencia es producto de un trabajo colectivo que tiene sentido en la medida que sale de los laboratorios y que puede ser útil y aplicable en la vida cotidiana de las personas.

La preocupación por difundir la ciencia y sus aplicaciones no es un esfuerzo sencillo, es una tarea que en muchas ocasiones no se materializa, es un reto que muchos científicos aún no resuelven, por lo que este libro es una obligada referencia que muestra con rigor cómo el conocimiento y su transmisión son felizmente posibles.

REFERENCIA

Juan Manuel Gutiérrez-Vázquez, 2004, *Con paso lento y agitadamente. Ensayos educativos*, México, DIE-CINVESTAV.



CONTRIBUCIONES

Las contribuciones para la revista *Avance y Perspectiva*, deberán enviarse a las oficinas del CINVESTAV en los siguientes formatos.

TEXTOS:

Word, Disco 3/5 ó CD-ROM

Cuando se trate de artículos de investigación la extensión máxima será de 15 cuartillas; en cuanto a los artículos de difusión se aceptará un máximo de 10 cuartillas.

Si el texto incluye tablas, éstas se entregarán en archivo por separado, en disquette en texto corrido y con una impresión adjunta que muestre la forma en que debe quedar la tabla. Además, se debe indicar en el original la ubicación de éstas. La indicación es también válida para esquemas y cuadros.

Las notas deberán incluirse al final del trabajo, antes de la bibliografía o de las referencias, debidamente numeradas.

Las referencias deben apegarse a los modelos siguientes:

Libro:

Wiener, Norbert, 1985, *Cibernética: o el control y la comunicación en animales y máquinas*, Barcelona, Tusquets.

Artículo de revista:

Ádem, José, 1991, "Algunas consideraciones sobre la pesca en México", en *Avance y Perspectiva*, vol.10, abril-junio, pp. 168-170.

Se sugiere que las referencias sean cuidadosamente revisadas por los autores y que los títulos de los artículos y los nombres de las publicaciones no se abrevien.

Todos los textos deben incluir el nombre del autor, grado académico, adscripción y cargo que desempeña, teléfono y correo electrónico.

IMÁGENES Y GRÁFICAS:

TIFF, 10 x 10 cm (mínimo), 300dpi, Blanco y negro, CD-ROM
Las imágenes se entregarán en forma física, de preferencia en transparencias. No se aceptarán imágenes de internet o cámara digital debido a que la resolución que presentan no es adecuada.

En caso de que el trabajo incluya diapositivas en power point, favor de remitirlas impresas en papel fotográfico y en máximo grado de resolución y no a través de archivos de office.



Centro de Investigación
y de Estudios Avanzados del IPN

CINVESTAV

avance@mail.cinvestav.mx

T/F 57 47 70 76

www.cinvestav.mx/publicaciones

Av. IPN 2508, Zacatenco, C. P. 07360

**AHORA NOS CONOCEN
COMO LA FLOTA MÁS MODERNA DEL
MUNDO, PERO SEGUIMOS SIENDO:
MEXICANA.**

La aerolínea mexicana líder en América
que además de toda su experiencia,
te ofrece viajar por una enorme red
de destinos en los aviones más modernos.



mexicana.com

MEXICANA 

Siente la experiencia



Centro de Investigación
y de estudios avanzados del IISI
CINVESTAV