

PROBLEMAS METODOLÓGICOS DEL DESEMPEÑO ACADÉMICO ASOCIADO A LOS PROGRAMAS DE ESTÍMULOS

GRACIELA CORDERO
ARROYO Y
EDUARDO BACKHOFF
ESCUADERO*

Resumen

El propósito del presente trabajo es analizar el proceso de construcción de los Programas de Estímulos al Desempeño del Personal Académico (PEDPA), con el fin de ilustrar algunos problemas metodológicos básicos del sistema de evaluación asociado y sus repercusiones en las puntuaciones y estímulos económicos del personal evaluado. Se toma el caso de una universidad pública mexicana, describiendo el contexto institucional, las distintas etapas para elaborar el sistema evaluativo y los resultados obtenidos. Se concluye que los PEDPA adolecen de problemas metodológicos debido a la poca congruencia del modelo evaluativo con los propósitos declarados, la ausencia de los grupos colegiados en la construcción del modelo y la falta de participación de especialistas en evaluación.

Palabras clave: evaluación educativa, educación superior, estímulos al personal académico.

Abstract

The purpose of this work is to analyze the creation process of Teaching Staff Performance Incentive Programs (PEDPA, Programas de Estímulos al Desempeño del Personal Académico), in order to point out some basic methodological problems of the evaluation system and their consequences in ratings and economic incentives granted to the evaluated staff. In order to do so, the author has chosen the case of a given public university in Mexico describing the institutional context, the stages of creation of the evaluation system and results obtained. The author concludes that the PEDPAs have methodological problems, in deed, due to the lack of coherence between the evaluation model and the corresponding purposes, absence of collegiated involvement in the creation of such evaluation and absence of evaluation specialists involved in the process.

Key words: education evaluation, higher education, incentives for the teaching staff.

* Investigadores del Instituto de Investigación y Desarrollo Educativo, Universidad Autónoma de Baja California. Correo-e: backhoff@faro.ens.uabc.mx
Recepción: 30 de enero de 2002.
Aceptación: 29 de mayo de 2002.

Introducción

En las últimas dos décadas las instituciones de educación superior (IES) mexicanas han experimentado las consecuencias, para bien y para mal, de un proyecto de modernización que pretende el logro de la excelencia y la calidad en las instituciones educativas. Así, el gobierno se ha decidido a transformar y reestructurar las instituciones públicas, bajo un esquema productivista y de rendición de cuentas (Díaz Barriga, 1997).

La idea de otorgar una retribución económica diferenciada al personal académico con base en su producción y agregada al salario base, tiene su origen en 1908 en los Estados Unidos de Norteamérica. En nuestro país, esta política se aplicó en la década de los ochenta, en una época de fuerte crisis económica que instituye esta vía como la única posible para recuperar el deterioro salarial de los académicos universitarios (Gil Antón, 2002).

Con la idea de evitar una “fuga de cerebros” en México, en medio de una crisis económica aguda, en 1984 se establece el Sistema Nacional de Investigadores (SNI) para retribuir económicamente a los científicos mexicanos que demostraran una producción académica adecuada, hecho que, en palabras de Ornelas (1995), constituye el punto de partida para establecer un sistema meritocrático en nuestro país, el cual descansa en las prácticas de evaluación centradas en la productividad.

Asimismo, para impulsar la calidad y excelencia de la educación superior, en 1990, el gobierno mexicano acuerda con la Asociación Nacional de Universidades e Instituciones de Educación Superior (ANUIES) crear el Fondo para la Modernización de la Educación Superior (FOMES), programa que funciona bajo la base de hacer competir a las IES públicas por la obtención de recursos financieros extraordinarios, presentando proyectos para su desarrollo y modernización.

Finalmente el gobierno mexicano implantó paulatinamente en todas las IES públicas el “Programa de Estímulos al Desempeño Académico (PEDPA)”¹. Este programa otorga un pago adicional al salario del profesor universitario, con base en una estimación de la calidad de su desempeño o productividad. Es importante decir que, esquemas similares para mejorar la planta docente de las IES también se encuentran en otros países latinoamericanos, como Venezuela, Brasil, Argentina y Chile (Gil Antón, 2000).

Como se puede apreciar, en las últimas dos décadas, la evaluación se ha convertido en uno de los ejes centrales de las políticas educativas mexicanas. Así, los programas de modernización educativa antes descritos exigen que se evalúen administraciones, funciones, programas, docentes y estudiantes, entre otros muchos componentes de la actividad universitaria.

La evaluación que se genera como parte de estos programas sirve, adicionalmente, para que las instituciones cuenten con información relevante y pertinente para tomar decisiones que las ayuden a identificar y resolver los problemas que las aquejan. Se parte de la idea de que el mejoramiento de una institución educativa sólo es posible a partir de contar con información válida y confiable de sus insumos, procesos y resultados.

Ahora bien, es de todos conocido que la calidad de las IES se sustenta fundamentalmente en el grado de consolidación de su planta académica. A través de la evaluación de los académicos se puede apreciar la madurez de una institución, así como sus capacidades de desarrollo. Por tal razón, la evaluación de los profesores universitarios constituye un punto de partida para mejorar la función sustantiva de sus instituciones (ANUIES, 2000).

Sin embargo, por múltiples razones la evaluación del docente es un campo poco explorado, muy controvertido y donde no se ha dicho la

¹ Otros nombres utilizados son: Beca al desempeño académico, Beca de productividad, etc.

última palabra, especialmente si sus resultados se asocian a un programa de recompensas. Entre los problemas detectados, destacan: la heterogeneidad de las instituciones educativas, la diversidad de funciones y actividades académicas, la multitud de indicadores posibles de evaluar y la diversidad de metodología de evaluación.

Sin considerar estos problemas y con una gran ignorancia e ingenuidad metodológica, en 1997 la Secretaría de Hacienda y Crédito Público (SHCP) establece los lineamientos específicos que deben utilizar las IES nacionales para evaluar a sus académicos y otorgar los estímulos económicos correspondientes. Estos lineamientos establecen, entre otras cosas, los criterios básicos de evaluación, los factores de productividad a evaluar y la ponderación de cada uno de ellos (calidad, dedicación y permanencia), así como la escala correspondiente de puntuaciones (del 0 al 1000). Igualmente, define el perfil del académico a beneficiar, la vigencia y montos de los estímulos, etcétera.

Si bien es cierto que estos lineamientos ayudaron a algunas universidades públicas a disminuir la discrecionalidad con que se evaluaba al académico y se otorgaba el estímulo económico, también es cierto que dichos lineamientos son muy pobres metodológicamente. En algunas universidades estos lineamientos se han convertido tal cual en el reglamento de sus programas de estímulo; en otros casos, las autoridades universitarias han adaptado los criterios federales a la concepción que la administración tiene sobre la forma de evaluar al docente, poniendo en juego la capacidad de negociación entre las autoridades universitarias en turno y las federales.

Por desgracia, ni en la SHCP ni en la mayoría de las IES nacionales quienes diseñan o adaptan los instrumentos de evaluación de los PEDPA tienen la preparación técnica deseable en materia de evaluación. En consecuencia, cuando se diseñan dichos instrumentos, generalmente, no se

sujetan a ningún tipo de estudio que compruebe su validez, confiabilidad y utilidad. Por ejemplo, es una realidad que se utilizan evaluaciones que en teoría son referidas a un criterio² y en la práctica se utilicen normativamente³. A pesar de estas deficiencias metodológicas, se evalúa a los académicos aún cuando los instrumentos no garanticen la validez de los resultados; y lo que es peor, éstos se utilizan para estimularlos o castigarlos, según sea el caso.

Los programas de pago al mérito en las IES son un objeto de estudio relativamente reciente. En los Estados Unidos de Norteamérica se han estudiado sistemáticamente desde inicios de la década de los ochenta. En un trabajo realizado hace casi veinte años, Lewis (1983) concluye que las definiciones del mérito varían de acuerdo a la disciplina académica y la categoría del profesorado, encontrando que la enseñanza es mucho menos valorada que la investigación o las actividades de servicio y vinculación. Por su parte, Kasten (1984) señala que existen diferentes tipos de vínculos entre las tareas de investigación, docencia y servicio en diferentes tipo de universidades y que esto tendría que ser tomado en cuenta para definir el sistema de pago al mérito.

En nuestro país, este es un objeto de estudio aún más joven. Aunque existen algunos trabajos conceptuales y empíricos al respecto (e.g. Díaz Barriga, 1996; Díaz Barriga y Pacheco Méndez, 1994; Grediaga, 1998; Hernández, 1998; Izquierdo, 1998; Krotz, 1993; López Zárate, 1991; Stefanovich, 1998), todavía representan un pequeño número, dada la relevancia del problema. Más aún, en el V Congreso Nacional de Investigación Educativa celebrado en 1999, si bien se presentaron algunos trabajos sobre académicos y sobre evaluación del desempeño docente, el programa no registra ningún trabajo sobre este objeto de estudio. En otro importante foro académico reciente en el ámbito nacional, el IV Foro Nacional de Evaluación organizado

² Tipo de evaluación que se realiza en relación con un criterio fijo, por ejemplo, determinado número de publicaciones.

³ Tipo de evaluación que se realiza en relación con las puntuaciones de las demás personas evaluadas.

por el CENEVAL, sólo se registró un trabajo dentro de esta línea de investigación (Cordero y Backhoff, 2000).

Hasta ahora, los investigadores mexicanos han centrado su atención en los efectos negativos que han producido las evaluaciones al desempeño docente en las comunidades académicas. Entre éstos, destacan los relacionados con la inconformidad de los docentes, el trabajo orientado a la búsqueda de “puntos”, la inflación meritocrática y el aislamiento académico (Díaz Barriga y Pacheco, 1997; Izquierdo, 1998; Rueda y Landesmann, 1999).

Desde nuestra perspectiva aún falta por abrir una línea de investigación que recupere la dimensión metodológica de los programas de estímulo al personal académico, entendidos como programas de evaluación educativa. Es importante recordar que, cuando se utilizan instrumentos de evaluación de alto impacto, como los diseñados para evaluar el desempeño académico de los docentes, su elaboración debe ajustarse a rigurosos estándares de calidad (Aiken, 1996). Asimismo, es recomendable elaborarlos en un proceso colegiado, teniendo en mente los objetivos que persiguen, sus presupuestos teóricos y sus indicadores técnicos. Todo ello definirá la calidad del instrumento educativo que se emplea.

Por lo anterior, el propósito del presente estudio es describir el proceso de construcción de un programa de estímulos al personal académico. El modelo resultante de este proceso se aleja de los lineamientos de la SHCP pero se acerca a las necesidades e intereses de la institución. El proceso ilustra algunos problemas metodológicos de todo sistema de evaluación del desempeño académico, así como sus repercusiones en las puntuaciones y estímulos económicos del personal que está sujeto a evaluación. Especialmente, nos interesa mostrar la falta de congruencia entre los propósitos de los programas de estímulos y el modelo evaluativo utilizado, así como la iniquidad que

produce una evaluación mal diseñada.

Para ello tomaremos el caso de una universidad pública mexicana, con gran prestigio y tradición en nuestro país⁴. Para lograr nuestro objetivo, describiremos el contexto institucional, las etapas que se siguieron para elaborar el programa evaluativo y los resultados que se obtuvieron. Finalizamos con una discusión de la metodología usada a la luz de los resultados obtenidos.

Contexto institucional

Este estudio se desarrolló en una de las universidades con más historia en el país, la cual se distingue por contar con una organización colegiada de la vida académica. El otorgamiento de los estímulos al personal académico está normado por el Estatuto del Personal Académico, el cual señala las funciones de los cuerpos colegiados responsables de la toma de decisiones en este proceso.

El Estatuto marca la existencia de dos instancias colegiadas que son responsables de la operación del programa: la Comisión Evaluadora y la Comisión Resolutoria (Véase Figura 1). Cada área del conocimiento (Ciencias Naturales, Ciencias Sociales, Ciencias de la Salud, Ciencias Económico-administrativas, Artes, Ingeniería y Nivel Medio Superior) tiene una comisión evaluadora conformada por cinco académicos (cuatro de ellos designados por otros cuerpos colegiados y un académico externo), y es presidida por el rector o su representante. Estas comisiones se encargan de revisar los expedientes de los solicitantes y otorgar la calificación correspondiente. La comisión resolutoria se encarga, por su parte, de expedir la convocatoria y tomar las decisiones generales sobre el proceso de otorgamiento de los estímulos. La comisión resolutoria está formada por un representante de cada comisión evaluadora, dos profesores designados por otros

⁴ Se omite el nombre de la universidad por razones de confidencialidad.

cuerpos colegiados y, también, es presidida por el rector. Además de estos dos cuerpos colegiados, la institución designa un coordinador general que se encarga de todas las cuestiones operativas del programa de estímulos.

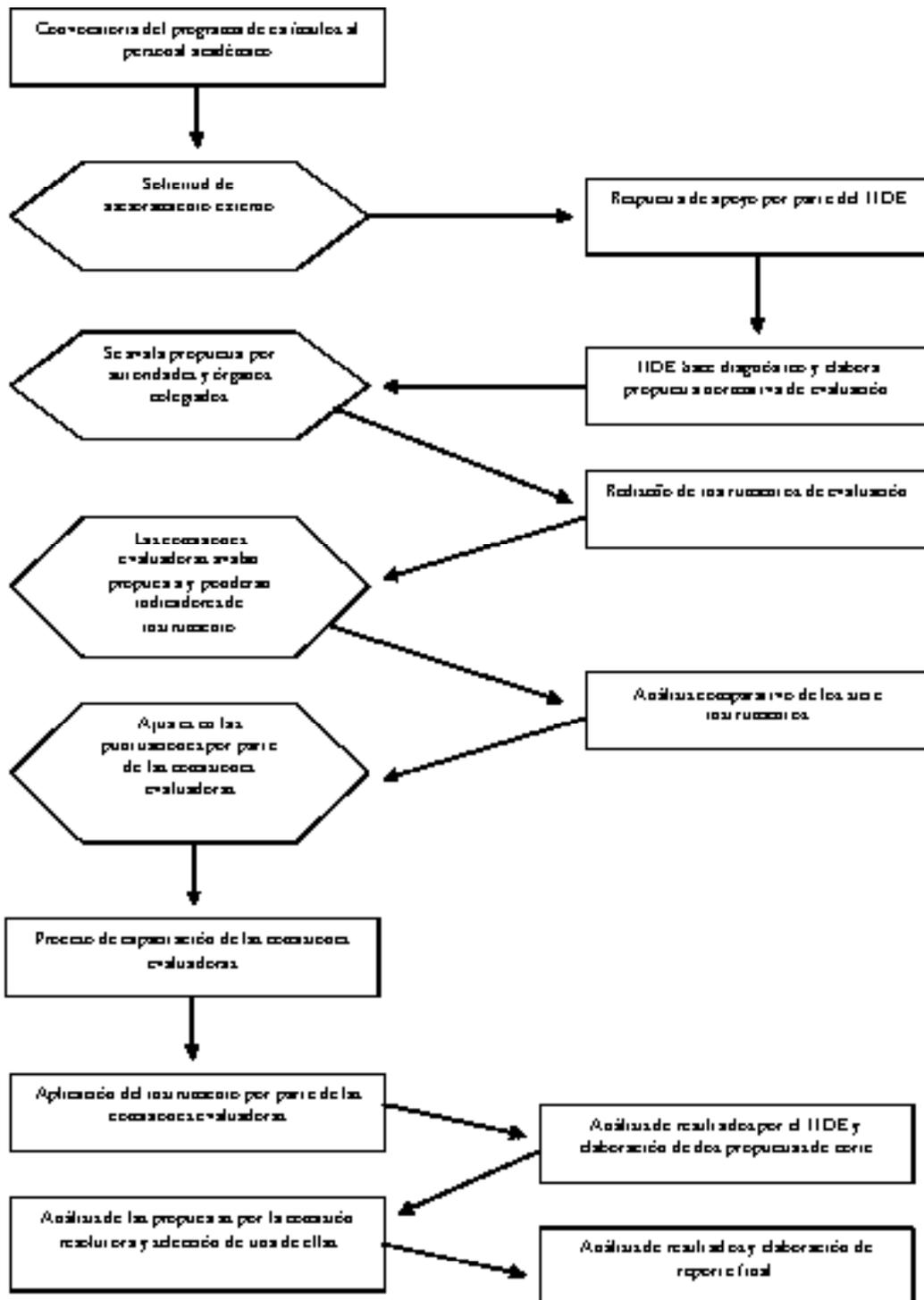
Figura 1
Comisiones responsables de la toma de decisiones del programa de otorgamiento de estímulos



En la experiencia que a continuación se describirá, el proceso de elaboración del programa de estímulos, pasó por una serie de fases que se muestran en la

Figura 2 y que se ha resumido en tres etapas.

Figura 2
Proceso de elaboración de un programa de estímulos al personal académico



Etapa 1. Etapa diagnóstica y preparatoria

De acuerdo con la opinión de los responsables del programa de estímulos de esta universidad, los resultados de la convocatoria anterior habían sido poco satisfactorios para la institución en general y para sus académicos en particular. Por esta razón y ante la inminencia de una nueva convocatoria del programa de estímulos al personal académico, las autoridades universitarias solicitaron al Instituto de Investigación y Desarrollo Educativo (IIDE) de la Universidad Autónoma de Baja California asesoría metodológica para desarrollar el programa que se aplicaría en la convocatoria de 1999. En la primera etapa de trabajo se formalizaron las condiciones y términos de dicha solicitud, se realizó un diagnóstico de los resultados de la convocatoria anterior y se ofreció a la universidad un modelo de evaluación para ser sometido a su valoración por los cuerpos colegiados responsables del programa.

Para elaborar el diagnóstico, el IIDE revisó la normativa interna para la operación del programa e identificó a los actores principales del proceso. Se revisaron, asimismo, los resultados de la convocatoria anterior, los instrumentos aplicados y la documentación generada por algunas de las comisiones evaluados, así como las sugerencias presentadas por la comunidad para mejorar el programa.

Como resultado de este diagnóstico se detectó la siguiente problemática:

1. Se utilizaba un esquema de evaluación de criterios que partía de un modelo ideal de académico para toda la universidad.
2. Se evaluaba a los maestros con indicadores ajenos a las actividades que desarrollaban en su área de conocimiento.
3. Los evaluadores modificaban el instrumento en el momento de evaluar para adecuarlo a las características de la producción académica de su área de conocimiento.
4. Se incluían topes en las puntuaciones para forzar las escalas a un puntaje predeterminado pensando que con este ajuste los puntajes serían equivalentes.

Ante esta problemática, el IIDE propuso trabajar con un modelo que se adaptase a las características de cada una de las áreas de conocimiento. Un modelo normativo (Popham, 1990) que permitiera evaluar el desempeño del académico en comparación con el de otros docentes de su misma área de conocimiento; es decir, en donde se evaluara la actuación del académico en relación con la actuación de otros docentes de su mismo grupo. La evaluación normativa permite mayor equidad en la valoración de profesores con carga y producción académicas diferentes, en tanto que no compara, por ejemplo, la producción de académicos de Nivel Medio Superior con los de Ciencias Naturales y exactas. El modelo de evaluación normativa permite evaluar la producción de las diferentes áreas académicas de acuerdo a los criterios propios de cada área y trabaja con escalas flexibles que posteriormente deben ser trabajadas estadísticamente.

Desde esta perspectiva se propuso diseñar un instrumento apropiado para cada área que incluyera indicadores de productividad adecuados al área y que podían ser diferentes entre sí. Se propuso, también, no establecer topes en la producción académica y no forzar las escalas para que se ajustaran a un puntaje predeterminado. Se consideró que esta propuesta de evaluación reunía las siguientes ventajas:

1. Era consistente con la forma de organización académica de la propia institución, ya que permitía a los consejos y comisiones por área ejercer su autonomía para definir aquellas actividades que consideraran importantes evaluar y asignar el puntaje correspondiente de acuerdo a los criterios propios de cada área.
2. Era sensible a reconocer la heterogeneidad de la práctica académica en función del área de conocimiento.

3. Era lo suficientemente flexible para dar respuesta a la mayoría de las propuestas de modificación del instrumento presentadas por las unidades académicas.
4. Permitía a los académicos ser evaluados con relación a sus pares y por lo tanto obtener calificaciones relativas de acuerdo al grupo de pertenencia.
5. Permitía hacer una comparación de los resultados obtenidos entre las distintas áreas.

Este modelo de evaluación fue presentado y aprobado por las autoridades universitarias, quienes, a su vez, lo expusieron ante las comisiones participantes en el proceso, siendo aprobado por ellas y constituyéndose en el marco general de trabajo del programa de otorgamiento de estímulo.

Etapa 2. Definición de instrumentos y procedimientos

El instrumento base de la convocatoria previa fue una tabla de actividades desarrolladas por función académica (docencia, investigación, extensión y administración). Tal como había sido acordado en la etapa anterior, el IIDE propuso elaborar un instrumento de evaluación para cada una de las áreas de conocimiento establecidas por la universidad, lo cual significó aumentar el número de instrumentos de uno a siete.

El IIDE diseñó estos siete instrumentos a partir del instrumento original utilizado en la convocatoria anterior. Se reorganizó su estructura y se incorporaron o eliminaron indicadores, en función de las sugerencias que los propios académicos habían realizado por escrito. Se tomaron como base las cuatro funciones establecidas en el instrumento anterior y se clasificaron en las subfunciones ya incluidas en el propio instrumento (véase Tabla I).

Tabla I
Esquema de organización de actividades

por función académica

Funciones	Subfunciones
Docencia	Actividades frente a grupo Actividades de apoyo a la docencia
Investigación	Publicaciones Ponencias Aplicación del conocimiento
Extensión	Difusión Gestión Arbitrajes Distinciones y premios
Participación en la vida colegiada	Comisiones y encomiendas

Las comisiones evaluadoras revisaron la estructura de los instrumentos y otorgaron puntaje a cada uno de los indicadores, de acuerdo a los criterios de calidad del área de conocimiento. Una vez que las siete comisiones evaluadoras presentaron su propuesta de puntaje, el IIDE las analizó en forma comparativa. Este análisis permitió observar que las comisiones evaluadoras no cambiaron la estructura del instrumento de evaluación. Sin embargo, tal como les fue indicado, sí otorgaron puntajes diferenciales a las actividades incluidas en la tabla de acuerdo a los criterios de calidad y trascendencia que dichas actividades tenían para el desarrollo de su área de conocimiento. Este análisis comparativo permitió identificar los pesos relativos que cada área otorgaba a las funciones a evaluar y las actividades que requerían mayor análisis por parte de las comisiones evaluadoras, pues su valor resultaba aparentemente desproporcionado con relación al otorgado por otras. Este análisis fue devuelto a las comisiones evaluadoras para ser revisado y hacer los cambios que juzgaran convenientes.

Etapa 3. Resultados

Una vez que las comisiones evaluadoras revisaron los expedientes y que dichos resultados fueron

avalados por la comisión resolutoria, los puntajes totales de los profesores de tiempo completo de las siete áreas académicas fueron enviados al IIDE.

El IIDE procedió a hacer un análisis inicial de las puntuaciones. Se observó, tal como se esperaba, que había siete escalas diferentes y que los rangos eran muy amplios, variando de 1,322 puntos en el área de Nivel Medio Superior, a 3,463 puntos en el área de la Salud. Tres áreas destacaron por sus puntajes extremos y por la

dispersión de sus puntuaciones: Ciencias Naturales, Ciencias Sociales y Salud, áreas donde la variabilidad de las puntuaciones presentaron los valores más altos (véase Tabla II).

A fin de resolver el problema de los cortes y la asignación de los estímulos, el IIDE presentó a los responsables del programa dos procedimientos generales de análisis: 1) transformación lineal con exclusión de valores extremos y 2)

Tabla II
Indicadores estadísticos básicos para cada una de las áreas

Área	# Acad.	Mínimo	Máximo	Rango	Media	Desv.Est.	Mediana
NMS	87	305.00	1627.25	1322.25	996.98	264.31	1020.00
Naturales	97	38.50	3501.90	3463.40	872.33	678.05	698.80
Sociales	55	75.75	2556.50	2480.75	802.55	538.26	620.00
Salud	116	73.90	3467.00	3393.10	764.20	471.86	685.50
CEA	40	325.38	1927.25	1671.87	1082.00	420.01	1082.78
Ingeniería	111	111.93	1580.00	1468.07	780.83	300.33	743.80
Artes	37	382.25	2298.50	1916.25	1150.48	410.07	1082.17

transformación percentilar. A continuación se describen estos procedimientos.

Transformación lineal con exclusión de valores extremos

Inicialmente se hizo una transformación lineal que permitiera interpretar y comparar las puntuaciones, a la vez que mantuviera la propiedad de intervalo de los puntajes. La transformación utilizada fue la llamada lineal, la cual consiste en utilizar el indicador central de la media y la desviación estándar de los datos para efectuar una conversión hacia otra distribución de frecuencias con una media y desviación estándar preestablecidas. Se utilizó esta técnica de transformación para obtener una caracterización estadística de los datos y efectuar una comparación directa entre las distintas áreas.

A fin de comparar las siete áreas, se utilizó una puntuación estándar que transformara los datos a una media deseada de 500 y una desviación estándar de 100. Esta transformación ubica

los datos en la región positiva y considera 3 ó 4 cifras positivas enteras (Larrazolo, s/f). Trabajar con una media de 500 y una desviación estándar de 100, permite ubicar en +- 3 desviaciones estándar al 99.9% de los datos y los puntajes oscilan entre 200 y 800. La relación matemática que se utiliza para la transformación es:

$$C.E. = (100 (x - m) / s) + 500$$

Donde :

C.E. es el valor estándar transformado

x es la puntuación cruda

m es la media de los datos

s es la desviación estándar de los datos

Con el fin de disminuir la dispersión de los datos en el caso de Ciencias Naturales, Ciencias Sociales y de la Salud, se excluyeron los valores extremos y se volvió a calcular la media y la desviación estándar. Para identificar el punto de exclusión de los valores extremos se ubicó el dato cuyo valor era notablemente mayor o menor con

relación a la dispersión observada.

Tal como puede verse en la Tabla III, la exclusión de valores extremos redujo el valor de la desviación estándar y el porcentaje de casos que se encontraba en la primera desviación.

Transformación percentilar

Tabla III

Comparación de los valores obtenidos sin y con exclusión de los valores extremos por área de conocimiento

Área	Sin exclusión de valores extremos		Con exclusión de valores extremos		% de datos en 1ª desv. sin exclusión	% de datos en 1ª desv. con exclusión
	Media	Desv. Est.	Media	Desv. Est.		
C. Naturales	872.33	678.05	819.70	402.86	78.3%	55.6%
C. Sociales	802.55	538.26	803.05	386.12	74.5%	58.1%
C. Salud	764.20	471.86	773.86	384.02	73.3%	65.5%

y de esta forma es posible conocer la posición relativa del profesor en el grupo sin importar la dispersión. Al darle peso al lugar, pero no a la calificación obtenida, se consideraba posible eliminar el problema de la variabilidad de los puntajes. En otras palabras, esta ubicación percentilar permite comparar a los docentes de las siete áreas, aún cuando su escala de calificación y su rango varíen, ya que no importa la distancia que existe entre ellos, sino el orden en que están colocados en su respectiva escala. Por otra parte, este método permite al profesor ubicar su puntuación en términos del porcentaje de profesores que están por arriba o por abajo de su resultado y, de esta forma, también puede comparar su actuación particular con los resultados obtenidos en otras áreas.

En virtud del comportamiento de las tres áreas mencionadas, se encontró que el procedimiento más recomendable era trabajar con el valor de las medianas y sus medidas derivadas –cuartiles, octiles, percentiles– ya que permiten la partición homogénea de los valores

En primer término, se determinó trabajar con números pares y cortar las escalas en 16 grupos (decasextiles ó 16tiles). Esto permitiría tener un número suficiente de grupos para hacer clasificaciones diversas de acuerdo a políticas institucionales. Se calcularon los 16tiles: 6.25, 12.50, 93.75 y cada escala se cortaron en 16 sectores.

En la Tabla IV se observa que es posible una comparación de los docentes de las siete áreas, aún cuando su escala de calificación y su rango varíen, ya que, como se dijo, no importa la distancia que existe entre ellos, sino el orden en que están colocados en su respectiva escala.

Otra ventaja de este método es que ayuda al profesor a comprender su posición en el programa, en tanto que le permite ubicar su puntuación

Tabla IV

Especificación de los 16tiles según los puntajes obtenidos para cada una de las áreas

16til	Áreas						
	NMS	Naturales	Sociales	Salud	CEA	Ingeniería	Artes
6.25	556.1	84.2	183.7	211.1	435.1	310.0	571.7
12.50	628.1	209.5	298.0	272.8	517.6	431.1	681.3
18.75	794.9	319.8	395.9	396.2	647.0	540.9	751.5
25.00	862.6	421.3	447.2	421.2	786.2	602.2	849.2
31.25	940.1	469.3	502.5	502.5	843.3	653.5	901.9
37.50	959.6	592.6	561.7	566.0	923.6	691.6	1033.6
43.75	981.0	653.2	603.2	641.8	1003.8	722.5	1062.7
50.00	1020.0	698.8	620.0	685.5	1082.7	743.8	1082.1
56.25	1033.3	827.1	723.5	747.7	1138.3	788.4	1172.7
62.50	1058.9	1005.3	775.5	846.1	1196.3	828.2	1323.2
68.75	1102.2	1057.3	862.7	909.5	1334.4	903.9	1342.5
75.00	1142.6	1094.0	1136.5	983.3	1414.5	970.6	1391.1
81.25	1170.9	1241.6	1264.5	1077.7	1477.6	1048.8	1497.4
87.50	1291.2	1488.5	1491.1	1256.5	1619.0	1115.1	1534.8
93.75	1440.4	2417.2	1858.0	1498.1	1776.4	1260.7	1964.7

en términos del porcentaje de profesores que están por arriba o por abajo de su resultado y de esta forma, puede también comparar su actuación particular con los resultados obtenidos en otras áreas.

Posteriormente, se hizo una distribución de los académicos que se ubicarían en cada uno de los 16tiles. El resultado de este ejercicio se muestra en la Tabla V.

Finalmente, se hizo un ejercicio de agrupación de 16tiles para determinar posibles cortes (véase tabla VI). En tanto que los cortes son una decisión de política educativa en función de lo que la institución quiere favorecer, este ejercicio *no debe entenderse*, de ninguna manera, como la mejor propuesta para establecer los niveles de beca. Es un ejercicio técnico para presentar la forma de utilizar los 16tiles.

Tabla V

Distribución del personal para cada uno de los 16tiles (R1, R2...R16) según su área de adscripción

Área	R1	R2	R3	R4	R5	R6	R7	R8	R9	R10	R11	R12	R13	R14	R15	R16	Total
NMS	5	5	6	5	6	5	6	5	6	5	6	5	6	5	6	5	87
Naturales	6	6	6	6	6	6	6	6	7	6	6	6	6	6	6	6	97
Sociales	3	3	4	3	4	3	4	3	4	3	4	3	4	3	4	3	55
Salud	7	7	7	8	7	7	8	7	7	8	7	7	8	7	7	7	116
CEA	2	3	2	3	2	3	2	3	3	2	3	2	3	2	3	2	40
Ingeniería	6	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	111
Artes	2	2	3	2	2	2	2	3	3	2	3	2	2	3	2	2	37
Total	31	33	35	34	34	33	35	34	37	33	36	32	36	33	35	32	

Tabla VI
Ejercicio de corte agrupando los 16tiles

Niveles	Agrupación de 16tiles
Nivel 0	R1+R2+R3+R4+R5
Nivel 1	R6+R7+R8
Nivel 2	R9+R10+R11
Nivel 3	R12+R13
Nivel 4	R14+R15
Nivel 5	R16

Finalmente la universidad tomó la decisión de agrupamientos y cortes a partir de la media y la desviación estándar y con exclusión de valores extremos en función de los criterios que consideró más adecuados y pensando en las actividades o grupos que deseaba estimular.

A fin de ilustrar las repercusiones que esta decisión tuvo en la asignación del estímulo al personal sujeto a evaluación, en este último apartado se hace una comparación de los resultados de ambos procedimientos para establecer los cortes.

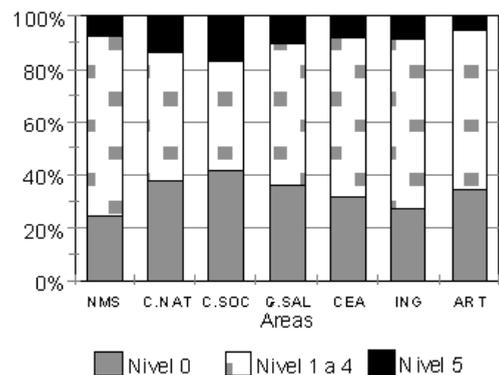
Comparación de resultados entre ambos procedimientos

A partir de los puntos de corte proporcionados por la universidad, el IIDE analizó las siete series de puntuaciones para aproximarse a la forma en que fueron otorgados los estímulos tomando como punto de partida la media y la desviación estándar con exclusión de valores extremos.

Tal como puede observarse en la Figura 3, el nivel medio superior sería el área que tendría un porcentaje más alto de académicos beneficiados (75%). Sin embargo, esta también sería una de las áreas donde se registraría un porcentaje menor en el nivel más alto del estímulo (correspondiente al nivel cinco). En segundo término, Ingeniería tendría también un alto porcentaje de académicos beneficiados (72.1%). En Ciencias Naturales, Ciencias Sociales y en Ciencias de la Salud, un porcentaje mayor de gente no tendría acceso al estímulo (38.1%, 41.8%, 37%, respectivamente); mientras que en estas mismas áreas se registrarían los porcentajes más altos de académicos bene-

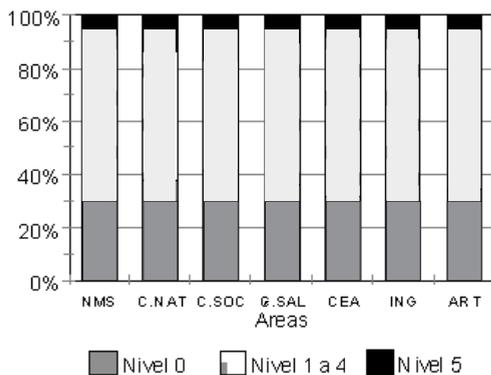
ficiados con el nivel 5. Ciencias Sociales sería el área donde se otorgaría estímulo a un porcentaje menor de académicos (tan sólo 58.1%) y sería también el área donde un mayor porcentaje recibiría el nivel 5 del estímulo (16.4%). Comparativamente con el resto de las áreas, podría decirse que artes sería poco beneficiada en tanto que el porcentaje de académicos que recibiría estímulo fue relativamente bajo (64.8%) siendo el porcentaje de académicos beneficiados con el nivel 5 el más bajo (5.4%) de todas las áreas.

Figura 3
Distribución de estímulos por un procedimiento de transformación lineal y exclusión de valores extremos



En la Figura 4 se muestra la forma en que se conformarían los grupos en función de una cierta agrupación de 16 tiles elaborada meramente con fines de un ejercicio comparativo. Tal como puede observarse en la gráfica, la distribución de los niveles de estímulo es equitativa en las siete áreas cuando se utiliza el método de la igualación percentilar, ya que en todos los casos el porcentaje en cada corte es equivalente.

Figura 4
Ejercicio de distribución de estímulos por un procedimiento de igualación percentilar



A partir de este ejercicio comparativo puede verse con claridad que el número de instrumentos de evaluación utilizados, pero también el procedimiento de corte puede privilegiar a un grupo sobre otro. Si se hubiera utilizado un solo instrumento (tal como plantea SHCP), el resultado privilegiaría a aquellos académicos cuya producción está mejor retratada en el instrumento propuesto por esta Secretaría, es decir, aquellos que realizan investigación. Utilizar un número mayor de instrumentos de evaluación que evalúa el tipo de productividad de cada área de conocimiento, permite una mayor equidad en tanto este modelo identifica y estimula a los académicos más productivos de cada área de conocimiento, sin poner a competir áreas que son completamente distintas, tales como Artís-

ticas y Ciencias Naturales. Finalmente, utilizar tantos instrumento como áreas de conocimiento estén definidas, pero hacer los cortes para asignación de beca por medio de un procedimiento percentilar permite una equidad mucho mayor en tanto que, independientemente del número de solicitantes por área, se estimula al mismo porcentaje de solicitantes por área.

Conclusiones

La actividad docente es el eje central de la vida académica de cualquier IES ya que en ella se cristalizan las funciones sustantivas de toda universidad: la docencia, la investigación, la difusión de la cultura y la vinculación con la sociedad. Por consiguiente, el profesor es la figura central del proceso educativo y el eje fundamental para impulsar la calidad de una universidad.

Las políticas gubernamentales actuales de modernización educativa están orientadas a impulsar y fortalecer la planta académica del las IES a través de estímulos económicos asociados a la producción del docente, política que hace indispensable la evaluación del desempeño académico.

Cualquier proceso de evaluación educativa puede ser de tres tipos: diagnóstica, formativa y sumaria. La primera sirve para conocer las deficiencias y fortalezas al inicio de un proceso educativo con el fin de tomar medidas correctivas pertinentes; la segunda se aplica durante los procesos educativos con el propósito de retroalimentarlos para mejorarlos, y la tercera busca conocer los resultados para certificarlos o premiarlos. Este último caso es la función que cumple la evaluación en el contexto de los PEDPA.

Sin embargo, este propósito no está claramente definido en los programas de incentivos; en muchas ocasiones se justifica la evaluación docente con el propósito, no solo de estimular económicamente al profesor, sino también de retroalimentar su actividad académica (función formativa). Esta ambigüedad en los fines que

persigue la evaluación, tratando de cumplir una doble función que no siempre es posible, hace que se presenten muchas dificultades en un proceso evaluativo (Canales, 2001). Por lo anterior, encontramos que los PEDPA han sido muy criticados por los resultados negativos que producen. Por ejemplo, se ha documentado que estos programas han estimulado el acopio indiscriminado de constancias de todo tipo, más que el interés por mejorar la calidad del trabajo docente (Díaz Barriga, 1997).

El análisis de la experiencia de construcción de este programa de estímulos permite puntualizar algunas de las orientaciones metodológicas más importantes:

- 1) Se definió con claridad el modelo evaluativo utilizado, que en este caso fue normativo;
- 2) Se analizaron las implicaciones de la elección de este modelo en términos de: a) las particularidades de cada área de conocimiento, b) el diseño del instrumento de evaluación, c) la escala de calificaciones utilizada y d) la forma de establecer las puntuaciones de cada nivel del estímulo. En el caso de esta universidad esto implicó el trabajar con tantos instrumentos de evaluación como áreas de conocimiento habían sido definidas, dejar libres las escalas de calificaciones en cada área de conocimiento y buscar procedimientos estadísticos que permitieran hacer la comparación entre los resultados de las siete áreas.

Los cuerpos académicos correspondientes tomaron las decisiones del proceso a partir de un análisis técnico y político de sus alcances y repercusiones. Evidentemente, en este modelo de evaluación del desempeño, la autoridad universitaria en turno y sus cuerpos colegiados se distanciaron de un modelo simple, como el que propone la SHCP en la búsqueda de un sistema de evaluación que intentara recuperar la diversidad y la complejidad de la práctica académica.

En el trabajo de colaborar en el diseño de un programa de estímulos en una universidad pública, encontramos diversos problemas bási-

cos en la metodología de la evaluación docente que se presentan con el esquema que impone la SHCP. A saber: 1) incongruencia entre el modelo evaluativo declarado (criterial) y el modelo utilizado, finalmente en virtud de que los recursos son limitados (normativo); 2) imposibilidad de evaluar con un solo instrumento (o escala) la diversidad de actividades y funciones académicas, 3) arbitrariedad en la determinación de ponderaciones de las actividades evaluadas y 4) elección errónea de la escala de puntuaciones. Todos estos errores metodológicos ocasionan que tanto la evaluación docente, como el uso que se le da para remunerar a los académicos no sean los más apropiados.

Desde nuestro punto de vista, podemos concluir que los sistemas de evaluación en el contexto de los PEDPA deben basarse en tres pilares fundamentales. El primero de ellos es la definición clara de lo que la institución espera del académico, considerando su giro profesional y toda su diversidad laboral (docentes, investigadores, gestores, artistas, deportistas, etc.), así como lo que se considera como evidencia apropiada para su desempeño exitoso. Dicha definición deberá estar de acuerdo con la misión de la institución y con las metas que se propone alcanzar, de tal manera que el instrumento no compita con los programas de desarrollo de la institución ni con los planes de trabajo de las unidades académicas (Izquierdo, 1998).

El segundo pilar es la participación de la comunidad académica que será evaluada, la cual tendría que darse en varios niveles: desde la elección del modelo de evaluación, el diseño del instrumento y la conformación de cuerpos evaluadores. La participación de los académicos es fundamental para contribuir a garantizar la operación de estos programas de forma tal que no atente contra su dignidad profesional. Pero esta participación no debe hacerse en forma individual y aislada, sino en cuerpos colegiados legitimados por la institución, que filtren y supervisen el proceso de operación del programa de estímulos.

El tercer pilar fundamental en el diseño y

operación de los PEDPA es la participación de especialistas en evaluación, quienes tienen el conocimiento y preparación profesional para asegurar que la toma de decisiones con base en los resultados de un proceso evaluativo sea válida y justa. Por muy buena intención que se tenga, el desconocimiento de los aspectos técnicos de la evaluación ocasiona graves errores que pueden distorsionar el propósito mismo de los programas de estímulos, programas que se consideran de gran escala e impacto social. La complejidad de la actividad docente obliga que estos sistemas evaluativos y remunerativos, en caso de utilizarse, deban ser estudiados y diseñados con mucho profesionalismo a fin de utilizar un instrumento cuyas consecuencias se desconocen y pueden

ser incluso contrarias al objetivo del programa.

Sintetizando, debemos mencionar que los esfuerzos para mejorar la actividad académica de las IES deben ser diseñados integralmente, atendiendo la diversidad de las actividades docentes, las diferencias de sus miembros, las condiciones actuales de la institución, así como sus planes futuros. Un sistema de esta naturaleza puede hacerse cuando las autoridades universitarias tienen claridad de los rumbos de su institución, cuando los grupos de académicos toman parte en las decisiones y cuando participan los especialistas en evaluación que pondrán a prueba que el sistema cumpla con lo que se propone.

Referencias

- AIKEN, L.R. (1996). *Tests psicológicos y evaluación*, México, Prentice Hall Hispanoamericana.
- ANUIES. *Evaluación del desempeño del personal académico*, México, ANUIES.
- CANALES, A. (2000). *La experiencia institucional con los programas de estímulo: La UNAM en el periodo 1990-1996*, México, DIE.
- CORDERO, G. y E. Backhoff (2000, noviembre). "Proceso de elaboración de un programa de estímulos al personal académico en una universidad pública mexicana". *Memorias del IV Foro de Evaluación Educativa*, México, CENEVAL.
- DÍAZ BARRIGA, A. (1996). "Los programas de evaluación (estímulos al rendimiento académico) en la comunidad de investigadores: Un estudio en la UNAM", *Revista Mexicana de Investigación Educativa*, 1(2), 408-423.
- DÍAZ BARRIGA, A. y T. Pacheco (Coords.) (1997). *Universitarios: institucionalización académica y evaluación*, México, Universidad Nacional Autónoma de México, Centro de Estudios sobre la Universidad.
- GIL ANTÓN, M. (2000). "Hay otro camino", *La crónica de hoy*, 20 de marzo de 2000.
- GIL ANTÓN, M. (2002). "Cuando el hielo cruje", *La crónica de hoy*, 14 de enero de 2002.
- GREDIAGA, R. (1998). "Cambios en el sistema de recompensa y reconocimiento en la profesión académica en México", *Revista de la Educación Superior*, Núm.108.
- HERNÁNDEZ, M.L. (1998). *El diseño gubernamental del programa "Carrera Docente" y su implementación en la Universidad de Guadalajara: Un estudio de caso*, Tesis doctoral, Aguascalientes, Universidad Autónoma de Aguascalientes, Centro de Ciencias Sociales y Humanidades.
- IZQUIERDO, M.A. (1998). *Estrategias, oposiciones y conflictos entre maestros de la Universidad Pedagógica Nacional en su competencia por los estímulos económicos*, Tesis doctoral,

Cuernavaca, Universidad Autónoma del Estado de Morelos, Instituto de Ciencias de la Educación.

KASTEN, K. (1984). "Tenure and merit pay as rewards for research, teaching and service at a Research University", *Journal of Higher Education*, 55 (4).

KOHN, A. (1994). "Por qué no funcionan los programas de incentivos", *Universidad Futura*, 5(15).

KROTZ, E. (1993). "Evaluación, puntos y sobresueldos en la Universidad Autónoma de Yucatán", *Universidad Futura*, 4(11), 58-68.

LARRAZOLO, N. (s/f). *Escala de calificación. Examen de habilidades y conocimientos básicos*. México, UABC-IIIDE.

LEWIS, L. (1983). "The definitions of academic merit", *Higher Education*, 12 (6).

LÓPEZ ZÁRATE, R. (1991). "Académicos: El caso de la UAM", *Universidad Futura*, 3(8-9).

MARTÍNEZ RIZO, F. (2000). "La evaluación del personal académico. En busca de sistemas de orientación sintética", en ANUIES (Comp.) *Evaluación del desempeño del personal académico*, México.

ORNELAS, C. (1995). *El sistema educativo mexicano. La transición de fin de siglo*, México, Fondo de Cultura Económica.

POPHAM, J. (1990). *Modern educational measurement*, Allyn and Bacon.

RUEDA, M. y M. Landesmann (Coord.) (1999). *¿Hacia una nueva cultura de la evaluación de los académicos?*, México, UNAM-CESU.

SECRETARÍA DE HACIENDA Y CRÉDITO PÚBLICO (1997). *Lineamientos generales para la operación del "Programa de estímulos al desempeño del personal docente" de educación media superior y superior*, México.

SECRETARÍA DE EDUCACIÓN PÚBLICA y Asociación Nacional de Universidades e Instituciones de Educación Superior (1997). "Programa de Mejoramiento del Profesorado de las Instituciones de Educación Superior", *Revista de la Educación Superior*, Núm. 101, Enero-Marzo.

STEFANOVICH, A.F. (1998). *Modelo general para la evaluación del trabajo de los académicos en las instituciones de educación superior*, tesis de doctor en educación, México, Universidad Anáhuac, Facultad de Educación.