

Participación de académicos y estudiantes en la evaluación de la docencia*

EDNA LUNA SERRANO Y MARIO RUEDA BELTRÁN**



El trabajo reporta el estado de los posgrados en la Universidad Autónoma de Baja California, con el propósito de construir instrumentos de evaluación de la docencia para este nivel que contemplen:

- a) las recomendaciones hechas por los docentes de los diferentes programas;
- b) las opiniones de los estudiantes sobre los reactivos y dimensiones de evaluación, y c) la especificidad de cada área de conocimiento. La investigación se diseñó como un estudio de caso a partir del planteamiento de los métodos combinados. Se utilizaron cuestionarios, entrevistas semiestructuradas, análisis de contenidos y cálculos estadísticos. Se identificaron las dimensiones de evaluación más importantes para la enseñanza en los posgrados de la UABC, se reconocieron los aspectos particulares de evaluación por área de conocimiento de los diferentes programas y se definió una jerarquía general de las dimensiones, así como las jerarquías correspondientes a los agrupamientos de los programas.

In this article, the author reviews the state of the postgraduate studies at the Autonomous University of Baja California (Universidad Autónoma de Baja California or UABC) in order to build assessment systems for teaching at postgraduate level which should include: a) the recommendation made by the teachers of the different programs; b) the students' opinion about the reagents and stature of assessment; and c) the specificity shown by each area of knowledge. This research was built as a case study which started with the planning of combined methods. It made use of questionnaires, semistructured interviews, contents analysis and statistic calculus. It also tried to identify the most important features of assessment with regard to teaching in the postgraduate studies at the UABC, recognized the particular aspects of assessment for each program and each area of knowledge and defined an overall hierarchy of those features, as well as the hierarchy that correspond to the grouping of the programs.

Evaluación / Docencia / Etnografía educativa / Educación superior / Posgrado / Participación del alumno
Assessment / Teaching / Educational ethnography / Higher education / Postgraduate studies / Student participation

La experiencia que aquí se relata parte de la convicción de que el tema de la evaluación de la docencia es sumamente complejo y que para enfrentarlo es necesario sistematizar y compartir los intentos realizados en situaciones concretas; tal es el presente caso que comunica lo ocurrido en el posgrado de la Universidad Autónoma de Baja California.

Los obstáculos que ofrece la evaluación de la docencia en el ámbito universitario tienen que ver con muy diferentes elementos: su reciente y rápido surgimiento en México, su asociación a políticas de compensación salarial, la discusión sobre las cambiantes funciones del profesor, la polémica de las teorías sobre la enseñanza y el aprendizaje, el desconocimiento y escasas iniciativas para diversificar los medios de evaluación de dicha actividad, por mencionar algunos de los más importantes. Por tal cúmulo de problemas que presenta la acción de evaluar la docencia en la Universidad, vemos como indispensable la participación de los involucrados, profesores y estudiantes, en dicho proceso como una condición mínima para atenuar la dificultad de la tarea.

ANTECEDENTES

Indiscutiblemente, la evaluación de la docencia por medio de la opinión externalizada por los estudiantes en un cuestionario se ha constituido como el medio más

frecuentemente empleado en la mayor parte de las universidades del mundo, con la aceptación explícita en muchas ocasiones de los propios académicos.

En el ámbito internacional, la utilización de los cuestionarios de evaluación de la docencia con base en los puntajes de evaluación de los estudiantes tiene una historia de más de setenta años y es sin duda la estrategia más empleada (Centra, 1993; Husbands y Fosh, 1993); en el nacional, los reportes sobre su uso sistemático datan de la segunda mitad de la década de los ochenta (García, 2000) y se considera que su empleo se generaliza en los noventa.

En México la evaluación de la docencia específicamente en posgrado también surge de forma generalizada después de 1990, debido al énfasis que las políticas gubernamentales otorgaron a la evaluación en general. La afirmación se sostiene por la información recopilada de once universidades públicas (Luna, 2000), en las que se identificó que sólo en tres de ellas la evaluación de la docencia se llevaba a cabo antes de 1990, con el propósito de mejorar el desempeño académico. En siete instituciones la evaluación de la docencia se establece después de ese año, con la doble intención de participar en los programas de compensación salarial y mejorar el desempeño docente. Todas las universidades utilizan más de un procedimiento de evaluación; sin embargo, el instrumento privilegiado son los cuestionarios de evaluación a partir de la opinión de los alumnos. Su procedencia es diversa, ya que en la mayor parte de los casos es el mismo de la licenciatura, o se retoma de la literatura especializada; sólo en dos universidades se diseñaron cuestionarios ex profeso para el nivel de posgrado.

* Este trabajo forma parte de uno más amplio que Edna Luna Serrano presentó con el título: "La evaluación de la docencia a partir de la opinión de los alumnos en posgrado: el caso de la Universidad Autónoma de Baja California", para obtener el grado de doctora en Educación en la Universidad Autónoma de Sinaloa en diciembre de 2000.

** Universidad Autónoma de Baja California y CESU-UNAM.
eluna@email.ens.uabc.mx
mariorb@servidor.unam.mx

Los cuestionarios de evaluación de la docencia con base en la opinión de los alumnos se utilizan tanto con propósitos de tipo administrativo como para mejorar la actividad docente (Ory, 1990). El uso de este medio con fines de reconstrucción y compensación salarial ha generado diversos cuestionamientos críticos; destacan los relacionados con el manejo administrativo, esto es, la aplicación, interpretación y utilización de los resultados (Theall y Franklin, 1990; Sproule, 2000). La práctica generalizada es cuidar los componentes psicométricos; sin embargo, esto resulta insuficiente en virtud de que en la evaluación de los docentes se interrelacionan otros aspectos, como los políticos, los contextuales, los organizacionales y de credibilidad hacia los evaluadores, los cuales pueden llegar a invalidar una evaluación, a pesar de que en términos de medición tenga un alto índice de confiabilidad y validez.

En los cuestionarios se asume la enseñanza como una actividad multidimensional que la mayoría de los instrumentos intenta representar; desde esta perspectiva se han investigado las dimensiones relacionadas con la efectividad docente. Feldman (1997) identificó como de mayor importancia la preparación y la organización del curso, la claridad y el entendimiento, el logro de los objetivos planteados, la estimulación del interés por la materia y la percepción del impacto de la instrucción por parte del alumno.

Un supuesto que ha dominado las prácticas de evaluación de la docencia a partir de la opinión del alumno es que el conocimiento de los docentes y sus habilidades pueden evaluarse sin considerar su contexto. Así, la mayoría de los instrumentos aplicados actualmente están basados en la noción de características genéri-

cas de la enseñanza (Marsh, 1984; Delandshere, 1994), donde se considera que las cualidades importantes de la enseñanza efectiva son sustancialmente invariables, esto es, que no varían de acuerdo con los cursos, disciplinas y universidades (Marsh, 1984; Abrami y d'Apollonia, 1990). Desde este enfoque se han realizado numerosos estudios sobre la validez de los puntajes de opinión de los estudiantes, y se ha encontrado que el nivel del curso –licenciatura o posgrado–, la naturaleza del mismo –obligatorio u optativo– y el área del contenido influyen sobre los puntajes asignados por los estudiantes al profesor (Feldman, 1978; Cohen, 1981; Abrami, d'Apollonia y Cohen, 1990). Por ejemplo, Smith y Cranton (1992) reportan que la organización del curso y la claridad en las exposiciones son los indicadores más significativos para los estudiantes de los primeros años de licenciatura, en tanto que el interés por el alumno y la evaluación del aprendizaje lo son para los de niveles avanzados de licenciatura y posgrado.

De acuerdo con Marsh (1987) y Marsh y Dunkin (1997), el desarrollo de un instrumento completo debe combinar tres estrategias: 1) realizar análisis empíricos utilizando técnicas estadísticas como análisis factorial y diseños multirrasgo-multimétodo (MTMM); 2) deducir los reactivos de un análisis lógico de los componentes de la efectividad docente, tomando en cuenta la retroalimentación de los profesores y los estudiantes, así como los propósitos de la evaluación, y 3) fundamentar los instrumentos de evaluación en una teoría de la enseñanza o el aprendizaje.

Un problema de investigación poco estudiado se centra en la ponderación de las distintas dimensiones que conforman los cuestionarios de opinión. Marsh (1984; 1987) señala la dificultad de asignar pesos

diferenciados a las dimensiones en función de los propósitos de la evaluación; sugiere que la ponderación se haga de manera empírica en función del tipo de curso y disciplina, y se asigne un promedio a cada dimensión. Se considera importante investigar este aspecto en el contexto de los sistemas de evaluación institucional.

El contexto del trabajo

La Universidad Autónoma de Baja California (UABC) ocupa a escala nacional el decimotercer lugar en la matrícula de licenciatura y el decimoquinto lugar en la matrícula de posgrado; respecto a la frontera norte, atiende a 7.4% de los estudiantes de licenciatura y a 3.2% de los de posgrado (ANUIES, 1997). En el estado es la única institución que ofrece estudios de posgrado en las diversas áreas del conocimiento.

La evaluación de la actividad docente en el posgrado de la UABC inició de manera sistemática en 1990, junto con la aplicación del programa de estímulos a la carrera docente. Uno de sus componentes son los puntajes otorgados por los estudiantes, los cuales representan 25% del total de los indicadores que componen el programa de estímulo. La evaluación de la enseñanza, a partir de la opinión de los estudiantes, se lleva a cabo por parte de la Dirección General de Investigación y Posgrado (DGIP), utilizando un formato único para todos los posgrados adaptado del de la licenciatura. Institucionalmente se utiliza con fines administrativos y de retroalimentación.

En un estudio previo realizado por personal de la DGIP, se identificaron diversos cuestionamientos a este formato de evaluación; la opinión generalizada de maestros y directivos respecto a su contenido fue que evaluaba parcialmente los

componentes de enseñanza derivados de las características propias de cada programa, y que las dimensiones de evaluación, dada su importancia desigual en el proceso de enseñanza, deberían tener un peso diferencial al valorarlas. Además, los docentes de la maestría y doctorado en Oceanografía Costera habían externado su deseo de participar en el proceso de elaboración de su instrumento de evaluación.

La disposición de docentes y directivos para participar en un proceso de construcción de instrumentos de evaluación, y la convicción de muchos de ellos sobre la necesidad de contar con cuestionarios acordes con las características de sus programas, fueron los elementos contextuales más relevantes que posibilitaron la realización de este trabajo.

Propósito

El propósito central que guió las acciones del estudio fue la construcción de instrumentos de evaluación docente para los posgrados de la UABC. Los instrumentos deberían incluir las recomendaciones formuladas por los docentes; las opiniones de los estudiantes sobre los reactivos y las dimensiones de evaluación, y los aspectos específicos señalados por cada área de conocimiento. Adicionalmente, se buscó determinar la jerarquía de las dimensiones de evaluación con base en la opinión de docentes y alumnos, así como conocer la valoración del sistema que administra la evaluación de la docencia desde la perspectiva de los docentes y estudiantes.

METODOLOGÍA

La investigación se abordó como un estudio de caso, en virtud de que, para nosotros, fue fundamental conocer el punto

de vista de los participantes y recabar las distintas opiniones de acuerdo con los propósitos pedagógicos y las diferencias entre las disciplinas, así como trabajar desde una perspectiva metodológica que posibilitara ser sensibles al contexto para identificar las particularidades de cada programa. Se utilizaron técnicas cualitativas y cuantitativas para cubrir los diferentes objetivos. Las más importantes fueron los cuestionarios, las entrevistas semiestructuradas, el análisis de contenido y los cálculos estadísticos (específicamente medidas de tendencia central y análisis de conglomerados). Así, el estudio se enmarcó en una metodología cualitativa de investigación a partir del planteamiento de los métodos combinados (Bryman, 1984).

En la primera parte se describe el proceso de construcción de los instrumentos de evaluación y la jerarquización de las dimensiones, a partir de la discusión realizada con maestros y alumnos mediante cuestionarios; en la segunda se pretende completar la información con la opinión de los participantes sobre los aspectos pedagógicos que caracterizan a un buen docente, así como las características deseables en el sistema que administra los cuestionarios de evaluación. Esto se llevó a cabo por medio de entrevistas semiestructuradas realizadas con dos muestras intencionales.

Aplicación de los cuestionarios

Esta sección constó de cuatro fases. En la primera, se elaboró el cuestionario (Evaluación de la actividad docente en posgrado) a partir de las dimensiones y reactivos que tanto la literatura anglosajona cuanto hispana marcan como relevantes (Landsheere, 1980; Shatz y Best, 1986; Tejedor,

1990; Centra, 1993; Marsh y Baily, 1993). Incluye las dimensiones: a) estructuración de objetivos y contenidos; b) claridad expositiva; c) organización de la clase; d) dominio de la asignatura; e) cualidades de interacción; f) evaluación del aprendizaje, y g) asesoría al proyecto de tesis, cuando procede. En la definición de su contenido se consideraron las sugerencias hechas por la DGIP, por ser la autoridad que regula la enseñanza en posgrado y la responsable de la administración de este tipo de evaluación en la UABC; además, se revisaron los documentos sobre el tema producidos por académicos de diferentes posgrados. Finalmente, se realizó un estudio piloto con los estudiantes para evaluar la claridad del cuestionario. Asimismo, se diseñó un cuadro de ponderación para identificar la jerarquía de las dimensiones. El propósito de estos instrumentos fue recoger la opinión de docentes y alumnos sobre los diferentes aspectos que cubre la evaluación de la docencia.

En la segunda, se proporcionó a los docentes el cuestionario diseñado para que señalaran al calce los comentarios y sugerencias a cada pregunta. También se incluyó el cuadro de ponderación de las dimensiones, en el que se les solicitó que, con base en su experiencia, asignaran el nivel de importancia que ellos propondrían para cada una de las dimensiones de la actividad docente. Como parte de esta etapa, el cuestionario se reconstruyó incorporando las sugerencias con mayor frecuencia.

En la tercera, se entregó a los estudiantes el cuestionario corregido y el cuadro de ponderación de las dimensiones con el fin de recoger sus opiniones. Por último, se elaboró la versión final del cuestionario al incluir los comentarios de los alum-

nos. En la cuarta se presentó a los maestros la versión final del cuestionario y los resultados más relevantes del trabajo de campo efectuado en las etapas anteriores.

El estudio se llevó a cabo con la población de docentes y estudiantes de once programas de maestría y uno de doctorado de la UABC; se encuestaron 127 docentes, 86% de la población total, y 125 estudiantes (66%); cabe hacer notar que las subpoblaciones encuestadas fueron representativas (véase cuadro 1).

CUADRO 1 • Posgrados por área del conocimiento, número de encuestados y porcentaje de maestros y estudiantes.				
Programas	Total de maestros	Encuestados	Total de alumnos	Encuestados
Ciencias exactas y naturales	53	49 = 92%	79	66 = 84%
M. en Ciencias Químicas	16	13 = 81%	25	20 = 80%
* M. en Zonas Áridas	8	7 = 81%	10	6 = 60%
* M. y Doc. en Oceanografía	29	29 = 100%	44	40 = 91%
Ciencias médicas	38	26 = 68%	37	28 = 76%
M. en Ciencias Médicas	8	8 = 100%	17	14 = 82%
M. en Nutrición	9	6 = 67%		
M. en Salud Pública	21	12 = 57%	20	14 = 70%
Ciencias agropecuarias	21	15 = 71%	26	8 = 31%
*M. en Producción Animal				
*M. en Ciencias Agrícolas				
Ciencias de la ingeniería	21	17 = 81%	22	14 = 64%
M. en Arquitectura	9	5 = 56%	7	6 = 86%
M. en Ingeniería de Sistemas			14	8 = 50%
M. en Ingeniería Termodinámica	12	12 = 100%	1	1 = 100%
Ciencias sociales	14	14 = 100%	24	9 = 38%
M. en Asuntos Internacionales	6	6 = 100%	17	5 = 29%
*M. en Ciencias Sociales	8	8 = 100%	7	4 = 57%

* Posgrados reconocidos por CONACYT en su Padrón de Excelencia

La aplicación del cuestionario a docentes y estudiantes se llevó a cabo con el apoyo de la DGIP, que promovió la participación de directivos y maestros e hizo posible el acceso a los salones de clase. El procedimiento general consistió en: a) explicar a los participantes los propósitos de la investigación y solicitar su colaboración en el proceso de construcción de los cuestionarios, así como aclarar que la investi-

gación era independiente de la DGIP, aunque contaba con su apoyo; *b*) solicitar la revisión de la propuesta de cuestionario y escribir al calce los comentarios y sugerencias a cada pregunta, y *c*) describir el cuadro de ponderación de las dimensiones, asignar un nivel jerárquico a las mismas y mencionar otros aspectos no incluidos cuya evaluación consideraran importante.

Las entrevistas

La muestra de los docentes fue de quince maestros, tres de las áreas de exactas y naturales, médicas, agropecuarias, ingeniería y sociales. Para su selección, se tomó como criterio identificar por lo menos un maestro de cada programa reconocido por alumnos y colegas como muy buen docente; cabe recalcar que se destacó el reconocimiento en la docencia y no en la investigación. La muestra de los estudiantes se formó con catorce alumnos, cuatro de las áreas de ciencias exactas y naturales e ingeniería; tres de ciencias sociales, dos de ciencias agrícolas y uno de ciencias médicas. Se entrevistaron nueve estudiantes catalogados como muy buenos y cinco como regulares; para identificar a los primeros se aplicó el criterio de contar con un promedio mínimo de 9 y no haber reprobado ninguna materia en el programa de posgrado; para los segundos, tener promedio de 8 o menor, y para ambos grupos, haber cursado por lo menos un año del programa.

Las entrevistas se efectuaron en cubículos o salones de clase y fueron grabadas. El propósito de entrevistar a los docentes fue documentar a profundidad las reflexiones de los maestros sobre su práctica docente. Para ello, se les preguntó sobre los criterios pedagógicos que la Facultad

les solicita respecto al proceso de enseñanza-aprendizaje y los criterios pedagógicos que los maestros consideran son los pertinentes en su práctica docente (se insistió en identificar lo que define su práctica pedagógica); preguntar sobre ambos aspectos posibilitó identificarlos y contrastarlos. Asimismo, se pidieron sugerencias sobre aspectos a integrar en los cuestionarios de evaluación de los alumnos y su opinión sobre el procedimiento de aplicación y sobre la utilización de los cuestionarios. El promedio de duración de cada entrevista fue de 45 minutos.

Las entrevistas a los estudiantes se llevaron a cabo con el propósito de conocer las opiniones de los alumnos sobre las características o cualidades que identifican a un buen docente, así como su valoración acerca de las funciones que cumple el cuestionario de evaluación de la docencia con base en la opinión del estudiante. En promedio las entrevistas fueron de 30 minutos.

RESULTADOS DE LOS CUESTIONARIOS

El procedimiento seguido para determinar los elementos de las dimensiones de evaluación del trabajo académico consistió en la elaboración de una base de datos que concentró los comentarios hechos a cada uno de los reactivos, las dimensiones, y la identificación del programa de procedencia del encuestado; los comentarios se agruparon por categorías de sugerencias y se computó su frecuencia. Se elaboró una tabla para cada reactivo y dimensión con el fin de integrar la información. La determinación de la frecuencia de sugerencias por categoría en cada reactivo permitió realizar el análisis exploratorio del comportamiento de los datos.

Dado que el número de participantes por programa de posgrado varió de manera importante, se decidió agruparlos por área de conocimiento y así trabajar con números mayores en las subpoblaciones.

De los docentes

Una de las características de los resultados, a pesar del número de cuestionarios aplicados, es lo reducido de las sugerencias; esto determinó atender observaciones con frecuencias de 2 o 3. El que un reactivo o dimensión no tuviera comentarios se consideró como aceptado; por tal motivo, la falta de respuesta fue interpretada como aprobación. Asimismo, cabe destacar que todas las dimensiones del cuestionario base fueron aceptadas, así como la mayor parte de los reactivos.

Las categorías de sugerencias para los reactivos fueron, por un lado, de forma –“redacción confusa y reactivo repetido”– y por otro, de fondo –una sola, “no procede evaluar”–, con diferentes argumentaciones según el caso: no se aplica para todos los cursos, el alumno no puede valorar si la bibliografía es actualizada, etc. Estas opiniones fueron hechas por los docentes de los diferentes posgrados.

En general, los docentes apelan a una redacción confusa en los reactivos donde explícitamente se solicita evaluar su ejecución en clase, como en los siguientes casos: “¿explica claramente los contenidos del curso?”, “¿presentó los fundamentos teóricos o metodológicos de los contenidos desarrollados en clase?”, y “¿resolvió los problemas planteados en clase?”; estas sugerencias provienen de manera principal de los posgrados en ciencias naturales y exactas, con un rango de frecuencias entre 3 y 4.

En la categoría “no procede evaluar”, la objeción “no se aplica para todos los

cursos” se identificó en dos tipos de reactivos; por un lado, en los que aluden a la ejecución del maestro con relación al dominio del contenido, y por otro, en los vinculados con la enseñanza de principios básicos de investigación. Entre los primeros se ubican los reactivos: “¿enfatisa los aspectos importantes del curso?”, “¿establece una clara diferenciación entre sus puntos de vista y los de los autores?”, y “¿relacionó el contenido de la materia con otras materias?”, con un rango de frecuencias entre 3 y 4; con referencia a los segundos se encuentran, “¿se definieron posibles problemas de investigación?”, “¿se realizaron actividades dirigidas a plantear problemas de investigación?”, “¿se comentaron las posibles metodologías a seguir de acuerdo con los problemas de investigación planteados?”; estas sugerencias las emitieron docentes de las áreas de ciencias exactas y naturales, ingeniería y ciencias sociales, con una frecuencia de 13.

En la misma situación se ubica “no procede evaluar por el alumno”, identificada en el reactivo que evalúa la pertinencia de la bibliografía; esto fue señalado por nueve docentes de las diferentes áreas del conocimiento. También, los docentes censuraron los reactivos que aluden a situaciones fuera de su control, como es el caso de “¿definió los objetivos del curso?”, por considerar que la asistencia de los alumnos los primeros días de clase, cuando generalmente se describen los objetivos del curso, no depende de ellos.

En las categorías de sugerencias a las dimensiones se identificaron opiniones de los docentes de ciertas áreas en particular, así como otras compartidas por todos. Específicamente, los docentes de ciencias agropecuarias y ciencias de la ingeniería indicaron que se evalúe “la aplicación de

los conocimientos a la solución de problemas reales” y “relacionar el contenido con la solución de problemas”; esta última fue mencionada, además, por los de ciencias médicas y ciencias sociales. Las sugerencias en las que coinciden los docentes de todas las áreas son evaluar el rubro de asesoría al proyecto de tesis por separado, y evaluar aspectos formales del cumplimiento docente, como puntualidad, asistencia, o si el profesor imparte el número de horas estipuladas; con el fin de ubicar estos comentarios se incorporó la dimensión método de trabajo. En todas las áreas se sugirió evaluar el uso de apoyos didácticos, con la excepción de los docentes del área de ciencias médicas. Por último, los comentarios relacionados con la evaluación del aprendizaje se hicieron en ciencias agropecuarias, ciencias de la ingeniería y ciencias médicas. Las frecuencias oscilaron entre 2 y 58.

El cuestionario inicial fue modificado a partir de los criterios: a) revisar los reactivos con frecuencia de sugerencias entre 2 y 4, sin que ello necesariamente modificara el reactivo; b) atender las sugerencias cuya frecuencia es de 5 o mayor, de acuerdo con el área de procedencia, lo que implicó modificar un reactivo, incluir uno nuevo e incluir una dimensión con sus respectivos reactivos.

Se efectuaron las modificaciones al cuestionario base de acuerdo con las observaciones y sugerencias de mayor frecuencia; de esta manera se diseñaron tres cuestionarios que difieren entre sí en la inclusión diferenciada de los siguientes reactivos: uno exclusivo para las áreas de ciencias agropecuarias e ingeniería – “¿relacionó el contenido de la materia con la solución de problemas del campo profesional?”–, y “¿se definieron posibles problemas de investigación?” sólo en los

cuestionarios de las áreas de ciencias exactas y naturales y ciencias médicas. Los cuestionarios correspondientes a las áreas del conocimiento quedaron agrupados en ciencias naturales y exactas y ciencias médicas; ciencias agropecuarias y ciencias de la ingeniería, y ciencias sociales.

De los estudiantes

Las respuestas de los estudiantes respecto a los rangos de frecuencias de sugerencias estuvieron por arriba de los docentes. Además, las participaciones de los alumnos presentaron la característica de ser propositivas, ya que explican la confusión y sugieren cómo mejorar el reactivo. Por ejemplo, en el caso: “¿ha explicado claramente los contenidos del curso?”, especificaron: “no se entiende contenido, ¿es el contenido temático?, habría que sustituir por temas o contenido temático”.

Las sugerencias para los reactivos fueron de forma; se identificó un tipo relacionado con la redacción de los reactivos expresado en las modalidades “redacción confusa, modificar redacción se sugiere [...] y son iguales”, que fue el de mayor ocurrencia; y “mover a otra dimensión” señalado en una ocasión. Se registraron comentarios de los alumnos de los diferentes posgrados, y llamó la atención la escasa participación de los de ciencias sociales. Las frecuencias de las participaciones oscilaron entre 0 y 12.

Los alumnos formularon sugerencias para cada una de las dimensiones de evaluación, con excepción de la de dominio de la asignatura. Sin embargo, sus participaciones se concentraron en la dimensión evaluación del aprendizaje, formularon cinco propuestas diferentes relacionadas con el propósito de que se realice una evaluación justa por parte del

maestro. Asimismo, sugirieron la evaluación de aspectos relacionados con la secuenciación lógica de los contenidos y la consideración de los conocimientos previos de los estudiantes al impartir los temas; en atención a ello, se incluyó la dimensión “estrategias de instrucción”. A diferencia de los docentes, no se detectó ningún planteamiento particular por área de conocimiento; por tal motivo, se consideró que, desde el punto de vista de los estudiantes, los cuestionarios responden a las características de su contexto de aprendizaje.

Se mantuvieron los mismos criterios generales utilizados con los docentes para integrar las sugerencias de los alumnos al cuestionario. Además, se siguió como criterio no incorporar las sugerencias de los alumnos que los docentes hubieran censurado con anterioridad, como fue el caso de evaluar si la bibliografía del curso era actualizada, y si se definieron y explicaron los objetivos del curso al inicio del mismo. En síntesis, en los instrumentos no se incluyó ningún reactivo o dimensión con el que los docentes no estuvieran de acuerdo.

Resultados: los cuestionarios

El diseño de los tres cuestionarios producto de este trabajo de investigación se organizó considerando las siguientes dimensiones: estructuración de objetivos y contenidos, que evalúa aspectos relacionados con la estructuración del programa del curso (como congruencia entre los elementos, la pertinencia de la bibliografía y del tiempo asignado a la enseñanza de cada tema); claridad expositiva, que evalúa las habilidades del docente para comunicar los contenidos de manera comprensible (cuestiona sobre la presen-

tación de los temas y si el maestro explicitó los aspectos importantes del curso); organización de la clase, que valora la preparación de la clase y la utilización de materiales que apoyen la enseñanza (como la programación y la realización de actividades); dominio de la asignatura, que evalúa el conocimiento del instructor sobre la materia (presentación de las implicaciones de distintos enfoques y metodologías, la solución de los problemas planteados en clase, y la relación del curso con otras materias o disciplinas); estrategias de instrucción, que valora las estrategias de enseñanza utilizadas por el docente (cuestiona sobre la secuenciación lógica de los contenidos y la consideración de los conocimientos previos de los estudiantes al inicio de cada tema); cualidades de interacción, que evalúa la disposición del maestro para ayudar al estudiante y motivarlo a participar en clase (como la atención al punto de vista, dudas y preguntas de los alumnos); evaluación del aprendizaje, que valora las estrategias del maestro para evaluar los logros del estudiante (incluye la explicitación de las estrategias de evaluación, la pertinencia de las mismas y la ponderación); por último, el método de trabajo valora los aspectos formales del cumplimiento docente (la entrega del programa del curso al inicio del mismo, si la bibliografía del curso se encontró disponible y sobre la asistencia y puntualidad del profesor).

Las variaciones entre los cuestionarios estriban en la inclusión diferenciada de dos reactivos más en la dimensión dominio de la asignatura. En las áreas de ciencias naturales y exactas y de ciencias médicas, se incorporó el reactivo, “¿se discutieron posibles problemas de investigación”. En las áreas de ciencias agropecua-

rias y ciencias de la ingeniería, se incluyó el reactivo, “relacionó el contenido de la materia con la solución de problemas del campo profesional?” (el anexo muestra la versión final de los cuestionarios).

Jerarquización de las dimensiones de evaluación

El análisis estadístico se realizó para determinar la jerarquía de las dimensiones de evaluación del trabajo académico para cada una de las poblaciones, docentes y estudiantes de posgrado; las tendencias de opinión entre los programas, y el nivel de discrepancia y similitud en la opinión de docentes y alumnos, en general y por programa. Los cálculos se efectuaron mediante técnicas de clasificación de análisis de conglomerados, dado que reconocen subconjuntos dentro de un grupo general de datos, tomando como base medidas que permiten identificar las distancias entre los elementos del conjunto (Johnsonbaugh, 1988). El tratamiento de la información fue el mismo para las dos poblaciones, lo que implicó trabajar de manera independiente cada grupo de datos. El procedimiento seguido consistió en las siguientes etapas:

1. A partir del cuadro de ponderación se elaboraron matrices de frecuencia de orden 7 dimensiones por 7 jerarquías por programa, y para cada población. Esto consistió en cuantificar el número total de encuestados que clasificó cada indicador de mayor a menor importancia, con la posibilidad de ubicar a más de un indicador en el mismo nivel. Este procedimiento permitió determinar la matriz de frecuencia del orden asignado a cada dimensión por

programa y para la población total, tanto de docentes como de alumnos.

2. Se normalizaron las matrices de frecuencia 7 por 7 por programa y para cada población. Esto consistió en dividir los componentes de cada matriz entre el número de individuos encuestados; con estos datos se ubicó el orden de importancia de los indicadores, en función de las dos poblaciones estudiadas.
3. Se elaboraron matrices de diferencias, se calcularon tomando el valor absoluto de la diferencia de componente a componente de las matrices de frecuencia normalizadas 2 a 2 (cada programa respecto a todos los demás programas).
4. Se determinó la distancia de opinión entre programas, a partir de la suma total de los componentes de la matriz de diferencias; con esta información se construyeron matrices de disimilitud que concentran la información de las distancias de opinión de cada población en general y de cada programa respecto a los demás; el cuadro 2 presenta la información correspondiente a los docentes. Sus componentes exhiben la diferencia de opinión de los académicos de un programa respecto a los de otro programa, por lo cual, tanto en los renglones como en las columnas del cuadro se ubican programas contra programas, con excepción de G, que se refiere a la población total; por ejemplo, el número 4.76, localizado en la coordenada G-MCS, cuantifica la diferencia de opinión de los profesores de la maestría en Ciencias Sociales respecto a la opinión general de todos los docentes. Con base en estas coordenadas se construyeron los diagramas de árbol que permiten detectar los conglomerados.

CUADRO 2 • Matriz de disimilitud de las dimensiones en relación con las jerarquías para la población de docentes												
	G	MCS	MCM	MA	MIS	MCA	MSP	MN	MCQ	MAI	DOC	MZA
G	0	4.76	3.82	6.37	3.41	3.98	4.14	4.55	3.51	5.6	3.9	6.64
MCS	4.76	0	5.98	7.8	5.75	6.58	5.4	5.8	5.7	7.41	6.92	8
MCM	3.82	5.98	0	6.6	5.94	4.64	5.77	5.38	5.9	5.85	4.7	7.81
MA	6.37	7.8	4.76	0	7.2	6.36	7.6	7.33	7.6	8	7.82	7.8
MIS	3.41	5.75	5.94	7.2	0	5.03	5.2	4.93	4.73	7.86	5.7	6.93
MCA	3.98	6.58	4.64	6.36	5.03	0	6	6.66	5.93	6.66	5.33	7.33
MSP	4.14	5.4	5.77	7.6	5.2	6	0	4.93	4.93	5.93	5.94	9.4
MN	4.55	6.25	5.38	7.33	4.93	6.66	4.93	0	4.26	6.66	6.52	7.33
MCQ	3.51	6.68	4.76	7.6	4.73	5.93	4.93	4.26	0	6.4	5.2	8.02
MAI	5.6	7.41	5.85	8	7.86	6.66	5.93	6.66	6.4	0	6.47	8
DOC	3.9	6.92	6.62	7.82	5.7	5.33	5.94	6.52	5.2	6.47	0	7.66
MZA	6.64	8	7.81	7.8	6.93	7.33	9.4	7.33	8.02	8	7.66	0

Codificación de cada programa de posgrado:

MCS: Maestría en Ciencias Sociales aplicadas, MN: Maestría en Nutrición, MCM: Maestría en Ciencias Médicas, MCQ: Maestría en Ciencias Químicas, MA: Maestría en Arquitectura, MAI: Maestría en Asuntos Internacionales, MIS: Maestría en Ingeniería de Sistemas, DOC: Maestría y Doctorado en Oceanografía Costera, MCA: Maestría en Ciencias Agrícolas, MZA: Maestría en Zonas Áridas, MSP: Maestría en Salud Pública, G: Población

5. Se construyeron las figuras 1 y 2 con la información del primer renglón de la matriz de disimilitud de cada población. La figura 1 presenta la comparación de la opinión general respecto a cada programa para los docentes y la figura 2 para los estudiantes. En los diagramas se ordenó la información a partir de una recta numérica que va del 0 al 10, lo cual permitió reconocer los conglomerados. Los resultados mostraron que sí existen tendencias de opinión entre los programas, tanto en la población de docentes como en la de estudiantes, en cuanto a la preferencia de las dimensiones, de tal manera que los grupos obtenidos reflejan un patrón de agrupamiento por área de conocimiento. En ambos
- casos, como se puede observar en las respectivas figuras, se ubican cuatro conglomerados. El primer conglomerado agrupa los programas cuya opinión tiene mayor similitud respecto a la opinión general; el segundo se encuentra moderadamente alejado, el tercero y cuarto comprenden a los que se alejan en mayor medida de la opinión general.

6. La figura 3 compara los resultados de las matrices de frecuencia normalizadas de académicos y estudiantes para cada programa. Presenta el contraste de opinión entre docentes y estudiantes, identificando los conglomerados y las diferencias. Los resultados identificaron tres conglomerados; el primero, ubica el programa con mayor

FIGURA 1 • Comparación de la opinión general respecto a cada programa de los docentes.

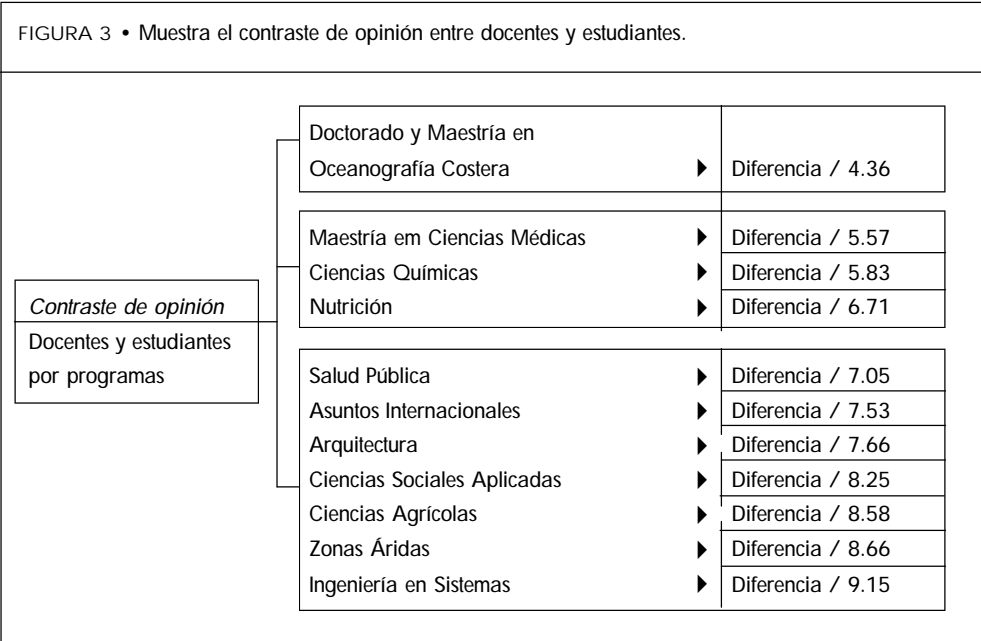
<div>Docentes</div> <div>Opinión general vs. Opinión por programa</div>	Maestría en Ingeniería en Sistemas		Diferencia / 3.41
	Maestría en Ciencias Químicas		Diferencia / 3.55
	Maestría en Ciencias Médicas		Diferencia / 3.82
	Maestría en Oceanografía Costera		Diferencia / 3.90
	Maestría en Ciencias Agrícolas		Diferencia / 3.98
	Maestría en Salud Pública		Diferencia / 4.14
	Maestría en Nutrición		Diferencia / 4.55
	Maestría en Ciencias Sociales Aplicadas		Diferencia / 4.76
	Maestría en Asuntos Internacionales		Diferencia / 5.6
	Maestría en Arquitectura		Diferencia / 6.37
	Maestría en Zonas Áridas		Diferencia / 6.64

FIGURA 2 • Comparación de la opinión general respecto a cada programa de los estudiantes.

<div>Estudiantes</div> <div>Opinión general vs. Opinión por programa</div>	Doctorado y Maestría en Oceanografía Costera		Diferencia / 1.78
	Maestría en Ciencias Químicas		Diferencia / 2.82
	Maestría en Ciencias Médicas		Diferencia / 3.33
	Maestría en Nutrición		Diferencia / 3.81
	Maestría en Salud Pública		Diferencia / 4.02
	Maestría en Arquitectura		Diferencia / 5.21
	Maestría en Zonas Áridas		Diferencia / 6.03
	Maestría en Ciencias Agrícolas		Diferencia / 6.55
	Maestría en Ingeniería de Sistemas		Diferencia / 7.05
	Maestría en Asuntos Internacionales		Diferencia / 7.35
	Maestría en Ciencias Sociales Aplicadas		Diferencia / 7.41

coincidencia de opinión entre maestros y estudiantes; seguido de los medianamente discrepantes, y un último conglomerado, cuya carac-

terística es la dispersión de opinión, sitúa los programas con mayor discrepancia entre la opinión de docentes y alumnos.



7. La jerarquización general de las dimensiones de evaluación se obtuvo de las matrices normalizadas para cada población. Las jerarquías (J) obtenidas se presentan en el cuadro 3; esta jerarquización permitió identificar, por un lado, el orden de importancia de las dimensiones para cada población, y por otro, las coincidencias y discrepancias de opinión. De esta manera, se identificó que el dominio del contenido es la dimensión más importante
- tanto para los docentes como para los estudiantes.

8. Las jerarquías de las dimensiones de evaluación de los conglomerados obtenidos en la población de docentes y de estudiantes se obtuvieron de las matrices normalizadas correspondientes a cada subpoblación; el procedimiento que se siguió para calcular los índices fue el mismo que para la jerarquización general. En el contraste entre la jerarquización general y la de los con-

CUADRO 3 • Jerarquías de las dimensiones de evaluación.		
Docentes	"J"	Estudiantes
Dominio de la asignatura	1°	Dominio de la asignatura
Estructuración de objetivo y contenidos	2°	Claridad expositiva
Organización de la clase	3°	Organización de la clase
Evaluación del aprendizaje	4°	Cualidades de interacción
Cualidades de interacción	5°	Evaluación del aprendizaje
Claridad expositiva	6°	Método de trabajo
Asesoría al proyecto de tesis	7°	Estructuración de objetivo y contenidos

glomerados de los docentes y los estudiantes se confirmó el dominio del contenido como la dimensión más importante.

RESULTADOS DE LAS ENTREVISTAS

El procedimiento seguido para analizar el contenido de las entrevistas consistió en primer lugar en su transcripción; en un segundo momento se hizo la organización del análisis que incluyó la lectura continua y repetida del discurso, el análisis previo, la exploración del material y la delimitación del tratamiento por seguir; en tercer lugar, se realizó la selección representativa del discurso que determinó las unidades de registro a partir del discurso textual y contextualizado de los participantes; en cuarto lugar, se elaboró la clasificación –con base en las dimensiones de análisis previamente establecidas en las entrevistas– en matrices de trabajo donde se realizó una transcripción textual del discurso seleccionado; en quinto lugar, se procedió a la categorización que comprendió nuevas categorías surgidas del propio discurso de docentes y alumnos, además de las ya establecidas previamente en el análisis de los cuestionarios; en sexto lugar, se hizo el análisis cuantitativo y cualitativo de las categorías, y finalmente se concluyó con la interpretación del contenido a partir de la categorización (Bardin, 1977).

Resultados de las entrevistas
a los docentes

Las características de los entrevistados muestran la situación académica y laboral de los mejores maestros por área de conocimiento de los posgrados de la UABC. En su mayoría ocupan la categoría de profe-

sor o investigador titular con definitividad y se dedican de tiempo completo a las actividades académicas; su experiencia docente oscila entre siete y veintiún años; la formación pedagógica es más común en algunas áreas del conocimiento. Así, los tres entrevistados de ciencias de la ingeniería han tomado cursos de formación docente, a diferencia de los de ciencias naturales y exactas, quienes no han contado con este tipo de cursos.

Las dimensiones de análisis exploradas fueron tres: 1) criterios pedagógicos del programa, que identificó, desde el punto de vista de los docentes, los principios pedagógicos señalados por el programa, 2) criterios pedagógicos personales, que hizo referencia a los principios de enseñanza-aprendizaje que el profesor considera son los pertinentes en su práctica docente, y 3) evaluación y sugerencias al sistema de evaluación, que indagó la opinión de los docentes sobre el procedimiento de aplicación y sobre la utilización de los cuestionarios, así como los aspectos por integrarse tanto a los cuestionarios como al sistema de evaluación.

Criterios pedagógicos del programa

De quince entrevistados, diez señalaron la inexistencia de criterios pedagógicos que el programa o la institución demanda, dos hicieron referencia a dominio del contenido, y uno a las políticas educativas que orientan la formación en posgrado (investigación *versus* profesionalización); los tres docentes del área de ciencias médicas fueron los únicos en mencionar los criterios pedagógicos señalados por el programa. Esta situación, sumada a la escasa formación para la docencia que se registró, indican que los docentes de posgrado enfrentan sin orien-

tación formal su quehacer docente; lo que proporciona la institución, en el mejor de los casos, es un listado de temas.

Criterios pedagógicos personales

Se identificaron dos grandes grupos de categorías en el discurso de los docentes; por un lado, las que no son exploradas en los cuestionarios de evaluación y consideramos originales, y por el otro, las que coinciden con las establecidas en los cuestionarios de evaluación que se aplican en la universidad.

En categorías originales se identificaron el tipo de enseñanza y el papel del maestro, las cuales son una reflexión de los docentes respecto a cómo caracterizan su clase y a sí mismos. Se identificaron dos formas que orientan el quehacer docente, flexibilidad con el programa atendiendo a las demandas del grupo, diagnóstico académico del alumno y dominio pedagógico del contenido. Estas características definen el trabajo docente de quienes las practican y se reconoce que se encuentran interrelacionadas. Atendiendo el discurso de los docentes, se observa cómo una situación lleva a la otra, si “el programa depende del perfil del estudiante y del grupo” [se requiere determinar sus habilidades e intereses o, en otras palabras] “ubicar el conocimiento de los alumnos”. Asimismo, en ellos resulta claro el manejo preciso que tienen sobre las acciones pedagógicas que siguen para propiciar el aprendizaje de sus estudiantes.

En categorías establecidas en los cuestionarios de evaluación se describe desde la visión de los docentes el valor que tienen en su práctica pedagógica el dominio del contenido, la estructuración del programa, la organización de la clase, las cualidades de interacción, la evaluación del

aprendizaje y el método de trabajo; sus comentarios confirman la pertinencia de considerarlas como parte de los cuestionarios de evaluación de la docencia. Asimismo, se corroboró la importancia que tiene, en particular, para los docentes de las áreas de ciencias de la ingeniería y ciencias agrícolas, la aplicación del conocimiento como una estrategia de enseñanza.

Evaluación del sistema y sugerencias

Las categorías recurrentes: propósitos de la evaluación, tamaño del grupo, oportunidad y frecuencia de la evaluación, hablan de las principales observaciones al sistema en conjunto. Es claro que el anhelo de los docentes es que este tipo de evaluación se lleve a cabo con el único propósito de mejorar la docencia.

En cuanto a la oportunidad y frecuencia de la evaluación, en general los docentes están de acuerdo con la aplicación de una evaluación al final del semestre, no así respecto a una intermedia o inicial; aquí se denota la polémica existente respecto a definir cuándo llevar a cabo la evaluación, decisión relacionada con los propósitos de la misma. La discusión alrededor del tamaño del grupo señala uno de los problemas técnicos que afectan la confiabilidad de los resultados, en particular cuando los grupos son de muy pocos alumnos (dos o tres), todavía más en un contexto de evaluación donde el puntaje de los alumnos al profesor tiene repercusiones económicas, como es en el caso de la UABC.

Resultados de las entrevistas
a los alumnos

Las dimensiones de análisis que se examinaron fueron las cualidades pedagógicas que caracterizan a un buen docente; es

decir, las características del proceder docente necesarias en el proceso enseñanza-aprendizaje desde el punto de vista de los estudiantes, y la evaluación y sugerencias al sistema que administra los cuestionarios que identificó la opinión de los alumnos sobre el procedimiento de aplicación y utilización de los cuestionarios.

Cualidades pedagógicas

Los estudiantes describieron de manera clara las cualidades que definen a un buen docente y a las que ellos dan prioridad al contestar los cuestionarios de evaluación de la docencia; tres son los aspectos fundamentales: que el maestro posea un conocimiento amplio, especializado y actualizado sobre la temática del curso; dominio del contenido (cabe hacer notar que para trece estudiantes de catorce entrevistados es la cualidad que define al docente de posgrado); que prepare la clase y haga uso de materiales de apoyo (organización de la clase), y que inspire confianza para interactuar con él, con el propósito de hacerle preguntas en el salón de clases o en asesorías (cualidades de interacción). Además, se pronunciaron porque lleve a cabo procedimientos instruccionales para facilitar el aprendizaje (estrategias de instrucción).

Evaluación del sistema y sugerencias

Respecto al sistema que administra los cuestionarios de evaluación de la docencia, la mayoría de los estudiantes considera que éstos no tienen una función importante, en la medida en que no se presentan cambios en los maestros mal evaluados. De manera aparente, sus comentarios no afectan a los docentes; a pesar de ello, los estudiantes desean que

sus evaluaciones sean anónimas. Desde su punto de vista, este tipo de evaluaciones debería llevarse a cabo con el propósito de mejorar la docencia e identificar a los malos maestros; respecto a cuándo es el mejor momento para aplicar los cuestionarios no se presentó acuerdo, hubo quien opinó sobre la pertinencia de hacerlo en etapas tempranas del semestre y quien opinó lo contrario.

Un hallazgo relevante de la investigación es el hecho de que mediante las entrevistas se confirmaron los resultados derivados de los análisis estadísticos en relación con las coincidencias y diferencias de opinión entre docentes y estudiantes; entre ellos destaca la identificación de dominio del contenido como la dimensión de mayor importancia para ambos actores.

DISCUSIÓN

La tradición e investigación del uso de cuestionarios de opinión de estudiantes como recurso para evaluar la labor de los profesores no ha sido suficiente para lograr un consenso en las comunidades de investigadores, y en el personal académico que tiene a su cargo esta función en las universidades. Por esta razón se consideró indispensable la participación de docentes y alumnos en la construcción de una iniciativa de evaluación de la docencia. Los resultados de este estudio identificaron las dimensiones de evaluación más importantes para la enseñanza en los posgrados de la UABC; reconocieron los aspectos particulares de evaluación por área de conocimiento de los diferentes programas, lo que nos llevó a construir tres cuestionarios diferentes, y se definió una jerarquía general de las dimensiones, así como las jerarquías correspondientes a los agrupamientos de los programas. Así, la evaluación de la docencia

por los alumnos en el ámbito de posgrado que actualmente se practica en la UABC es un resultado directo de esta investigación, ya que utiliza los instrumentos generados a partir de este estudio. De esta manera, los instrumentos toman en cuenta los aspectos críticos del proceso de instrucción para cada área del conocimiento; en este sentido nuestra experiencia resultó exitosa y puede sugerirse para su empleo en otras instituciones de educación superior.

Por otro lado, los supuestos sobre las habilidades genéricas de la docencia efectiva se refuerzan con los resultados obtenidos en este estudio; sin embargo, esta postura no debe desatender el reconocimiento de la especificidad de la enseñanza disciplinaria o por lo menos de las distintas situaciones de aprendizaje en los programas de formación profesional. Es importante realizar estudios que aporten información sobre los cambios de criterio en las expectativas de los estudiantes hacia sus profesores durante su recorrido en el proceso de formación profesional.

En cuanto a las jerarquías de las dimensiones empleadas en los cuestionarios, la similitud en las opiniones expresadas por profesores y estudiantes muestra también la semejanza de criterios en cuanto a la importancia que cada actor les otorga en el proceso de enseñanza-aprendizaje; sin embargo, este hecho podría interpretarse como la necesidad de ampliar el conocimiento sobre la discusión de las funciones del profesor y de los nuevos roles que los estudiantes deben cumplir, en un sistema de educación superior que responda a las necesidades de una sociedad en plena evolución.

Aunque el proceso de construcción del instrumento para evaluar la docencia se apejó a los lineamientos sugeridos por Marsh (1987), consideramos que un ele-

mento clave que deberá agregarse a dichos lineamientos es la comunicación directa con los profesores y estudiantes que emplearán el instrumento diseñado; esto sin duda repercute en el clima favorable con el que la evaluación puede tener lugar, pero sobre todo por la oportunidad que se brinda a todos los actores del proceso al reflexionar colectivamente sobre la problemática de la docencia.

La perspectiva metodológica empleada (Bryman, 1984) permitió conocer el problema de la evaluación con una actitud de sensibilidad y respeto, al tratar de rescatar el punto de vista de los propios involucrados en el proceso de evaluación y de acercar los intereses de la administración central de la Universidad con los de los académicos y estudiantes; esto fue posible por la postura del trabajo como una acción identificada por todos los participantes como un proyecto de investigación.

Un elemento adicional que ha arrojado esta experiencia de evaluación de la docencia es que en el momento que se convoca a los profesores y estudiantes a participar en ella, se establece la oportunidad de descubrir aspectos importantes. Uno de ellos es el hecho de que la institución o las personas encargadas de la dirección académica de los programas no han contemplado la orientación precisa sobre qué es lo que se espera que los profesores cumplan en cada situación de enseñanza prevista curricularmente, referido como criterios pedagógicos del programa. Otro es la oportunidad de conocer los criterios personales que cada profesor, de manera consciente o no, emplea como recurso para orientar su práctica en el momento de la interacción con sus estudiantes; dicho conocimiento puede sentar las bases para la caracterización de los rasgos esenciales de la enseñanza de cada disciplina.

El estudio plantea la discusión del uso de los resultados de la evaluación, por una parte los académicos y estudiantes están de acuerdo con la aplicación de procedimientos de evaluación, sin embargo, unos y otros muestran sus reservas al no existir correspondencia de su participación con acciones concretas derivadas de los resultados obtenidos. Si se declara que la evaluación es realizada para mejorar la calidad de la enseñanza, un resultado desfavorable en la evaluación debería conducir hacia programas de formación permanente o mínimamente al acceso de información o ayuda especializada para que el profesor pueda solventar el problema.

Un elemento central de los resultados de esta investigación es la necesidad de encaminar los procesos de evaluación hacia la autoevaluación asistida o acompañada, para hacer frente a las situaciones de "simulación" que se propician cuando las acciones de evaluación son ajenas a los académicos o impuestas desde una iniciativa exterior. La participación de los académicos en los procesos de evaluación ya ha sido explorada por otras iniciativas (Gilio, 2000; Rueda *et al.*, 1996; Loredó, *et al.*, 1999) y con los resultados de este estudio se refuerza su factibilidad y su necesidad, si se quiere retomar el rumbo del uso de la evaluación para el mejoramiento de la actividad.

La literatura sobre evaluación señala de

manera contundente la necesidad de separar propósitos de premio o recompensa de los de mejora de la actividad; sin embargo, en el contexto mexicano, debido a la asociación de las políticas de evaluación con el salario, predomina la mezcla de estos dos propósitos, lo que deja a los procesos de evaluación en una situación poco clara.

La perspectiva de comparar los puntos de vista de profesores y estudiantes también arrojó información adicional sobre la evaluación. Por una parte los profesores, en cuanto la evaluación puede afectar sus salarios, son muy cuidadosos en eliminar todos aquellos aspectos evaluados que están fuera de su control. Por su parte, los estudiantes otorgan mayor peso en sus opiniones a las formas de evaluación empleadas por sus profesores, muy probablemente preocupados porque se atienda un aspecto clave de su formación profesional.

Finalmente, consideramos importante señalar que la evaluación de la docencia –vía cuestionarios contestados por estudiantes– tiene, como toda evaluación, límites que será necesario reconocer, sobre todo ante la complejidad de una tarea tan importante y polémica como lo es la docencia universitaria, de manera que han de ser bienvenidas todas las iniciativas por explorar otros medios en su evaluación, más aún si con ellos se contribuye a su mejoramiento.

REFERENCIAS

- ABRAMI, P. C. y d'Apollonia, S. (1990), "The dimensionality of ratings and their use in personnel decisions", en M. Theall y J. Franklin (eds.), *Student ratings of instruction: Issues for improving practice*, *New Directions for Teaching and Learning*, núm. 43, pp. 97-112, San Francisco, Jossey Bass.
- ABRAMI, P. C., d'Apollonia, S., y Cohen, P. A. (1990), "Validity of student ratings of instruction what we know and what we do not", en *Journal of Educational Psychology*, vol. 82, núm. 2, pp. 219-231.
- ASOCIACIÓN Nacional de Universidades e Instituciones de Educación Superior (1997), *Anuario estadístico. Personal de*

licenciatura y posgrado en universidades e institutos tecnológicos ANUIES, México.

- BARDIN, L. (1977), *L'analyse de contenu*, Francia, Presses Universitaires.
- BRYMAN, A. (1984), "The debate about quantitative and qualitative research: a question of method or epistemology?", en *The British Journal of Sociology*, vol. XXXV, núm. 1, 75-92.
- CENTRA, J. A. (1993), *Reflective faculty evaluation: enhancing teaching and determining faculty effectiveness*, San Francisco, Jossey-Bass.
- DELANDSHERE, G. (1994), "The assessment of teachers in the United States", en *Assessment in Education*, vol. 1, núm. 1, pp. 95-113.

- COHEN, P. A. (1981), "Student ratings of instruction and student achievement: A meta-analysis of multisection validity studies", en *Review of Educational Research*, núm. 51, núm. 281-309.
- FELMAN, K. A. (1978), "Course characteristics and college students' ratings of their teachers and courses: what we know and what we don't", en *Research in Higher Education*, núm. 9, 199-242.
- (1997), "Identifying exemplary teachers and teaching: evidence from student ratings", en R. Perry y J. Smart (eds.), *Effective teaching in higher education: Research and practice*, Nueva York, Agathon Press.
- GARCÍA, G. J. (2000), "Las dimensiones de la efectividad docente, validez y confiabilidad de los cuestionarios de evaluación de la docencia: una síntesis de la investigación internacional", en M. Rueda y F. Díaz-Barriga (eds.), *Evaluación de la docencia. Perspectivas actuales*, México, Paidós Educador.
- GILIO, C. (2000), "La evaluación docente desde una perspectiva participativa: El caso de la licenciatura en Química de la Universidad Autónoma de Querétaro", en M. Rueda y F. Díaz Barriga (eds.), *Evaluación de la docencia*, México, Paidós Educador.
- HUSBAND, C. T., y Fosh, P. (1993), "Students' evaluation of teaching in higher education: experiences from four european countries and some implications of the practice", en *Assessment and Evaluation in Higher Education*, vol. 18, núm. 2, pp. 95-114.
- JOHNSONBAUGH, R. (1998), *Matemáticas discretas*, México, Iberoamérica.
- LANDSHEERE, G. (1980), "La evaluación de los enseñantes", en M. Debesse y G. Mialaret (eds.), *La función docente*, Barcelona, Oikos-Tau.
- LOREDO, J. (1999), "Evaluación docente en la Universidad Anáhuac", en M. Rueda y M. Landesmann (coords.), *¿Hacia una nueva cultura de la evaluación de los académicos?*, México, CESU-UNAM.
- LUNA, E. (2000), "La evaluación de la docencia a partir de la opinión de los alumnos en posgrado: el caso de la Universidad Autónoma de Baja California", tesis de doctorado, Universidad Autónoma de Sinaloa, México.
- MARSH, H. (1984), "Students' evaluations of university teaching: dimensionality, reliability, validity, potential biases, and utility", en *Journal of Educational Psychology*, vol. 76, núm. 5, pp. 707-754.
- MARSH, H. W. (1987), "Students' evaluations of university teaching: Research findings, methodological issues, and directions for future research", en *International Journal of Educational Research*, núm. 11, pp. 253-388.
- MARSH, H. W., y Bailey, M. (1993), "Multidimensional students' evaluations of teaching effectiveness", en *Journal of Higher Education*, vol. 64, núm. 1, pp. 1-18.
- MARSH, H.W., y Dunkin, M. J. (1997), "Students' evaluations of university teaching: A multidimensional perspective", en R. Perry y J. Smart (eds.), en *Effective teaching in higher education: Research and practice* Nueva York, Agathon Press, pp. 241-320.
- ORY, J. C. (1990), "Student ratings of instruction: Ethics and practice", en M. Theall y J. Franklin (eds.), *Student ratings of instruction: Issues for improving practice, New Directions for Teaching and Learning*, núm. 43, Francisco, Jossey Bass.
- RUEDA, M., y Rodríguez, L. (1996), "La evaluación de la docencia en el posgrado de Psicología de la UNAM", en M. Rueda y J. Nieto (comps.), *La evaluación de la docencia universitaria*, México, UNAM.
- SHATZ, A. M. y Best, B. J. (1986), "Selection of items for course evaluation by faculty and students", en *Psychological Reports*, núm. 58, 239-242.
- SMITH, R. A. y Cranton, P. A. (1992), "Students perceptions of teaching skills and overall effectiveness across instructional settings", en *Research in Higher Education*, núm. 33, pp. 747-764.
- SPROULE, R. (2000), "Student evaluation of teaching: A methodological critique of conventional practices", en *Education Policy Analysis Archives*, vol. 8, núm. 50, pp. 1-34.
- TEJEDOR, F. J. (1990), "La evaluación del profesorado en la universidad de Santiago", en *Revista Española de Pedagogía*, vol. XLVIII, núm. 186, pp. 337- 362.
- THEALL, M., y Franklin J. (1990), "Student ratings in the context of complex evaluation systems", en M. Theall, and J. Franklin (eds.), *Student ratings of instruction: Issues for improving practice, New Directions for Teaching and Learning*, núm. 43, San Francisco, Jossey Bass.

ANEXO • Cuestionario sobre la actividad docente en posgrado (Revisado por docentes y estudiantes).

Nombre de la unidad académica _____

Nombre del posgrado _____ Nombre de la asignatura _____

Nombre del docente _____ Fecha _____

Lee cuidadosamente las siguientes instrucciones: El propósito del presente cuestionario es conocer tu punto de vista respecto al docente y su materia. Es importante que tus respuestas sean objetivas. Por tal motivo, este cuestionario es anónimo. Para contestar cada una de las preguntas utiliza una escala descendente de 4 (más favorable) a 1 (menos favorable).

Cuatro representa: SIEMPRE O MUY BIEN / Tres representa: LA MAYORÍA DE LAS VECES O BIEN / Dos representa: ALGUNAS OCASIONES O MAL / Uno representa: NUNCA O MUY MAL

Método de trabajo

1. ¿Entregó a los estudiantes el programa del curso al inicio del mismo?

Estructuración de objetivos y contenidos

2. ¿Los contenidos del curso fueron congruentes con los objetivos?
3. ¿La bibliografía utilizada fue adecuada para cubrir los temas del curso?
4. ¿Los tiempos asignados a los temas fueron adecuados?

Claridad expositiva

5. ¿Explicó con claridad cuáles son los aspectos importantes?
6. ¿Explicó con claridad los temas?

Organización de la clase

7. ¿Utilizó recursos didácticos para facilitar la comprensión de los temas?
8. ¿Las tareas y actividades desarrolladas facilitaron el aprendizaje de los temas?
9. ¿El método de enseñanza fue el adecuado para lograr los objetivos?

Dominio de la asignatura

10. ¿Presentó los fundamentos teóricos y metodológicos de los contenidos desarrollados en clase?
11. ¿Resolvió las dudas y problemas planteados en clase?
12. ¿Relacionó el contenido de la materia con otras materias o disciplinas?

Estrategias de instrucción

13. ¿Los contenidos presentados en clase estuvieron secuenciados de manera lógica?
14. ¿Tomó en cuenta los conocimientos previos de los estudiantes al inicio de cada tema?

Cualidades de interacción

15. ¿Escuchó los puntos de vista de los alumnos?
16. ¿Ayudó a los estudiantes dentro y fuera de clase?
17. ¿Motivó la participación de los estudiantes?
18. ¿Permitió que los estudiantes expresaran sus dudas y preguntas?

Evaluación del aprendizaje

19. ¿Explicó los requisitos mínimos para aprobar el curso al inicio del mismo?
20. ¿Los métodos de evaluación fueron acordes a lo cubierto en el curso?
21. ¿Ponderó de manera adecuada los diferentes aspectos de evaluación?

Método de trabajo

22. ¿La bibliografía del curso se encontró disponible?
23. ¿El instructor asistió con regularidad y puntualidad a sus clases?
24. ¿El profesor cubrió el número de horas estipuladas en el programa?

Reactivos específicos en la dimensión dominio de la asignatura:

En las áreas de ciencias naturales y exactas y de ciencias médicas:

25. ¿Se discutieron posibles problemas de investigación?

En las áreas de tecnologías y ciencias agropecuarias e ingeniería:

26. ¿Relacionó el contenido de la materia con la solución de problemas del campo profesional?