

# EVALUACIÓN DE LA DOCENCIA: PARADOJAS DE UN PROCESO INSTITUCIONAL

EDNA LUNA  
SERRANO\*  
MA. CONSUELO  
VALLE ESPINOZA\*  
GUADALUPE  
TINAJERO  
VILLAVICENCIO\*

## Resumen

La importancia de la evaluación estriba en el mejoramiento del desempeño docente en el proceso de enseñanza-aprendizaje del alumno. Se presenta el caso de la instrumentación, con sus características y problemas que enfrentó la evaluación docente en la Universidad Autónoma de Baja California con base en la opinión de sus alumnos a través de los Cuestionarios de Evaluación Docente

Palabras clave: Evaluación, docentes, alumnos, CEDA.

## Abstract

The relevance of evaluation is found in the improvement of teaching in the teaching-learning process of the student. The implementation case at the Universidad Autónoma de Baja California is presented, with the characteristics and problems encountered when teachers were evaluated based on student opinion through Teaching Evaluation Questionnaires (CEDAs).

Key words: Evaluation, teachers, students, CEDA.

.....

\* Universidad Autónoma  
de Baja California  
Correo-e:  
eluna@uabc.mx

La evaluación de la docencia universitaria se reconoce como una actividad compleja, entre otras razones, debido a que múltiples factores inciden en el proceso enseñanza-aprendizaje. Su relevancia estriba en la posibilidad de mejorar el desempeño de los docentes para la formación profesional de los estudiantes. Dos de los elementos que se aprecian como indispensables para quienes realizan la evaluación son la capacidad técnica, y la autoridad y legitimidad académica. De ahí la importancia de contar con experiencias sistemáticas de evaluación que respondan a las necesidades específicas de las universidades en México. El presente trabajo se propone analizar el proceso de instrumentación del sistema de evaluación de la docencia con base en la opinión de los alumnos llevado a cabo en la Universidad Autónoma de Baja California (UABC), con el propósito de dar cuenta de los principales logros, dificultades y retos que se enfrentan al desarrollar un sistema confiable, válido y útil para mejorar las prácticas docentes en la universidad.

---

## Antecedentes

A partir de 1990 se concretaron diversas acciones de evaluación institucional; la que atañe a los docentes es el Programa de Estímulos al Desempeño Académico. Con su aplicación se apoyó de manera directa la política de deshomologación salarial, dirigida a incrementar la remuneración del personal académico al margen del salario tabular base. Con la implantación de los programas de compensación salarial se institucionaliza un reto de siempre, la evaluación de la docencia.

En México se acepta el uso generalizado de los cuestionarios de evaluación de la docencia por los alumnos en las instituciones de educación superior (IES); sin embargo, la investigación sobre el tema se considera incipiente (García, 2000). Son pocos los trabajos que dan cuenta del proceso de construcción de los cuestionarios, lo que permitiría conocer sobre su pertinencia, y mucho menos se conocen las características del

sistema que los administra en términos de la confiabilidad de los procedimientos para recolectar información y el uso de los resultados.

Por décadas, la discusión se ha centrado en la confiabilidad y validez de los puntajes de opinión de los estudiantes como una medida de la efectividad docente; por un lado se puede decir que son confiables, estables y relativamente válidos (Marsh y Dunkin, 1997). Sin embargo, en los últimos años se han hecho investigaciones sobre los procedimientos de aplicación, interpretación y utilización de los resultados; con ello se ha demostrado que un procedimiento incorrecto puede invalidarlos (*v. gr.* Theall y Franklin, 1990). Así, aspectos básicos del tema, como la definición de los propósitos de la evaluación, las estrategias de aplicación de los cuestionarios y el uso de los resultados han cobrado relevancia en la discusión actual.

La literatura señala que el sistema que administra los cuestionarios, en principio, debe determinar y hacer explícitos los propósitos de la evaluación. Es fundamental precisar qué tipo de información se va a obtener y qué se va a hacer con ella. Al mismo tiempo, se requiere tener claridad sobre las necesidades de los docentes, los estudiantes y administrativos, en muchos de los casos contradictorias, por lo que es importante que se equilibren de manera justa y pertinente los intereses de los tres actores. Los procedimientos administrativos de acopio de la información deben incorporar los controles necesarios para asegurar que la información generada sea confiable y válida. Existen pruebas de que los puntajes se modifican si se contestan dentro o fuera del salón de clases, si el maestro permanece en el salón, si explica la importancia de la evaluación para su promoción o permanencia en la institución (Centra, 1976; Marsh, Overall y Kesler, 1979). En el procesamiento de la información, uno de los aspectos más importantes es el anonimato de los evaluadores ya que los estudiantes deben tener confianza en que los puntajes no afectarán su relación con el instructor (Arreola, Lawrence y Alemoni, 1990).

Un uso inadecuado de los resultados puede provenir de las necesidades institucionales que responden a políticas poco claras. En muchas universidades se ha promovido la práctica generalizada de efectuar evaluaciones con fines de control o compensación salarial y de retroalimentación al mismo tiempo. La literatura señala las inconveniencias de este proceder; muestra el conflicto que tienen los maestros en discutir aspectos de su quehacer docente en una situación de evaluación administrativa (Centra, 1993). La dificultad en obtener información relevante para mejorar la enseñanza con instrumentos de uso administrativo, dado que para perfeccionar la docencia se requiere de información a nivel micro del proceder docente, mientras que para la toma de decisiones administrativas a nivel macro. En este sentido, al plantear la evaluación de manera inseparable al control, se confunden funciones completamente heterogéneas que no pueden realizarse de manera simultánea con un mismo procedimiento, dado que responden a situaciones paradigmáticas y epistemológicas distintas (Ardoino, 2000).

En la implantación del sistema que administra los cuestionarios se precisa considerar una serie de elementos prácticos. La oportunidad y frecuencia de la evaluación significa acordar los momentos para llevar a cabo la aplicación de los cuestionarios; la decisión respecto a la etapa del semestre (al inicio, a la mitad, al final) en la cual aplicar el cuestionario y el número de veces depende de los propósitos de la evaluación y de las necesidades de los docentes. El tamaño del grupo es un factor determinante para la confiabilidad de los resultados, por lo tanto, es fundamental clarificar las acciones a seguir cuando los grupos tienen tamaños muy variables; los casos extremos son los grupos menores a diez estudiantes, porque el coeficiente de confiabilidad resulta muy por abajo de lo considerado útil (Centra, 1973).

La confidencialidad de los puntajes es abordada a la luz de los intereses de los docentes y de

los estudiantes. A los docentes, el sistema debe asegurar información detallada y disponible sólo para el profesor preocupado por el diagnóstico y retroalimentación de su desempeño. Los alumnos deben sentir plena confianza en que su relación con el profesor no se verá afectada por las opiniones que ellos emiten. En este orden de ideas, la confianza en el uso adecuado de los resultados es una característica fundamental del sistema (Theall y Franklin, 1990).

---

### Contexto inmediato: el caso de la UABC

La Universidad Autónoma de Baja California (UABC), en la actualidad ofrece 65 programas de licenciatura y 49 de posgrado (tres de doctorado, 26 de maestría y 20 de especialidad) distribuidos en las diferentes áreas del conocimiento. La evaluación de la enseñanza en licenciatura, a partir de la opinión de los estudiantes, inició de manera sistemática en 1988 con el propósito de identificar las áreas del quehacer docente que requerían preparación pedagógica; con la información recabada se pretendía orientar los programas de formación docente. La instancia responsable del diseño del cuestionario de evaluación de la docencia por los alumnos y del programa para procesar la información fue la Dirección General de Asuntos Académicos (DGAA); la aplicación y análisis de los cuestionarios estuvo a cargo de las unidades académicas. En esta etapa se delegó en cada escuela o facultad la responsabilidad de analizar los puntajes; asimismo, el uso de los resultados se dejó a discreción de los directivos de las unidades académicas. Se esperaba que los directivos de cada unidad académica reportaran a la DGAA los aspectos con menor puntuación sobre los cuales requerían que se ofertaran cursos de formación docente. Sin embargo, gran parte del trabajo en los primeros años estuvo dirigido a introducir en la cultura de evaluación a directivos, académicos y estudiantes; en este sentido,

a convencer a los involucrados de los beneficios de la evaluación<sup>1</sup>.

Un cambio que desvió la intención original de la evaluación de la docencia en licenciatura fue la introducción, a partir de 1990, de los programas de compensación salarial en la UABC, donde se estipuló como uno de los indicadores de evaluación el puntaje de los estudiantes; no obstante, su incorporación no fue posible debido a la insuficiencia de información. Los puntajes otorgados por los estudiantes se integraron como uno de los rubros que componen el programa de estímulo en 1994, y sólo para aquellos docentes que contaban con sus reportes de evaluación semestral.

A partir de ese año se computarizó la captura de los puntajes y se delegó en la DGAA la responsabilidad de procesar la información y elaborar un reporte semestral de cada maestro. A partir de esta fecha, los reportes se destinan por un lado, a cada unidad académica para conocimiento de los docentes y directivos, y por otro, a la Secretaría General para ser integrados en el expediente del programa de estímulos de cada profesor<sup>2</sup>. De esta manera, las unidades académicas continúan con la responsabilidad de aplicar los cuestionarios y capturar la información, y una dependencia central tomó el control del manejo de la información para asegurar, de alguna manera, que la institución cuente con los reportes necesarios que demanda el programa de estímulos al desempeño académico.

La aplicación de los cuestionarios se automatizó a partir del año 2000, como una medida para mejorar la administración del sistema; la DGAA gira instrucciones a las unidades académicas para que los estudiantes de licenciatura acudan en un periodo establecido (hacia el final del semestre) a contestar el cuestionario a un centro de cómputo. Si bien el sistema computarizado ofrece una

mayor confianza tanto a los docentes como a los estudiantes en el manejo de la información, también impuso el reto de lograr que los usuarios asistan en el periodo estipulado a contestar el cuestionario; hasta la fecha, el lapso de tiempo se ha ampliado de un mes a tres meses, como una medida para lograr un mayor número de respuesta por parte de los estudiantes. Un dato que ilustra el nivel de participación es el primer semestre del 2002, en el cual 17,035 alumnos (73% de la población de licenciatura) evaluaron a sus profesores.

Este trabajo parte de la convicción de la importancia de hacer partícipes a los actores principales de la evaluación de la docencia, los docentes, en el proceso de construcción de los cuestionarios para su evaluación. La metodología seguida replica la estrategia desarrollada para el diseño de cuestionarios para posgrado (Luna y Rueda, 2001).

---

### Propósito

El propósito de investigación fue la construcción de instrumentos de evaluación docente para el nivel de licenciatura que contemplen: a) las recomendaciones hechas por los docentes; b) la especificidad de cada área de conocimiento. Así como, determinar la jerarquía de las dimensiones de evaluación con base en la opinión de los docentes.

---

### Metodología

El diseño del trabajo corresponde a un estudio de caso y se abordó desde la perspectiva de los métodos combinados la cual sostiene que, tanto la investigación cualitativa como cuantitativa, se refuerzan si se trabajan como complementarias

---

<sup>1</sup> Reconstruir parte de la historia del sistema de evaluación en licenciatura en la UABC fue posible gracias a la información proporcionada por Julio César Reyes y Lourdes Loza, ex-jefes de departamento de la misma dirección, quienes coordinaron en los primeros años la implantación del sistema.

<sup>2</sup> Información proporcionada por la Lic. Aída León, jefe de departamento de la DGAA.

(Brymann, 1984). Se desarrolló en dos etapas: en la primera se elaboró un cuestionario base que se propuso a la consideración de los profesores, propiciando su participación; en la segunda, se reconstruyó el cuestionario al incorporar las principales sugerencias de los docentes y se determinó la jerarquía de las dimensiones de evaluación para la muestra en general y por áreas de conocimiento.

El estudio se llevó a cabo con una muestra representativa por bloques aleatorios de la población de docentes de las licenciaturas de la UABC. Para ubicar el universo objeto de estudio

se identificó a los maestros de carrera que tienen horas asignadas frente a grupo, descartando a los profesores que gozan de alguna licencia, año sabático, formación docente; para los profesores de asignatura, el criterio de selección fue que impartieran un mínimo de 10 horas clase. La información corresponde al primer periodo semestral del año 2001. En la muestra se distinguieron los siguientes factores: área de conocimiento y tipo de nombramiento. Se encuestó a 263 docentes, 18 % de la población total, 147 de carrera (22 %) y 116 de asignatura (15 %); la Tabla 1 muestra la distribución de las submuestras.

**Tabla 1**  
**Población de docentes de carrera y de asignatura,**  
**así como número y porcentaje de encuestados**

Programas por área de conocimiento	Total de maestros		Encuestados		Total
	Carrera	Asignatura	Carrera	Asignatura	
<b>Ciencias Agropecuarias</b>	<b>70</b>	<b>8</b>	<b>10= 14%</b>	<b>2= 25%</b>	<b>12= 15%</b>
agronomía	38	2	5= 13%	1= 50%	6= 15%
veterinaria y zootecnia	32	6	5= 16%	1= 16%	6= 16%
<b>Ingeniería y Tecnología</b>	<b>155</b>	<b>248</b>	<b>5 7 =</b>	<b>3 7 %</b>	<b>99= 25%</b>
arquitectura	29	22	<b>42=17%</b>		11=22%
ciencias computacionales	13	2	7= 24%	4= 18%	7= 47%
ingeniería	113	224	5= 38%	2= 100%	81= 24%
			45= 29%	36= 16%	
<b>Ciencias Naturales y Exactas</b>	<b>91</b>	<b>16</b>			<b>19= 18%</b>
biología, física y matemáticas	23	7	<b>16= 18%</b>	<b>3= 19%</b>	4= 13%
ciencias marinas	51	3	4= 17%	0= 0%	13= 23%
ciencias químicas	14	6	10= 19%	3= 100%	2= 10%
			2= 14%	0= 0%	
<b>Ciencias de la Salud</b>	<b>97</b>	<b>151</b>			<b>47= 19%</b>
medicina	38	83	<b>23= 24%</b>	<b>24= 16%</b>	49= 40%
odontología	59	68	38= 26%	11= 13%	22= 17%
			13=22%	9=13%	
<b>Ciencias Sociales y Administrativas</b>	<b>265</b>	<b>354</b>			<b>86= 14%</b>
ciencias políticas	20	12	<b>41= 15%</b>	<b>54= 13%</b>	5= 16%
contabilidad y administración	103	203	3= 15%	2= 17%	47= 15%
comunicación, sociología y psicología	55	62	16= 16%	31= 15%	17= 15%
derecho	40	55	8= 15%	9= 15%	15= 16%
economía	30	15	6= 15%	9= 16%	7= 16%
turismo	17		5= 17%	2= 13%	4= 17%
	7		3= 18%	1= 14%	
<b>Totales</b>	<b>678</b>	<b>777</b>	<b>147= 22%</b>	<b>116= 15%</b>	<b>263= 18%</b>

En la primera etapa se diseñó el cuestionario “Evaluación de la docencia: teoría” de acuerdo con las dimensiones y reactivos que tanto la literatura anglosajona como hispana marcan como relevantes (Landshere, 1980; Shatz y Best, 1986; Tejedor, 1990; Centra, 1993; Feldman, 1997; Luna y Rueda, 2001; Rueda, 2001). Las dimensiones que incluye son: a) estructuración de objetivos y contenidos; b) claridad en la instrucción; c) organización de la clase; d) dominio de la asignatura; e) estrategias de instrucción; f) cualidades de interacción; g) evaluación del aprendizaje y h) aspectos formales del cumplimiento docente. Para determinar su contenido se consideraron las sugerencias hechas por la Dirección General de Asuntos Académicos (DGAA), dada su experiencia en este tipo de evaluación. Asimismo, se consultó a los estudiantes mediante un estudio piloto para evaluar la claridad del cuestionario.

La segunda etapa consistió en proporcionar a los docentes el cuestionario diseñado para que señalaran los comentarios y sugerencias a cada reactivo y dimensión de evaluación y de esta manera dar respuesta a las preguntas de investigación: 1) cuáles aspectos incluir en el cuestionario y 2) cuáles se deben excluir. Asimismo, se incluyó un cuadro de ponderación de las dimensiones y se les solicitó que, con base en su experiencia, asignaran el nivel de importancia a cada una de las dimensiones que cubre la actividad docente. El cuestionario se reconstruyó incorporando las sugerencias de mayor frecuencia de los docentes. Por último, se puso a disposición de los profesores el cuestionario reconstruido para su valoración.

El cuestionario se aplicó en dos modalidades; primero, en línea, dando un mes a partir de la fecha de envío para recuperar los cuestionarios; en las áreas donde no se cumplió con el número especificado en las submuestras se procedió a hacer la aplicación de manera presencial. En ambas modalidades el procedimiento general fue: a) solicitar la revisión de cada reactivo y escribir al calce su aceptación o rechazo así como sus

comentarios y sugerencias, especificando los aspectos no evaluados que consideren importante incluir, y b) contestar el formato diseñado para asignar un nivel jerárquico a las dimensiones que componen el cuestionario.

---

## Resultados

El análisis cualitativo de las respuestas a los cuestionarios se basó en técnicas de análisis de contenido, la codificación se derivó de las cláusulas textuales que responden a las preguntas de investigación (qué incluir y qué excluir del cuestionario de evaluación de la docencia con base en la opinión de los alumnos), de acuerdo con las siguientes reglas: 1) exhaustividad, todos los elementos relevantes de la muestra deben quedar incluidos en las categorías y 2) exclusividad mutua, el contenido no puede ser ubicado en más de un lugar; y 3) generalidad y especificidad, las categorías generales se relacionaron con las dimensiones y las específicas con los reactivos.

La información se reunió en una base de datos con el fin de clasificarla de acuerdo con el área de conocimiento del profesor, categoría administrativa (de carrera o asignatura) y tipo de respuesta, de aceptación o rechazo, para cada reactivo y dimensión. Se elaboró una tabla por reactivo y dimensión que reúne las cláusulas textuales a partir de las cuales se derivaron las categorías de sugerencias, y se computó su frecuencia.

La mayor frecuencia de las respuestas de los docentes se concentraron en aceptar los reactivos y las dimensiones, ya sea sin comentarios o explicitando las razones de su acuerdo. El análisis que a continuación se describe se basa en los comentarios de aceptación condicional y rechazo a los reactivos. Las categorías de sugerencias que se identificaron fueron: 1) cuestionar el lenguaje del reactivo, agrupa todos los comentarios que discuten los términos utilizados, y en algunas ocasiones señalan de manera precisa las palabras a modificar, así como las razones; 2) cuestionar



el contenido implícito del reactivo, reúne las sugerencias de eliminar los términos que asumen una actitud pasiva de parte de los estudiantes; y 3) cuestionar la evaluación, incluye los comentarios que descalifican a los alumnos como evaluadores, así como los que desaprueban la inclusión de aspectos que a juicio de los profesores no dependen de ellos.

En la categoría “cuestionar el lenguaje del reactivo”, los docentes discutieron la pertinencia de incluir expresiones con significados precisos, que no se presten a confusión, como en los siguientes casos: “¿los contenidos del curso fueron congruentes con los objetivos?”; “¿utilizó recursos didácticos para facilitar la comprensión de los temas?”; “¿presentó los fundamentos teóricos y metodológicos de los contenidos desarrollados en clase?”; se solicitó aclarar el significado de congruencia y recursos didácticos, así como dar ejemplos. En “¿la bibliografía utilizada fue adecuada para cubrir los temas del curso?”, se sugirió sustituir bibliografía por fuentes de información para incluir a las nuevas tecnologías. Estos comentarios provienen de los docentes de los programas de las diferentes áreas del conocimiento estudiadas, con una frecuencia que oscila por área entre 2 y 10.

En general, en la categoría “cuestionar el contenido implícito del reactivo”, los profesores señalaron cuidar no incluir vocablos que promuevan un comportamiento pasivo de parte de los estudiantes. Incluye a los reactivos: “¿explicó con claridad los temas?”; “¿resolvió las dudas y problemas planteados en clase?”; y “¿ayudó a los estudiantes dentro de clase?”; las argumentaciones según el caso fueron: no siempre el profesor es quien expone el tema y no promover una pedagogía centrada en la enseñanza, en la transmisión del conocimiento. Asimismo puntualizaron que ayudar no es resolver, es orientar al estudiante para que él construya las soluciones. Las sugerencias proceden de los maestros de los diversos programas estudiados, con una frecuencia que oscila entre 2 y 7 por área.

En la categoría “cuestionar la evaluación”, los docentes por un lado, desacreditaron a los estudiantes como evaluadores en los reactivos: “¿los tiempos asignados a los temas fueron adecuados?” donde argumentaron que el alumno no conoce lo extenso de cada tema ni lo que es suficiente para los propósitos de un curso, y en “¿los contenidos presentados en clase estuvieron secuenciados de manera lógica?” señalaron sus dudas en relación con la preparación del alumno para evaluar ese aspecto; estas opiniones proceden de los maestros de las diferentes áreas con un rango de frecuencia que fluctúa entre 3 y 13. Por otro lado, los maestros discutieron los reactivos que evalúan elementos que a su juicio no atañen a la materia, tal es el caso de: “¿tomó en cuenta los conocimientos previos de los estudiantes al inicio de cada tema” donde los argumentos fueron que los huecos en el conocimiento previo no son responsabilidad directa del docente y no dependen de la asignatura, dado que cada profesor debe cumplir un programa, este comentario lo realizaron los profesores de todas las áreas con un rango de frecuencias de 2 a 15 comentarios.

El cuestionario base se modificó de acuerdo con los siguientes criterios: a) se revisaron los reactivos con frecuencia de sugerencias entre dos y cuatro, sin necesariamente realizar cambios en los mismos; y b) se atendieron las observaciones con frecuencia de cinco o más en función del área de procedencia. De esta manera, se realizaron las siguientes modificaciones al cuestionario original: se modificó la redacción de algunos reactivos con el propósito de hacerlos más precisos; se eliminaron los términos que de manera implícita promueven un comportamiento pasivo de parte de los estudiantes en el proceso de aprendizaje y se sustituyeron con el fin de propiciar un papel activo en la regulación de su instrucción; asimismo, se suprimió el reactivo que valora el tiempo dedicado a cada tema por considerarlo que no procedía su evaluación por parte de los alumnos; por

último, se incluyeron aspectos relativos al uso de las nuevas tecnologías. Cabe mencionar que en este estudio no se identificaron sugerencias particulares por área de conocimiento que nos

lleven a postular valoraciones específicas. El cuestionario modificado de acuerdo con las sugerencias de los docentes se presenta a continuación.

### Cuestionario de evaluación de la docencia: teoría

El (la) maestro (a):

1. ¿Entregó a los estudiantes el programa del curso al inicio del mismo? Si ( ) No( )
2. ¿Asistió con regularidad a sus clases? Si ( ) No( )
3. ¿Explicó los requisitos para aprobar el curso al inicio del mismo? Si ( ) No( )

Escala de calificación para el siguiente conjunto de preguntas

1. Nunca 2. Algunas veces 3. Casi siempre 4. Siempre

4. ¿Los contenidos del curso apoyaron el logro de los objetivos?
5. ¿Las fuentes de información (por ejemplo: bibliografía, sitios en red) fueron adecuadas para cubrir los temas del curso?
6. ¿Explicó con claridad cuáles son los aspectos importantes del curso?
7. ¿Los temas fueron tratados de manera clara?
8. ¿Los materiales de apoyo (por ejemplo: diapositivas, diagramas, programas) utilizados para facilitar la comprensión de los temas fueron adecuados?
9. ¿Las tareas y actividades desarrolladas facilitaron el aprendizaje de los temas?
10. ¿Se trataron los principios teóricos o metodológicos de los contenidos desarrollados en clase?
11. ¿Ayudó a resolver las dudas y problemas planteados en clase?
12. ¿Relacionó el contenido de la materia con otras materias?
13. ¿Mostró dominio de la materia?
14. ¿Los contenidos trabajados en clase estuvieron secuenciados de manera lógica?
15. ¿Tomó en cuenta los conocimientos previos de los estudiantes al inicio de cada tema?
16. ¿Utilizó ejemplos para definir algún concepto o principio general?
17. ¿Indicó algunas aplicaciones de los conocimientos para el trabajo profesional?
18. ¿Escuchó los puntos de vista de los alumnos?
19. ¿Motivó a los alumnos para ser responsables de su proceso de aprendizaje?
20. ¿La evaluación fue acorde con los contenidos trabajados en el curso?

Con base en una escala en la que 1 es la calificación más baja y 10 la más alta, califique:

21. Evaluación global del desempeño del profesor (a)	1	2	3	4	5	6	7	8	9	10
22. Evaluación de mi desempeño como estudiante en esta materia	1	2	3	4	5	6	7	8	9	10

El análisis cuantitativo se realizó con el propósito de determinar la jerarquía de las dimensiones de evaluación del trabajo académico y las tendencias de opinión de los docentes de los diversos programas agrupados por área de conocimiento. Los cálculos se realizaron mediante técnicas de análisis de conglomerados. La información se derivó del cuadro de ponderación y el procedimiento seguido consistió en: a) elaborar matrices de frecuencia de orden 8 jerarquías por 8 dimen-

siones por área de conocimiento y para la muestra total, con el propósito de determinar el número de opiniones en la jerarquía de cada dimensión; b) normalizar las matrices de frecuencia al dividir los componentes de cada matriz entre el número de individuos encuestados. Los resultados permitieron identificar la jerarquía de las dimensiones por área y para la muestra al reconocer el número máximo de votación de cada dimensión respecto a las jerarquías; de esta manera, se identificó a



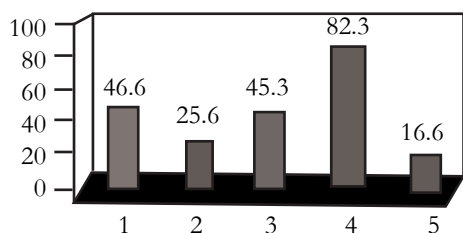
dominio del contenido como la dimensión más importante y como las de menor importancia a estrategias de instrucción, cualidades de interacción y evaluación del aprendizaje tanto en la jerarquización general como en las particulares por área (Tabla 2); c) construir matrices de diferencia al tomar el valor absoluto de los componentes de la matriz de frecuencia normalizada por área de conocimiento menos los componentes de la matriz de frecuencia normalizada de la opinión

general; la suma de todos los componentes de estas matrices de diferencia permitió determinar la similitud de opinión de los profesores de una área respecto a la muestra total (Gráfica); de esta manera se identificaron las tendencias de opinión entre los docentes de los programas por áreas de conocimiento, encontrando que los profesores de Ciencias Sociales y Administrativas son los que más difieren con relación a las opiniones del total de la muestra.

**Tabla 2**  
**Jerarquización de las dimensiones de evaluación para la muestra general y por área de conocimiento**

J	General	Agropecuarias	Ingeniería	Salud	Naturales	Sociales
1°	Dominio	Dominio	Dominio	Dominio	Dominio	Dominio
2°	Estructuración	Método	Organización	Estructuración	Estructuración	Estructuración
3°	Método	Organización	Método	Organización	Organización	Método
4°	Organización	Claridad	Estructuración	Método	Método	Claridad
5°	Claridad	Estructuración	Claridad	Claridad	Claridad	Organización
6°	Estrategias Inst.	Estrategias	Estrategias	Estrategias	Estrategias	Estrategias
7°	Cualidades	Cualidades	Cualidades	Cualidades	Cualidades	Cualidades
8°	Evaluación	Evaluación	Evaluación	Evaluación	Evaluación	Evaluación

**Gráfica**  
**Diferencias de opinión de los profesores por área de conocimiento respecto al total de la muestra**



- 1 Ciencias Agropecuarias
- 2 Ciencias de la Salud
- 3 Ciencias Naturales y Exactas
- 4 Ciencias Sociales y Administrativas
- 5 Ciencias de la Ingeniería y Tecnología

## Discusión y conclusiones

Al analizar la historia de las prácticas de evaluación de la docencia por los alumnos en la UABC es posible dar cuenta de los logros, las dificultades, así como de los retos que se tienen al implantar un sistema de evaluación. La reconstrucción del proceso de implantación del sistema de evaluación de la docencia por los alumnos en la UABC hace evidente que introducir en la cultura de la evaluación a una comunidad académica es un trabajo a largo plazo, sujeto a factores de orden político y económico. La evaluación de la docencia en licenciatura inició con el propósito de mejorar la enseñanza; sin embargo, posteriormente, se introdujo como un indicador importante en los

programas de compensación salarial en porcentaje, uno de los efectos positivos de este hecho es el haber fomentado la sistematización del proceso de evaluación. Más de 17,000 estudiantes (73% de la población de la población total de la UABC) evaluaron a sus maestros en el primer periodo del semestre 2002, dado que se convirtió en una demanda institucional el solicitar reportes regulares sobre los resultados de la evaluación. Esto hizo necesario hacer eficiente el proceso de captura y procesamiento de la información, lo cual se ha logrado a través de la automatización del sistema. Por otro lado, la asociación de la evaluación de la docencia por los alumnos a los programas de estímulo académico ha ocasionado una gran discrepancia entre el discurso de las autoridades y lo que pasa en la práctica. En el discurso se maneja un doble propósito, por un lado, el de retroalimentar al docente para mejorar la enseñanza, y por otro, el de control administrativo vía el programa de estímulos; sin embargo, en los hechos, el principal uso de los resultados es el de control, dado que no se cuenta con los mecanismos de retroalimentación, ni con las estrategias de ayuda que posibiliten a los docentes transformar su quehacer de enseñanza a partir de la información que proporcionan los puntajes. Cabe tener presente el hecho documentado de que la mejora en la enseñanza es mínima cuando se reduce la retroalimentación a sólo entregar los resultados de los puntajes. En este sentido, uno de los grandes desafíos de la institución es el de instrumentar las condiciones necesarias para el uso de los resultados con el propósito de mejorar las actividades de docencia. Para ello, es necesario conceptualizar la evaluación de la docencia como una actividad compleja que involucra: aspectos instrumentales relacionados con los métodos y técnicas de evaluación, componentes políticos que inducen un conjunto de acciones, así como estrategias de análisis de los resultados de la evaluación. En este sentido, analizar los hechos de la evaluación como prácticas sociales, que no se reducen a los elementos puramente técnicos o sociales, sino que representan la posibilidad de un

proceso de transformación del cual forman parte agentes sociales concretos que mantienen entre sí relaciones específicas, cuyo producto debería de orientar el mejoramiento de la calidad de la enseñanza.

En relación con el sistema que administra los cuestionarios, la literatura recomienda resolver situaciones elementales relacionadas con la confiabilidad y validez de los resultados, entre ellas destacan: tomar en consideración el tamaño del grupo, los puntajes obtenidos en grupos pequeños (menos de diez) son poco confiables; si se van a realizar comparaciones, estas deben realizarse entre los profesores que comparten características similares, agruparlos por disciplina o materia, ya que se sabe que los estudiantes otorgan puntajes diferenciales en función del área de conocimiento. En este sentido, uno de los retos de la institución es que el sistema cuente con los mecanismos necesarios para asegurar resultados equitativos, pertinentes y útiles para los actores principales de los procesos de enseñanza y aprendizaje, los docentes y los estudiantes.

Los maestros postularon propiciar un papel más activo y responsable por parte de los estudiantes en su proceso de formación, en este sentido, la teoría implícita en el cuestionario se orientó hacia la aproximación constructivista del aprendizaje; asimismo, se incorporó el uso de las nuevas tecnologías, lo cual puede fomentar una pedagogía innovadora. Por último, se suprimieron los reactivos que, a juicio de los docentes, los estudiantes no están en condiciones de evaluar por falta de información. La jerarquización general así como todas las jerarquías por área de conocimiento ubicaron como la dimensión de evaluación más importante a dominio del contenido, lo cual concuerda con lo obtenido en estudios anteriores con los docentes de posgrado (Luna y Rueda, 2001).

Desde la perspectiva de los profesores, el principal uso de los cuestionarios debería ser para el mejoramiento de las actividades de docencia; sin embargo, para ellos es muy clara la función de control que estos tienen en la actualidad, y san-

cionaron los componentes del cuestionario que pudieran estar fuera de su control o ser difíciles de evaluar por los alumnos. En este sentido, per-

siste la necesidad de contar con procedimientos e instrumentos de evaluación independientes para cada propósito de evaluación.

## Referencias

- ARDOINO, J. (2000). "Consideraciones teóricas sobre la evaluación en educación", en M. Rueda, y F. Díaz (Eds.), *Evaluación de la docencia. Perspectivas actuales*, México, Paidós Educador.
- ARREOLA, R. A., y Aleamoni, L. M. (1990). "Practical decisions in developing and operating a faculty evaluation system", en M. Theall, J. Franklin (Eds.), *Student ratings of instruction: Issues for improving practice*, No 43, New Directions for Teaching and Learning, San Francisco, Jossey Bass.
- BRYMAN, A. (1984). "The debate about quantitative and qualitative research: a question of method or epistemology?", *The British Journal of Sociology*, XXXV, 1.
- CENTRA, J. A. (1973). *Item reliability, the factor structure rating, comparison with alumni ratings*. Student Instructional (Report, no. 3). Princeton, N.J., Educational Testing Service.
- CENTRA, J. A. (1976). "The influence of different directions on student ratings of Instruction", *Journal of Educational Measurement*, 13.
- CENTRA, J. A. (1993). *Reflective faculty evaluation: enhancing teaching and determining faculty effectiveness*, San Francisco, The Jossey-Bass Higher and Adult Education Series.
- FELDMAN, K. A. (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.
- LANDSHEERE, G. (1980). "La evaluación de los enseñantes", en M. Debesse y G. Mialaret (Eds.), *La función docente*, Barcelona, Oikos-tau.
- LUNA S., E., y Rueda B. M. (2001). "Participación de académicos y estudiantes en la evaluación de la docencia", *Perfiles Educativos*, (23) 93.
- MARSH, H. W., et al. (1979). "Multidimensionality of students' evaluation of teaching effectiveness", *American Educational Research Journal*, 16.
- MARSH, H.W., y Dunkin, M. J. (1997). "Student's evaluations of university teaching: A multidimensional perspective", en R. Perry y J. Smart (Eds.), *Effective teaching in higher education: Research and practice*, Nueva York, Agathon Press.
- RUEDA B., M. (2001). "Una propuesta de cuestionario dirigido a los docentes para evaluar la función docente en la universidad", en M. Rueda y F. Díaz-Barriga (Eds.), *Evaluación de la docencia. Perspectivas actuales*, México, Paidós Educador.
- SHATZ, A. M., y Best, B. J. (1986). "Selection of items for course evaluation by faculty and students", *Psychological Reports*, (58).
- TEJEDOR, F. J. (1990). "La evaluación del profesorado en la universidad de Santiago", *Revista Española de Pedagogía*, XLVIII, 186.

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*, No. 43, New Directions for Teaching and Learning, San Francisco, Jossey Bass.