

Profesores universitarios y su efectividad docente. Un estudio comparativo entre México y Estados Unidos *

JOSÉ MARÍA GARCÍA GARDUÑO **



El propósito de este estudio es explorar, mediante cuestionarios de evaluación de la docencia respondidos por alumnos en una institución mexicana de educación superior, si la investigación sobre efectividad docente arroja resultados similares a los estudios realizados en Estados Unidos.

Las variables consideradas fueron: edad, género, años de experiencia, antigüedad, categoría académica, tipo de contrato, grado académico, principal área de actividad y área de conocimiento donde impartían clases los docentes. La población fue de 1 391 profesores de licenciatura y posgrado de una universidad privada de la ciudad de México; 365 de ellos tenían contrato de tiempo completo y 1 026, de asignatura.

Los datos se obtuvieron de los resultados globales del cuestionario de evaluación de la docencia por los alumnos que la institución empleaba para evaluar semestralmente a sus profesores. Incluye los resultados de la evaluación de seis semestres y 8 745 grupos-materia. Los datos se analizaron por medio de una serie de ANOVA unifactoriales; se utilizó la prueba de Scheffe, como medida de comparación *post hoc*. El estudio concluye que los resultados en las variables mencionadas fueron semejantes a los hallazgos producidos por la investigación estadounidense.

The purpose of the present study was to explore whether the research findings on student ratings in a Mexican higher education institution were similar to those yielded by international research, chiefly that carried out in the United States. Variables considered in the current study were related to the characteristics of the instructor and the discipline: age, gender, years of teaching experience, type of contract (full time vs. part-time) academic rank, academic degree, professor's main area of activity (teaching vs. research, and type of discipline taught (hard vs. soft sciences). The entire population of instructors of a private Mexican university participated which comprised: student ratings of 1 361 instructors, 365 full-time and 1026 part-time who taught 8 745 course-groups during 6 semesters, were analyzed by a series of one-way ANOVAS. Results suggest that the findings of the current study are similar to the findings yielded by American research. Hence, March's (1987) presumption regarding that the findings of research on American student ratings might be generalized to other countries and cultures is supported by this study.

Efectividad docente / Cuestionarios de evaluación / Profesores universitarios / México / Estados Unidos
Teaching efficiency / Evaluation questionnaire / University teachers / Mexico / United States

INTRODUCCIÓN

El objeto central de este estudio es la docencia y la efectividad docente. La docencia es un acto complejo; parte de su complejidad radica en el hecho de que no sólo es una ciencia o técnica, sino también es un arte. Para Gage (1985) la docencia es un arte con bases científicas. Existen tantas definiciones de ella como enfoques. De acuerdo con Fenstermacher (1986, p. 38), la manera más general de definirla es como una relación que se da entre una persona (*P*) que posee algún contenido (*C*), la cual intenta impartirlo a otra persona (*R*), quien inicialmente carece de *C*, de tal manera que *P* y *R* se involucran en un relación con el propósito de que *R* adquiriera *C*.

Por otro lado, en este estudio se ha adoptado una definición de efectividad docente congruente con el enfoque de investigación, el cual se inscribe dentro de lo que se ha denominado investigación proceso-producto. Según Abrami, d' Apollonia y Rosenfeld (1997, p. 326), en términos de este enfoque, la efectividad docente se define como las actividades del docente que ocurren antes y durante el acto de la enseñanza, las cuales producen cambios en los alumnos en dominios relevantes, incluyendo el cognitivo, el afectivo y, ocasionalmente, el psicomotor. Las limitaciones de este enfoque de inves-

tigación son que se centra en los procesos y productos empírico-cuantitativos de la docencia y deja de lado su significado y comprensión más amplia.

Otra limitación de este trabajo es que la efectividad docente sólo fue abordada tomando en cuenta la opinión de los alumnos vertida en los cuestionarios de evaluación de la docencia (CEDA) y se dejaron de lado otros criterios para estimarla. Estos cuestionarios tienen sesgos o limitaciones inherentes a su uso y aplicación (García, 2000). Sin embargo, la mayoría de los estudios señala que son instrumentos válidos y confiables para medir ciertos aspectos de la docencia. Por ejemplo, algunos señalan que los criterios para juzgar la efectividad docente que emplean los profesores coinciden en gran medida con los criterios de los alumnos (Feldman, 1988).

La investigación sobre efectividad docente en la educación superior empezó en 1927 en la Universidad de Purdue con Herman Remmers, el creador de los cuestionarios de evaluación de la docencia (CEDA) (Centra, 1993). Sin embargo, el auge de la investigación comenzó a partir de la década de los setenta y alcanzó su zenit en la de los ochenta (Greenwald, 1997). En México, la primera institución que empleó los CEDA fue la Universidad Iberoamericana (UIA), a finales de los sesenta y su primera publicación sobre el tema salió a la luz en 1972 (Universidad Iberoamericana, 1972). Durante los años setenta, la Universidad Nacional Autónoma de México (UNAM) realizó investigaciones ocasionales sobre el tema, y posiblemente una de las primeras investigaciones publicadas sobre los CEDA fue la de Arias Galicia (1984), que se refirió a la validez de un cuestionario para medir la eficacia docente. Sin embargo,

* El presente trabajo es resultado de la investigación efectuada en el Departamento de Educación de la Universidad Iberoamericana. El autor desea agradecer a Alejandro Márquez la revisión y las valiosas sugerencias que hizo al manuscrito.

** Profesor en el Departamento de Comunicación del Instituto Tecnológico y de Estudios Superiores de Monterrey (ITESM), campus Ciudad de México.
josemariagarduno@yahoo.com.mx

tomando como indicador los estudios publicados en México, no hubo continuidad sobre el trabajo de este autor durante esa década.

A partir de la segunda mitad de los años noventa la investigación empezó a tener una presencia más notable en publicaciones y foros. Parte de ese interés estuvo determinado por las nuevas políticas de evaluación del trabajo académico promovidas por el Programa para la Modernización Educativa del Gobierno Federal (Poder Ejecutivo Federal, 1989) y por el liderazgo de académicos de la UNAM, quienes han aglutinado los esfuerzos de investigadores dedicados a la evaluación de la docencia de varias instituciones universitarias públicas y privadas. Los productos se han plasmado en diversos coloquios nacionales e internacionales y en varias publicaciones (por ejemplo, Rueda y Nieto, 1996; Rueda y Landesman, 1999; Rueda y Díaz-Barriga, 2000; Rueda, Díaz-Barriga y Díaz, 2001), en las cuales se han reportado investigaciones y analizado temas tales como: diseño y validación de instrumentos para evaluar la docencia, evaluación académica vía los programas de compensación salarial, dimensiones de la evaluación, la evaluación formativa de la docencia, estrategias para evaluar, evaluación y formación docente, entre otros.

A pesar del avance logrado en los últimos años en la investigación empírica sobre la evaluación de la docencia universitaria, la cantidad de los estudios realizados es aún mínima. En el medio internacional, la mayor parte de la investigación sobre la evaluación de los resultados del desempeño del profesor en los cursos que imparte se ha efectuado por medio del empleo de los resultados de los CEDA. Marsh (1984) señala que uno de los pro-

pósitos de los CEDA, además de diagnosticar y retroalimentar a los profesores sobre su desempeño, medir la efectividad docente y tomar decisiones sobre el otorgamiento de definitividad en el empleo, es investigar sobre los resultados y procesos docentes.

En el medio anglosajón, principalmente en Estados Unidos, se ha investigado y probado profusamente la validez, confiabilidad y limitaciones de este tipo de instrumentos. Asimismo, ha sido uno de los temas más investigados en la educación superior. Feldman (1990) calcula que se han publicado alrededor de 2 000 artículos y libros sobre ese tópico. Desde sus inicios y hasta la década de los setenta, la mayor parte de la investigación estuvo dedicada a estudiar la validez y la confiabilidad de los instrumentos; posteriormente se encaminó a investigar la relación del desempeño o efectividad docente con otras variables; por ejemplo, con las características del profesor, las de los alumnos y del curso, entre otras.

A pesar de que México, al igual que otros países, importó los cuestionarios de evaluación de Estados Unidos y que su empleo es cada vez más frecuente en las instituciones de educación superior, no se ha estudiado cuán semejantes son los resultados de la investigación estadounidense sobre la relación entre efectividad docente y algunas variables que tienen que ver con las características de los profesores y la disciplina que enseñan. El propósito de este estudio fue explorar si algunas variables relacionadas con la efectividad docente, medida a través de los resultados de los CEDA, en una institución mexicana, arroja resultados similares a los estudios realizados en Estados Unidos. Se considera en particular la estabilidad de los puntajes a lo largo del tiempo y en va-

riables que se relacionan con las características de los o las docentes, y si existen diferencias en la efectividad docente de acuerdo con las áreas o disciplinas de conocimiento. El criterio de selección de este grupo de variables obedeció a que éstas han sido investigadas ampliamente en Estados Unidos y que existió información en la institución mexicana donde se llevó a cabo el estudio.

La pertinencia e importancia de este estudio no sólo radica en realizar una comparación transcultural, entre algunos determinantes de la efectividad docente, que hasta ahora no se había hecho en el medio nacional ni internacional, y que se trata de la investigación longitudinal más grande llevada a cabo en un país hispanoamericano, sino también porque los resultados de la comparación podrían contribuir a iluminar el estado del arte en dos ámbitos. El primero tiene que ver con la hipótesis de que existen similitudes transculturales en los determinantes de la efectividad docente. En el ámbito internacional existen pocos estudios que hayan investigado este tema. Marsh (1987) presume que es posible generalizar los resultados de la investigación estadounidense a otros países o culturas de educación superior diferentes de la anglosajona. La presunción de Marsh ha tenido cierta confirmación. Ting (2000) realizó una investigación con 11 190 estudiantes de sociología de la Universidad de Hong Kong. El autor concluye que los resultados en los cuestionarios de evaluación para juzgar la calidad de la docencia son semejantes a los encontrados en países occidentales. Sin embargo, señala que las tradiciones culturales influyen en la evaluación de las experiencias de aprendizaje de los estudiantes chinos. A partir de los escasos estudios realizados, la hipótesis de

este trabajo es que los resultados de esta investigación pueden ser similares a los producidos en Estados Unidos.

El segundo ámbito es de carácter práctico; Estados Unidos ha sido el mayor productor de investigación sobre esta temática. A lo largo de más de siete décadas, en particular durante las tres últimas, han estudiado virtualmente todos los aspectos derivados del uso y de los resultados de los cuestionarios de evaluación. Las posibles similitudes en los determinantes de la efectividad docente entre México y Estados Unidos podrían permitir un empleo más confiable en México, y tal vez en otros países iberoamericanos, de los hallazgos producidos por la investigación realizada internacionalmente, y en particular en Estados Unidos.

Síntesis de los resultados
de la investigación producida
en Estados Unidos sobre las variables incluidas
en este estudio

Durante las décadas de los setenta y los ochenta, se publicaron en los países anglosajones, principalmente en Estados Unidos, cientos de estudios que probaron la confiabilidad y validez de los CEDA para medir la efectividad docente en el salón de clase (Feldman, 1989; Overall y Marsh, 1980; Marsh, 1987). En este estudio sólo se analizaron los resultados de las investigaciones que tienen que ver directamente con las variables que se exploraron, a saber: edad de los profesores, género, experiencia docente, antigüedad, categoría académica, tipo de contrato (tiempo o asignatura), grado académico, área de conocimiento o disciplina en que los profesores imparten clases, y principal área de actividad del profesor (docencia o investigación).

Las investigaciones de Marsh y colaboradores (Overall y Marsh, 1980; Marsh, 1984;) y de Feldman (1977) señalan que los promedios de los puntajes obtenidos en los CEDA presentan gran estabilidad a lo largo del tiempo. En un estudio realizado durante un periodo de 13 años, Marsh y Hocevar (1991) encontraron que casi no hubo cambios en los alumnos y exalumnos que evaluaron los cursos. Por otro lado, estos estudios indican que los CEDA, más que evaluar el curso, califican al profesor.

El meta-análisis de Feldman (1983) y la síntesis de Marsh (1987) indican que no existen diferencias significativas entre edad, experiencia y categoría académica de los instructores, y la efectividad o desempeño docente. Pocos de los estudios reseñados por los autores citados indican que existe una correlación negativa entre esas variables y la efectividad docente, y otros señalan que existe una correlación positiva. En relación con la edad y experiencia docente, Feldman (1983) halló que en algunos estudios se presenta una relación con la forma de una J invertida, esto es, a menor edad y experiencia, mayor efectividad docente. Sin embargo, no existen patrones consistentes entre estas variables; la investigación no ha sido capaz de explicar las causas de estas inconsistencias. Sin embargo, cuando se analizan las dimensiones de efectividad docente incluidas en los CEDA, hay diferencias en el desempeño de acuerdo con la edad, experiencia y categoría académica; los profesores más jóvenes *versus* los de más experiencia, edad y categoría académica superior, son evaluados de manera diferente.

El mismo patrón se observa en relación con el género. Los meta-análisis de Feldman (1992; 1993) y otros estudios (Goodwin y Stevens, 1993) señalan que, en gene-

ral, no existen diferencias entre el desempeño de los docentes del sexo masculino y los del femenino. Sin embargo, como en las variables señaladas anteriormente, no hay patrones consistentes; algunos estudios muestran que las profesoras son mejor evaluadas (Feldman, 1993), y otros más indican que los alumnos tienden a evaluar mejor a los profesores del mismo sexo (Kierstad, D'Agostino y Dill, 1988). Por otro lado, Freeman (1994) encontró que los alumnos evalúan mejor a los docentes que poseen tanto características femeninas como masculinas en su estilo docente.

Desde hace tiempo, tanto en la literatura especializada como en los pasillos y salones universitarios, ha existido un debate acerca de si los profesores-investigadores son mejores o peores docentes que aquellos que se dedican exclusivamente a la docencia. Feldman (1987) realizó un meta-análisis de 43 investigaciones publicadas a lo largo de 33 años; el autor encontró que existe, en promedio, un coeficiente de correlación pequeño (.13), pero significativo entre productividad y efectividad docente. En un meta-análisis más reciente de 58 investigaciones, Hattie y Marsh (1996) encontraron que la correlación entre investigación y docencia es sólo de .06. Estos investigadores también concluyen que en los estudios más recientes se ha encontrado una correlación más pequeña entre estas dos variables.

El tipo de disciplina juega un papel importante en los puntajes o promedio de evaluación obtenidos por los profesores. Las investigaciones realizadas han demostrado consistentemente que los cursos relacionados con las ciencias y las ingenierías obtienen promedios más bajos en los CEDA que los de las humanidades (evaluadas con los más altos promedios), y que las ciencias sociales, las cuales

tienen promedios intermedios entre las ciencias y las humanidades (Feldman, 1978; Neumann y Neumann, 1983; Barnes y Barnes, 1993). Cashin (1990) señala que mientras más cuantitativa sea la orientación del curso, existe la probabilidad de que el promedio de la evaluación grupal sea menor.

Sobre los grados académicos de los profesores y el tipo de contratación (asignatura o por horas, o tiempo completo) sólo se encontró un estudio realizado. Es probable que no existan muchas investigaciones porque la mayoría de los profesores de las universidades estadounidenses ostentan el grado de doctor y son de tiempo completo (con excepción de los colegios de dos años en Estados Unidos). En una muestra de 42 profesores de un colegio de dos años, Pierce (1986) encontró que no existe relación entre efectividad docente y los grados académicos de los profesores, ni entre el tipo de dedicación (de tiempo completo o asignatura) y efectividad docente; sin embargo, una de las limitaciones del estudio es que sólo clasificó a los profesores en dos categorías: *bachelor* (grado equivalente a la licenciatura) y maestría o grado superior; es decir, no diferenció a los profesores con grado de doctorado de los de maestría. Por otro lado, el tamaño de la muestra fue pequeño. Lo más parecido a la investigación sobre esta última variable en las universidades estadounidenses es comparar la efectividad docente de los ayudantes o asistentes docentes (*teaching assistants*) con los profesores de mayor rango y grado académico. Como en los casos anteriores, la investigación no es concluyente. La síntesis de los estudios que realizó Marsh (1987) y el meta-análisis de Feldman (1983) indican que los docentes ayudantes, quienes regularmente son estu-

diantes de posgrado de la propia institución, tienden a ser evaluados con puntuaciones más bajas que los profesores titulares. Sin embargo, los estudios de Marsh (1987) encontraron que los asistentes son mejor evaluados que los profesores titulares en cuanto a la calidad de la interacción que sostienen con los alumnos y que los profesores obtienen puntajes más altos en el grado de conocimiento y cobertura del curso.

MÉTODO

Diseño de la investigación

Se empleó un diseño correlacional-comparativo *ex post facto*.

Variables

Las variables utilizadas fueron las siguientes:

1. Estabilidad de los CEDA a lo largo del tiempo.
2. Edad de los profesores.
3. Género (femenino o masculino).
4. Años de experiencia docente.
5. Antigüedad en la institución.
6. Categoría académica (ayudante, adjunto, asociado, titular y numerario).¹
7. Tipo de contrato (tiempo o asignatura). Los profesores de tiempo son los que trabajan de 20 a 40 horas semanales para la institución y tienen una oficina permanente dentro de la institución. Los profesores de asignatura imparten clases por horas, un máximo de 19 horas por semana, y usualmente no tienen oficina dentro de la institución.
8. Máximo grado académico (licenciatura, maestría o doctorado).²
9. Área de conocimiento. Los profesores se agruparon de acuerdo con el área de

conocimiento en la que impartían clases, bajo la agrupación de las disciplinas que tenía la institución:

- a) *División de Ciencias del Hombre (ciencias sociales)*. Incluye los departamentos de psicología, comunicación, ciencias sociales y políticas, derecho y educación, y desarrollo humano.
- b) *División de Humanidades*. Comprende los departamentos de historia, filosofía, literatura latinoamericana y ciencias teológicas.
- c) *División de Ciencias e Ingeniería*. Reúne los departamentos, carreras de licenciatura y programas de posgrado de matemáticas, ingeniería mecánica y eléctrica, ingeniería química, ingeniería civil, ingeniería física, ingeniería en sistemas y nutrición, y ciencias de los alimentos.
- d) *División de Ciencias Económico-Administrativas*, agrupa los departamentos, carreras de licenciatura y programas de posgrado de economía, administración de empresas y contaduría pública.
- e) *División de Arte*. Incluye los departamentos, carreras y programas de posgrado de arquitectura y urbanismo, historia del arte, diseño industrial y gráfico.

10. Principal área de dedicación del profesor o profesora (docencia o investigación). Se consideró que el o la profesora se dedicaba principalmente a la investigación si tenía registrado y había recibido financiamiento interno o externo para realizar un proyecto de investigación o estaba inscrito en el Sistema Nacional de Investigadores (SNI). Los profesores dedicados a la docencia no pertenecían al SNI ni habían registrado algún proyecto de investigación.

11. Efectividad docente. Se midió a través del promedio global que los profesores obtuvieron en el CEDA a lo largo de los seis semestres analizados; a mayor promedio en los CEDA, mayor efectividad docente.

Población

Participaron 1 391 profesores de licenciatura y posgrado de una universidad privada de la ciudad de México de aproximadamente 10 900 estudiantes (10 000 de licenciatura y 900 de posgrado), los cuales constituyen el universo de la población de la que se pudo obtener información sobre las variables del estudio. Estos profesores impartían clases en 32 licenciaturas, 21 programas de maestría, cinco de doctorado y dos especializaciones durante uno o hasta seis periodos semestrales de primavera y otoño, de 1993 a 1995. La institución objeto de estudio fue la primera en México en emplear los CEDA, a finales de la década de los sesenta.

El estudio abarca los resultados de la evaluación de seis semestres, los cuales corresponden a 8 745 grupos-materia impartidos entre 1993 y 1995; el tamaño promedio del grupo en esta institución es de aproximadamente 16 alumnos, lo que significa que el estudio incluyó la respuesta de más de 139 000 cuestionarios de evaluación de la docencia. Un total de 2 283 cursos fueron impartidos por profesores de tiempo y 6 462 por profesores de asignatura. Los criterios para seleccionar la muestra fueron: a) que en los archivos de la institución existiera el registro de por lo menos un curso del profesor evaluado por los alumnos; b) que en los registros de la dirección de personal de la institución estuviera disponible información sobre las características de los profes-

sos a las que aluden las variables seleccionadas para el estudio.

Las principales características demográficas de la población fueron las siguientes: 479 del sexo femenino (34.4%) y 912 del masculino (65.6%); el rango de edad fue de 23 a 86 años, con un promedio de 45.06 años (DE = 10.86) de los profesores de tiempo y 42.28 años (DE = 10.82) de los profesores de asignatura; 365 de ellos tenían contrato de planta o de tiempo (298 de tiempo completo y 67 de medio tiempo) y 1 026 de asignatura; el rango de experiencia docente fue de uno a 43 años, con un promedio de 14.04 años (DE = 8.88).

Instrumentos y fuentes de información

La efectividad docente se obtuvo de los resultados del cuestionario de evaluación de la docencia por los alumnos denominado EMA-4 (cuatro, porque fue la cuarta versión del instrumento). Este instrumento está conformado por 15 reactivos tipo Likert dedicados a evaluar el desempeño del profesor en las siguientes dimensiones: 1) conocimiento de la materia, 2) objetivos y contenidos del curso, 3) interacción con los alumnos, 4) organización de las actividades de aprendizaje y 5) evaluación. En la década de los ochenta la institución donde se llevó a cabo la investigación realizó un estudio preliminar de la confiabilidad del instrumento y encontró un Alfa de Cronbach de 0.86, de lo cual puede desprenderse que el instrumento tiene una confiabilidad adecuada. Asimismo, los resultados también sugieren que es un instrumento confiable, ya que, en general, hubo una alta estabilidad en los promedios globales a través de los semestres analizados.

La información sobre la edad, experiencia, grado y categoría académica de los profesores fue proporcionada por la dirección de recursos humanos de la institución.

Procedimiento

Los datos se analizaron por medio de una serie de ANOVA unifactorial. Se eligió este tipo de prueba debido a que la mayoría de los estudios realizados en Estados Unidos ha empleado estadística inferencial como ANOVA, prueba *t* y *z*. Se utilizó la prueba de Scheffe, como medida de comparación *post hoc* del análisis de varianza en el nivel de significancia de .05%. Se empleó el ANOVA en las comparaciones de dos grupos, ya que el ANOVA unifactorial y la prueba *t* pueden ser empleados indistintamente al comparar dos medias (Shavelson, 1988). La efectividad docente se calculó considerando el promedio global del total de los cursos impartidos por el profesor o profesora; el promedio de cursos analizados por profesor fue de 6.29. La edad, experiencia y antigüedad de los docentes se agruparon en cuartiles. La categoría académica de los profesores de asignatura se hizo equivalente a la que la institución emplea para clasificar a los de tiempo completo.

El análisis se limita al conjunto de los profesores de la institución agrupados por área de conocimiento o división. Una limitación es que no se analizó el desempeño de los profesores por departamento ni por programa académico; tampoco se realizó un análisis por separado de los profesores de medio tiempo y completo ni de los cursos de licenciatura y posgrado. Sin embargo, los cursos de posgrado representaron menos de 10% del total de los analizados y los profesores de medio tiempo constitu-

yeron menos de 20% del total de la población de los de tiempo.

Resultados

En el cuadro 1 se observa que el instrumento presenta estabilidad a lo largo de los seis semestres. A partir del semestre otoño-93 los promedios son consistentes y muestran escasa variabilidad. No obstante, se puede apreciar que el promedio global de primavera-93 varió, de manera abrupta, en los semestres posteriores; ello probablemente se debió a que la institución modificó uno de los reactivos originales del instrumento a partir del semestre de otoño-1993.

El cuadro 2 resume las comparaciones de la variables del estudio. En general, se observa que los resultados de las comparaciones de las variables son semejantes a los hallazgos de los estudios internacionales. Se confirma que la edad, experiencia y antigüedad en la institución no están relacionadas con la efectividad docente.

Por otro lado, se encontraron diferencias significativas en el género; globalmente, las mujeres docentes son mejor evaluadas que los hombres. Sin embargo, la diferencia en el desempeño entre ambos grupos es sólo de cinco centésimas. Asimismo, se encontraron diferencias significativas en la categoría académica; de acuerdo con el criterio de Scheffe, las diferencias están determinadas por los profesores numerarios, los de más alta categoría académica dentro de la institución, y los profesores adjuntos, pero, como se puede observar, no existe un patrón consistente en los promedios globales de acuerdo con la categoría académica de los profesores; los ayudantes obtuvieron promedios más altos en los CEDA que los adjuntos y asociados.

En el mismo cuadro se puede observar que existen diferencias significativas en la efectividad docente según el tipo de contrato; los profesores de tiempo se desempeñan mejor que sus pares de asignatura. Asimismo, existen diferencias en el desempeño de acuerdo con el grado académico de los docentes; el criterio de Scheffe indica que los docentes con grado o estudios de maestría y doctorado son mejor evaluados por sus alumnos que sus pares con licenciatura. En el cuadro 2 también se observa que existen diferencias significativas en el desempeño docente según el área de conocimiento o disciplina. La prueba de Scheffe señala que las diferencias son atribuibles a que el conjunto de profesores de las divisiones de humanidades y ciencias del hombre obtienen promedios significativamente más altos en efectividad docente que sus pares del resto de las divisiones.

Por otro lado, hay diferencias entre el desempeño o efectividad docente de acuerdo con la principal área de actividad de los profesores; los investigadores dedicados a la docencia se desempeñan mejor que los profesores dedicados exclusivamente a ella.

Discusión

El propósito de este estudio fue indagar la presunción de Marsh (1987) en lo relativo a la posible generalización de los hallazgos de la investigación estadounidense sobre los CEDA a otras culturas de educación superior. Específicamente se buscó averiguar si la investigación en una institución mexicana de educación superior sobre algunas variables relacionadas con las características del profesor y la disciplina, arroja resultados similares a los estudios realizados en Estados Unidos.

Los hallazgos encontrados en el medio internacional son similares a los resultados de este estudio en las siguientes variables: 1) los CEDA presentaron estabilidad a lo largo del tiempo; 2) el desempeño docente no está relacionado o no presenta patrones consistentes con la edad, la experiencia, ni con la antigüedad de los o las docentes.

Existieron diferencias significativas de género en favor de las docentes. Sin embargo, a pesar de esas diferencias encontradas, sería difícil afirmar que éstas son concluyentes, debido a que la asociación entre estas variables no es muy fuerte, si se toma en consideración el tamaño del valor F y los altos grados de libertad que tiene la muestra. Por lo tanto, puede afirmarse que los resultados son semejantes a los de las investigaciones estadounidenses, las cuales no han encontrado una relación consistente entre género y efectividad docente.

Asimismo, las investigaciones estadounidenses no han encontrado patrones consistentes entre categoría académica y desempeño docente; la mayoría de los estudios señalan que no existe relación entre estas dos variables. A pesar de que en la presente investigación se encontraron diferencias que indican que los profesores de la mayor categoría (numerarios) obtienen promedios significativamente más altos en efectividad que los adjuntos, no existe un patrón que permita inferir que la categoría académica es una variable que determina la efectividad docente. Los docentes con categoría de ayudante obtuvieron un promedio global similar al de los adjuntos y semejante al de los asociados.

Otra variable que no ofreció resultados definitivos fue la de principal área de dedicación del profesor (investigación *vs.* docencia). Los hallazgos de la investiga-

ción estadounidense indican que existe una diferencia pequeña, pero significativa, entre estas dos variables; los investigadores se desempeñan ligeramente mejor que los profesores dedicados sólo a la docencia. Este mismo patrón se encontró en esta investigación: la diferencia en efectividad docente entre estos dos grupos tiene un valor F pequeño pero significativo, por lo que es posible que exista una relación débil entre estas variables debido a los altos grados de libertad de la muestra y al tamaño del valor F .

De acuerdo con el valor F y su probabilidad, los resultados más significativos se dan al contrastar a los docentes según el tipo de contrato, grado académico y área o disciplina de conocimiento. Las diferencias significativas entre las disciplinas reportadas por los estudios internacionales fueron confirmadas en esta investigación. Como lo señalan los hallazgos de Cashin (1990), las humanidades obtuvieron los puntajes más altos y la ciencias sociales (división de ciencias del hombre en este estudio) obtuvieron puntajes intermedios. En el caso particular de esta investigación, según el criterio de Scheffe, las diferencias fueron determinadas por las divisiones de humanidades y ciencias sociales (ciencias del hombre). Fue sorprendente que la división de arte obtuviera promedios globales similares a los de las ciencias e ingeniería; ello puede explicarse por el hecho de que los estudiantes de las licenciaturas en Diseño Gráfico, Industrial y Arquitectura que pertenecen a esta división obtienen promedios de aprovechamiento más bajos que en otras divisiones, lo que podría indicar que son sujetos de una evaluación más rigurosa por parte de sus profesores.

Las variables por tipo de contrato y grado académico de los profesores, las

cuales han sido escasamente investigadas en Estados Unidos y sólo en colegios de dos años (*community colleges*), muestran diferencias significativas y claras, de acuerdo con el tamaño del valor F , en la efectividad docente en favor de los profesores de tiempo completo *vs.* los de asignatura. Esto podría indicar que, en efecto, la calidad de la docencia en la educación superior tiene que ver con la proporción de profesores de tiempo completo. A pesar de la experiencia profesional y conocimientos que los docentes de asignatura llevan al ámbito universitario, el cuerpo de profesores de tiempo, dedicados con mayor consistencia al cultivo de su disciplina, constituye un grupo con mejor desempeño docente, de acuerdo con el juicio de los alumnos.

Este estudio confirmó la hipótesis de que los profesores con grado de doctorado se desempeñan mejor en la docencia, conforme al juicio de los alumnos, pero las diferencias sólo se dan –según Scheffe– entre los que tienen nivel de licenciatura y doctorado; no existen diferencias entre los o las profesoras con grado de doctor y los de maestro. Cabe recordar que en este estudio no se diferenció el estatus académico de los profesores en términos de si eran graduados o candidatos al grado (pasantes) debido a que en un análisis preliminar no se encontraron diferencias significativas entre estas variables. Por consiguiente, es probable que el requisito de la obtención

del grado, para efectos de la docencia, sea un indicador de la constancia y disciplina de los académicos, pero no necesariamente de su eficacia docente. Sin embargo, los resultados deben tomarse con cautela, pues es necesario realizar estudios más específicos que comparen grados académicos y áreas de conocimiento o carreras en las que los profesores imparten docencia.

A pesar de que este estudio abarcó toda la población docente de la que se tenía información en la institución, no es posible afirmar que sus hallazgos puedan generalizarse a otras instituciones de educación superior del país. Es recomendable que se replique en otros ámbitos universitarios, públicos y privados, y que se hagan investigaciones más finas que comparen los resultados de los CEDA con áreas de conocimiento, carreras y niveles de estudios de los alumnos (posgrado y licenciatura). Sin embargo, si se toma en consideración que la investigación de Ting (2000), realizada en Hong Kong, encontró resultados semejantes a los estadounidenses, es de presumirse que la investigación estadounidense sobre los CEDA pueda generalizarse al ámbito de la educación superior mexicana. A pesar de las limitaciones de la investigación proceso-producto y del método cuantitativo que se empleó en este estudio, se obtuvieron hallazgos iniciales que se podrían profundizar por medio de enfoques cualitativos en el estudio de la docencia.

| CUADRO 1 • Promedios globales en los CEDA del conjunto de la población de docentes de tiempo y asignatura de la institución. | | | | | | | |
|---|-------|-------|-------|-------|-------|-------|---------|
| Periodo | P93 | O93 | P94 | O94 | P95 | O95 | Global* |
| N | 778 | 850 | 910 | 921 | 871 | 1074 | 1304 |
| Media | 3.914 | 4.029 | 4.092 | 4.118 | 4.088 | 4.097 | 4.068 |
| Desviación estándar | 0.438 | 0.434 | 0.395 | 0.362 | 0.396 | 0.416 | 0.342 |

P = Primavera / O = Otoño

* La N global no corresponde con el total de la población, ya que no todos los docentes impartieron clases en todos los semestres incluidos en el estudio.

| CUADRO 2 • Resultados globales de los ANOVA en las variables del estudio. | | | | | |
|--|--|---|---------|--------------|--|
| Variable | Grupos (n por grupo) | Efectividad docente (media global en el CEDA) | Valor F | Probabilidad | Grupos que determinan las diferencias significativas (Scheffe) |
| | | 4.072 | | | |
| Edad (años) N = 1303 | 1) 23-33 (298) 2) 34-40 (359) 3) 40-48 (356) 4) 40-48 (290) | 4.041 4.082 4.078 | 1.03 | 0.38 | - |
| Género N = 1304 | 1) Femenino (422) 2) Masculino (882) | 4.102 4.052 | 6.06 | 0.014* | - |
| Años de experiencia docente N = 1300 | 1) 1-3 (320) 2) 4-7 (368) 3) 8-13 (296) 4) 14 o más (316) | 4.076 4.058 4.051 4.083 | 0.80 | 0.50 | - |
| Antigüedad en la institución (años) N = 1301 | 1) 0-3 (334) 2) 4-6 (313) 3) 7-12 (335) 4) 13 o más (319) | 4.078 4.052 4.066 4.073 | 0.36 | 0.782 | - |

| Variable | Grupos (n por grupo) | Efectividad docente (media global en el CEDA) | Valor F | Probabilidad | Grupos que determinan las dife- rencias significativas (Scheffe) |
|--------------------------------------|---|--|---------|--------------|--|
| Categoría académica N = 1211 | 1) Ayudante (362) 2) Adjunto (235) 3) Asociado (302) 4) Titular (209) 5) Numerario (103) | 4.059 4.021 4.055 4.109 4.152 | 3.97 | 0.003** | Numerario (5) y adjunto (2) |
| Tipo de contrato N = 1304 | 1) Tiempo (314) 2) Asignatura (990) | 4.138 4.046 | 17.32 | 0.000*** | |
| Grado académico N = 1302 | 1) Licenciatura (838) 2) Maestría (320) 3) Doctorado (144) | 4.116 4.347 4.4353 | 12.24 | 0.000*** | Licenciatura (1) con grupos 2 y 3 |
| Área de conocimiento/ división | 1) Ciencias del hombre (366) 2) Humanidades (95) 3) Ciencias e ingeniería (343) 4) Ciencias económico- administrativas (222) 5) Arte (194) | 4.146 4.242 3.990 4.010 3.997 | 19.64 | 0.000*** | 1) Ciencias del hombre (1) con gru- pos: 3, 4 y 5 |
| Principal área de dedicación | 1) Docencia (1248) 2) Investigación (56) | 4.064 4.165 | 3.84 | 0.05* | 2) Humanidades (2) con grupos: 3, 4 y 5 |

*p <.05, **p< .01, ***p< .001

REFERENCIAS

- ABRAMI, P., S. d'Apollonia y S. Rosenfeld (1997), "The dimensionality of student ratings of instruction: What we know and what we do not know", en R. Perry y J. Smart (eds.), *Effective teaching in higher education: Research and practice*, Nueva York, Agathon Press, pp. 321-367
- ARIAS Galicia, F. (1984), "El inventario de comportamiento docentes (ICD): Un instrumento para evaluar la calidad de la enseñanza", en *Perfiles Educativos*, núm. 4, pp. 14-22.
- BARNES, L. L. y M. W. Barnes (1993), "Academic discipline

- and generalizability of student evaluations of instruction", en *Research in Higher Education*, núm. 34, pp. 135-149.
- CASHIN Cashin, W. E. (1990), "Students do rate different academic fields differently", en M. Theall y J. Franklin (eds.), *Student ratings of instruction, New Directions for Teaching and Learning*, núm. 43, San Francisco, Jossey-Bass, pp. 113-121.
- CENTRA, J. A. (1993), *Reflective faculty evaluation*, San Francisco, Jossey-Bass.
- FELDMAN, K. A. (1977), "Consistency and variability among college students in rating their teachers and courses: A re-

- view analysis", en *Research in Higher Education*, núm. 6, pp. 223-274.
- (1978), "Course characteristics among college students' Fe in ratings of their teachers and courses: What we know and what we don't?", en *Research in Higher Education* núm. 9, pp. 199-242.
- (1983), "Seniority and experience of college teachers as related to evaluations they receive from students", en *Research in Higher Education*, núm. 18, pp. 3-123.
- (1987), "Research productivity and scholarly accomplishments of college teachers as related to their instructional effectiveness", en *Research in Higher Education*, núm. 26, pp. 227-298.
- (1988), "Effective teaching from the students' and faculty's view: Matched or mismatched priorities?", en *Research in Higher Education*, núm. 28, pp. 291-344.
- (1989), "Instruccional effectiveness of college teachers as judged by teacher themselves, current and former students, colleagues, administrators, and external (neutral) observers", en *Research in Higher Education*, núm. 30, pp. 137-194.
- (1990), "An afterword for 'the association between student ratings of specific instructional dimensions and student achievement': Refining and extending the synthesis of data from multisection validity students", en *Research in Higher Education*, núm. 31, pp. 315-317.
- (1992), "College students' views of male and female college teachers: Part I-evidence from the social laboratory experiments", en *Research in Higher Education*, núm. 33, pp. 317-375.
- (1993), "College students' views of male and female teachers: Part II-evidence from students' evaluations of their classroom teachers", en *Research in Higher Education*, núm. 34, pp. 151-211.
- FENSTERMACHER, Gary D. (1986), "Philosophy of research on teaching: Three aspects", en M. Wittrock (ed.), *Handbook of research on teaching*, 3a ed., Nueva York, Macmillan, pp. 37-49.
- FREEMAN, H. R. (1994), "Student evaluations of college instructors: Effects of type of course taught, instructor gender role, and student gender", en *Journal of Educational Psychology*, núm. 86, pp. 627-630.
- GAGE, N. L. (1985), *Hard gains in the soft sciences. The case of pedagogy*, Bloomington, Indiana, Phi Delta Kappa.
- GARCÍA, José María (2000), "¿Qué factores extraclase o sesgos afectan la evaluación de la educación superior?", en *Revista Mexicana de Investigación Educativa*, núm. 5, pp. 303-325.
- GREENWALD, A. G. (1997), "Validity concerns and usefulness of student ratings of instruction", en *American Psychologist*, núm. 52, pp. 1182-1186.
- GOODWIN, L. D. y E. A. Stevens (1993), "The influence of gender on university faculty members' perceptions of 'good teaching'", en *Journal of Higher Education*, núm. 64, pp. 166-185.
- HATTIE, J. y H. W. Marsh (1996), "The relationship between research and teaching: A meta-analysis", en *Review of Educational Research*, núm. 66, pp. 507-542.
- KIERSTEAD, D., D'Agostino y H. Dill (1988), "Sex role stereotyping of college professors: Bias in students' ratings of instructors", en *Journal of Educational Psychology*, núm. 80, pp. 342-344.
- MARSH, H. W. (1984), "Students' evaluations of university teaching: dimensionality, reliability, validity, potential biases, and utility", en *Journal of Educational Psychology*, núm. 76, pp. 707-754.
- (1987), "Students' evaluations of university teaching: Research findings, methodological issues, and directions for future research" [número completo], en *International Journal of Educational Research*, vol. 11, núm. 3.
- MARSH, H. W. y D. Hocevar (1991), "Students' evaluation of teaching effectiveness: The stability of mean ratings of the same teacher over a 13-year period", en *Teaching & Teacher Education*, núm. 7, pp. 303-314.
- NEUMANN, Y. y L. Neumann (1983), "Characteristics of academic areas and students' evaluation of instruction", en *Research in Higher Education*, núm. 19, pp. 323-334.
- OVERALL, J. U. y H. W. Marsh (1980), "Students' evaluations of instruction: A longitudinal study of their stability", en *Journal of Educational Psychology*, núm. 72, pp. 321-325.
- PIERCE, S. T. (1986), "A comparative analysis of part-time versus full-time community college faculty effectiveness", tesis doctoral, North Carolina State University, resumen de: *Dissertation Abstracts International*, núm. 47, p. 374-A.
- PODER Ejecutivo Federal (1989), "Programa para la modernización educativa", México, Poder Ejecutivo Federal.
- RUEDA, M. y F. Díaz-Barriga (2000), *Evaluación de la docencia*, México, Paidós.
- RUEDA, M., F. Díaz-Barriga y M. Díaz (2001), *Evaluar para comprender y mejorar la docencia en la educación superior*, México, Universidad Autónoma Metropolitana/ Universidad Nacional Autónoma de México/ Universidad Autónoma Benito Juárez de Oaxaca.
- RUEDA, M. y M. Landesman (coords.) (1999), *¿Hacia una nueva cultura de evaluación de los académicos?*, México, CESU-UNAM, *Pensamiento Universitario* núm. 88.
- RUEDA, M. y J. Nieto (coords.) (1996), *Evaluación de la docencia universitaria*, México, Facultad de Psicología, UNAM.
- UNIVERSIDAD Iberoamericana (1972), "Evaluación del personal docente de la Universidad Iberoamericana", en *Didac*, núm. 8.
- SHAVELSON, R. J. (1988), *Statistical reasoning for the behavioral sciences*, 2a. ed., Boston, Allyn and Bacon.
- TING, Kwok-fai (2000), "A multilevel perspective on student ratings of instruction: Lessons from the Chinese experience", en *Research in Higher Education*, núm. 41, pp. 637-653.
- WATKINS, D. (1994), "Student evaluations of university teaching: A cross-cultural perspective", en *Research in Higher Education*, núm. 35, pp. 251-266.

NOTAS

1. Ésta es la máxima categoría académica que otorga la institución y podría ser equivalente al Titular C de otras instituciones de educación superior mexicanas o al *full professor* de las universidades estadounidenses.
2. Debido a que los análisis preliminares no revelaron diferencias en la efectividad docente entre pasantes y graduados, este grupo incluye a todos los profesores de ambos grupos (pasantes y graduados).