



Para citar este artículo, le recomendamos el siguiente formato:

Cisneros-Cohernour, E. (2008). Una propuesta interpretativa para la evaluación de la enseñanza en educación superior. *Revista Electrónica de Investigación Educativa, Especial*. Consultado el día de mes de año, en: <http://redie.uabc.mx/NumEsp1/contenido-cisneros.html>

Revista Electrónica de Investigación Educativa

Número Especial, 2008

Una propuesta interpretativa para la evaluación de la enseñanza en educación superior

An Interpretive Proposal for Evaluating Teaching in Higher Education

Edith J. Cisneros-Cohernour
cchacon@tunku.uady.mx

Facultad de Educación
Universidad Autónoma de Yucatán

Calle 50 No. 56 entre 29 y 33
Colonia Francisco de Montejo, 97200
Mérida, Yucatán, México

(Recibido: 10 de junio de 2008; aceptado para su publicación: 11 de septiembre de 2008)

Resumen

Este artículo examina las limitaciones del paradigma positivista tradicional utilizado en la evaluación de la enseñanza en educación superior y propone el uso de una perspectiva interpretivista como alternativa para evaluarla. Se utiliza el estudio de caso de un profesor no experimentado, para ilustrar las debilidades del paradigma tradicional e identificar cuestiones críticas que surgen cuando la evaluación se basa en opiniones de los estudiantes, como única fuente para determinar la calidad de la enseñanza.

Palabras clave: Evaluación del docente, estudio de caso, enseñanza superior.

Abstract

This article examines the limitations of the traditional positivist paradigm in the assessment of teaching quality in higher education and proposes the use of an Interpretivist perspective as an evaluative alternative. A case study of a novice college instructor is used for illustrating the limitations of the traditional paradigm and identifying critical issues that take place when the evaluation bases only on student ratings as the only source for determining teaching quality.

Key words: Teacher evaluation, case studies, higher education.

Introducción

El interés por la evaluación de la enseñanza en educación superior se incrementó en los últimos años debido, en parte, a las presiones crecientes de rendición de cuentas, cambios en las tendencias y requisitos de ingreso a la educación superior, así como a las restricciones financieras que enfrentan estas instituciones desde hace unos años, y que forzaron a las universidades a prestar mayor atención a los procedimientos formales para evaluar la calidad de la enseñanza en este nivel educativo (Wilson, 2001; Cisneros-Cohernour, 2001). Esta tendencia global se incrementó en países en vías en desarrollo, ya que la mayoría de los sistemas institucionales actuales consideran la evaluación de la enseñanza una parte inseparable de la calidad de las universidades y otras instituciones de educación superior.

En el caso de México, el interés por la creación de estos sistemas de evaluación de la enseñanza se incrementó a partir de la consideración del desempeño docente como un indicador de calidad de los programas educativos y de los cuerpos académicos. Como indican Rueda y Landesmann (1999), esta tendencia refleja “una nueva orientación hacia las actividades académicas y una forma de control de las prácticas laborales del personal académico” (p. 150). Esto resulta congruente con las tendencias globales, ya que a medida que las presiones por rendición de cuentas se incrementan, se enfatiza el uso de los resultados de la evaluación para la toma de decisiones acerca de la contratación, otorgamiento de la definitividad, promoción o el otorgamiento de estímulos por el desempeño docente (Altbach, 2004; Braskamp y Ory, 1994).

Aunque la preocupación por la calidad de la enseñanza aumentó, la evaluación aún se lleva a cabo de acuerdo con un paradigma positivista tradicional, que resalta importancia de la generalización y el establecimiento de relaciones causales. En la mayor parte de los casos, la evaluación se realiza mediante el uso de escalas, entrevistas semiestructuradas, tests de personalidad y, principalmente, cuestionarios (Feldman, 1986, 1989; Falk y Dow, 1971). El análisis de la literatura referente a evaluación docente, revela que desde los años sesenta, la mayor parte de las evaluaciones de la enseñanza se realizan únicamente a través de cuestionarios administrados a los estudiantes y constituyen la única fuente de información sobre las evaluaciones del desempeño docente (Kinney y Smith,

1992; Braskamp, Brandeburg y Ory, 1984; Cashin, 1988; Marsh, 1987; El-Hassan, 1995).

En la mayoría de los casos, los cuestionarios incluyen un conjunto de ítems predeterminados y estandarizados acerca de las características de enseñanza y sus dimensiones (Tagomori y Bishop, 1995); la forma en que se administran a los estudiantes es también estandarizada. Los resultados de las encuestas usualmente son analizados e interpretados de manera que se reducen a un puntaje numérico. Una vez obtenido, los resultados de un docente son comparados con los de otros profesores o en relación con ciertos criterios o estándares predeterminados. Algunos investigadores realizan observaciones en el aula para evaluar a los docentes, pero la tendencia es usar criterios estandarizados y cuantificables.

I. Método

El objetivo de este artículo es examinar las limitaciones del paradigma positivista en la evaluación de la enseñanza en educación superior y proponer el uso de una perspectiva interpretivista, como alternativa para un estudio más profundo de la calidad de la enseñanza, en el contexto en que ésta tiene lugar.

Para ello se llevó a cabo un estudio de caso desde la perspectiva interpretivista, para ilustrar las debilidades del paradigma tradicional e identificar cuestiones críticas que surgen cuando la evaluación se basa en opiniones de los estudiantes como única fuente para determinar la calidad de la enseñanza.

A continuación se examinan las limitaciones de las evaluaciones de la enseñanza bajo un paradigma positivista, seguidas de una propuesta de evaluación de acuerdo con la perspectiva interpretivista, a través de un estudio de caso y sus implicaciones para la evaluación de la enseñanza, principalmente en cuanto a dos aspectos fundamentales vinculados con la validez de constructor: el uso de los resultados de la evaluación y sus consecuencias.

Para ilustrar las limitaciones del paradigma positivista, principalmente en cuanto al uso y consecuencias de la evaluación, se presenta un fragmento de una investigación para un estudio de caso¹, realizada bajo una perspectiva interpretivista.

El estudio se enfocó en David, un profesor universitario en proceso de definitividad de la Facultad de Ciencias de una prestigiada universidad norteamericana,² de aproximadamente cuarenta años, complexión atlética, con varios premios y reconocimientos por su investigación, así como becas para realizar estancias en el extranjero y de investigación por la Fundación Nacional de las Ciencias de su país.

David fue recomendado por un especialista de la Oficina de Recursos Instruccionales de la Universidad, como un profesor ideal para llevar a cabo un estudio de caso porque a pesar de obtener evaluaciones pobres por su enseñanza

al inicio de su contratación, mejoró significativamente, después de recibir apoyo de esa oficina. Era un profesor que impartía cursos de matemáticas, principalmente en el nivel de licenciatura.

Durante el semestre en que se realizó el estudio de caso, David fue incluido en la lista de mejores docentes, publicada al final de cada periodo escolar por su universidad.

Una vez que David aceptó participar en el estudio de caso, comenzó el proceso de recolección de datos que se llevó a cabo durante ese semestre. Esto involucró entrevistas semiestructuradas y abiertas con otros docentes que impartían la misma materia que David, con su jefe de departamento, el director de la facultad, el vicerrector académico, con el jefe de la Oficina de Apoyo Instruccional y con el especialista de esa oficina, mismo que le proporcionó apoyo para que mejorara su enseñanza. Asimismo, se organizaron grupos de enfoque con los estudiantes y observaciones en el aula. Estas últimas tuvieron lugar durante todo el semestre; proveyeron información valiosa acerca del estilo de enseñanza de David, sus métodos instruccionales, su interacción con sus estudiantes, así como de las interacciones entre estos a lo largo del semestre.

Además de lo anterior, se llevó a cabo el análisis documental de todas las formas de evaluación de su docencia, que sus estudiantes llenaron a lo largo de los últimos cinco años, así como un análisis de sus materiales de clase y recursos didácticos.

1.1 Limitaciones de la evaluación de la enseñanza sobre la base de un paradigma positivista

La validez de la investigación acerca de la evaluación de la enseñanza en educación superior basada en un paradigma positivista, se centra en cuestiones de medición, principalmente en cuanto a las fortalezas de diferentes fuentes de información y a cuestiones de confiabilidad y de validez externa e interna de las evaluaciones basadas en opiniones de los estudiantes.

En los años noventa, las revisiones de la literatura de investigación de la década de los setenta, apoyaban la confiabilidad de los resultados obtenidos por medio de estas evaluaciones (Marsh, 1987; Centra, 1993; Costin, Greenough y Menges, 1971; McKiechie y Lynn, 1979; Cashin, 1988, 1995). De acuerdo con Marsh y Bailey (1993), las evaluaciones basadas en opiniones de estudiantes acerca de la efectividad de la enseñanza, son multidimensionales, confiables, estables y relativamente válidas en contra de ciertos indicadores de enseñanza efectiva.

Sin embargo, en una revisión de la literatura de investigación realizada a fines de los años noventa por Brodie (1998), encontró una correlación positiva entre el otorgamiento de altas calificaciones por parte de los docentes y los puntajes conferidos por los estudiantes a la calidad de su enseñanza. De acuerdo con este autor, "los maestros que otorgan calificaciones más altas y exigen menos tareas de los estudiantes reciben los más altos puntajes en las evaluaciones que éstos

hacen de sus profesores” (p.17). Asimismo, encontró que las correlaciones positivas entre los puntajes otorgados por los estudiantes y ciertas características de “efectividad docente” se exageraron en las publicaciones y que “algunos investigadores han borrado correlaciones bajas o negativas o creado correlaciones positivas al invertir la escala de calificaciones” (Brodie, 1999, p.1). Lo anterior pone en entredicho la validez de los resultados de las evaluaciones, basadas exclusivamente en la opinión de los estudiantes.

Otro problema identificado con la validez de contenido de esas evaluaciones, se refiere a la definición del constructo “buena enseñanza”. Un buen número de estudios utilizan definiciones generales y restrictivas de este constructo, que pueden ser inapropiadas para algunos contextos y circunstancias (Stake y Cisneros-Cohernour, 2004). Como afirma Doyle (1982):

No es posible pensar que un solo conjunto de características pueda aplicarse de la misma forma para definir todos los tipos de enseñanza y que sean aplicables a todo tipo de estudiantes...crear tal lista, implica un riesgo muy serio (p. 27).

El uso de listas de cotejo de características o conductas en la mayoría de los estudios, limita la creatividad docente, es contraria a la diversidad de estilos de enseñanza de los instructores y puede resultar en un castigo para aquellos cuyo estilo “no cae dentro de la norma” (Stake y Cisneros-Cohernour, 2004). En particular, es importante examinar las consecuencias y posibles injusticias resultantes de la evaluación, debido a la tendencia de comparar a docentes que imparten diferentes cursos en diferentes contextos, sin contar con un riguroso control de variables (Stake y Cisneros-Cohernour, 2004; Cisneros-Cohernour, 2001).

Aún quienes afirman que los resultados de las evaluaciones realizadas bajo el paradigma positivista son defendibles desde el punto de vista psicométrico, expresan la necesidad de elaborar más estudios acerca de la validez de constructo, como el caso de los aspectos de validez de contenido, substantiva, estructural y de consecuencias (Ory y Ryan, 2001).

También son necesarios estudios referentes a la generalización de los resultados de las evaluaciones, particularmente cuando los instrumentos utilizados son replicados de otro, desarrollados en y para contextos y culturas diferentes. Evaluaciones docentes realizadas en Hong-Kong reportan que los cuestionarios traducidos de sus versiones originales de países como los Estados Unidos, Gran Bretaña, Australia y Canadá no son útiles para evaluar efectivamente la calidad de la enseñanza en países orientales. Como indican Pratt, Kelly y Wong (1999):

La práctica de usar cuestionarios y escalas desarrollados en otros países para obtener información acerca de la calidad de la enseñanza en Hong-Kong o China puede resultar en la imposición de los valores de otra cultura, especialmente en relación con los papeles apropiados, las responsabilidades, y la forma en que los maestros de educación superior se relacionan con sus alumnos (p.1).

Debido a las serias implicaciones de validez de los resultados de las evaluaciones, basados en las opiniones de los estudiantes y las consecuencias negativas que puedan derivarse del mal uso de los resultados de la evaluación para la toma de decisiones, es necesario contar con estudios que permitan examinar con más profundidad la forma en que los resultados de estas evaluaciones se utilizan y las consecuencias de estos usos (Ory y Ryan, 2001).

II. Una propuesta interpretivista

Aunque la opinión de los estudiantes es sumamente importante, el hecho de incluir una sola fuente de información y excluir de la evaluación otros elementos que conforman la complejidad de la docencia y su contexto, pone en entredicho la validez de los resultados para representar la calidad del desempeño docente. Como afirman Stake y Cisneros-Cohernour (2004) y Stake y Burke (2000), la calidad de la docencia sólo puede ser juzgada propiamente en el contexto en que ésta tiene lugar y considerando todos los factores que en ella influyen. Si no se hace un esfuerzo por examinarla en su complejidad, la evaluación es inválida.

Con base en lo anterior y en virtud de las limitaciones del paradigma tradicional, y debido al creciente interés en el uso de resultados de la evaluación para la toma de decisiones, es necesario contar con nuevas perspectivas como la Interpretivista, que provee una mayor comprensión de la calidad de la enseñanza, su complejidad y contexto. Sobre esta base pueden comprenderse más ampliamente las metas institucionales, el ambiente del aula, la organización administrativa, los procesos, los contenidos curriculares, el aprovechamiento de los estudiantes y el impacto de los programas en la sociedad (Shulman, 1986; Cave, Hanney y Travett, 1988).

La perspectiva interpretivista se basa en el supuesto de que la evaluación es contextualizada y apropiada para medir el constructo evaluado y es suficientemente flexible para determinar la calidad de la enseñanza de diferentes docentes con diferentes estilos y orientaciones. De acuerdo con esta perspectiva, el evaluador realiza su trabajo en diferentes condiciones y con el tiempo suficiente para entender y evaluar la calidad de la enseñanza en forma justa. Este tipo de evaluación es sensible al contexto general y particular de la enseñanza. Entender la enseñanza desde una perspectiva interpretivista, implica ser capaz de “captar los significados de lo que constituyen estas acciones” (Schwandt, 1994, p.91). Para ilustrar las limitaciones del paradigma positivista, principalmente en cuanto al uso y consecuencias de la evaluación, se presenta un fragmento de una investigación por estudio de caso realizada bajo una perspectiva interpretivista.

David, el profesor sobre el que se llevó a cabo el estudio de caso, trabaja para un departamento de la Facultad de Ciencias de una universidad norteamericana desde 1993. La facultad tiene alrededor de 5,500 estudiantes y 400 profesores, la mayoría son hombres. El departamento donde trabaja tiene alto prestigio en cuanto a investigación, y el jefe de departamento está orgulloso, además, por el

excelente nivel de enseñanza de su personal: más de tres cuartas partes de los profesores de este departamento han recibido premios por la calidad de su instrucción en el nivel de licenciatura.

La Facultad de Ciencias proveía a sus docentes con todos los recursos que estos pudieran necesitar para impartir sus cursos. Las clases de David, por ejemplo, tenían lugar en un aula amplia, con excelente iluminación, aire acondicionado y calefacción, ubicada en el norte del campus. Al curso de Cálculo que impartió ese semestre,³ asistían doce estudiantes, uno de los cuáles era mujer. A continuación se muestra un extracto del estudio de caso de David (Cisneros-Cohernour, 1997), que describe algunos aspectos de su enseñanza:

David usualmente llega al aula temprano y espera a sus estudiantes en la puerta del aula para saludarlos a medida que entran al salón de clase. Su tono siempre es amable e informal. Trata de mantener este tono a lo largo de toda la sesión.

Él sigue una rutina mientras enseña su curso. Usualmente pasa la primer parte exponiendo un tema y la última parte supervisando a sus estudiantes mientras éstos trabajan en equipos. A lo largo de la exposición recuerda a sus estudiantes lo visto en la sesión anterior, los objetivos de la clase del día, indica los puntos principales de su presentación en el pizarrón y hace preguntas a sus estudiantes para verificar su comprensión acerca de la exposición. Mientras él escribe en el pizarrón sus estudiantes toman notas.

Al iniciar su presentación, David divide el pizarrón en tres secciones y escribe en perfecto orden de izquierda a derecha los principales puntos que acaban de presentar. Entonces, pregunta a los estudiantes si tienen dudas acerca de este tema. Si ningún estudiante tiene dudas, usa un ejemplo relacionado con el tema y hace preguntas acerca de éste a fin de verificar la comprensión de sus estudiantes. Sólo cuando todos sus estudiantes han participado y respondido correctamente a sus preguntas, provee más información.

A lo largo de la sesión, David continúa enseñando de esta forma, continuamente resume los puntos principales, ya sea por medio de diagramas, dibujos, gráficas, o utilizando ejemplos para ilustrar los temas que enseña. No utiliza diapositivas ni otras ayudas audiovisuales.

Cada vez que termina de explicar un concepto borra esta información del pizarrón, siempre de arriba abajo y de izquierda a derecha tal como le fue enseñando en la Oficina de Apoyo Instruccional. Una vez que ha terminado de explicar, él siempre pide a los estudiantes que trabajen en equipos resolviendo un problema que les permite aplicar los conceptos aprendidos. Él supervisa el trabajo, provee retroalimentación a los equipos y les deja tarea adicional para la casa. La clase se reúne una vez más durante la semana para trabajar en el laboratorio. Entre sesiones, David se mantiene en contacto con sus estudiantes vía e-mail y provee asesorías para quienes lo requieran.

Durante el grupo de enfoque, algunos estudiantes indicaron que el curso era difícil y que en ocasiones tenían problemas durante las sesiones de laboratorio. En esas sesiones, David contaba con un ayudante, a quien algunos estudiantes

tenían problemas en comprender cuando les daba indicaciones pero agregaron que como David clarificaba posteriormente los principales puntos estudiados en clase, esto evitaba confusiones.

Durante las observaciones en el aula, fue evidente que David es un hábil comunicador, muy organizado, que mantiene contacto visual con sus estudiantes y que trata de crear una atmósfera casual pero propicia para el aprendizaje. Posee muchos de los atributos que los programas de formación docente identifican como sinónimos de buena enseñanza: organización, claridad, buen tono de voz, buen grado de avance en la cobertura de los temas, etc.

Durante las entrevistas, David indicó que no poseía estudios de pedagogía ni había cursado ningún tipo de formación docente. Estos conocimientos los adquirió en la Oficina de Recursos Instruccionales de su universidad, después de buscar apoyo para mejorar su enseñanza, a raíz de los pobres puntajes que recibió en sus primeras evaluaciones docentes.

En los primeros años de su período de definitividad, tanto David como su jefe de departamento estaban preocupados por el bajo puntaje de esas evaluaciones. El jefe de departamento indicó que en la facultad se tomaban en cuenta seriamente los resultados de las evaluaciones que los estudiantes hacen de sus profesores al final de los cursos, y que cuando un profesor obtenía un puntaje por debajo de la media del campus, se le pedía que trabajara arduamente a fin de incrementarlo. Si para ello era necesario que el profesor recibiera algún apoyo, se le aconsejaba que acudiera a la Oficina de Apoyo Instruccional o a cualquier otra unidad de la universidad que pudiese auxiliarlo para mejorar su enseñanza.

De acuerdo con el especialista de la Oficina de Apoyo Instruccional, al principio David era un profesor tenía escaso contacto con sus estudiantes, se limitaba a explicar y a pedirles que resolvieran problemas similares a los revisados en clase fuera del aula. Por este motivo, le recomendó ser más personal en su relación con los estudiantes, utilizar más ejemplos y resumir los principales puntos de la clase en diferentes momentos durante la sesión. Pensaba que dado que la universidad trataba de incrementar el número de mujeres y estudiantes de grupos minoritarios, era importante que los docentes promovieran un ambiente cordial en el aula, para generar una sensación de comodidad en los alumnos.

Contrario a lo que David esperaba, sus puntajes no se incrementaron significativamente una vez que incluyó estos cambios en su enseñanza, sin embargo, el semestre anterior a la realización del estudio de caso, logró mejorar sus puntajes por encima de la media del campus y fue seleccionado entre los mejores instructores de la universidad.

El análisis de las evaluaciones docentes que los estudiantes hicieron del desempeño de David durante un período de cuatro años, evidenció que éste había recibido evaluaciones bajas por su enseñanza durante más de tres años y se apreciaba un ligero incremento en los puntajes a raíz las asesorías de la Oficina de Apoyo Instruccional. Los estudiantes rara vez hacían comentarios acerca de

las debilidades de David como maestro, sin embargo durante los tres primeros años comentaban que existían ciertos contenidos que les eran difíciles de entender y pedían que fueran eliminados, así como un software que David utilizaba para apoyar su curso. Durante sus primeros cuatro años en el departamento, 98% de las formas de evaluación llenadas por sus estudiantes, incluían comentarios de este tipo.

Durante el año en que David fue evaluado para determinar si procedía otorgarle su definitividad, el jefe de departamento se encontró en un dilema, por los bajos puntajes obtenidos; los esfuerzos por mejorar no implicaban puntajes suficientes. Los estudiantes aún se quejaban de los contenidos y el uso de un software cuyo uso se les dificultaba. Ante la presión, David tomó la decisión de eliminarlos. No se sentía cómodo con esta decisión porque sabía que los estudiantes requerían esos conocimientos que servían de base para avanzar en sus estudios, pero sentía que de no hacerlo sus puntajes no se elevarían. Después de esto la universidad le otorgó su definitividad, pues los puntajes de sus evaluaciones docentes se incrementaron significativamente, razón por la cual fue seleccionado ente los mejores instructores de la universidad.

Las presiones por obtener un puntaje por arriba de la media del campus, también eran motivo de preocupación para los colegas de David, tanto para aquellos que impartían los mismos cursos que él como los que estaban a cargo de otros en el departamento. Dos de sus colegas incluso sentían que se trataba a los estudiantes como clientes y que se estaba descuidando la responsabilidad de la institución para con la sociedad. Más de un profesor en el departamento expresó su preocupación por el uso de los puntajes de las evaluaciones para la toma de decisiones, capaces de impactar en la enseñanza. Uno de ellos expresó:

Los resultados de las evaluaciones y los comentarios que hacen los estudiantes no son siempre útiles para mejorar. Ellos no siempre están en la mejor posición para proveernos con retroalimentación. Infortunadamente, nuestro departamento se basa solamente en estos puntajes para evaluar la instrucción. Yo creo que este es un error porque se basa en el supuesto que el puntaje otorgado por los estudiantes es un sinónimo de la calidad instruccional del profesor. El hecho de que un docente mejore su puntaje no necesariamente significa que ha mejorado como maestro. Hay profesores que cubren el contenido en el nivel y con el rigor esperado y obtienen una evaluación promedio o ligeramente arriba del promedio, pero hay otros que cubren menos del 80% del contenido del curso y de forma superficial, se enfocan más en dinámicas y actividades que les agradan a los estudiantes. Estos últimos pueden salir mejor evaluados, pero ¿realmente significa que son mejores maestros?

Darle todo este peso al puntaje numérico crea otros problemas, puede haber el peligro de que algunos instructores manipulen a los estudiantes para obtener “buenas evaluaciones” de su enseñanza. Los estudiantes han comenzado a darse cuenta del poder que tienen en el proceso de evaluación y lo están usando.

Otro colega de David agregó:

Uno de los problemas de comparar instructores es que diferentes cursos pueden necesitar ser evaluados de forma diferente. Es injusto comparar profesores que enseñan diferentes disciplinas.

Durante el desarrollo de los grupos de enfoque con los estudiantes, aquellos que habían seleccionado una carrera en el área de ciencias sociales y humanidades expresaron que el curso les parecía difícil porque carecían de los conocimientos previos y que lo habían cursado sólo porque era un requisito para sus programas de estudio. Los estudiantes que seleccionaron el curso sin que fuera un prerrequisito tenían mejor percepción de éste y de su utilidad. En relación con el estilo de enseñanza de David, los estudiantes indicaron que era muy claro aunque algo repetitivo, pero que les agradaba la forma en que trabajaba con ellos.

III. Resultados

El estudio de caso de David permite entender muchos aspectos complejos de la evaluación de la enseñanza que difícilmente pueden ser identificados con otros enfoques de evaluación. El concepto de “buena enseñanza”, utilizado por la Oficina de Apoyo Instruccional enfatizaba la necesidad del uso de técnicas pedagógicas, mejorar la organización y estructura del curso, su planeación y la interacción del docente con los estudiantes. Muchos de estos aspectos son consistentes con los *Principios para la buena enseñanza de estudiantes de licenciatura* de Chickering y Gamson (1987), pero existen otros elementos que los especialistas de esta oficina no examinaron debido a que no tenían la misma formación en el área de contenido que el docente. Esta información es aquella vinculada con los otros cursos (que será parte del perfil del egresado) y a la profundidad con la que debían cubrirse algunos temas en el curso de David para proveer a los estudiantes de las bases necesarias para sus estudios posteriores.

Asimismo, el especialista instruccional observó las clases de David en un limitado número de ocasiones, por lo que no pudo juzgar la conexión entre las sesiones, la relación entre los objetivos o las condiciones del contexto que afectaban el aprendizaje de los estudiantes. Aunque revisó los puntajes otorgados por los estudiantes a lo largo de tres años e hizo recomendaciones de mejora, éstas se basaron en información insuficiente del contexto de enseñanza.

Si las decisiones acerca de la calidad de la enseñanza sólo se fundamentan en el uso de los resultados de las encuestas de opinión de los estudiantes, se presentan problemas. Las decisiones acerca de la definitividad, contratación o reconocimiento del desempeño docente de profesores como David, con base en resultados de una sola fuente de información son preocupantes, dadas las limitaciones que ésta presenta. El puntaje numérico por si solo no refleja la forma en que los diferentes intereses, antecedentes y necesidades de los estudiantes influyen en sus percepciones acerca de la importancia del curso y la calidad de la enseñanza del docente. Al centrarse únicamente en el resultado numérico, el jefe de departamento y las autoridades de la Facultad de Ciencias crearon una situación que condujo al profesor a eliminar un contenido importante del programa

de estudio. Este uso de los resultados de la evaluación puede traer consecuencias serias, como la disminución de la calidad en la formación de los estudiantes, no mejorarla. Asimismo, la tendencia de resumir los resultados de la evaluación en un puntaje numérico, originó que David se preocupara más por elevar sus puntajes que por mejorar su enseñanza (Cisneros-Cohernour, 2007).

En general, los resultados del estudio permitieron identificar problemas con la forma en que la Oficina de Apoyo Instruccional utiliza los resultados de la evaluación. La decisión de esta oficina de usar informes numéricos acerca del puntaje obtenido por el docente y compararlo con el puntaje promedio obtenido por otros docentes en el campus, tenía poca utilidad y podía crear confusiones. El director del departamento donde David laboraba, lo mismo que una gran parte de los administradores, suponía que si un profesor obtenía un puntaje de dos en la escala del instrumento de evaluación, significaba que era la mitad de bueno que un docente que obtenía cuatro puntos en la misma escala. Para los docentes, este tipo de informes tampoco era útil, porque tenía poca utilidad para propósitos de mejora.

El estudio de caso permitió entender con mayor profundidad el contexto de enseñanza e identificar cuestiones críticas para la formación del docente que no pudieron ser detectadas en las evaluaciones hechas por los estudiantes ni por el evaluador externo.

Conclusiones

En este trabajo se examinaron las limitaciones del paradigma positivista en la evaluación de la enseñanza en educación superior, principalmente en cuanto al uso de los resultados de la evaluación y sus consecuencias, para lo cual se elaboró un diseño de investigación para un estudio de caso, desde la perspectiva interpretivista. La recolección de datos se llevó a cabo durante seis meses e involucró métodos múltiples, tales como entrevistas semiestructuradas y abiertas, grupos de enfoque, análisis documental, observaciones cualitativas y análisis de materiales de clase y recursos didácticos.

Los resultados fueron consistentes con los de otras investigaciones, en cuanto a las limitaciones de las evaluaciones de la enseñanza basadas en las opiniones de los estudiantes, principalmente en cuanto al problema del significado de los puntajes y a las interpretaciones de los resultados para la toma de decisiones. Esto se debe a que los puntajes otorgados por los estudiantes no necesariamente son sinónimo de “buena enseñanza”, sino resultado de decisiones del docente con la intención de aumentar su puntaje numérico, sin importar la calidad de la formación de los estudiantes. Por ello es necesario cuestionar seriamente los resultados de la evaluación y sus implicaciones para la toma de decisiones.

Asimismo, es necesario considerar que los estudiantes, los especialistas instruccionales, los colegas, los administradores o el docente mismo, son fuentes de información del desempeño docente, pero tienen limitaciones y sesgos. Ignorar

estas limitaciones puede afectar seriamente la validez de la evaluación y sus consecuencias.

El estudio demuestra las fortalezas de la investigación realizada a partir de una perspectiva interpretivista para la evaluación docente, porque permite comprender con mayor profundidad el significado del constructo *buena enseñanza* dentro del contexto en que esta tiene lugar, las fortalezas y debilidades del desempeño del docente en su contexto, así como las consecuencias del uso de los resultados de la evaluación basados en la opinión de los estudiantes.

La investigación permitió entender también cómo ignorar importantes elementos del constructo *buena enseñanza*, en el contexto del estudio de caso, comprometió la validez de constructo de la evaluación, principalmente en cuanto a los aspectos de validez de contenido y de consecuencias. Asimismo, permitió verificar que el alto énfasis en los puntajes numéricos de la evaluación docente, con poca atención en su significado, originó serias consecuencias que pudieron deteriorar la calidad de la educación y promover prácticas poco éticas entre los profesores, con la intención de incrementar el puntaje de sus evaluaciones.

Estos aspectos son serios porque comprometen los resultados de la evaluación y su validez, y porque pierde su objetivo: identificar las fortalezas y debilidades en la enseñanza de los profesores.

Contar con estudios de evaluación bajo una perspectiva interpretivista, presenta numerosas ventajas en la evaluación de la calidad del desempeño docente, principalmente por ser la única que ahonda en la conceptualización del constructo *buena enseñanza* en el contexto en que ésta tiene lugar y es sensible a las percepciones de diferentes actores. El uso de múltiples fuentes de información y múltiples métodos de recolección de datos contribuye a una mayor comprensión de los significados e interpretaciones de los resultados de la evaluación. Este procedimiento gana complejidad, validez y aporta mayor comprensión sobre las fortalezas y debilidades del docente, y aunque se podría argumentar una desventaja en el costo y el tiempo que conlleva, la riqueza de la información y su utilidad para el mejoramiento de la calidad de la enseñanza, así como los beneficios que brinda para los docentes, constituye la opción más valiosa para evaluar la calidad de la enseñanza.

Referencias

Altbach, P. (2004). Globalization and the university: Myths and realities in an unequal world. *Tertiary Education and Management I*, 1-20. Disponible en: http://www.bc.edu/bc_org/avp/soe/cihe/publications/pub_pdf/Globalization.pdf

Braskamp, L. A., Brandeburg, D. y Ory, J. (1984). *Evaluating teaching effectiveness*. London: Sage.

Braskamp, L. A. y Ory, J.C. (1994). *Assessing Faculty Work*. San Francisco: Jossey-Bass.

Brodie, D. A. (1998). Do students report that easy professors are excellent teachers? *The Canadian Journal of Higher Education*, 23 (1), 1-20.

Brodie, D. A. (1999, abril). *Has publication bias inflated the reported correlation between student achievement and ratings of instructors?* Trabajo presentado en la Annual Meeting of the American Educational Research Association, Montreal, CA.

Cashin, W. E. (1988). Student ratings of teaching: A summary of the research. *IDEA paper 20*, 1-6.

Cashin, W. E. (1995, septiembre). Student ratings of teaching: The research revisited. *IDEA Paper 32*.

Cave, M., Hanney, S., Kogan, M., y Travett, G. (1988). *The use of performance indicators in higher education: A critical analysis of developing practice*. London: Kingsley.

Centra, J. A. (1993). *Reflective faculty evaluation: Enhancing teaching and determining faculty effectiveness*. San Francisco: Jossey-Bass Inc.

Chickering, A. W. y Gamson, Z. F. (1987, marzo). Seven principles for good practice in undergraduate education. *AAHE Bulletin 39* (7), 3-7.

Cisneros-Cohernour, E. J. (1997). *Trade-offs: The use of student ratings results and its possible impact on instructional improvement*. [Informe no publicado]. Illinois, EE. UU: University of Illinois, Center for Instructional Research and Curriculum Evaluation.

Cisneros-Cohernour, E. J. (2001). *The evaluation of teaching in the context of a research university: Different meanings, trade-offs and equity concerns*. Tesis doctoral no publicada, University of Illinois, Urbana-Champaign, EE.UU.

Costin, F., Greenough, W. T. y Menges, R. J. (1971). Student ratings of college teaching: reliability, validity, and usefulness. *Review of Educational Research*, 41 (5), 511-535.

Doyle Jr., K. O. (1982). *Evaluating teaching*. Lexington, MA: Lexington Books.

El-Hassan, K.(1995). Student ratings of instruction: Generalizability of findings. *Studies in Education*, 21, 411-429.

Falk, B. y Dow, K. L. (1971). *The assessment of university teaching*. London: Society for Research into Higher Education Ltd.

Feldman, K. A. (1986). The perceived instructional effectiveness of college teachers as related to their personality and attitudinal characteristics. *Research in Higher Education*, 24, 139-170.

Feldman, K. A. (1989). Instructional effectiveness of college teachers as judged by teachers themselves, current and former students, colleagues, administrators, and external (neutral observers). *Research in Higher Education*, 30, 137-172.

Kinney, D. P. y Smith, S. P. (1992). Age and teaching performance. *Journal of Higher Education*, 63, 282-302.

Marsh, H. W. (1987). Student evaluations of university teaching: Research findings, methodological issues and directions for future research. *International Journal of Educational Research*, 11, 253-388.

Marsh, H. W. y Bailey, M. (1993). Multidimensional students' evaluations of teaching effectiveness. *Journal of Higher Education*, 64, 1-18.

McKeachie, W. J. y Lynn, Y. G. (1979). A Note on validity of student ratings of teaching. *Educational Research Quarterly*, 4 (3), 45-47.

Ory, J. y Ryan, K. (2001). How do student ratings measure up to a new validity framework? *New Directions in Institutional Research*, 109.

Pratt, D. D., Kelly, M. y Wong, W. S. S. (1999). Chinese conceptions of 'effective teaching' in Hong Kong: Towards culturally sensitive evaluation of teaching. *International Journal of Lifelong Education*, 18 (4).

Rueda, M. y Landesmann, M. (1999). Hacia una nueva cultura de la evaluación de los académicos. *Perfiles Educativos México*, 21, 83-84.

Schwandt, T. A. (1994). Constructivist, interpretivist approaches to human inquiry. En Norman K. Denzin e Yvonna Lincoln (Eds.), *Handbook of Qualitative Research*. Thousand Oaks, (pp. 118-137). CA: Sage.

Shulman, L. (1986). Knowledge and teaching: Foundations of the new reform. *Harvard Education Review*, 57 (1), 1-22.

Stake, R. E. (1998, abril). Representing quality in evaluation. Ponencia presentada en el congreso anual de la American Educational Research Association, Montreal, Canadá.

Stake, R. E. (2005). Qualitative case studies. En Norman Denzin e Ivonna Lincoln (Eds.), *Handbook of Qualitative Research*, (3a. ed., pp.443-466). Thousand Oaks, CA: Sage.

Stake, R. E. y Burke, M. (2000). *Evaluating teaching*. [Informe de Investigación]. Champaign, Illinois, EE. UU.: Center for Instructional Research and Curriculum Evaluation [CIRCE].

Stake, R. E. y Cisneros-Cohernour, E. J. (2004). The Quality of Teaching in Higher Education. *Lithuanian Journal of Higher Education* 1, 94-107. Lituania.

Tagomori, H. T. y Bishop, L. A. (1995). Student evaluations of teaching: Flaws in the instruments. *Thought & Action*, 11, 63-78.

Wilson, R. (2001, 2 de febrero). It's 10 a.m. do you know where your professors are? Faculty members at Boston U. fume over plan to require them to be in their offices at least 4 days a week. *The Chronicle of Higher Education, The Faculty* [sección], A10-A12.

¹ Estudio de caso se refiere, en este caso, a una investigación acerca de un objeto definido, en este caso un docente, que involucró la exploración detallada a lo largo de un período de tiempo, lo suficientemente extenso para obtener un entendimiento profundo de de la complejidad de la enseñanza y su evaluación en contexto (Stake, 2005).

² El caso de David es parte del estudio "Tradeoffs: The use of student ratings results and its posible impact on instructional improvement" (Cisneros-Cohernour, 1997). Para garantizar la confidencialidad, se utilizaron pseudónimos para reemplazar los nombres de los departamentos, miembros del personal académico y jefes de departamento que participaron en este estudio.

³ Los colegas de David y su jefe de departamento indicaron que 12 es el número promedio de estudiantes en varios de los cursos de la facultad, pero que existen cursos con treinta y cinco o con más de cien estudiantes.