

EVALUACIÓN DE LOS ELEMENTOS DE ENSEÑANZA Y DE APRENDIZAJE EN EL AULA SEGÚN LO PERCIBEN MAESTROS Y ESTUDIANTES

Tevni Grajales Guerra
Universidad de Montemorelos, México

RESUMEN

La supervisión de la labor docente ha conducido al uso de diversas estrategias y criterios para evaluar el desempeño: (a) la aplicación de pruebas estandarizadas a nivel nacional a fin de evaluar el desempeño docente, (b) el juicio de expertos, (c) las calificaciones que el alumno obtiene en un curso en particular y (d) la evaluación que los estudiantes hacen de sus maestros. La evaluación del docente por parte del alumno es estudiada en esta investigación en comparación con la autoevaluación que hace el docente de su desempeño personal. De un total de 107 clases de diversas especialidades se observa que, aunque los alumnos (N = 1455) y maestros (N = 56) parecen percibir el proceso de enseñanza y de aprendizaje de la misma manera, no necesariamente lo valoran igual. Características tales como la personalidad del maestro y el grado de conocimiento pedagógico que tienen el maestro o el alumno son variables que pueden explicar cómo los estudiantes y los maestros evalúan el desempeño docente. El estudio recomienda la evaluación de los estudiantes respecto al desempeño de su maestro como un recurso valioso pero no suficiente para calificar el desempeño docente, por lo que propone la combinación de otras estrategias, tales como la evaluación entre colegas, la autoevaluación y el juicio de expertos, como criterios complementarios al momento de hacer una evaluación sumativa del desempeño profesional.

La función administrativa en general ha sido explicada de diversas maneras por teóricos como Fayol, Koontz, Urwick, Terry y la Asociación Norteamericana de Administradores, entre otros (cf. Sánchez Guzmán, 1981, p. 12). En sus esfuerzos por describir el proceso administrativo han propuesto que éste se constituye de cierto número de elementos básicos, en la enumeración de los cuales difieren los autores. Pero algunos de estos elementos son incuestionables. Un caso particular es el que se conoce como control, que para la Asociación Norteamericana de Administradores se subdivide en organización y supervisión.

La supervisión, como factor impor-

tante en el control administrativo, procura el perfeccionamiento de los procesos y de las personas en ellos involucrados. A medida que las teorías administrativas han dirigido su atención hacia las personas en lugar de hacerlo hacia los procesos, se ha despertado mayor interés por hacer de la supervisión una actividad más participativa. Una tarea indispensable en la supervisión es la evaluación, sea ésta de tipo diagnóstica, formativa o sumativa. Por medio de la evaluación se pretende identificar en qué medida se está logrando alcanzar la norma establecida en el desempeño o los objetivos específicos. De manera que la evaluación es democrática y participativa en la

EVALUACIÓN DE ELEMENTOS DE ENSEÑANZA Y DE APRENDIZAJE

medida en que las personas que son evaluadas sean también las encargadas de determinar el grado de logro que obtienen.

En lo que a la evaluación de los procesos de enseñanza y aprendizaje se refiere, aunque en sus inicios se orientó hacia los resultados obtenidos por los estudiantes, sea en pruebas de aula o pruebas estandarizadas, más tarde se vio la necesidad de evaluar el desempeño del docente. El desempeño del docente inicialmente era evaluado por un experto, quien se suponía conocía las técnicas y los procedimientos apropiados. Con el correr del tiempo primó la idea de hacer más democrático el proceso e integrar a otros docentes e incluso a estudiantes, quienes, de una manera u otra, se pensó, tendrían algo que decir respecto al desempeño de su maestro. También se consideró que las calificaciones que obtienen los alumnos evidencian el desempeño del maestro; de modo que cuando el alumno fracasa, también fracasa el maestro. Lamentablemente esto puede conducir a que los criterios de los maestros bajen a fin de mantener una buena imagen. Otra forma más directa de participación del estudiante consiste en darle la oportunidad de evaluar el desempeño de su docente por medio de encuestas o formularios aplicados al final del curso escolar.

La participación de estudiantes en el proceso de evaluación de los docentes introdujo un acalorado e interesante debate. Entre los estudiosos del asunto que han publicado sus resultados se incluyen Howard y Maxwell (1982), Cohen (1987), Feldman (1989), Abrami, d'Apollonia y Cohen (1990), Marsh y Dunkin (1992), Greenwald y Gillmore (1997), Marsh y Roche (1997), Greenwald (1997) y d'Apollonia, Lou y Abrami (1998). Al respecto, Blatchley, profesor en la Pittsburg State University

(comunicación personal, 14 de octubre de 1998), dice que la participación de los estudiantes en la evaluación de la enseñanza ha llegado a ser un factor crítico, siendo que los métodos objetivos son cuestionables en su validez y pueden ser duplicados con facilidad por cualquier persona. Además, afirma que le gustaría utilizar encuestas de estudiantes en sus propias clases aunque no fueran requeridas como un mandato y las recomienda como medios para la retroalimentación y el desarrollo profesional del maestro.

Por su parte, Robert Brown, de la Lincoln Nebraska University (comunicación personal, 13 de octubre de 1998) afirma que las evaluaciones que hacen los estudiantes de sus profesores son evaluaciones multidimensionales. En ellas, los alumnos consideran otros factores además de su propio aprendizaje. Un estudiante puede evaluar a su maestro por debajo de otros maestros, aunque al mismo tiempo obtenga buenas calificaciones con ese profesor. Ello significa que las calificaciones que obtiene el alumno no necesariamente corresponden a la evaluación que el alumno hace de su maestro.

La importancia y el uso que se le da a la evaluación que los alumnos hacen de sus profesores puede estar generando resistencia en los docentes que no consideran al alumno calificado para hacer un juicio al respecto, especialmente cuando los resultados se utilizan para determinar el estatus laboral del maestro. Otros estudiosos del asunto se preocupan por la facilidad con la cual un docente puede manipular, a fin de obtener una buena puntuación de parte de sus alumnos. Bob Henry del Lambton College (comunicación personal, 8 de octubre de 1998) afirma que Weissburg enumera las siguientes formas de manipular resultados: permanecer en el aula

mientras se aplica el cuestionario, asegurarse de que los alumnos sepan que sus puntuaciones serán utilizados para considerar su promoción, no aplicar el cuestionario durante un examen final, hacer creer a los estudiantes que obtendrán mejores calificaciones de las que en realidad van a obtener, enseñar cursos que tengan pocos estudiantes, enseñar a grupos de alumnos en cursos de su especialidad en los cuales ellos tienen mucho interés, mostrar entusiasmo respecto al material que presenta. Todo esto puede contribuir a un mejor resultado en la evaluación que hace el estudiante, sin requerir por ello un mejoramiento real del desempeño docente.

Scriven (1995) afirma que, entre los argumentos fuertes que se ofrecen para usar la evaluación que hacen los estudiantes de sus maestros, se encuentra el hecho de que los estudiantes se encuentran en una posición única para determinar el incremento personal de su conocimiento y comprensión, así como para determinar los cambios de motivación que hayan tenido respecto a la materia. Pueden observar y valorar hechos como la puntualidad del instructor, la comprensión de sus exposiciones y escritos y el entusiasmo del maestro, entre otros. Además, afirma que se ha encontrado una correlación estadísticamente significativa entre las puntuaciones que otorgan los alumnos en la evaluación de sus maestros y los incrementos de aprendizaje.

Este estudio se realizó con el propósito de contribuir a la comprensión del fenómeno de la evaluación de los procesos de enseñanza y aprendizaje, desde la perspectiva de los docentes y los estudiantes en un contexto latinoamericano. Procuró determinar qué coincidencias y diferencias pueden existir en la forma como evalúan el desarrollo del curso al

comparar alumnos con maestros. Se trató de determinar si el género del docente y el número de alumnos en la clase son predictores de la evaluación que hacen los estudiantes.

Este estudio contribuye a clarificar el uso y los beneficios que puede tener la evaluación que hacen los estudiantes, al final del curso, del desempeño de su maestro. Ofrece la oportunidad a los administradores y evaluadores educativos de refinar su concepto del papel de la evaluación que hace el estudiante.

Método

El estudio se realizó en la Universidad de Montemorelos en Nuevo León, México, al finalizar el segundo semestre del año escolar en el mes de mayo de 1998. Fueron incluidos en este estudio cursos de educación, ciencias computacionales, artes visuales, música, enfermería, administración de empresas y de salud. El cuestionario utilizado consistió de un total de 26 preguntas relacionadas con las actividades de clase, organizadas respecto a los siguientes temas: el docente (ocho ítems), la materia (cinco ítems), la metodología del curso (ocho ítems), el alumno (cuatro ítems). Durante la última semana de clase del semestre se hizo llegar a cada profesor un sobre que contenía un cuestionario para que el maestro se autoevaluara. Además, en el sobre se incluyeron suficientes cuestionarios para ser aplicados entre los alumnos del curso. Las instrucciones escritas consistieron en dos documentos. Uno para el profesor, en el que se le instruyó respecto a cómo realizar su autoevaluación y se le pidió que entregara a uno de los estudiantes del curso el resto del contenido del sobre. El estudiante asignado recibió instrucciones escritas de cómo aplicar la encuesta en el aula, en ausencia del docente, y de

EVALUACIÓN DE ELEMENTOS DE ENSEÑANZA Y DE APRENDIZAJE

recoger después los formularios contestados por sus compañeros, para entregarlos dentro del sobre al coordinador de la carrera en cuestión. Por su parte, el docente también entregó su autoevaluación al coordinador de la carrera.

Resultados

Los resultados del estudio representan 107 clases dictadas por un total de 56 profesores, 29 de los cuales dictaron más de una de las clases observadas. Respecto al género de los docentes participantes, 19 eran mujeres y 37 varones. El 68% de las clases estuvieron a cargo de docentes varones. El total de formularios de evaluación recuperados por parte de los alumnos fue de 1455. El 21.5% de las clases evaluadas fue de la Facultad de Ingeniería y Tecnología, 15% de la Facultad de Ciencias de la Salud, 57% de la Facultad de Ciencias, Artes y Humanidades y 6.5% de la Facultad de Administración. La clase más pequeña estuvo constituida por un alumno y la más grande por 46; observándose un promedio de 13 alumnos por clase.

Para organizar los resultados, se procedió a levantar dos bases de datos independientes. La primera contiene las res-

puestas que cada uno de los 1455 estudiantes ofreció a cada uno de los 25 indicadores de la encuesta. Estas respuestas se computaron por clase y se obtuvo un promedio de los resultados por cada grupo o clase. También se elaboró una segunda base de datos en la que se registró la autoevaluación que ofreció cada maestro y el promedio de la evaluación del grupo de estudiantes por indicador y por clase. Con esta segunda base de datos se procedió a identificar los casos en los que existe correlación significativa entre la repuesta del maestro y la de sus respectivos alumnos. Además, se identificó las clases en las cuales la evaluación de los alumnos difiere significativamente de la autoevaluación del maestro.

Como puede observarse en la Tabla 1, las 107 clases observadas fueron clasificadas en cuatro categorías, a saber:

1. Acuerdo total: Se refiere a que la calificación otorgada por el maestro a cada ítem observado está correlacionada de manera significativa ($p < .05$) con la de sus alumnos y, al mismo tiempo, el promedio de calificación general que otorgó el maestro a su clase no difiere significativamente del que le otorgaron los estudiantes.

Tabla 1

Frecuencia y porcentaje de clases según grado de acuerdo maestro-alumno

Categoría	Clases	Porcentaje del total de clases
Acuerdo total	25	23.4
Acuerdo parcial en la valoración de ítemes	23	21.5
Acuerdo parcial en la calificación general de la clase	20	18.7
Desacuerdo total	39	36.4
Total	107	100

2. Acuerdo parcial en la correlación: Esta categoría está formada por las clases en las que, aunque los maestros y alumnos difieren de manera significativa en el promedio de calificación general que otorgan a la clase, los valores que

otorgan a cada uno de los ítemes evaluados están correlacionados.

3. Acuerdo parcial en la calificación general de la clase: Esta categoría está formada por las clases en las que, aunque los maestros y alumnos no difieren

de manera significativa en el promedio de calificación que otorgan a la clase, los valores que otorgan a los ítemes evaluados no están correlacionados.

4. Desacuerdo total: En este grupo están las clases en que no se encontró correlación significativa al comparar la calificación otorgada por maestro y alumnos a cada ítem observado y al mismo tiempo difieren significativamente en la calificación general que se otorga a la clase.

Con el propósito de probar la hipótesis según la cual la frecuencia de casos por categorías es significativamente la misma se aplicó la chi cuadrada obteniéndose un valor de 7.953 con 3 grados de libertad y un nivel de significación de .047, observándose que el porcentaje más elevado de casos corresponde a los de desacuerdo total (36.4%).

Un segundo paso en el proceso de analizar los resultados consistió en determinar en qué dirección se manifiesta el desacuerdo entre los docentes y alumnos. Con este fin se utilizó la prueba *t* de Student para muestras independientes y se observó la diferencia entre la media de la calificación total otorgada por cada

maestro, respecto a la media de la calificación total otorgada por el respectivo grupo de alumnos. Del total de 106 clases, en cinco de ellas el maestro y los alumnos no difirieron de manera significativa ($p < .05$) en la calificación promedio que otorgaron al proceso evaluado. Por otro lado, en 39 de las clases el maestro calificó el proceso con un promedio significativamente superior ($p < .05$) al que le otorgaron sus alumnos y en 62 clases el maestro calificó el proceso significativamente ($p < .05$) por debajo de la calificación promedio que otorgan los alumnos. Al aplicar una prueba de chi cuadrada para determinar qué tipo de diferencia predomina, se encontró que los maestros tienden a autoevaluar su desempeño por debajo del valor que le otorgan sus alumnos ($\chi^2_{(1)} = 5.238$; $p = .022$).

Una observación puntual, considerando únicamente los casos en que existe un total desacuerdo entre maestros y alumnos, evidencia que una mayoría significativa de maestros tiende a calificar los procesos del aula con valores inferiores a los que otorgan los estudiantes (ver Tabla 2).

Tabla 2
Comparación de la dirección en la que maestros y alumnos califican el proceso en los casos en que se da total desacuerdo

Situación observada	Número de clases	Número de maestros
Casos en que el maestro otorga una calificación superior a la que le otorgan sus alumnos.	13	10
Casos en que el maestro otorga una calificación menor a la que otorgan los alumnos.	26	19
Total de casos	39	29

Al considerar los casos en que maestros y alumnos están en total desacuerdo, los resultados de la prueba chi cuadrada ($\chi^2_{(1)} = 4.333$; $p = .037$) conducen a la aceptación de la hipótesis que afirma

que la mayoría de los maestros tienden a ser conservadores al evaluar el proceso de enseñanza-aprendizaje, con respecto a los valores que otorgan los alumnos en los casos cuando hay total desacuerdo entre

EVALUACIÓN DE ELEMENTOS DE ENSEÑANZA Y DE APRENDIZAJE

los mismos.

Se utilizó una regresión lineal (stepwise) para identificar las variables predictoras de la diferencia entre los resultados ofrecidos por los maestros y los que ofrecen los estudiantes, tomando como variables predictoras el género del maestro y el número de alumnos en el aula. Se encontró que el género del maestro podría ser una variable predictora ($\beta = .211$; $p = .029$) con un coeficiente de determinación al cuadrado de .04. Se requiere de un estudio más amplio a fin de verificar si en efecto este coeficiente de determinación es aceptable al incluir otras variables no observadas en este estudio, como pueden ser la experiencia del docente, su nivel de preparación y el grado de afectividad docente-alumno, entre otras.

Discusión

Aunque maestro y alumnos comparten la misma experiencia al reunirse en actividades académicas en la misma aula de clase, difieren en la forma de percibir el proceso.

Al comparar los valores que otorgan a cada uno de los aspectos observados, tanto maestros como alumnos obtienen una correlación significativa en el 40% de los casos y en un 64% de casos se presenta una diferencia significativa al comparar la calificación total que otorgan maestros y alumnos al proceso observado.

Una correlación significativa entre maestros y alumnos en sus respuestas a cada uno de los indicadores del proceso no necesariamente significa que valoran los indicadores observados de igual manera sino que coinciden en la forma como tienden a ubicar cada indicador con respecto a los otros indicadores. Por ejemplo, a continuación, en la Tabla 3 aparecen diez indicadores hipotéticos

que son valorados de manera diferente pero con una alta correlación.

Tabla 3
Ejemplo de distribución de resultado en un aula hipotética

Indicador	Valor A	Valor B
A	3	5
B	1	3
C	3	4
D	1	3
E	2	4
F	3	5
G	2	4
H	1	3
I	3	5
J	2	3
Promedio total	2.1	3.9

Si la Tabla 3 representara la evaluación que hace el grupo de alumnos (valor A) y el maestro (valor B) de la clase, podríamos decir que existe una correlación significativa ($r = .88$; $p = .001$) en su forma de valorar cada uno de los indicadores, pero al mismo tiempo difieren significativamente en la calificación general que otorgan al proceso educativo evaluado ($t_{(9)} = -13.500$; $p = .000$), lo que indica que una correlación significativa no evidencia necesariamente que los maestros y los estudiantes otorguen la misma puntuación o valor al proceso evaluado. De la misma manera, en el caso de las 107 clases evaluadas, del 40% de las evaluaciones correlacionadas de manera significativa, la mitad (21) son evaluaciones significativamente diferentes. Y del 60% de las clases en las que la evaluación del maestro no estaba

correlacionada con la del alumno, dos terceras partes (40 casos) no muestran diferencia significativa entre alumnos y maestros (véase Tabla 4).

Esto significa que el administrador educativo que utiliza formularios de encuestas, entre alumnos, para evaluar el desempeño del maestro, tiene que tomar en cuenta que aunque exista una correlación significativa en la forma como valoran maestros y alumnos cada uno de los aspectos evaluados, esto no significa que ambos otorguen la misma calificación tanto a los aspectos particulares como al proceso en general.

El criterio que siguen los alumnos para evaluar al maestro tiende a ser más coincidente con el del maestro en la medida en que el estudiante desconoce la teoría del proceso de enseñanza-aprendizaje. Los maestros sin preparación pedagógica y los alumnos en carreras no relacionadas con la pedagogía tienden a coincidir más frecuentemente en sus evaluaciones. Como se puede ver en la Tabla 5, las clases observadas y constituidas por grupos de carreras pertene-

cientes a los programas de educación son las que presentan frecuencias de correlaciones y pruebas *t* significativamente diferentes, lo que evidencia que en estos casos los maestros y los estudiantes tienden a diferir en su forma de valorar los procesos observados. En cambio, las clases formadas por estudiantes y maestros en carreras no relacionadas con la pedagogía no evidencian falta de armonía en la forma como maestros y estudiantes evalúan los trabajos observados.

Tabla 4
Tabulación cruzada de pares de casos según significatividad de la correlación y de la prueba t

	t signifi- cativa	t no signi- ficativa	Totales
Correlación no significativa	21	40	61
Correlación significativa	20	20	40
Totales*	41	60	101

* no incluye 6 casos con información incompleta

Tabla 5
Comparación de los niveles de significación de las correlaciones y las pruebas t según el tipo de contexto pedagógico de la clase

Técnica estadística	Clases según tipo de carrera	Significativas	No significativas	Valor de chi cuadrado (gl=1)
Correlación	Pedagógicas	19	40	7.475*
	No pedagógicas	21	21	0
Prueba t	Pedagógicas	40	21	5.918*
	No pedagógicas	24	22	0.087

**p* < .05

En 18 clases pedagógicas, el maestro se sobrevalora y en 41 se subvalora. Ello significa que, en el caso de la institución estudiada, los maestros tienden a ser conservadores en su forma de valorar el trabajo en clase ($\chi^2_{(2)} = 37.803$; *p* = .000).

Conclusiones

La participación de los alumnos en el proceso de evaluación del desempeño del maestro es recomendable, porque (a) responde a los intereses de una administración democrática y participativa, (b)

EVALUACIÓN DE ELEMENTOS DE ENSEÑANZA Y DE APRENDIZAJE

ofrece la oportunidad al estudiante de conocer los aspectos del proceso de enseñanza-aprendizaje que la institución enfatiza y (c) ofrece una perspectiva diferente e importante respecto a lo que acontece en el aula. Pero la evaluación que hacen los estudiantes respecto del desempeño de una clase no es suficiente para determinar un criterio que conduzca a decisiones administrativas.

El proceso de evaluación del docente debe ser complementado por la autoevaluación que haga de su labor y, de ser posible, por la evaluación que realicen compañeros profesionales del docente, la cual ofrece la oportunidad de equilibrar los puntos de vista de docentes y alumnos, siendo que en algunos casos el docente tiende a ser muy conservador o liberal en su autoevaluación al compararla con la evaluación que hacen sus alumnos.

Aunque los alumnos y maestros parecen percibir el proceso de enseñanza-aprendizaje de la misma manera, no necesariamente lo valoran igual; existe la posibilidad de que el maestro tienda a sobrevalorarlo o subvalorarlo respecto al valor que otorgan los alumnos. Esta conducta del docente parece estar más relacionada con su personalidad, siendo que este estudio encontró que maestros que fueron evaluados en más de una clase, mostraron la misma posición de subvaloración y sobrevaloración ante los diferentes grupos de alumnos. Es decir, no se encontró siquiera un caso en el que el maestro se subvalorara en una clase y se sobrevalorara en otra.

Los resultados de este trabajo de investigación no permiten afirmar que el tamaño de la clase o el género del maestro sean predictores de la evaluación que los alumnos hacen del proceso enseñanza-aprendizaje.

La evaluación del docente debe ser

participativa y multifacética, lo cual permite una comprensión más amplia e integrada de lo que sucede en el aula de clase. Especialmente cuando se trata de evaluación sumativa relacionada con el futuro del maestro.

El mejor uso que puede darse a la evaluación del maestro por parte del estudiante es el de facilitar al docente información relevante que le permita conocer la forma como el alumno percibe su desempeño. Este tipo de evaluación debe tener un enfoque formativo para el docente y el proceso. Griffin y Pool (1998) concluyeron que evaluaciones repetidas con participación de los estudiantes y consultas a lo largo del proceso conducen a una mejor respuesta a las necesidades de los estudiantes y al mejoramiento del desempeño instruccional. Queda en el ambiente la pregunta respecto a la validez de las evaluaciones del estudiante respecto a la instrucción y si la autonomía del estudiante y otras variables influyen en la puntuación que otorgan los alumnos.

Referencias

- Abrami, P. C., d'Appolina, S. y Cohen, P. A. (1990). Validity of student ratings of instruction: What we do not. *Journal of Educational Psychology*, 82, 219-231.
- Cohen, Peter A. (1987, abril). *A critical analysis and reanalysis of the multisection validity meta-analysis*. Documento presentado en la reunión anual de la American Educational Research Association, Washington, D.C.
- d'Apollonia, S., Lou, Y. y Abrami, P. C. (1998, abril). *Making the grade: A meta-analysis on the influence of grade inflation on student ratings*. Documento presentado en la reunión anual de la American Educational Research Association, New York.
- Feldman, K. A. (1989). The Association between student ratings of specific instructional dimensions and student achievement: Redefining and extending the synthesis of data from multisection validity studies. *Research in Higher Education*, 30, 583-645.
- Greenwald, A. G. (1997). Validity concerns and

GRAJALES GUERRA

- usefulness of student ratings of instruction. *American Psychologist*, 52, 1182-1186.
- Greenwald, A. G. y Gillmore, G. M. (1997). Grading leniency is a removable contaminant of student ratings. *American Psychologist*, 52, 1209-1217.
- Griffin, Bryan W. y Pool, Harbison. (1998). Monitoring and improving instructional practices: Are the student evaluation valid? *Journal of Research and Development in Education*, 32(1), 1-8.
- Howard, G. S. y Maxwell, S.E.(1982). Do grades contaminate student evaluations of instruction? *Research in Higher Education*, 16, 175-188.
- Marsh, H. W. y Dunkin, M. (1992). Students' evaluations of university teaching: A multidimensional perspective. En *Higher Education: Handbook on theory and research* (Vol. 8, pp. 143-234). New York: Agathon.
- Marsh, H. W. and Roche. L. A. (1997). Making students' evaluations of teaching effectiveness effective: The critical issues of validity, bias, and utility. *American Psychologist*, 52, 1187-1197.
- Sánchez Guzmán, Francisco. (1981). *Introducción al estudio de la administración*. México: Limusa.
- Scriven, Michael. (1995). *Student ratings offer useful input to teacher evaluations*. (Report N° EDO-TM-95-8). Washington, DC: The Catholic University of America, Department of Education. (ERIC Document Reproduction Service N° ED 398 240).