Casigal García
José Luis.
Administración

ISSN 1390-7719

REVISTA PUCE 102

PONTIFICIA UNIVERSIDAD CATÓLICA DEL ECUADOR

NÚMERO GENERAL

3 DE NOVIEMBRE 2015-3 DE MAYO 2016
QUITO, ECUADOR

Centro de Publicaciones

PONTIFICIA UNIVERSIDAD
CATÓLICA DEL ECUADOR

Centro de Publicaciones
Revista PUCE
Número General
Quito, Ecuador

Rector
Dr. Fernando Ponce León, S. J.

Vicerrector
Dr. Fernando Barredo Heinert, S. J.

Directora General Académica
Dra. Graciela Monesterolo Lencioni

Director del Centro de Publicaciones
Mtr. César Eduardo Carrión

Comité Ejecutivo del Centro de Publicaciones
(Comité Editorial):
Dra. Mercedes Malla Simon
Dr. León Espinosa O.
Dr. Álvaro Mejía S.
Mtr. César Eduardo Carrión

Presidente (Editor en Jefe)
Dr. Alberto Rengifo A.

Secretario (Coordinador del Comité Editorial)
Lic. Walter Jiménez S.

Autores que pertenecen a la PUCE
Andrade Susana
Argüello Susana
Buitrón María Elena
Cagigal José Luis
Dávila Ana Gabriela
De la Torre David
Díaz Ana
Estrella Ana
García María Verónica
Gea-Izquierdo Enrique
Gordillo Ruth
Granja Jorge
Lyco Halina
Maldonado Priscila
Nicolalde Damián
Paredes Ana
Pareja Daniela
Robayo Aníbal
Sarmiento Marisol
Segovia Fausto
Schweider Peter
Stefos Efstathios
Wong María Belén

Autores externos a la PUCE
Estrella Andrés
Gavidia José (Ministerio del Interior, Comandancia General de Policía Nacional, Quito, Ecuador).
Gea Izquierdo Enrique (Universidad de Málaga, Málaga, España).
Nicolalde Daniel
Ortega José Luis (Universidad de Granada, Facultad de Ciencias de la Educación, Granada, España).
Palacios Teresa (Universidad Central del Ecuador, Facultad de Ingeniería, Minas, Petróleos y Ambiental, Quito, Ecuador).
Peña Sonia (Universidad de Granada, Facultad de Ciencias de la Educación, Granada, España).
Valencia Mónica (Universidad de Carabobo, Facultad de Ciencias de la Educación, Carabobo, Venezuela).

Corrector de estilo y ortografía
Alfonso Sánchez

Colección n.° 102
3 de mayo de 2016
Publicación semestral
ISSN: 13907719
Registro de Derecho Autoral: 010645
Ingresada al Catálogo Latindex. Folio: 21880

La Revista de la Pontificia Universidad Católica del Ecuador es una publicación semestral (mayo y noviembre) de su Centro de Publicaciones. Difunde trabajos académicos y científicos, estrictamente originales en español, en las área de Leyes, Pedagogía, Ingeniería, Economía, Biología, Química, Historia, Geografía, Antropología, Sociología, Filosofía, Teología, Comunicación, Lingüística, Literatura, Medicina, Administración, Arquitectura, Gestión Social, Psicología y Diseño. El propósito de la Revista PUCE es difundir conocimientos, intercambiar experiencias e incentivar la producción del pensamiento especializado. El contenido de esta revista está dirigido a docentes, investigadores, estudiantes universitarios y público en general. Esta revista es arbitrada por especialistas de indiscutible valor, cuyos nombres se mantienen en absoluta confidencialidad.

Los artículos son de responsabilidad exclusiva de sus autores.
Los derechos de autor son exclusivos de la PUCE.
Se prohíbe la reimpresión parcial o total con cualquier finalidad ***sin el previo consentimiento por escrito del Centro de Publicaciones de la PUCE.***

Editorial (Punto de venta)
Centro de Publicaciones PUCE
Av. 12 de Octubre y Robles
Apartado n.° 1701-2184
Telèf.: 593-02-291700
2991 700 (Tronical). Extensiones: 1013, 1711, 1122
Telèf.: 593-02-2991711, ext.: 1927
atengifo@puce.edu.ec
Quito, Ecuador

Impresión
PPL Impresores. 2529762
Quito, Ecuador
pplimpresores@gmail.com

Revista PUCE. ISSN 1012-389X. Núm. 102
3 nov 2015-3 mayo 2016. Cagigal, pp. 169-198

RELACIONES ENTRE NOTAS DEL ALUMNO Y EVALUACIONES AL PROFESOR

José Luis Cagigal G.[1]

RESUMEN

Para definir si existe reciprocidad entre notas que da el profesor y evaluaciones que sus alumnos le conceden, se compararon estas variables con muestras independientes de tres semestres. A pesar que algunos docentes piensan que las calificaciones dadas a los estudiantes del curso inciden en su evaluación, este estudio muestra que no existe dicha reciprocidad.

Sin embargo, sí existe reciprocidad entre la expectativa que tiene el estudiante de sus notas y la evaluación que da al profesor, definida esa expectativa como la diferencia entre la nota en la materia evaluada y el promedio de sus notas en la carrera.

El estudio descubre una implícita aceptación, por parte de los alumnos, que se prefiere la dispersión en las notas y no la uniformidad. Del análisis de las distribuciones de las Evaluaciones al Docente (EAD), se sugieren algunas medidas para mejorar su aplicación.

Palabras clave: evaluaciones a docentes, notas, calidad docente, gestión académica

ABSTRACT

To determine whether there is reciprocity between grades and the student evaluations to teachers, both variables were compared with independent samples over three semesters. Although some teachers think that the course grades given to the students influence their assessment, this study shows that there is no such reciprocity.

However, this reciprocity exists between the expectation that the student

[1] Pontificia Universidad Católica del Ecuador, Facultad de Ciencias Administrativas, Quito, Ecuador (jlcagigal@hotmail.com).

has on his grades and the assessment given to the teacher, expectations defined as the difference between the grade on the subject tested and the average of his grades in the school have been witnessed.

The study finds an implicit acceptance by the students for dispersion in the grades and not for uniformity. From the analysis of the distributions of the EAD some measures are suggested to improve its implementation

Key words: student evaluations, grades, quality of teaching, academic management

INTRODUCCIÓN

Las evaluaciones de los estudiantes a los docentes (EAD), introducidas en las universidades norteamericanas en la segunda década del siglo pasado, como mecanismo de medición de la calidad de la docencia, se generalizaron rápidamente en ese país (Seldim, 1993) y han sido utilizadas como un instrumento para impulsar el mejoramiento de la calidad docente, particularmente en las universidades.

El concepto de calidad de la docencia, en sí mismo, es una idea abstracta bastante ambigua, por sus variadas percepciones, y por ello el uso de las EAD, como instrumento de mejoramiento de la calidad docente, puede ser ineficaz. Para ello, las EAD deben ajustarse a procesos correctos y ser usadas cuidadosamente.

Han trascurrido casi nueve décadas desde que Remmers y Brandenburg (1927) publicaron la primera investigación sobre el tema, sin que se haya podido establecer conclusiones definitivas sobre la validez y confiabilidad de las EAD para el fin esperado de medir la calidad profesional de los docentes. Sin embargo, ninguna otra modalidad de evaluación a los docentes, tales como las visitas al aula, las autoevaluaciones, el análisis de la distribución de notas, el análisis del portafolio, ha sido usada y analizada con tal esmero como el procedimiento de encuestas a los alumnos, sobre todo en el ámbito universitario.

A medida que el uso de los puntajes resultantes de las EAD, sobrepasaron los fines primarios relacionados con el monitoreo administrativo y de retroalimentación al docente, para convertirse en indicadores orientados al reconocimiento, selección, promoción o asignación de cursos a los docentes; la controversia en torno a la validez de los instrumentos y de los resultados de estas evaluaciones ha ido creciendo y

ello ha generado una polémica en la que las EAD son cuestionadas por unos y defendidas por otros, a través de investigaciones que intentan probar una posición a favor o en contra de su uso, como instrumento de medición de la calidad docente. La principal objeción en contra de los resultados de las evaluaciones asocia una posible reciprocidad entre las notas que el docente entrega a los alumnos y las evaluaciones que el alumno otorga al profesor. Los estudios presentados se muestran contradictorios en sus conclusiones y, a pesar de la cantidad de datos acumulados y analizados durante años por diversos investigadores, aún se mantiene el debate y el cuestionamiento, probablemente intuitivo, a su significado y a su uso, particularmente para los fines de reconocimiento de méritos académicos.

La falta de acuerdos en las conclusiones generales ha originado nuevas investigaciones posteriores derivadas de la diversidad de los resultados presentados, en las que se analizan algunos de los diversos factores adicionales que pueden incidir en las EAD, tales como la dificultad del contenido de las materias, el tamaño de clases, el nivel del curso, el tipo de curso (obligatorio u optativo), el interés del tema, la carga de trabajo para el estudiante, las formas de evaluación en la materia y otras variables que diferencian los cursos y los grupos de materias, al igual que el análisis del contenido de los instrumentos de medición y sus características en los que puede darse una sensibilidad considerable de las respuestas a la formulación del texto, como también al momento de aplicación del cuestionario y la forma de implementación. La variedad de estudios realizados en torno a las EAD y los factores que en ellas inciden, solo resaltan la complejidad del tema, pues su análisis sobrepasa a la simple asociación de una potencial entrega recíproca de facilidades entre los profesores y los estudiantes.

Es importante destacar que, en cualquier evaluación, sea la de estudiantes a docentes, o la recíproca, de profesores a alumnos, el propósito de la misma y su forma de implementación son elementos claves que inciden tanto en los enfoques usados, como en los resultados que muestran. En el proceso educativo, la evaluación se orienta a la búsqueda del mejoramiento de la persona evaluada (evaluación formativa). Sin embargo, las evaluaciones terminan en una calificación del desempeño (evaluación sumativa), que se utiliza para comparaciones entre los evaluados o también con respecto a un estándar predefinido (evaluación basada en criterios).

Mientras las EAD se limiten a su carácter formativo, como un instrumento para que el profesor, gracias a una retroalimentación de parte de sus alumnos, le permita mejorar sus formas de enseñar,

es más fácil que los resultados puedan ser aceptados por el docente y que pueda acoger las sugerencias sin asumir una posición defensiva frente a los mismos.

La polémica sobre las EAD se agudiza aún más cuando los directivos de la universidad, recogen estos resultados y establecen comparaciones con los otros docentes o realizan la divulgación de los resultados, tratando con ello de impulsar el mejoramiento de la eficacia y la eficiencia de la función docente, y, peor aún si se establecen normas en las que los resultados de las evaluaciones sirven para fines administrativos de escalafón docente, lo que afecta a los programas establecidos de carrera universitaria de los docentes y, a la larga, a sus remuneraciones. En este entorno de estímulo-respuesta, es probable que los docentes quieran buscar las formas de mejorar estos resultados e intenten ajustar sus prácticas en el aula, no necesariamente hacia un mejoramiento del proceso de enseñanza-aprendizaje, sino a lo que ellos creen que puedan ser factores que incidan en una mejor o peor evaluación por parte de los educandos.

La idea primaria de explicar las diferencias, entre otras causas, bajo el supuesto no probado que los alumnos otorgan a sus profesores mejores evaluaciones cuando estos conceden en sus cursos altas calificaciones, puede reflejarse en menor exigencia en el trabajo de clase, lo cual ocasiona un ascenso general de las notas asignadas a los alumnos, sin que con ello mejore ni el nivel de aprendizaje en el curso ni la pedagogía empleada.

En la línea de la reciprocidad – del "toma y daca" o del "quid pro quo" – varios autores (Baldwin & Blattner, 2003; Krautmann & Sander, 1999; Sojka, Gupta, & Deeter-Schmelz, 2002) infieren la existencia de la reciprocidad entre notas y EAD, al probar que hay una correlación positiva entre las evaluaciones docentes con las notas de los alumnos.

Más allá de ello, hay investigadores que asocian este posible trueque de altas notas con una degradación de la calidad denominada "inflación de las notas" (Rosovsky y Hartley, 2002); (McDonald y Johnson, 2003), esto es, un aumento sistemático de los puntajes promedios de los estudiantes en las universidades, bajo el supuesto de que la presunta reciprocidad incentiva a los profesores a ajustar las calificaciones dadas a los alumnos con el fin de mejorar sus evaluaciones como maestros.

Otras investigaciones, en un sentido paralelo al de la reciprocidad, asocian los resultados de mejores evaluaciones a una aceptación y a una supuesta búsqueda de un mayor facilismo en las notas, por parte de los alumnos. (Rodin

y Rodin, 1972; Greenwald y Gillmore, 1997) o también la preferencia del alumno por menores cargas de trabajo en las materias (Miron y Segal, 1986).

En otra línea de investigación experimental, se intenta relacionar el grado de aceptación alcanzado ante los estudiantes, con los factores que demuestran una agradable personalidad (Cabezas, Christian y Mancheno Paulina, 2008)[2], o el uso del entretenimiento en el aula y otros similares, lo que puede inducir un tipo de diversión en el aula orientada a generar un efecto de espectáculo o efecto "payaso", denominado "Dr. Fox efect"[3] (Naftulin, Ware, y Donnelly, 1973).

En contraposición con los estudios citados, otros autores (Centra, 2003; Cashin, 1995), sostienen que numerosas investigaciones sobre EAD, demuestran que las evaluaciones de los alumnos a los profesores son confiables, válidas y útiles para mejorar la enseñanza, además de ser multidimensionales en lo que miden.

Dada la complejidad real de medir la calidad de la docencia, los académicos han elaborado variados cuestionarios para realizar las EAD y han incorporado la ponderación de diversos factores asociados a las percepciones variadas sobre el concepto de calidad en la enseñanza. March (1987) y Feldman (1997) y otros, identificaron múltiples dimensiones de la calidad en la docencia, entre ellos el valor de lo aprendido, la organización, el entusiasmo, la interacción del grupo, la relación con el docente, la cobertura y profundidad de lo enseñado, las formas de calificación, la carga de trabajo y su dificultad, como variables que han servido para el diseño de estos cuestionarios. Pocos estudios se han realizado sobre la elevada colinearidad entre las variables utilizadas, aunque en su formulación, las afirmaciones dadas parecen independientes entre sí.[4]

Para algunos, el criterio esencial de calidad de la docencia debería ser el aprendizaje logrado por el alumno. Sin embargo, si se intentase medir la calidad de un docente por el grado de aprendi-

[2] Christian Cabezas señala que los docentes de alto rendimiento en la PUCE presentan un pensamiento abstracto, son más aventureros (extrovertidos) que conservadores, y son más emotivos que prácticos; en cambio, los de bajo rendimiento presentan un pensamiento concreto, son más conservadores que aventureros y son más prácticos que emotivos.

[3] http://www.er.uqam.ca/nobel/r30034/PSY4180/Pages/Naftulin.html

[4] En un estudio anterior realizado por el autor se puede observar que las correlaciones entre los promedios en las preguntas evaluadas en un grupo de alumnos, con inclusión de aquellas cuestiones relativas a variables objetivas como asistencia y puntualidad, comparadas con preguntas sobre metodología, presentaban coeficientes de correlación superiores a 0,76 en casi todos los grupos. Esto apunta a que la percepción sobre la calidad del docente frente a la que el grupo evaluador tiende a ser global y no diferencia mucho el significado específico de cada pregunta, por lo que evalúa todas las facetas de la actividad del docente bajo un prisma común de bueno, mediocre o malo, independientemente de su realidad.

zaje alcanzado por el curso, se da el caso, que el temor generado por un profesor muy exigente logra mejores resultados de aprendizaje, debido al gran esfuerzo de los alumnos, sin que medie una forma de enseñanza que facilite en forma efectiva dicho aprendizaje.

Fuera de las universidades norteamericanas, el tema de las EAD ha sido poco estudiado. Sin embargo, las políticas de calidad centradas en las evaluaciones a los docentes se han generalizado en las universidades del mundo y se ha enfocado su uso para la promoción del profesorado y, no tanto para el mejoramiento de calidad, pues existe el riesgo que su uso indiscriminado pueda contribuir a una disminución de la misma en el largo plazo.

ANTECEDENTES DE LA INVESTIGACIÓN

La presente investigación que se realiza con datos obtenidos de la Facultad de Ciencias Administrativas y Contables (FCAC) de la PUCE en los años 2006 y 2007. Esta facultad introdujo las EAD en 1986.

El Dr. Jaime Moreno, al recoger opiniones de docentes y alumnos en grupos focales de varias facultades, realiza un análisis estadístico de los resultados de las evaluaciones y construye instrumentos de evaluación docente que se incorporan a la web para su aplicación general en la PUCE, basados en grupos de preguntas asociadas a varias áreas de competencias pedagógicas, según puede observarse en el Cuadro n.º 1. El alumno califica al docente contestando el cuestionario según una escala de Likert de 5 a 0, según su acuerdo o desacuerdo con las afirmaciones dadas, todas expresadas en positivo, lo cual puede inducir a una calificación uniforme por columnas. En ese momento solo conoce notas parciales. La nota obtenida por el docente es el promedio simple de los puntajes recibidos.

Desde el año 2002, para el ascenso de categoría de los profesores, consta en el capítulo seis del Reglamento General Académico, como uno de los requisitos el haber obtenido puntajes por lo menos iguales al promedio de la respectiva unidad académica en la evaluación del desempeño, la cual incluye como parte de la misma las evaluaciones de los estudiantes.

METODOLOGÍA DEL ESTUDIO

El estudio recoge los datos de las EAD y de las notas de la población de estudiantes de la FCAC durante tres semestres y cada uno se analiza en forma independiente, para que se pueda verificar la consistencia de resultados en distintos períodos.

Se asume como hipótesis inicial la *no* dependencia entre las dos variables indicadas, notas de los alumnos en una materia y evaluaciones recibidas por el docente de parte de los alumnos en esa materia. El análisis inicial relaciona indicadores estadísticos de las notas globales en cada materia con el promedio de notas obtenido por el docente en la evaluación de cada curso, considerada la variable dependiente. Como variables dependientes se usan los indicadores de notas de alumnos, tales como el promedio del curso en la materia, pero se amplían a medidas de la distribución de notas, tal como se describen a continuación, en la definición de indicadores estadísticos utilizados. Las variables independientes se obtienen de las distribuciones de notas publicadas que el profesor ha asignado a los alumnos del curso, en la nota parcial sobre 30 (S30) antes del examen final y próxima a la evaluación al docente, nota del examen final sobre 20 (S20), cuando la evaluación al docente ya se ha realizado y la suma de notas anteriores (S50), sobre las cuales se calculan los indicadores estadísticos del Cuadro n.º 2.

Para evitar la incidencia del tamaño de los grupos y de la posible deserción de alumnos, que no completan el curso, se realiza una depuración de datos. Se eliminaron los cursos con un número de alumnos inferior a 12, para excluir el posible impacto del tamaño del curso en los resultados y se eliminaron también de las listas de notas, aquellas de valor nulo en alguno de los datos, y se asumió el retiro de la materia. Sin embargo, esto no garantiza que algunos de estos alumnos eliminados en la lista de notas no hayan participado en la evaluación al profesor.[5]

Para establecer la relación de asociación, sea positiva, nula o negativa, entre la variable dependiente *(promedio de la evaluación dada por el curso al do-*

[5] La eliminación de las notas con valor nulo, incide también en los resultados obtenidos de las distribuciones, medidas de tendencia central y de dispersión. Se asume para justificar esta decisión en la investigación que una nota nula sobre 20 o sobre 30, más se debe al retiro del estudiante de la materia y no a una nota real dada por el docente. Sin embargo, si aquel alumno que se ha retirado realiza la evaluación al docente y lo califica superficialmente, sus notas totales bajas, asociadas a una evaluación al profesor recíproca, aumentaría el valor de los coeficientes de correlación entre notas y evaluaciones, más aún considerando el efecto en el promedio de ambas variables.

cente) y cada una de las distintas variables independientes anteriormente indicadas se calcula el coeficiente de correlación, esto es un número *entre -1 y +1*, expresado en tanto por ciento. Por ello, el coeficiente de correlación mide la variación conjunta de datos, esto es el grado en el que la variable dependiente es mayor o menor en similar proporción a la variable independiente.

En cada semestre, después de depurar los datos, las muestras seleccionadas son superiores a los 294 docentes, lo que les hace suficientemente amplias para realizar inferencias estadísticamente válidas con un nivel de confianza de 95%, para coeficientes de correlación superiores a 9,64 % y para aquellas muestras de la tercera parte, alrededor de cien datos, se debe superar el 16,9 % como coeficiente de correlación, basados en la fórmula[6] del valor t correspondiente al nivel de confianza esperado

$$t = r/[(1-r^2)/(n-2)]^{1/2}$$

HIPÓTESIS SOBRE LOS RESULTADOS ESPERADOS

De conformidad con la relación esperada entre las variables estadísticas escogidas con la variable independiente que es el valor de la evaluación obtenida por el docente, se asume que a un mayor valor en las notas promedio de los grupos de alumnos debería corresponder a una mejor evaluación a los docentes, esto es, debería obtenerse un valor significativo de correlación positiva; por ello, todas las medidas de posición, promedio, cuartiles, Q1 (cuarti1 1), Q2 (mediana) y Q3 (cuartil3) deberían tener una correlación positiva con la evaluación dada a los docentes. Esto es, cuanto mayor sea el valor de la nota promedio del curso, en la media, mediana y cuartiles, en las notas de los alumnos, se espera una mayor "generosidad" de los estudiantes en la evaluación al docente.

Igualmente, si las notas de los alumnos se agrupan en mayor proporción (%) por encima de los parámetros fijos establecidos como porcentaje del límite superior M, (sea este 30, 20 o 50, respectivamente), P>0,6M (M=Máxima), P>0,7M, P>0,8M, P>0,9M, en particular los más altos, se esperaría una correlación significativamente positiva, pues mayor número de alumnos se agrupan en los intervalos superiores. Estos supuestos se resumen en el Cuadro n.º 3.

A pesar de ser muestras independientes para cada semestre, se pudiera asumir la consistencia en los signos de los valores de la correlación obtenida.

[6] Ver Downie N. y Heath, R. *Métodos estadísticos aplicados*, p. 249.

No se asumen supuestos previos referentes a las medidas de dispersión, como son la desviación estándar y la desviación intercuartílica, ni tampoco para los coeficientes de asimetría y la curtosis, indicadores utilizados para el análisis descriptivo de las distribuciones de notas y de las EAD.

ANÁLISIS DE LOS RESULTADOS

Características de las distribuciones de las notas

Los indicadores de tendencia central, dispersión y momentos de la distribución de notas S30, S20 y S50 dadas por los profesores a los alumnos, presentan valores y características comunes, tal como puede verse en la Tabla n.º 1. El promedio y la mediana (cuartil 2) de las notas con valores de 39,4 y 45,0 respectivamente, se ubican por encima del 79% de la escala del rango de las notas (entre 0 y la máxima) y notablemente por encima de la nota mínima de pase que es 30 sobre 50 en la nota final. El 92 % de las calificaciones finales se ubican por encima de la nota de pase (30/50), y el 62 % de las notas finales están entre 40 y 50 puntos. Lo anterior muestra una asimetría hacia notas altas, –mediana superior al promedio– y una distribución inclinada hacia la derecha respecto a la curva normal (curtosis positiva), tal como puede observarse en los gráficos 1 y 2 anexos.

Por ello, todas las distribuciones de notas tienen valores negativos en los coeficientes de asimetría, lo que muestra una acumulación de puntajes en las notas altas, si se compara la distribución con la curva normal centrada en el rango de pase, tal como se puede apreciar en los gráficos de las notas de uno de los semestres.

Predictibilidad entre las notas de los alumnos

Al disponer de la nota parcial (S30) y la nota (S20) del examen final se realizó un análisis del índice de predictibilidad entre ambas mediante la correlación entre ellas. Esta se ubica consistentemente entre 42 % y 43 % en cada uno de los tres períodos académicos examinados. Pudiera esperarse que la nota del examen final (S20) está asociada más estrechamente con la nota previa acumulada (S30). Una correlación de 43 % significaría que la nota anterior S30 apenas explica un 18 % de la nota del examen final, lo que implica que el 82 % restante corresponde a otros factores y no se relaciona con el trabajo valorado en las tres primeras notas sobre diez, que corresponden a las mediciones del aprendizaje y el esfuerzo durante el semestre. Puede contribuir a ello, el hecho

frecuente que los estudiantes pueden concentrar más su esfuerzo en el examen final en aquellas materias en las que tienen notas bajas y pueden descuidar aquellas en las que ya están seguros. También puede contribuir a ello el hecho que la nota S20 suele ser fundamentalmente un examen escrito y no tiene componentes de trabajos colectivos.

Si se comparan los datos que constan en la Tabla n.º 2, se observa que los indicadores de posición, promedio y mediana, se mantienen en similares niveles, en ambas notas sobre 30 y sobre 20. Sin embargo, en forma consistente los dos indicadores de dispersión relativa, el coeficiente de variación y la razón de desviación intercuartílica/mediana, en la nota del examen final S20, son superiores a los de la nota S30, lo cual indica que los docentes dispersan más las notas en el examen final que en la nota S30. Esto puede deberse a que, en la nota S30, por las formas de evaluar, tales como notas de grupo y talleres, las calificaciones se hacen más homogéneas la nota única del examen final o también pudiera ser resultante del menor condicionamiento del docente, al final del semestre, cuando ya han sido realizadas las EAD.

Características de las distribuciones de las EAD

Los indicadores de tendencia central, dispersión y momentos de las EAD en los tres semestres analizados presentan valores y características comunes y consistentes, tal como puede verse en la Tabla n.º 3. El promedio de las EAD está cercano a 90/100 en la escala total; más de las tres cuartas partes de los docentes reciben notas superiores a 86/100, una cuarta parte de los docentes supera el 95,7/100 y un 4 % recibe notas máximas superiores de 99/100. El 25 % de los peores evaluados recibe notas entre 38 y 86 sobre 100, con una amplia dispersión entre ellos. Existe una elevada asimetría (entre -1,52 y -1,94) en la distribución de las EAD, como consta en la tabla n.º 3 y se ve en el gráfico n.º 3. La concentración de las EAD en puntajes altos incide mucho en el grupo asignado al docente según cuartiles. Por ejemplo, una nota nula en un grupo de 20 alumnos, baja en 5 puntos porcentuales la calificación promedio del profesor, pasándolo al grupo siguiente inferior.[7]

Relación entre las EAD y las notas del curso

Los resultados obtenidos de la aplicación del coeficiente de correlación

[7] El impacto de la incidencia de las evaluaciones extremas debería ser filtrado como "outliers"- fuera de rango- por el sistema informático, eliminando calificaciones uniformes en el último rango. Puede darse, entre otros motivos, por problemas de relación, disciplina, copia en pruebas, hechos no académicos, y deberían eliminarse del cálculo del promedio por su marcado sesgo emocional.

de Pearson aplicada al promedio de todas las EAD para cada docente con cada uno de los indicadores de las distribuciones de las tres notas, S20, S30 y S50 en cada semestre consta en la tabla No 4 muestran que las correlaciones no son consistentes ni en valores ni en el signo esperado de la correlación, en las tres muestras semestrales independientes analizadas, por lo cual puede concluirse que la relación entre EAD y notas promedio de los cursos, es claramente aleatoria.

En contraste, las EAD en el examen final S20, llevan signos contrarios a los esperados en 22 de 24 datos, –aunque sin valores significativos– esto es, mayores EAD corresponden a menores notas en esa última nota. Si se considera que la nota S20 es posterior a la fecha de realización de las EAD, la contradicción puede sugerir que el docente actúa con más independencia en el examen final o puede deberse a las formas de calificación, examen versus trabajos. Ya se mencionó que las notas S20 del examen final, muestran mayor dispersión que las S30 y tienen baja correlación entre ellas.

Por último, dado que las distribuciones de EAD y de notas no tienen las características de una distribución normal, para completar el análisis de resultados se verificó, por técnicas no paramétricas, la consistencia de signos y valores con lo esperado, y para ello se hizo el análisis de correlaciones mediante el coeficiente de correlación de Spearman[8], denominado correlación de puestos, en vez de coeficiente de correlación de Pearson, utilizado antes. En los resultados se mantiene la inconsistencia citada en los signos y valores, por lo que se mantienen las conclusiones anteriores.

Del análisis se concluye que no hay evidencia estadística para inferir que en las EAD exista una asociación positiva entre las notas del curso y las EAD. Esto no descarta que a nivel individual puedan existir casos aislados de reciprocidad, pero su incidencia no es general en el conjunto de notas impartidas.

En la tabla n.º 4, aparece un resultado interesante al relacionar el promedio de las EAD con las dos medidas de dispersión, la desviación intercuartílica (DIQ) y la desviación estándar (SIGMA), pues aparece una consistencia completa en los signos en todos los nueve casos de notas y semestres, aunque solo en tres valores en la DIQ y uno en la SIGMA los valores son significativos a un 95 % de confianza. La interpretación de esta consistencia apunta a que en la evaluación a los docentes los alumnos pre-

[8] El coeficiente ρ de Spearman –estadístico no paramétrico– no requiere que la distribución de las variables sea normal y que la relación entre ellas sea lineal. Mide la existencia de una tendencia monótona creciente o decreciente entre las variables al ordenarlas por puestos.

fieren dispersión en las notas de los alumnos del grupo más que notas altas uniformes e indiferenciadas, lo que de alguna manera indica un reconocimiento a una diferencia equitativa en las notas del docente.

Análisis de las EAD según grupos de docentes

Se investigó también si las EAD mantenían sus relaciones por grupos de docentes, los mejor calificados (SUP), los intermedios (MED) y los peor calificados (INF).

En las notas del examen final S20 (Tabla n.º 5) de grupo INF –profesores con menores evaluaciones– aparece claramente una consistencia en la correlación negativa entre EAD y notas en los tres semestres con la mayoría de los valores significativos al 95 %. Esta consistencia especial no ocurre en el grupo medio ni en el superior. Este grupo INF representa una tercera parte de los docentes. Es el grupo en el que más se dispersan los puntajes de las EAD, entre 90 % y 37 % de la escala. Este resultado contradice en el supuesto de una reciprocidad entre notas dadas y evaluaciones recibidas y plantea un interrogante complejo, pues significaría que, en el grupo inferior, los profesores que tienen una mejor evaluación han otorgado notas finales S20 más bajas. Esta paradoja plantea un interrogante: ¿La mayor dispersión y exigencia en notas es percibida antes y valorada por los alumnos?

Del análisis de grupos se mantiene que no hay existencia de reciprocidad entre las EAD y las notas. tanto para el grupo general de todos los docentes (TT) como para los grupos mejor, medio y peor evaluados, aunque aparece un interrogante sobre exigencia y EAD.

Opiniones de los alumnos y de los docentes[9]

En la marcha de la investigación, se hizo un sondeo a docentes (20) y estudiantes (72) sobre su percepción del valor de las EAD, como instrumento válido y relevante para medir la calidad docente, la objetividad de las respuestas de alumnos y la supuesta reciprocidad.

Los resultados de esta encuesta (Tabla No 6), indican la duda de alumnos y de los docentes sobre la reflexividad y validez de las EAD. Un 64 % de los alumnos encuestados considera que las EAD rara vez son válidas para medir la calidad del docente y un 50 % de los profesores opina lo mismo. El 64 % de los alumnos y el 75 % de los docentes cree que rara

[9] Como mencionan los autores Miron, Mordechai y Segal, Esther, en su artículo "Opinión de los estudiantes sobre la validez de las EAD", poca investigación se ha hecho de las percepciones de los que evalúan. Ver http://www.jstor.org/stable/3446689

vez se realizan en forma objetiva y reflexiva. Este resultado se relaciona con las inconsistencias en los resultados estadísticos analizados.

En el tema de reciprocidad entre notas y evaluaciones, difiere la posición de los alumnos y docentes. Los estudiantes niegan que esta exista en la mayoría de los casos (46 %), aunque un 47 % de los alumnos aceptan que hay bastante reciprocidad. En contraste, un 50 % de los profesores considera que hay un exceso de reciprocidad y un 35 % considera que hay bastante reciprocidad. La percepción de exceso o bastante reciprocidad suma 54 % en alumnos y 85 % en profesores.

Conocer las causas del pobre significado dado por los alumnos a las EAD permitiría enfocar las razones de las dudas de su validez y, a la par, saber lo que lo ellos realmente evalúan y por qué lo hacen así y definir qué se requiere para que este importante insumo de información a docentes y autoridades se realice con la debida reflexión y criterio por parte de los alumnos.

Ampliación del estudio

Para contrastar los resultados globales obtenidos, se obtuvo información adicional sobre las EAD individuales de los alumnos en el segundo semestre de 2008 y de sus notas (6145 asignatura-alumno) tanto en la materia en ese semestre y como del promedio de la carrera, con el fin de confirmar o no la relación de reciprocidad y analizar la incidencia que pudiera tener la expectativa de la nota por parte del alumno y su evaluación al docente.

Los resultados generales de correlación entre notas y las EAD confirmó lo anteriormente dicho, esto es, que no hay una correlación significativa entre las notas que el alumno recibe y las evaluaciones de los alumnos al docente.

Para analizar la expectativa de nota, se introdujo una nueva variable en el análisis que fue *la diferencia entre la nota del alumno en la materia y su promedio en la carrera*.[10] Esta variable denominada diferencia en expectativas, arroja un resultado interesante de una correlación positiva (13,6 %) consistente en todos los niveles entre esta variable y las EAD, estadísticamente significativa. Esto representaría que si el alumno espera lograr en la materia una nota superior a su promedio en la carrera tiende a calificar mejor al profesor y si este alumno prevé en ese curso una nota inferior a su promedio general, tiende a calificar peor al profesor, al considerar que la realización de las EAD precede a la nota final completa.

[10] Cabe destacar la alta correlación entre las notas de los alumnos en las materias del semestre analizado y la nota promedio de la carrera que es de 70,9 %.

La información recopilada permitió analizar el nivel de homogeneidad o de dispersión en las notas por preguntas o ítem dadas a los docentes por cada alumno. En un cuestionario de 15 ítem evaluados, un 60,2 % de los alumnos asignó una única calificación al profesor en la escala de 1 a 5. En esta homogeneidad destaca la asignación de la nota más alta (5) a todas las preguntas por parte del 57,0 % de los alumnos y de la nota mínima (1) a todas las preguntas por parte de un 1,2 % de los alumnos. Este resultado indica que la percepción del alumno hacia su profesor tiende a ser holística y que al contestar el cuestionario rara vez discrimina entre distintas preguntas al hacer su evaluación. Dada la escasa diferenciación en las respuestas y la percepción holística de su calidad se pueden acortar los cuestionarios o deben ser estructurados de modo que no todas las afirmaciones concuerden con la escala de mejor a peor.

CONCLUSIONES Y RECOMENDACIONES

Las EAD se establecen hace más de ocho décadas, en las universidades, como un instrumento para el mejoramiento de la calidad de la docencia. Su finalidad primaria de mejoramiento o información al docente, pero desde que las EAD se usan para las decisiones de promoción de los profesores, surgen numerosos estudios que cuestionan los resultados y los usos de las mismas. En la literatura investigativa sobre este tema existe una alta contradicción entre quienes objetan la validez de las EAD y los que la defienden.

Las opiniones de alumnos y docentes revelan coincidencias al cuestionar la validez y la seriedad en la realización de las EAD - aspecto que requiere ser investigado en sus razones - pero señalan posiciones distintas en el tema de la reciprocidad, aunque la mayoría de docentes y de alumnos reconoce la existencia de cierta reciprocidad entre las notas y las EAD.

Sin embargo, del análisis de los indicadores y notas en este estudio se concluye que no hay evidencia estadística para inferir que en las EAD exista una asociación positiva con las notas otorgadas a los alumnos del curso. Esto no descarta que a nivel individual puedan existir casos de reciprocidad, pero su incidencia no se refleja en los indicadores estadísticos generales, sino en la evaluación a docentes, en los que se da una calificación mínima a todas las preguntas, sesgo del resentimiento, que debe ser eliminado en el procesamiento de datos.

Un 60,2 % de los alumnos usa un solo número de la escala de 1 a 5 para las 15 preguntas del cuestionario. Dada

la percepción holística que predomina en los alumnos con relación a sus docentes, la amplitud del cuestionario es poco relevante. Existen factores más allá de las notas entregadas, tales como la simpatía, personalidad, dinamismo, tipo de asignatura que inciden en los resultados de las EAD y en esa percepción global que el alumno construye de su profesor.

Gracias a los sistemas informáticos actuales se puede mejorar tanto la presentación de la encuesta como su procesamiento para aumentar la calidad de la información que se genera. Así, para incitar a la reflexión sobre las preguntas del cuestionario, estas deben presentarse en forma aleatoria al alumno, una por una, que varíen la posición en la escala de calificación, repetición para incitar la consistencia y la reflexión. Debe haber una introducción que motive la reflexión, la objetividad y la diferenciación en las respuestas dadas. En el procesamiento, además de eliminar el sesgo del resentimiento, se puede ponderar las evaluaciones con el rendimiento global del estudiante en la carrera.

Las EAD constituyen un instrumento eficaz que debe mantenerse como medio para que el propio docente, si se hacen bien. Es un criterio parcial de la calidad de un docente y requiere ser complementada con otros indicadores de la calidad, como el nivel de aprendizaje alcanzado en valores, conocimientos, actitudes y destrezas, resultado que es altamente complejo evaluar.

ANEXOS:
CUADROS , TABLAS Y GRÁFICOS CITADOS

CUADRO n.° 1: CUESTIONARIO DE EVALUACIÓN A LOS PROFESORES

Preguntas por tipo de competencia

COMPETENCIA SOCIABILIDAD PEDAGÓGICA

1. Asiste a la actividad académica en los horarios previstos
2. Empieza la actividad académica con puntualidad
3. Trata a los estudiantes con cortesía y respeto
4. Aprovecha el tiempo sin extenderse más allá del término estipulado para la actividad académica
5. Hay un clima de confianza propicio a la comunicación

COMPETENCIA GESTIÓN DE CLASES

1. Estimula la participación de los estudiantes
2. Desarrolla de manera coherente las actividades académicas
3. Explica los criterios de evaluación que se emplean (pruebas, exámenes y otros)
4. Analiza con los estudiantes las calificaciones asignadas en las evaluaciones (pruebas, exámenes y otros)

COMPETENCIA IMPACTO EMOCIONAL

1. Promueve un ambiente de trabajo favorable al aprendizaje
2. Promueve actividades académicas como: consultas, investigación, salidas de campo, observaciones u otras
3. Su nivel de exigencia incentiva mi aprendizaje
4. Me gustaría volver a trabajar con este(a) profesor(a) en el futuro

COMPETENCIA REFUERZO DE HABILIDADES BÁSICAS

1. Estimula el análisis y la reflexión
2. Propicia la generación de nuevas ideas, enfoques, alternativas, opciones, soluciones, innovaciones.
3. Adquiero nuevos conocimientos o habilidades

CUADRO n.° 2: INDICADORES DE LAS VARIABLES INDEPENDIENTES UTILIZADOS EN LA INVESTIGACIÓN

Medidas de posición

NOMBRE	SÍMBOLO	CONCEPTO
MEDIA	MEAN	Promedio de notas de los alumnos del curso
MEDIANA	Q2	Nota que separa el 50% inferior en el curso
CUARTIL1	Q1	Nota que separa el 25% inferior en el curso
CUARTIL3	Q3	Nota que separa el 25% superior en el curso
P>.6	P>.6	% de alumnos con nota superior al 60% de la nota total. Corresponde al % del curso con notas iguales o superiores a 18/30; 12/20; 30/50 respectivamente.
P>.7	P>.7	% de alumnos con nota superior al 70% de la nota total. Corresponde al % del curso con notas iguales o superiores a 21/30; 14/20; 35/50 respectivamente.
P>.8	P>.8	% de alumnos con nota superior al 80% de la nota total. Corresponde al % del curso con notas iguales o superiores a 24/30; 16/20; 40/50 respectivamente.
P>.9	P>.9	% de alumnos con nota superior al 90% de la nota total. Corresponde al % del curso con notas iguales o superiores a 27/30; 18/20; 45/50 respectivamente.

Medidas de dispersión

NOMBRE	SÍMBOLO	CONCEPTO
DESVIACIÓN ESTÁNDAR	SIGMA	Raíz cuadrada de media de los cuadrados de las diferencias de notas con respecto al promedio
DESVIACIÓN INTERCUARTÍLICA	DIQ	Diferencia entre los valores del cuartil3 y el cuartil1 (Límites del 50% central)

Otras medidas de la distribución

NOMBRE	SÍMBOLO	CONCEPTO
COEFICIENTE DE ASIMETRÍA	ASIM.	Cociente de la media de los cubos de las desviaciones respecto del producto del promedio y el cubo de la desviación típica y describe el lado más extendido respecto al promedio. Si es positiva la distribución se extiende hacia el lado derecho y es negativa cuando se alarga en la dirección opuesta.
CURTOSIS	CURT.	Momento 3 que mide si la distribución es elevada o aplanada en comparación con la distribución normal. Es leptocúrtica si lleva el signo positivo y platocúrtica si es negativo.

CUADRO n.° 3
CORRELACIÓN ESPERADA ENTRE EL PROMEDIO DE LAS EAD CON INDICADORES ESTADÍSTICOS DE LA DISTRIBUCIÓN DE NOTAS

MEAN	Q1	Q2	Q3	P>.6	P>.7	P>.8	P>.9	SIGMA	DIQ	ASIM	CURT
+	+	+	+	+	+	+	+	?	?	?	?

CUADRO n.° 4: ENCUESTA DE OPINIÓN A DOCENTES Y ALUMNOS

Este es un sondeo de opiniones sobre la evaluación a los docentes por parte de los alumnos. Agradecemos su aporte.

Marque su condición de docente o de alumno y los años en la Facultad:

[ ] Profesor con _____ años como docente universitario en la Facultad	[ ] Alumno del _____ nivel (semestre)

Marque X en la casilla [ ] correspondiente a su más sentida opinión.

1. Considera que las evaluaciones al docente por los alumnos es un instrumento válido para medir la calidad del docente.

[ ] Sí, en la mayoría de casos	[ ] Son bastante válidas	[ ] Rara vez son válidas

2. Considera que los alumnos realizan la evaluación a los docentes en forma reflexiva y objetiva.

[ ] Sí, en la mayoría de casos	[ ] Son bastante reflexivas y objetivas	[ ] Rara vez son reflexivas y objetivas

3. Considera que existe una reciprocidad entre notas del profesor a los alumnos y las evaluaciones que éstos dan al profesor, esto es, a mejores notas mejores evaluaciones.

[ ] No, en la mayoría de casos	[ ] Hay bastante reciprocidad en ello	[ ] Hay un exceso de reciprocidad en ello

TABLA n.° 1: RESUMEN DE INDICADORES DE LAS NOTAS Y % DE LA NOTA MÁXIMA									
	AÑO 2006-2007 2° S			AÑO 2007-2008 1° S			AÑO 2007-2008 2° S		
MEDIDA	S 30	S 20	S 5O	S 30	S 20	S 5O	S 30	S 20	S 5O
Moda	27.0	20.0	42.0	27.0	20.0	43.0	27.0	20.0	43.0
Promedio	24.1	16.1	40.3	24.0	16.0	39.4	23.9	15.8	39.7
Desv. estándar	4.0	3.3	6.1	4.3	3.2	7.6	4.2	3.5	6.3
Cuartil 1	22.0	14.1	36.0	21.5	14.0	35.0	21.5	13.9	35.0
Cuartil 2	25.0	17.0	41.0	25.0	16.5	41.0	24.9	16.4	41.0
Cuartil 3	27.0	19.0	45.0	27.0	18.4	45.0	27.0	18.2	45.0
Asimetría	-1.1	-1.0	-0.8	-1.1	-0.9	-1.5	-1.0	-1.1	-0.6
Curtosis	1.6	0.7	0.5	3.1	0.8	3.9	1.3	1.7	0.2
VALOR EXPRESADO EN % DE LA NOTA MÁXIMA									
	AÑO 2006-2007 2° S			AÑO 2007-2008 1° S			AÑO 2007-2008 2° S		
MEDIDA	S 30	S 20	S 5O	S 30	S 20	S 5O	S 30	S 20	S 5O
Moda	90%	100%	84%	90%	100%	86%	90%	100%	86%
Promedio	80%	81%	81%	80%	80%	79%	80%	79%	79%
Desv. estándar	13%	17%	12%	14%	16%	15%	14%	18%	13%
Cuartil 1	73%	71%	72%	72%	70%	70%	72%	70%	70%
Cuartil 2	83%	85%	82%	83%	83%	82%	83%	82%	82%
Cuartil 3	90%	95%	90%	90%	92%	90%	90%	91%	90%
Asimetría	-4%	-5%	-2%	-4%	-5%	-3%	-3%	-6%	-1%
Curtosis	5%	4%	1%	10%	4%	8%	4%	9%	0%

TABLA n.° 2: INDICADORES COMPARADOS DE NOTAS S30 y S20						
SEMESTRE:	2006 - 2		2007 - 1		2007 - 2	
NOTAS SOBRE:	S30	S20	S30	S20	S30	S20
Promedio	24.10	16.10	24.00	16.00	23.90	15.80
Promedio / Máximo	80%	81%	80%	80%	80%	79%
Mediana - Q2	25.00	17.00	25.00	16.50	24.90	16.40
Mediana / Máximo	83%	85%	83%	83%	83%	82%
Desviación estándar	4.00	3.30	4.30	3.20	4.20	3.50
Desviación Intercuartílica	5.00	4.90	5.50	4.40	5.50	4.30
Coeficiente de Variación	17%	20%	18%	20%	18%	22%
Desviación Intercuartílica / Q2	20%	29%	22%	27%	22%	26%

TABLA n.° 3: INDICADORES DE LA DISTRIBUCIÓN DE LAS EVALUACIONES A DOCENTES			
MEDIDA	**2006-2**	**2007-1**	**2007-2**
Moda	100.00	93.36	93.75
Promedio	89.40	89.48	89.38
Desviación estándar	9.03	8.63	9.65
Cuartil 1	86.31	86.72	86.79
Cuartil 2	92.00	91.74	92.64
Cuartil 3	95.65	95.72	95.78
Asimetría	-1.63	-1.52	-1.94
Curtosis	3.11	2.40	4.66
Mínimo	37.77	51.86	37.77
Máximo	100.00	100.00	100.00

Tabla n.° 4: CORRELACIÓN ENTRE LAS EADS Y LOS INDICADORES DE NOTAL

AÑO S	EXA	No	Q1	Q2	Q3	DIQ	MEAN	SIGMA	ASIM.	CURT.	P>.6	P>.7	P>.8	P>.9
2006 2	S20	315	-3,6%	-2,4%	0,3%	5,6%	-2,5%	4,6%	3,5%	-3,6%	-2,6%	-3,2%	-2,5%	-1,0%
2007 1	S20	368	-10%	-11%	-7%	4%	-10%	0%	2%	-8%	2%	-4%	-14%	-8%
2007 2	S20	295	-14,1%	-12,2%	-8,8%	12,8%	-12,1%	8,2%	5,5%	-1,5%	-16,6%	-15,9%	-12,0%	3,9%
2006 2	S30	315	0,5%	4,5%	8,3%	6,5%	3,2%	5,8%	0,4%	1,8%	-3,7%	-4,4%	5,9%	4,4%
2007 1	S30	368	-17%	-13%	-14%	1%	-17%	6%	-4%	4%	-7%	-13%	-16%	-11%
2007 2	S30	295	6,1%	11,7%	15,7%	7,3%	11,5%	4,1%	-6,5%	3,9%	8,0%	0,1%	4,8%	-7,2%
2006 2	S50	315	-1,5%	-0,1%	2,5%	4,5%	0,5%	4,1%	4,1%	-3,5%	-0,5%	-4,8%	-1,3%	3,4%
2007 1	S50	368	-21%	-16%	-12%	15%	-19%	20%	9%	-5%	-12%	-13%	-20%	-11%
2007 2	S50	295	-2,8%	-0,2%	5,1%	11,9%	0,7%	5,1%	-1,2%	-6,9%	-7,6%	-4,1%	-4,2%	-2,6%

TABLA n.° 5: CORRELACIÓN POR GRUPOS DE LA NOTA S20 CON INDICADORES														
			No	Q1	Q2	Q3	DIQ	MEAN	SIGMA	ASIM.	CURT.	P>.6	P>.7	P>.8	P>.9
2006 2	S20	INF	105	-30,1%	-29,2%	-25,9%	18,0%	-29,7%	19,2%	19,8%	-20,3%	-23,1%	-23,7%	-24,7%	-21,4%
2006 2	S20	MED	105	3,3%	0,6%	-4,5%	-11,7%	1,2%	-7,3%	4,6%	2,3%	0,0%	-0,7%	-2,1%	6,6%
2006 2	S20	SUP	105	-16,3%	-12,9%	-4,2%	20,7%	-12,6%	19,4%	1,9%	0,7%	-17,3%	-14,3%	-7,0%	-7,6%
2006 2	S20	TT	315	-3,6%	-2,4%	0,3%	5,6%	-2,5%	4,6%	3,5%	-3,6%	-2,6%	-3,2%	-2,5%	-1,0%
2007 1	S20	INF	122	-18,1%	-16,2%	-17,4%	2,7%	-17,5%	-3,0%	7,3%	-21,3%	-0,2%	-5,4%	-11,6%	-10,3%
2007 1	S20	MED	124	-16,6%	-18,9%	-23,1%	-2,3%	-20,4%	-4,4%	15,2%	-13,3%	-10,5%	-24,1%	-26,9%	-3,6%
2007 1	S20	SUP	122	8,6%	11,8%	16,6%	5,8%	13,7%	2,4%	-3,8%	-17,2%	12,1%	11,8%	3,5%	0,8%
2007 1	S20	TT	368	-10,0%	-10,6%	-7,0%	4,4%	-10,5%	0,4%	1,9%	-7,8%	2,5%	-4,4%	-14,5%	-7,7%
2007 2	S20	INF	99	-20,3%	-20,1%	-22,2%	7,0%	-21,5%	9,6%	2,2%	2,4%	-29,2%	-27,8%	-19,5%	-3,5%
2007 2	S20	MED	96	-7,3%	-9,1%	-7,4%	3,9%	-7,1%	0,1%	5,8%	-9,7%	2,1%	-4,0%	-6,5%	2,8%
2007 2	S20	SUP	100	-4,5%	-1,5%	-4,6%	2,2%	-4,8%	6,4%	-2,0%	1,9%	-6,7%	-2,5%	0,1%	8,8%
2007 2	S20	TT	295	-14,1%	-12,2%	-8,8%	12,8%	-12,1%	8,2%	5,5%	-1,5%	-16,6%	-15,9%	-12,0%	3,9%

TABLA n.° 6: OPINIONES DE ALUMNOS Y PROFESORES SOBRE LAS EVALUACIONES A DOCENTES (EAD)				
OPINION DE ALUMNOS(A) Y PROFESORES (P)	A = 72		P = 20	
	n.°	%	n.°	%
¿LAS EAD SON VÁLIDAS PARA MEDIR LA CALIDAD DEL DOCENTE?				
SÍ EN LA MAYORÍA DE LOS CASOS	16	22%	3	15%
SON BASTANTE VÁLIDAS	10	14%	7	35%
RARA VEZ SON VÁLIDAS	46	64%	10	50%
¿LAS EAD SE REALIZAN EN FORMA REFLEXIVA Y OBJETIVA?				
SÍ EN LA MAYORÍA DE LOS CASOS	14	19%	2	10%
SON BASTANTE REFLEXIVAS Y OBJETIVAS	12	17%	3	15%
RARA VEZ SON REFLEXIVAS Y OBJETIVAS	46	64%	15	75%
¿HAY RECIPROCIDAD ENTRE NOTAS Y EVALUACIONES?				
NO EN LA MAYORÍA DE LOS CASOS	33	46%	3	15%
HAY BASTANTE RECIPROCIDAD	34	47%	7	35%
HAY UN EXCESO DE RECIPROCIDAD EN ELLO	5	7%	10	50%

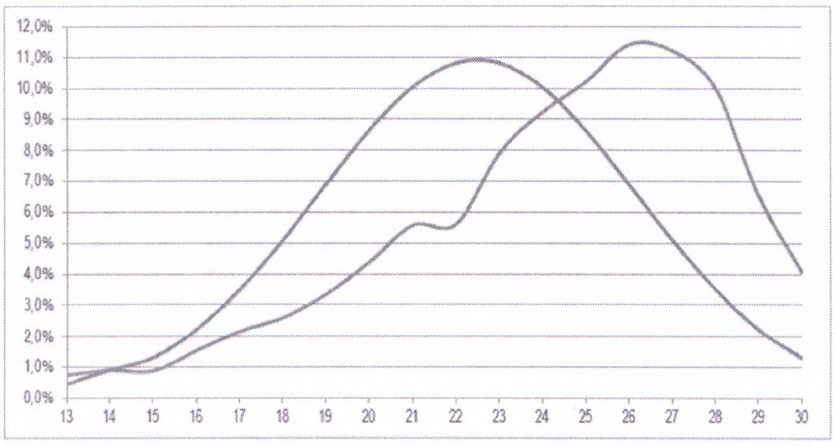


Gráfico n.º 1. Notas S30 comparadas con distribución normal

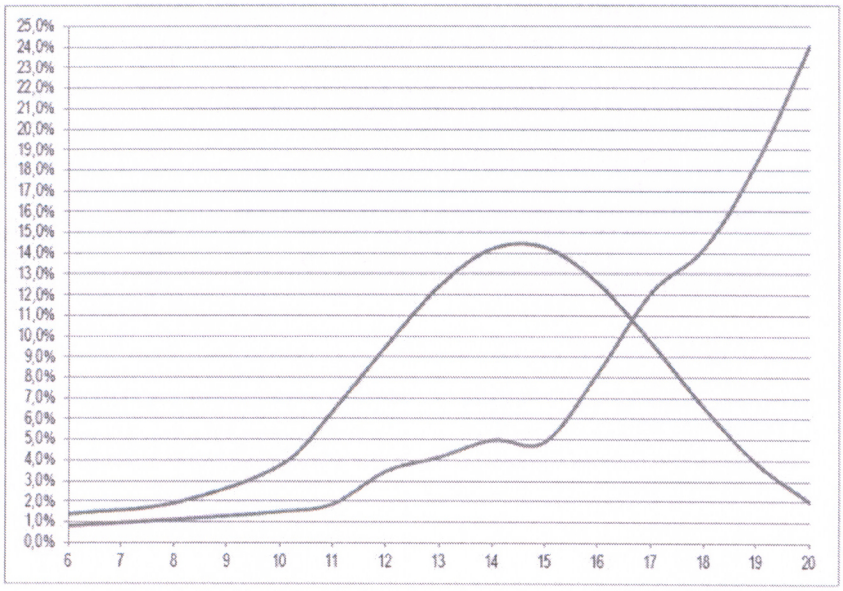


Gráfico n.º 2. Notas S20 (examen final) versus distribución normal

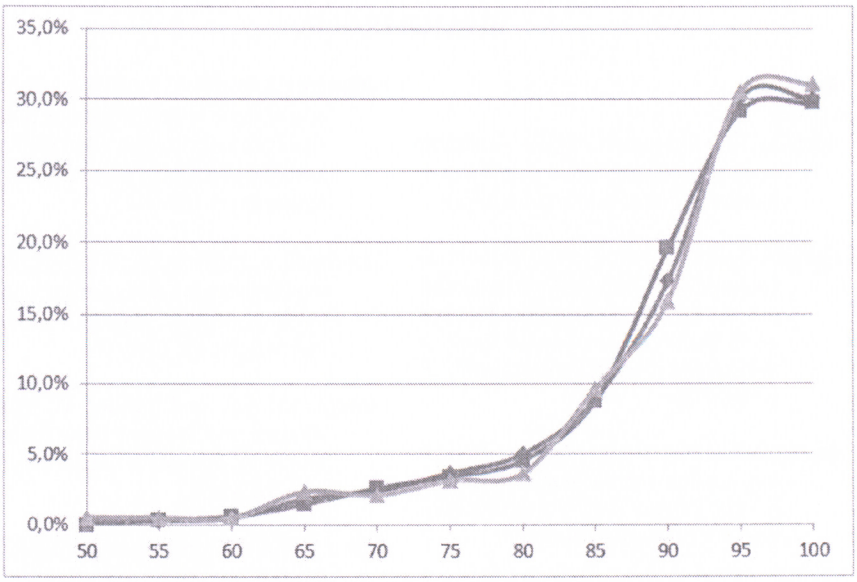


Gráfico n.° 3. Distribución de las evaluaciones a docentes

BIBLIOGRAFÍA

Revistas

Baldwin, T., & Blattner, N. (2003). Guarding against potential bias in student evaluations. *College Teaching, 51*, 27-31.

Cabezas, C. y Mancheno, P. (2008). "Estudio de Factores de la Personalidad que influyen en un alto desempeño docente en la Pontificia Universidad Católica del Ecuador (PUCE)", *Revista de la PUCE*, n.º 84, 9-20.

Cashin, W. E. (1988). Student ratings of teaching. A summary of research *(IDEA Rep. n.º. 20)*, Kansas State University, Division of Continuing Education.

Cashin, W. E. and Downey, R.G. (1992). "Using Global Student Rating Items for Summative Evaluation." *Journal of Educational Psychology*, 84(4), 563-572.

Centra, J. (2015). *Will teachers receive higher students evaluations by giving higher grades and less coursework?* ETS, SIR II. Recuperado de http://www.ets.org/Media/Products/StudentEval.pdf

Centra, J.A. and Creech, F.R. (1976). "The Relationship between Student, Teacher, and Course Characteristics and Student Ratings of Teacher Effectiveness." (*Project Rep. n.º 76-1*). Princeton, NJ: Educational Testing Service.

Feldman, K.A. (1978). "Course Characteristics and College Students' Ratings of Their Teachers and Courses: What We Know and What We Don't." *Research in Higher Education*, 9, 199-242.

Feldman, K.A. (1978). "Course Characteristics and College Students' Ratings of Their Teachers and Courses: What We Know and What We Don't." *Research in Higher Education*, 9, 199-242.

Greenwald, A. and Gillmore, G. (1997). "Grading Leniency is a Removable Contaminant of Student Ratings." *American Psychologist*, 52, 1209-17.

Krautmann, A.C. and Sander, W. (1999). "Grades and Student Evaluations of Teachers." *Economics of Education Review*, 18, 59-63.

Kulik, J. A. (2001). *"Student Ratings: Validity Utility and Controversy." New Directions for Institutional Research*, 109, 9-25. Recuperado de http://deepblue.lib.umich.edu/bitstream/2027.42/34805/1/1ftp.pdf

Marsh, H.W. (1982). "Factors Affecting Students' Evaluations of the Same Course Taught by the Same Instructor on Different Occasions." *American Educational Research Journal*, 19(4), 485-497.

Marsh, H.W. (1983). "Students' Evaluations of University Teaching: Research Findings, Methodological Issues, and Directions for Future Research." *International Journal of Educational Research*, 11, 253-388.

McDonald, R., Johnson, R. (2003) Grade Distribution and its impact on CIS Faculty evaluations: 1992-2002, Information System Education Journal, Vol. 1 n.º 42.

Recuperado dehttp://www.isedi.org/1/42

Miron, M. and Segal E. (1986). Student Opinion of the Value of Student Evaluations. *Higher Education*, Vol. 15, No. 3/4 (1986), pp. 259-265. Recuperado de http://www.jstor.org.stable/3446689

Naftulin, D. Ware, J., and Donnelly, F. (1973). *The Doctor Fox Lecture: A Paradigm Of Educational Seduction*,.Journal of Medical Education, vol. 48, p. 630-635. Recuperado de http://www.er.uqam.ca/nobel/r30034/PSY4180/Pages/Naftulin.html

Seldin, P. (1993a). "How Colleges Evaluate Professors: 1983 versus 1993." *American Academy of Higher Education Bulletin* (oct.): 6-8, 12.

Remmers, H.H. and G.C. Brandenburg (1927). 'Experimental data on the Purdue Rating Scale for Instructors," *Educational Administration and Supervision*, 13, 519-527. Manhattan, KS: Kansas State University

Rodin, M. and Rodin, B. (1972). "Student Evaluations of Teachers." *Science*, 177, 1164-1166.

Rosovsky, R. and Hartley, M. *Evaluation and the Academy: Are We Doing the Right Thing? Grade Inflation and Letters of Recommendation*. American Academy of Arts and Sciences, 2002

Sojka, J., Gupta, A. K., & Deeter-Schmelz, D. R. (2002). Student and faculty perceptions of student evaluations of teaching. *College Teaching, 50*, 44-49.

Revistas

Fuentes secundarias

Aleamoni, L.M. and Hexner, P.Z. (1980). "A Review of Research on Student Evaluation and a Report on the Effect of Different Sets of Instructions on Student Course and Instructor Evaluation." *Instructional Science*, 9, 67-84.

Bosshardt, W. and Broder, J. *Comparing Student and Instructor Evaluations of Teaching (2001)*. The Journal of Economic Education, Vol. 32, n.º 1, 3-17.

Daniel E. Ho Timothy H. Shapiro. *Evaluating Course Evaluations: An Empirical Analysis of a Quasi-Experiment at the Stanford Law School, 2000-2007*, March 31, 2008

Eley, M. and Stecher, E. (1997). "A Comparison of Two Response Scale Formats Used in Teaching Evaluation Questionnaires." *Assessment and Evaluation in Higher Education*, 22(1), 65-79.

Ellis, L., Burke, D.M., Lomire, P., and McCormack, D.R. (2003). "Student Grades and Average Ratings of Instructional Quality: The Need for Adjustment." *Journal of Educational Research*, 97(1), 35-40.

Esperanza E. & Covadonga Villasol, Sobre la opinión que los alumnos tienen de la efectividad de la docencia, Universidad Politécnica de Madrid. *Revista Iberoamericana de Educación (ISSN: 1681-5653)*

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, 137–189.

Koon, J., & Murray, H. S. (1996). Using multiple outcomes to validate student ratings of overall teacher effectiveness. *Journal of Higher Education*, 66, 61-81.

Kulik, J.A. and McKeachie, W.J. (1975). "The Evaluation of Teachers in Higher Education." *Review of Research in Education*, 3, 210-240.

Marsh, H.W. and Roche, L.A. (1997). "Making Students' Evaluations of Teaching Effectiveness Effective." *American Psychologist*, 52(11), 1187-1197

Maslow, A.H. and Zimmerman, W. (1956). "College Teaching Ability, Scholarly Activity and Personality." *Journal of Educational Psychology*, 47, 185-189.

Murray, H.G. (1983). "Low-Inference Classroom Teaching Behaviors and Student Ratings of College Teaching Effectiveness." *Journal of Educational Psychology*, 71, 856-865.

Powell, R. W. (1977). Grades, learning, and student evaluation of instruction. *Research in Higher Education*, 7, 193-205.

Rayder, N.F. (1968). "College Student Ratings of Instructors." *Journal of Experimental Education*, 37, 76-81.

Reardon, M. and Waters, L.K. (1979). "Leniency and Halo in Student Ratings of College Instructors: A Comparison of Three Rating Procedures with Implications for Scale Validity." *Educational and Psychological Measurement*, 39(1), 159-162.

Remmers, H. H. (1927). The Purdue Rating Scale for Instructors. *Educational Administration and Supervision*, *6*, 399-406.

Rodin, M. and Rodin, B. *Evaluation of Teaching.* (1973). Journal of Economic Education, Vol.5, No 1, pp. 5-9. Recuperado de http://www.jstor.org.stable/1182829

Shapiro, E.G. (1989). *Effect of instructor and class characteristics on student's class evaluations.* Research in Higher Education. 31: 135.148.

Souce, B. *The Development of Teacher Appraisal: A Recent History* (2000). The British Journal of Educational Studies, Vol. 48, No. 1 pp. 22-37. Recuperado de http://www.jstor.org./stable/1555845

Tang, T.L.-P. (1997). "Teaching Evaluation at a Public Institution of Higher Education: Factors Related to the Overall Teaching Effectiveness." *Public Personnel Management*, 26(3), 379-389.

Vasta, R., & Sarmiento, R. F. (1979). Liberal grading improves evaluations but not performance. *Journal of Educational Psychology*, 71, 207-211.

Williams, W. M., and Ceci, S. J. (1997). "How I'm Doing? Problems with Student Ratings of Instructors and Courses." *Change*, 29(5), 13–23.

Young, P., et al. (1999). "Student Evaluation of Faculty: Effects of Purpose on Pattern." *Journal of Personnel Evaluation in Education*, 13(2), 179-190.

Libros

Downie N. y Heath, R. (1973) *Métodos estadísticos aplicados*, Harper and Row Publishers Inc., México.

Neter, J., Kutner, M. H., Nachtsheim, C. J. Wasserman, W. (1996) *Applied linear regression models* (3ª ed). Irwin Co., Chicago.