

Interacción entre las expectativas académicas del alumno y la evaluación del profesorado

Juan Cándido Gómez-Gallego*, María Gómez-Gallego*,
María Concepción Pérez-Cárceles*, Alfonso Palazón-Pérez de los Cobos*
y Juan Gómez-García**

*Universidad Católica San Antonio (España) y **Universidad de Murcia (España)

Las valoraciones que hacen los estudiantes universitarios sobre la docencia impartida por sus profesores constituyen un indicador de calidad docente que puede tener importantes efectos académicos y administrativos para alumnos, profesores e instituciones universitarias. En consecuencia, es importante analizar la existencia de posibles factores que pudieran sesgar tales evaluaciones. En este trabajo se investiga sobre la relación entre la valoración que hacen los alumnos de sus profesores y la discrepancia entre las calificaciones esperadas y las realmente obtenidas por los alumnos. El estudio es de tipo transversal y exploratorio. La muestra está constituida por 1967 encuestas de evaluación que realizan 396 estudiantes sobre 60 profesores. Se ha aplicado el Cuestionario de Evaluación Docente, que mide aspectos de planificación, docencia y evaluación según los requerimientos del Espacio Europeo de Educación Superior. Se comprueba que existe correlación positiva significativa entre las valoraciones de la calidad docente del profesor y las calificaciones esperadas y obtenidas por los alumnos. Sin embargo, la asociación es negativa en el caso del cumplimiento de las expectativas del alumno. Es decir, ante un menor cumplimiento de las expectativas, menor es la valoración que el alumno hace del profesor.

Palabras clave: Nota esperada, nota obtenida, evaluación docente, expectativa académica, sesgo en la evaluación.

Interaction between student academic expectations and teaching evaluation. Students' evaluation of university lectures stands for an indicator of teaching quality which can have a relevant impact on academic and administrative issues for students, lecturers and academic institutions. Therefore, it is important to analyze the existence of possible factors that could bias such assessments. This paper investigates the relationship between lecturers' assessment by students and the discrepancy between the expected and the actual grades obtained by university students. The research was carried out following a quantitative methodology ("ex post facto"). The study is cross-sectional and exploratory. The sample consists of 1,967 evaluation surveys conducted on 396 students and 60 lecturers. Teaching evaluation scale includes twenty-two items about planning, teaching and evaluation adapted to the requirements of the European High Education Area. There is significant positive correlation between the assessments of teaching quality and the qualifications expected and obtained by the students. The association is negative in the case of the fulfillment student expectation: low rates of achievement of students expectations translates in a poorer assessment on university lecturers.

Keywords: Expected grade, actual grade, teaching evaluation, academic expectancy, evaluation bias.

Fecha de recepción: 18/02/2013 • Fecha de aceptación: 27/03/2013
Correspondencia: Juan Gómez García
Universidad de Murcia
Campus de Espinardo, C.P. 30100, Murcia (España)
Correo electrónico: jgomezg@um.es

Actualmente, las valoraciones de los estudiantes universitarios sobre la docencia de los profesores son un indicador muy utilizado de calidad docente (Worthington, 2002). La enseñanza es una variable multidimensional

que incluye aspectos interrelacionados como la planificación, claridad, entusiasmo, interacción y evaluación (Marsh, 2007). La valoración que realizan los estudiantes debe reflejar, por tanto, estos dominios (Apodaca & Grad, 2005). Palazón (2011) valida el Cuestionario de Evaluación Docente (CED) en una población de 1500 estudiantes de grado de la Universidad Católica San Antonio de Murcia. Este instrumento de evaluación tiene tres factores: Planificación, Docencia y Evaluación, que explican un 76.6% de la varianza total de las puntuaciones de la escala. El factor planificación engloba ítems que se relacionan con la capacidad del profesor de organizar y planificar la asignatura; el factor docencia incluye aspectos didácticos, métodos de enseñanza y clima de interacción en el aula; el factor evaluación hace referencia al sistema de evaluación utilizado, la objetividad en la aplicación de criterios y al conocimiento de los mismos por los alumnos.

Para evaluar y mejorar la efectividad docente se utilizan, entre otros, las valoraciones de los estudiantes (Daud & Kassim, 2011), teniendo incluso relevancia en los procesos de promoción. Aunque aún no está aclarado si estas valoraciones reflejan la efectividad docente o simplemente se asocian con los estilos de enseñanza (Marsh, 2007). Además, se han reportado distintos factores, no relacionados con la efectividad docente, que influyen sobre las valoraciones (Clayson, Frost & Sheffet, 2006). Se trata de variables que podrían sesgar las valoraciones y, según Cashing (1995), pueden relacionarse con características propias del estudiante, del profesor y de tipo situacional o administrativo. Hay pocos trabajos que determinen en qué medida influyen estos posibles factores y tampoco si existe algún dominio de las valoraciones sobre la docencia que resulte más afectado (Beran & Violato, 2005; Marsh, 1987). A continuación se revisan los principales factores de sesgo según su fuente de procedencia.

Características del profesor

Feldman (1993) encontró una baja correlación estadísticamente significativa (en favor de las mujeres, $r=.02$) entre el género del pro-

fesor y la valoración que hacen los estudiantes de su docencia. Los pocos estudios que ofrecen resultados significativos muestran que los profesores mujeres reciben una valoración más favorable que los profesores hombres (Griffin, 2004). Sin embargo, podría haber un efecto interacción que indica que los estudiantes tienden a dar mayores valoraciones a los profesores de su género (Wachtel, 1998). Lo mismo puede decirse de la relación ranking del profesor-evaluación docente. Muchos estudios muestran correlaciones lineales no significativas entre el ranking y las puntuaciones de evaluación docente (Ting, 2000) y, cuando se encuentran resultados significativos, los profesores con mayores rankings reciben mayores puntuaciones (Feldman, 1983). En lo concerniente a la relación entre la edad del profesor y la evaluación docente, la mayoría de los estudios muestran de nuevo correlaciones no significativas. En los pocos estudios que sí reportan resultados significativos, se encuentra un efecto inverso (Wachtel, 1998): los profesores con mayor edad reciben evaluaciones menos favorables.

Características del curso

Marsh (1987) y Theall & Franklin (2001) sugieren que los cursos de ciencias naturales ofrecen menores valoraciones docentes que los cursos de humanidades, aunque esta diferencia no explica mucha varianza (menos del 1%) en los resultados (Centra & Creech, 1976). La literatura sobre la relación entre el tamaño de clase y la evaluación docente muestra resultados mixtos. Mientras algunos autores muestran correlaciones negativas (Bedard & Kuhn, 2008; Centra, 2003; Isely & Singh, 2005), otros no muestran correlaciones significativas (Krauttman & Sander, 1999). La relación entre el tamaño de clase y la valoración docente parece no ser lineal (Fernández, Mateo & Muñoz, 1998), indicando que hay clases muy grandes así como clases pequeñas evaluadas favorablemente comparadas con las clases medianas.

Características del estudiante

Una cuestión interesante a tener en cuenta en relación al sesgo del estudiante es la con-

cerniente al período de exámenes en el que la nota del curso se recibe. En el sistema de educación superior, los estudiantes suelen tener dos oportunidades de aprobar en un curso. La hipótesis consiste en que los estudiantes que aprueban en la primera convocatoria tenderían a conceder evaluaciones más positivas a los profesores. No se conoce ningún trabajo donde la relación entre el período de exámenes y la valoración docente se haya estudiado.

Otro sesgo analizado en la literatura ha sido la asistencia a clase. Parece plausible que los estudiantes que asisten a clase regularmente presentan mayor motivación y están más interesados por las habilidades y actividades educacionales que lleva a cabo el profesor (Devadoss & Foltz, 1996; Fjortoft, 2005). Sin embargo, Beran & Violato (2005) y Davidovitch & Soen (2006) encuentran relaciones significativas y Guinn & Vincent (2006) no significativas. Probablemente estos resultados entran en conflicto en parte debido a las diferencias en el diseño de la investigación.

La misma conclusión se puede obtener en lo concerniente a la relación entre el género del estudiante y la valoración docente (Centra & Gaubatz, 2000). Recientes investigaciones, sin embargo, muestran que debería haber un efecto interacción (o mejor, un patrón de género) entre el género de los profesores y estudiantes con respecto a la valoración docente. Los estudiantes masculinos tienden a escoger con menor frecuencia de lo esperado a una mujer como mejor profesora (Basow, Phelan & Capotosto, 2006).

Con respecto a la característica edad del estudiante, no se ha estudiado mucho hasta ahora. No está claro, por tanto, si mayores notas de estudiantes de los cursos superiores son resultado de asignaturas más especializadas o de mayor madurez de los alumnos (Wachtel, 1998).

Finalmente, el sesgo potencial que presenta una mayor controversia en su interpretación es el efecto de la nota esperada por el alumno sobre la evaluación del profesorado. Por ello, se trata de la característica más estudiada de los alumnos en este área de investigación; concretamente, la relación entre la nota esperada de los estudiantes y el resultado de

la evaluación del profesor. Como resultado de la revisión de la literatura realizada por Feldman (1976, 1997), se obtiene la existencia de una correlación significativa (entre .10 y .30) entre la nota esperada y el resultado de la evaluación. Posteriormente, Centra (2003) aporta un resultado de .20. A partir de estos análisis, se genera un debate sobre si esta relación debería considerarse una prueba de la validez de las percepciones de un estudiante sobre su propio proceso de formación (Marsh & Roche, 2000; Griffin, 2004; Gump, 2007).

Existen tres planteamientos que desarrollan los posibles efectos de esta relación (Marsh, 2001; Marsh & Dunkin, 1997). En primer lugar, la hipótesis de indulgencia considera que los profesores que dan notas superiores a las merecidas serán recompensados con resultados de su evaluación superiores a los merecidos (Greenwald & Gillmore, 1997; Krautmann & Sander, 1999; McPherson, 2006). La hipótesis de validez relaciona mejores notas esperadas con mejor aprendizaje del alumno y, por tanto, con una mejor evaluación del profesor (Marsh, 1987). Por último, la hipótesis de características previas del alumno defiende que éstas podrían afectar a su aprendizaje, notas, así como a la efectividad del profesor (Marsh, 1987).

Según Marsh (2007), independientemente de las implicaciones del efecto de la nota esperada; las notas, ya sean obtenidas o esperadas, reflejan una combinación de los sistemas de evaluación del profesor y el aprendizaje y características del estudiante. En este sentido, el estudio propone considerar las variables nota obtenida (Spooren, 2010) y nota esperada. Además, se incluye la variable cumplimiento expectativa del alumno, que se define como la diferencia entre la nota esperada y la nota recibida, y pretende cuantificar el efecto de la indulgencia del profesor ya que, según esta hipótesis, no son las notas absolutas las que influyen en la evaluación del profesor, sino la indulgencia con la que son asignadas.

En este trabajo se investiga sobre la relación entre la valoración que hacen los alumnos de sus profesores y el grado de acuerdo entre las calificaciones otorgadas por los profesores a los alumnos y las esperadas por los es-

Tabla 1. *Estadísticos Descriptivos.*

	N	Mínimo	Máximo	Media	Desv. típ.	Interv. Confianza 95%	
						Lím. inf.	Lím. sup.
Edad Alumno	396	18	41	20.79	3.98	20.39	21.17
Nota Obtenida	1967	2	10	6.86	1.37	6.79	6.91
Nota Esperada	1967	5	10	7.38	1.28	7.32	7.43
Difer. NE-NO	1967	-4.3	6.3	.52	1.12	.47	.57
Valoración Global	1967	1	5	4.17	.96	4.13	4.22
Puntuación Total	1967	22	110	85.77	15.42	85.10	86.46
Planificación	1967	4	20	15.79	3.15	15.65	15.93
Docencia	1967	12	60	46.07	9.35	45.66	46.48
Evaluación	1967	5	25	19.74	4.50	19.54	19.94

tudiantes. Tal relación se estima para toda la población y en los subgrupos definidos por los diferentes niveles de calificaciones obtenidas por el alumno y atendiendo a características personales como género y edad.

Método

Participantes

La muestra está constituida por 396 estudiantes de grado en el año académico 2010-2011 y 60 profesores, habiéndose obtenido 1967 encuestas de evaluación sobre la calidad docente en la Universidad Católica San Antonio (UCAM). La edad media de los estudiantes que participaron en el estudio fue de 20.79 años; el rango de edades fue de 18-41 años. El porcentaje de mujeres fue de un 47.4%.

Instrumento

Se administró a los estudiantes el Cuestionario de Evaluación Docente (CED) validado por Palazón (2011). El cuestionario consta de 22 ítems que se refieren a las dimensiones docentes de planificación, interacción y clima de clase, motivación, didáctica, tutorías y evaluación. Los ítems saturan tres factores que se interpretan y denominan factor planificación, factor docencia y factor evaluación.

Los ítems se valoran de 1 a 5; en todos los casos, mayor puntuación indica más calidad en la docencia. El ítem 22 es un indicador de la calidad docente en su globalidad. La puntuación total de la escala se obtiene como suma

de las puntuaciones de los ítems y su rango es 22-110. Los rangos de los factores son 3-15, 14-70 y 5-25 para planificación, docencia y evaluación, respectivamente.

Diseño y procedimiento

La investigación se ha realizado siguiendo una metodología cuantitativa, modalidad *ex post facto*. El estudio es de tipo transversal y exploratorio. Para recoger la información se utilizaron cuestionarios autocumplimentados. Los datos se analizaron mediante la metodología cuantitativa.

El análisis estadístico se realiza mediante el programa SPSS-19. Se aplican el test de correlación de Pearson, el test t para muestras relacionadas y el procedimiento ANOVA.

Resultados

En la Tabla 1 se observan los estadísticos descriptivos de las variables utilizadas en el estudio. Los promedios de nota esperada y nota obtenida difieren significativamente. El tamaño del efecto de la diferencia se estima en .52 puntos. En los indicadores de evaluación docente, los promedios son representativos ya que los coeficientes de variación oscilan entre .18 y .23.

En cuanto a las características del estudiante, existen diferencias significativas en la valoración global, la puntuación total, los factores (planificación, docencia y evaluación) y en el cumplimiento de las expectativas según el género del alumno. De hecho, las alumnas eva-

Tabla 2. Valoración del profesor según género del alumno.

		N	Media	Desviación típica	Intervalo de confianza para la media al 95%		Sig.
					Límite inferior	Límite superior	
Valoración Global	M	932	4.36	.86	4.30	4.41	.000
	H	1035	4.01	1.01	3.95	4.07	
Puntuación Total	M	932	88.98	15.23	88.00	89.96	.000
	H	1035	82.89	15.03	81.97	83.81	
Planificación	M	932	16.27	3.05	16.08	16.47	.000
	H	1035	15.35	3.17	15.16	15.54	
Docencia	M	932	47.71	9.28	47.11	48.31	.000
	H	1035	44.60	9.17	44.04	45.16	
Evaluación	M	932	20.63	4.13	20.37	20.90	.000
	H	1035	18.93	4.67	18.65	19.22	
Difer. NE-NO	M	932	.11	.94	.05	.17	.000
	H	1035	.89	1.13	.82	.96	

lúan con una mayor puntuación a sus profesores en todas las variables analizadas en el estudio (ver Tabla 2), lo que puede explicarse por el mayor cumplimiento de sus expectativas.

Este estudio centra su análisis en la influencia de la calificación del alumno en la evaluación del profesor. Las variables utilizadas para caracterizar esa calificación son la nota obtenida, la nota esperada y el cumplimiento de la expectativa del alumno. Las correlaciones entre estas variables y los indicadores de calidad docente del profesor se presentan en la Tabla 3.

El análisis de la nota esperada presenta correlaciones que oscilan entre .103 y .128 en las

distintas variables que evalúan al profesor. Aunque la nota esperada es la variable sobre calificación del alumno que más estudios han analizado, esta investigación demuestra una mayor influencia de la nota obtenida. Concretamente, se observa una correlación positiva significativa que oscila entre .159 y .276 en las distintas valoraciones de la calidad docente analizadas. De entre los factores diferenciados, la mayor correlación se encuentra con el factor evaluación y la menor con el factor planificación.

En la Tabla 4 se observa que, cuando se categoriza la nota obtenida en los niveles suspenso (menor que 5), aprobado (de 5 a menor

Tabla 3. Matriz de correlaciones.

		Valoración Global	Puntuación Total	Planificación	Docencia	Evaluación
Edad alumno	r	-.021	.071	.066	.094	.006
	Sig.	.356	.002	.003	.000	.798
Nota Obtenida	r	.268	.245	.159	.190	.276
	Sig.	.000	.000	.000	.000	.000
Nota Esperada	r	.125	.128	.107	.103	.124
	Sig.	.000	.000	.000	.000	.000
Difer. NE-NO	r	-.185	-.153	-.072	-.114	-.195
	Sig.	.000	.000	.001	.000	.000

Tabla 4. Valoración del profesor según el nivel de calificación del alumno.

		N	Media	Desv. típica	Interv. Confianza 95%		Sig.
					Límite inferior	Límite superior	
Valoración Global	Suspenso	125	3.43	1.21	3.22	3.65	.000
	Aprobado	759	4.02	1.00	3.95	4.09	
	Notable	851	4.33	.85	4.27	4.38	
	Sobresaliente	232	4.50	.72	4.41	4.60	
Puntuación Total	Suspenso	125	78.27	14.70	75.67	80.87	.000
	Aprobado	759	82.57	15.20	81.48	83.65	
	Notable	851	88.19	14.99	87.18	89.20	
	Sobresaliente	232	91.44	14.31	89.59	93.29	
Planificación	Suspenso	125	15.07	2.99	14.54	15.60	.000
	Aprobado	759	15.26	3.22	15.03	15.49	
	Notable	851	16.18	3.07	15.97	16.39	
	Sobresaliente	232	16.47	2.96	16.08	16.85	
Docencia	Suspenso	125	43.44	8.69	41.90	44.98	.000
	Aprobado	759	44.28	9.34	43.62	44.95	
	Notable	851	47.24	9.18	46.63	47.86	
	Sobresaliente	232	49.05	8.90	47.90	50.20	
Evaluación	Suspenso	125	16.33	5.32	15.39	17.27	.000
	Aprobado	759	19.00	4.61	18.67	19.33	
	Notable	851	20.44	4.06	20.16	20.71	
	Sobresaliente	232	21.42	3.73	20.94	21.90	

que 7), notable (de 7 a menor que 8.5) y sobresaliente (mayor igual que 8.5), se obtiene que la valoración del profesor en los cinco indicadores de calidad docente utilizados es creciente, en todos los casos. A más nivel de ca-

lificación del alumno más valoración del profesor, siendo las diferencias significativas en los cinco indicadores.

En la Figura 1 se representa el comportamiento medio de los indicadores de calidad do-

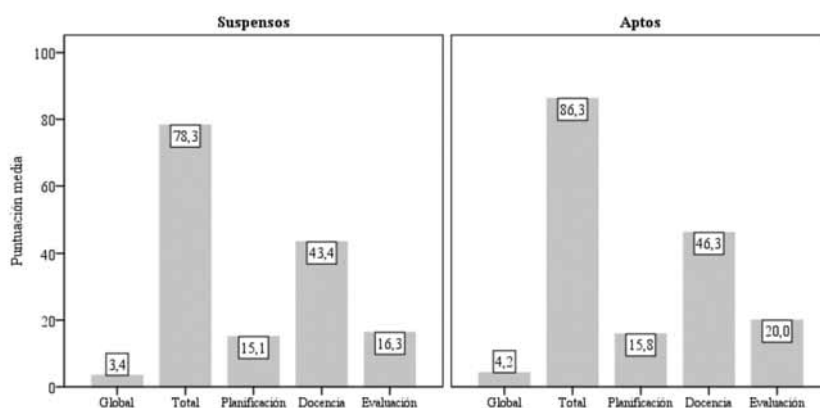


Figura 1. Valoración del profesor por alumnos suspensos y aptos.

Tabla 5. Valoración del profesor por alumnos optimistas y pesimistas.

		N	Media	Desv. típica	Interv. Confianza 95%		Sig.
					Límite inferior	Límite superior	
Valoración Global	Pesimista	328	4.32	.91	4.22	4.42	.002
	Optimista	1639	4.14	.97	4.10	4.19	
Puntuación Total	Pesimista	328	88.20	13.03	86.78	89.61	.002
	Optimista	1639	85.29	15.82	84.52	86.05	
Planificación	Pesimista	328	16.11	2.85	15.80	16.42	.041
	Optimista	1639	15.72	3.20	15.57	15.88	
Docencia	Pesimista	328	47.35	7.94	46.49	48.22	.007
	Optimista	1639	45.82	9.59	45.35	46.28	
Evaluación	Pesimista	328	20.41	4.19	19.95	20.86	.003
	Optimista	1639	19.60	4.55	19.38	19.82	

cente en cada uno de los dos grupos de alumnos: suspensos y aptos. Se observa que en todos los indicadores de calidad docente, el promedio es superior en el grupo de alumnos aptos que en el grupo de alumnos suspensos.

Por último, la variable que representa el cumplimiento de la expectativa del alumno presenta una correlación negativa significativa con la evaluación del profesor en todas las variables observadas (ver Tabla 3). Es decir, cuanto mayor es la diferencia entre la nota esperada y la obtenida o, lo que es lo mismo, ante un menor cumplimiento de las expectativas del alumno, menor es la valoración que éste realiza de su profesor. Si se observan los resultados para cada factor, la correlación es mayor con la evaluación y menor con la planificación del profesor.

A partir de estos resultados, se plantea el análisis de la existencia de diferencias significativas en la evaluación del profesor realizada por alumnos pesimistas, que obtienen una nota superior a la esperada, y alumnos optimistas, con una nota esperada superior a la obtenida. Como puede observarse en la Tabla 5, la evaluación media del profesor realizada por los alumnos pesimistas es siempre superior a la de los alumnos optimistas, lo que indica que los alumnos que obtienen una nota superior a la merecida, según sus expectativas, otorgan

una mayor nota a sus profesores. Este comportamiento justifica las hipótesis de indulgencia y validez detalladas anteriormente.

La Figura 2 presenta las valoraciones medias de los indicadores de calidad docente cuando se desagregan las clases de alumnos pesimistas y optimistas por nivel de calificación. Se observa que el vector de puntuaciones de calidad docente según los cinco indicadores difiere según la categoría de calificación obtenida. Así, por ejemplo, los alumnos que obtienen sobresaliente, aun siendo optimistas, valoran más alto en el indicador de evaluación que los pesimistas. Y este resultado difiere del comportamiento en la muestra total y en las clases de aprobados y notables.

Discusión

En las valoraciones que los alumnos hacen sobre la enseñanza impartida por sus profesores existen sesgos asociados a diferentes características del estudiante: nota esperada, nota obtenida, cumplimiento de expectativa y género del alumno.

La magnitud de las correlaciones entre la nota esperada y los distintos indicadores de calidad de la docencia es similar a las reportadas por Feldman (1976, 1997), que determi-

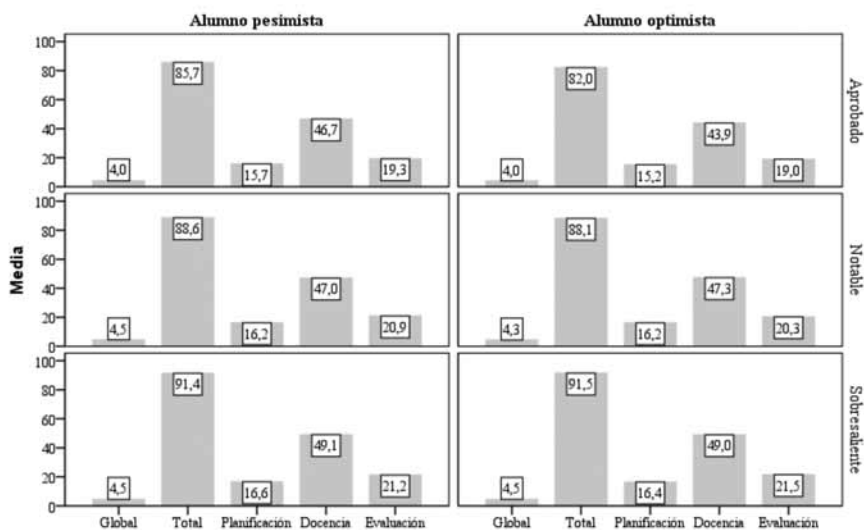


Figura 2. Valoración del profesor por alumnos pesimistas y optimistas.

na un intervalo de .10-.30, y cercana al .20 de Centra (2003).

Los resultados sobre la relación entre la nota obtenida por el alumno y la calidad docente del profesor están de acuerdo con los publicados por Cohen (1981) y Sporeen (2010). La influencia de esta variable sobre la puntuación total del Cuestionario de Evaluación Docente no es alta, tal y como afirman Beran & Violato (2005). Este sesgo es ligeramente mayor en los casos del factor evaluación y la valoración global.

Por lo tanto, existe asociación positiva entre las valoraciones docentes del profesor y las calificaciones esperadas y obtenidas por los alumnos. El efecto es mayor en el caso de notas obtenidas. Si se categoriza la calificación obtenida en los niveles suspenso, aprobado,

notable y sobresaliente, se obtienen valoraciones del profesor crecientes en función del nivel de calificación. También se observa que los promedios de todos los indicadores de calidad docente son superiores en los alumnos aptos respecto de los alumnos suspensos.

Por otra parte, la variable diferencia entre nota esperada y nota obtenida se asocia negativamente a la evaluación docente. Esto indica que los alumnos que obtienen una nota superior a la merecida, según sus expectativas, otorgan mayor valoración a sus profesores, es decir, la valoración de alumnos pesimistas es siempre superior a la que realizan alumnos optimistas.

En cuanto al género del alumno, las alumnas evalúan mejor a sus profesores que los alumnos, lo que podría justificarse por el mayor cumplimiento de sus expectativas.

Referencias

- Acevedo, R. (2003). *Factores que inciden en la competencia docente universitaria*. Tesis doctoral. Universidad Complutense de Madrid.
- Apodaca, P., & Grad, H. (2005). The dimensionality of student ratings of teaching: integration of uni- and multidimensional models. *Studies in Higher Education*, 30(6), 723-748. doi: 10.1080/03075070500340101
- Basow, S.A., Phelan, J.E., & Capotosto, L. (2006). Gender patterns in college students' choices of their best and worst professors. *Psychology of Women Quarterly*, 30, 25-35. doi: 10.1111/j.1471-6402.2006.00259.x
- Bedard, K., & Kuhn, P. (2008). Where class size really matters: class size and student ratings of instructor effectiveness. *Economics of Education*

- Review*, 27(3), 253-265. doi: 10.1016/j.econedurev.2006.08007
- Beran, T., & Violato, C. (2005). Ratings of university teacher instruction: how much do student and course characteristics really matter? *Assessment & Evaluation in Higher Education*, 30(6), 593-601. doi: 10.1080/02602930500260688
- Cashing, W.E. (1995). Student ratings of teaching: the research revisited. Center for Faculty Evaluation and Development. *Idea Paper*, 32.
- Centra, J.A. (2003). Will teachers receive higher student evaluations by giving higher grades and less course work? *Research in Higher Education*, 44, 495-518. doi: 10.1023/A:1025492407752
- Centra, J.A., & Creech, F.R. (1976). *The relationship between student, teacher and course characteristics and student ratings of teacher effectiveness*. Princeton, NJ: Educational Testing Service.
- Centra, J.A., & Gaubatz, N.B. (2000). Is there gender bias in student evaluations of teaching? *The Journal of Higher Education*, 71, 17-33.
- Clayson, D.E., Frost, T.F., & Sheffet, M.J. (2006). Grades and the student evaluation of instruction: a test of the reciprocity effect. *Academy of Management Learning & Education*, 5(1), 52-65.
- Cohen, P.A. (1981). Student ratings of instruction and student achievement: A meta-analysis of multisection validity studies. *Review of Educational Research*, 51, 281-309. doi: 10.3102/00346543051003281
- Daud, N., & Kassim, N.L. (2011). Examining student rating of teaching effectiveness using FACETS. *Journal of Applied Measurement*, 12(2), 135-143.
- Davidovitch, N., & Soen, D. (2006). Class attendance and students' evaluation of their college instructors. *College Student Journal*, 40, 691-703.
- Devadoss, S., & Foltz, J. (1996). Evaluation of factors influencing student class attendance and performance. *American Journal of Agricultural Economics*, 78, 499-507. doi: 10.2307/1243268
- Feldman, K.A. (1976). Grades and college students, evaluations of their courses and teacher. *Research in Higher Education*, 4, 69-111. doi:10.1007/BF00991462
- Feldman, K.A. (1983). Seniority and experience of college teachers as related to evaluations they receive. *Research in Higher Education*, 18, 3-12. doi: 10.1007/BF00992080
- Feldman, K.A. (1993). College student's views of male and female college teachers: Part II. Evidence from students' evaluations of their classroom teachers. *Research in Higher Education*, 34, 151-211. doi: 10.1007/BF00992161
- Feldman, K.A. (1997). Identifying exemplary teachers and teaching. Evidence from student ratings. En R. Perry, & J. Smart (Eds.), *Effective Teaching in Higher Education. Research and Practice* (pp. 368-395). New York: Agathon.
- Fernández, J., Mateo, M.A., & Muñoz, J. (1998). Is there a relationship between class size and student ratings of teaching quality? *Educational and Psychological Measurement*, 58, 596-604. doi: 10.1177/0013164498058004003
- Fjortoft, N. (2005). Students' motivation for class attendance. *American Journal of Pharmaceutical Education*, 69, 107-112.
- Greenwald, A.G., & Gillmore, G.M. (1997). Grading leniency is a removable contaminant of student ratings. *American Psychologist*, 52, 1209-1217. doi: 10.1037/0003-066X.52.11.1209
- Griffin, B.W. (2004). Grading leniency, grade discrepancy and student ratings of instruction. *Contemporary Educational Psychology*, 29, 410-425. doi: 10.1016/j.cedp-sych.2003.11.001
- Guinn, B., & Vincent, V. (2006). The influence of grades on teaching effectiveness rating sat a Hispanic-serving institution. *Journal of Hispanic Higher Education*, 5, 313-321. doi: 10.1177/1538192706291138
- Gump, S.E. (2007). Student evaluations of teaching effectiveness and the leniency hypothesis: A literature review. *Educational Research Quarterly*, 30(3), 56-69.
- Isely, P., & Singh, H. (2005). Do higher grades lead to favorable student evaluations? *Journal of Economic Education*, 36, 29-42. doi: 10.3200/JECE.36.1.29-42
- Krautmann, A.C., & Sander, W. (1999). Grades and student evaluation of teachers. *Economics of Education Review*, 18, 59-63. doi: 10.1016/S0272-7757(98)00004-1
- Marsh, H.W. (1987). Student's evaluations of university teaching: research findings, methodological issues and directions for further research. *International Journal of Educational Research*, 11, 253-388. doi: 10.1016/0883-0355(87)90001-2
- Marsh, H.W. (2001). Distinguishing between good (useful) and bad workload on students' evaluations of teaching. *American Educational Research Journal*, 38(1), 183-212.
- Marsh, H.W. (2007). Students' evaluations of university teaching: dimensionality, reliability, validity, potential biases and usefulness. En R.P. Perry, & J.C. Smart (Eds.), *The Scholarship of Teaching and Learning in Higher Education:*

- An Evidence-Based Perspective* (pp. 319-383). New York: Springer.
- Marsh, H.W., & Dunkin, M.J. (1997). Students' evaluations of university teaching: a multidimensional perspective. En R.P. Perry, & J.C. Smart (Eds.), *Effective Teaching in Higher Education: Research and Practice* (pp. 241-320). New York: Agathon.
- Marsh, H.W., & Roche, L.A. (2000). Effects of grading leniency and low workload on students' evaluation of teaching: popular myth, bias, validity or innocent bystanders? *Journal of Educational Psychology*, 92, 202-228. doi: 10.1037/0022-0663.92.1.202
- McPherson, M.A. (2006). Determinants of how students evaluate teachers. *Journal of Economic Education*, 37, 3-20. doi: 10.3200/JECE.37.1.3-20
- Palazón, A. (2011). *Evaluación de la docencia y del aprendizaje en el marco del Espacio Europeo de Educación Superior*. Tesis doctoral. Universidad Católica San Antonio de Murcia.
- Spooren, P. (2010). On the credibility of the judge. A cross-classified multilevel analysis on students' evaluation of teaching. *Studies in Educational Evaluation*, 36, 121-131. doi: 10.1016/j.stueduc.2011.02.001
- Theall, M., & Franklin, J. (2001). Looking for bias in all the wrong places: A search for truth or a witch hunt in student ratings of instruction? *New Directions for Institutional Research*, 27, 45-56. doi: 10.1002/ir.3
- Ting, K. (2000). A multilevel perspective on student ratings of instruction: Lessons from the Chinese experience. *Research in Higher Education*, 41, 637-666. doi: 10.1023/A:1007075516271.
- Wachtel, H.K. (1998). Student evaluation of college teaching effectiveness: A brief review. *Assessment and Evaluation in Higher Education*, 23, 191-210. doi: 10.1080/0260293980230207
- Worthington, A. (2002). The impact of student perceptions and characteristics on teaching evaluations: a case study in finance education. *Assessment and Evaluation in Higher Education*, 27(1), 49-64. doi: 10.1080/02602930120105054