

Tabla 4: *Medias, desviaciones estándar y coeficientes de correlación entre las variables estudiadas (segundo semestre de 2007-2008)*

Variables	M	DE	1	2	3	4	5	6	7	8	9	10	11	12	13
1. Puntuación total (29603)	4.55	.55	—												
2. F1: Enseñanza(35641)	4.37	.73	<u>.95</u>	—											
3. F2: Cumplimiento(35304)	4.71	.47	<u>.87</u>	<u>.70</u>	—										
4. Evaluación general (39356)	4.19	.99	<u>.76</u>	<u>.76</u>	<u>.58</u>	—									
5. Tamaño del grupo (41053)	30.47	15.15	-.04	-.06	-.01	-.05	—								
6. Nivel del curso (41053)	1.02	.13	.02	.02	.02	.02	-.14	—							
7. Año de estudios 38535)	3.41	1.38	.07	.07	.04	.03	.14	.09	—						
8. Curso requerido (40928)	.77	.42	-.07	-.09	-.05	-.08	.05	-.07	.14	—					
9. Motivación (40603)	3.89	1.03	.29	.32	.20	.36	-.05	.04	.03	-.11	—				
10. Hábitos de estudio (40855)	3.55	.97	.09	.11	.05	.12	-.02	.01	-.01	.00	.33	—			
11. Esfuerzo (40866)	3.62	1.08	.01	.02	.02	-.02	.05	.02	.02	.13	.19	.36	—		
12. Dificultad (40846)	3.49	1.03	-.11	-.11	-.06	-.16	.09	.02	-.01	.18	.00	.12	<u>.59</u>	—	
13. Nota esperada (40884)	4.46	.74	.27	.29	.19	.34	-.14	.05	.01	-.11	.31	.26	-.09	-.29	—

Nota. N = número de observaciones; M = media; DE = desviación estándar. ^a Curso subgraduado = 1 ; curso graduado = 2. ^b Curso requerido = 1; curso electivo = 0. Coeficientes subrayados indican una correlación moderada o alta.

En las Tablas 7, 8, 9 y 10 se resumen los resultados de los análisis de regresión múltiple en los cuales las medidas del desempeño general del profesor (la puntuación total del COE, la puntuación del Factor 1: calidad de la enseñanza, el Factor 2: cumplimiento de deberes/responsabilidades y la evaluación general del profesor/a) fueron las variables criterio o dependientes y el resto de las variables fueron las variables explicativas o independientes. El propósito de este tipo de análisis fue determinar el grado en que cada una de las variables explicativas se asociaba con el criterio, más allá de la contribución de las otras variables en el análisis. De esta forma, se pretendía identificar variables que aparentaban amenazar la validez de las medidas generales del COE. El análisis se hizo en forma jerárquica. Las dos variables que tuvieron la correlación más alta con las variables criterio se incluyeron en un primer modelo (nota esperada y motivación) y luego en un segundo modelo se añadieron las otras variables.

La estadística de tolerancia para las variables explicativas en los modelos de regresión fluctuó desde .54 hasta .96. Estos valores son mucho más altos que el criterio de una tolerancia de menos de .20 como indicador de un problema de multicolinealidad para el análisis de regresión (Menard, 2002).

El primer modelo en la regresión jerárquica, que incluyó sólo la nota esperada y la motivación para tomar el curso como variables explicativas, fue estadísticamente significativo en ambos semestres y respecto a las cuatro variables criterio: puntuación general del COE, el Factor de calidad de la enseñanza, el Factor de cumplimiento de deberes/responsabilidades y la evaluación general del profesor(a). La combinación de las dos variables explicativas estuvo asociada con el 12% y el 14% de la variabilidad de la puntuación total del COE en el primer y segundo semestre, respectivamente; con el 16% y el 14% de la variabilidad del factor calidad de la enseñanza en el primer y segundo semestre, respectivamente; y con el 6% de la variabilidad del factor cumplimiento de deberes y responsabilidades, en ambos semestres. La asociación más fuerte fue entre la combinación de la nota esperada y la motivación para tomar el curso con la evaluación general del profesor(a): 19% y 20% en el primer y segundo semestre, respectivamente. Los coeficientes beta (β) sugieren que la nota esperada y la motivación para el curso tuvieron una importancia relativa similar en dichas asociaciones. En el segundo modelo, se añadieron las otras variables explicativas. Se observó que, en

ambos semestres, éstas aportaron sólo uno por ciento adicional a la magnitud de la asociación con las cuatro variables criterio.

Tabla 7: Resultados del análisis de regresión múltiple jerárquico de la puntuación total del COE en siete variables explicativas

Variables explicativas	Primer semestre (N = 27667)			Segundo semestre (N = 27356)		
	<i>R</i>	<i>R</i> ²	<i>β</i>	<i>R</i>	<i>R</i> ²	<i>β</i>
	Modelo 1	.38,	.14*		.35	.12*
* Nota esperada			.24			.20
Motivación			.23			.23
Modelo 2	.39	.15*		.36	.13*	
Nota esperada			.22			.20
Motivación			.23			.23
Tamaño del grupo			-.03			-.01
Nivel del curso			.00			-.01
Año de estudios			.02			.06
Curso requerido			-.01			-.02
Hábitos de estudio			-.03			-.05
Esfuerzo			.03			.05
Dificultad			-.10			-.07

Nota. *β* = peso estandarizado de la variable explicativa. **P* < .001

Tabla 8: Resultados del análisis de regresión múltiple jerárquico de la puntuación del factor 1 (Enseñanza) en siete variables explicativas

Variables explicativas	Primer semestre (N = 34346)			Segundo semestre (N = 32900)		
	<i>R</i>	<i>R</i> ²	<i>β</i>	<i>R</i>	<i>R</i> ²	<i>β</i>
	Modelo 1	.40	.16		.37	.14
Nota esperada			.25			.21
Motivación			.25			.25
Modelo 2	.41	.17*		.38	.15*	
Nota esperada			.22			.19
Motivación			.25			.25
Tamaño del grupo			-.06			-.02
Nivel del curso			-.01			-.01
Año de estudios			.01			.04
Curso requerido			-.04			-.04
Hábitos de estudio			-.02			-.03
Esfuerzo			.04			.06
Dificultad			-.10			-.08

Nota. *β* = peso estandarizado de la variable explicativa. **P* < .001

Tabla 9: Resultados del análisis de regresión múltiple jerárquico de la puntuación del factor 2: (Cumplimiento) en siete variables explicativas

Variables explicativas	Primer semestre (N = 33853)			Segundo semestre (N = 32552)		
	<i>R</i>	<i>R</i> ²	<i>β</i>	<i>R</i>	<i>R</i> ²	<i>β</i>
	Modelo 1	.24	.06		.24	.06
Nota esperada			.15			.14
Motivación			.15			.15
Modelo 2	.25	.06*		.25	.06*	
Nota esperada			.14			.14
Motivación			.16			.16
Tamaño del grupo			.01			.02
Nivel del curso			.01			.00
Año de estudios			.01			.03
Curso requerido			.00			-.02
Hábitos de estudio			-.05			-.06
Esfuerzo			.04			.05
Dificultad			-.05			-.05

Nota. *β* = peso estandarizado de la variable explicativa. **P* < .001

Tabla 10: Resultados del análisis de regresión múltiple jerárquico de la evaluación general del profesor/a (combinación de ítems 41 y 42 del COE) en siete variables explicativas

Variables explicativas	Primer semestre (N = 37,963)			Segundo semestre (N = 36,247)		
	R	R ²	β	R	R ²	β
Modelo 1	.44	.20*		.43	.19*	
Nota esperada			.29			.26
Motivación			.27			.28
Modelo 2	.46*	.21		.44*	.20	
Nota esperada			.25			.23
Motivación			.29			.29
Tamaño del grupo			.01			.01
Nivel del curso			.00			.00
Año de estudios			.03			.02
Curso requerido			.00			-.01
Hábitos de estudio			-.04			-.04
Esfuerzo			.01			.02
Dificultad			-.13			-.10

Nota. β = peso estandarizado de la variable explicativa. * $P < .001$

Debido a que la nota esperada y la motivación para tomar el curso pueden depender tanto de la labor del profesor como de factores ajenos a su control, se decidió explorar que parte de la variación en estas variables podía estar asociada con las ejecuciones del profesor incluidas en el COE. Con este fin, se llevó a cabo un análisis de regresión por pasos (*stepwise*), según utilizado por Hoyt y Lee (2002) en un estudio psicométrico del sistema IDEA, que se emplea para obtener la percepción de los estudiantes del proceso de enseñanza y sus logros. En el análisis, la nota esperada y la motivación

constituyeron las variables criterio o dependientes y las variables explicativas fueron las 26 ejecuciones del profesor medidas en el COE (ítems 14 al 39).

En el análisis inicial, nueve ejecuciones del profesor contribuyeron a explicar la motivación para el curso en el primer semestre y 13 ejecuciones en el segundo semestre. Esta asociación fue estadísticamente significativa. Al examinar la aportación adicional que hacían las variables al entrar en la ecuación de regresión, se encontró que, en ambos semestres, sólo tres de ellas aportaban uno por ciento o más en la explicación de la variación tanto en la motivación como en la nota esperada. En el caso de la motivación, fueron las mismas ejecuciones en los dos semestres mientras que respecto a la nota esperada dos de las tres prácticas fueron comunes en ambos semestres. Las ejecuciones y los resultados de este análisis se resumen en las Tablas 11 y 12. Sólo el 10% de la variación en la motivación y en la nota esperada estuvo asociada con ciertas ejecuciones del profesor(a) al enseñar. Ningún ítem acerca del cumplimiento de deberes mostró relación importante con la nota esperada o con la motivación para el curso.

Tabla 11: *Resumen de los resultados del análisis de regresión, por pasos, de la motivación en 26 prácticas de enseñanza*

Núm. Item/Práctica	Primer semestre		Segundo semestre	
	ΔR^2	β	ΔR^2	β
	R = .31 * R ² = .10*		R = .31 * R ² = .10*	
14. Presenta el material con claridad.	.08	.17	.07	.15
22. Estimula al estudiante a mantenerse al día en las lecturas o tareas asignadas.	.02	.12	.02	.13
20. Hace referencia a nuevos hallazgos, publicaciones u obras recientes en áreas relacionadas con el curso.	.01	.09	.01	.10

*P < .001

Tabla 12: Resumen de los resultados del análisis de regresión por pasos de la nota esperada en 26 prácticas de enseñanza

Núm. Item/Práctica	Primer semestre		Segundo semestre	
	ΔR^2	β	ΔR^2	β
	R = .31 * R ² = .10*		R = .31 * R ² = .10*	
29. Contribuye a que los estudiantes logren los objetivos del curso.	.09	.18	.07	.16
14. Presenta el material con claridad.	.01	.11		
38. Facilita el aprendizaje mediante el uso de métodos de enseñanza tales como...	.01	.10	.01	.10
33. Redacta con claridad los exámenes u otras actividades de evaluación.			.01	.10

*P < .001

Por último, se determinaron las ejecuciones del profesor(a) que explicaban mejor la evaluación general hecha por los estudiantes. Con este fin, se estudiaron las correlaciones bivariadas entre los ítems 14 al 39 con la evaluación general (combinación de los ítems 41 y 42) y luego se hizo un análisis de regresión múltiple en el que la evaluación general fue la variable criterio y los ítems 14 al 39 fueron las variables explicativas. Los resultados se presentan en la Tabla 11. El conjunto de ítems estuvo asociado con el 67% y el 64% de la varianza en la evaluación general de profesor(a) en el primero y el segundo semestre, respectivamente. Las ejecuciones que mayor contribución hicieron en esta asociación fueron: (1) la presentación clara del material (ítem 14; $\beta = .26$ en el primer semestre y $\beta = .24$ en el segundo semestre, $ps < .001$); la contribución del profesor para que los estudiantes logren los objetivos del curso (ítem 29; $\beta = .20$ en el primer semestre y

$\beta = .19$ en el segundo semestre, $ps < .001$) y la redacción con claridad de los exámenes y otras actividades de evaluación (ítem 33; $\beta = .11$ en el primer semestre y $\beta = .09$ en el segundo semestre, $ps < .001$).

Tabla 13: Resultados del análisis de regresión de la evaluación general en las ejecuciones del profesor(a)

Ítem	Primer semestre		Segundo semestre	
	r_{xy}	β	r_{xy}	β
	$R = .82, R^2 = .67$		$R = .80, R^2 = .64$	
14. Presenta material con claridad	.72	.26	.69	.24
15. Presenta material en orden	.62	.04	.60	.02
16. Resume las ideas presentadas	.61	.05	.60	.05
17. Promueve la participación	.57	.06	.55	.04
18. Aclara dudas	.62	.05	.60	.04
19. Demuestra conocimiento	.48	.00 (ns)	.50	.02
20. Hace referencia a nuevos hallazgos	.46	.01	.47	.02
21. Verifica el nivel de entendimiento	.58	.01	.56	.01
22. Estimula a mantenerse al día	.60	.06	.59	.07
23. Acepta puntos de vista diferentes	.59	.06	.56	.06
24. Promueve ambiente de respeto	.47	.03	.46	.03
25. Asiste a clase o hace arreglos	.35	-.02	.35	-.03
26. Cumple con duración de clase	.34	.01	.36	.01
27. Utiliza el tiempo de clase	.39	-.02	.39	-.01
28. Sigue el prontuario	.47	-.03	.46	-.03
29. Contribuye al logro de los objetivos	.71	.20	.69	.19
30. Hace referencia al texto o lecturas	.41	-.01	.42	-.01
31. Anuncia fecha de exámenes	.37	-.01 ns	.35	-.02
32. Prepara exámenes enfocados	.43	-.02	.44	.00
33. Redacta con claridad exámenes	.60	.11	.59	.09

Tabla 13: (Continuación)

Ítem	Primer semestre		Segundo semestre	
	r_{xy}	β	r_{xy}	β
	$R = .82, R^2 = .67, EE = .57$		$R = .80, R^2 = .64$	
34. Ofrece con claridad instrucciones	.61	.0	.60	.05
35. Entrega los resultados de los exámenes	.40	.03	.42	.03
36. Recalca áreas de mayor dificultad	.56	.05	.56	.07
37. Facilita el aprendizaje mediante recursos	.46	.01	.46	.01
38. Facilita el aprendizaje mediante métodos	.50	.04	.49	.03
39. Hace posible la comunicación fuera	.50	.04	.51	.05

Nota. N = 28934 en el primer semestre; N = 28580 en el segundo semestre. r_{xy} = correlación bivariada entre la ejecución y la evaluación general, β = peso estandarizado de la ejecución, *ns* = no significativo estadísticamente. Todas las correlaciones bivariadas fueron estadísticamente significativas ($p < .01$). Los coeficientes β ennegrecidos fueron los de mayor peso.

Discusión de los Hallazgos

Los análisis realizados en este estudio revelaron información útil para interpretar adecuadamente las puntuaciones obtenidas con el COE, sostener en cierto grado su validez y mejorar su utilización. En general, los estudiantes le asignaron puntuaciones altas a los profesores; opinaron que éstos llevaban a cabo, con alta frecuencia, las acciones indicadoras de una enseñanza de calidad y que cumplían siempre o casi siempre con sus deberes y responsabilidades. Esta tendencia es cónsona con los resultados de otros estudios con muestras e instrumentos distintos pero con el mismo fin (por ejemplo: Hoyt y Lee, 2002; Toland y De Ayala, 2005). Por lo menos se pueden considerar dos posibles explicaciones para este hallazgo. La primera es que, en efecto, la mayor parte de la facultad ejecute la enseñanza en niveles de excelencia según éstos se definen en el COE. La segunda explicación es que los estudiantes sean indulgentes al opinar acerca de sus profesores, tendencia que ha sido observada y planteada anteriormente (Hoyt y Lee, 2002). Al presente, no

contamos con datos para determinar empíricamente el peso relativo que puedan tener estas dos explicaciones acerca de las altas puntuaciones observadas al administrar el COE.

El análisis factorial exploratorio reveló dos factores o dimensiones subyacentes a las puntuaciones del COE. El primer factor incluyó 10 ítems relacionados con la calidad del proceso de enseñanza y el segundo factor incluyó ocho ítems acerca del cumplimiento de deberes y responsabilidades del profesor en la enseñanza. Los ítems correspondientes a estos factores revelaron una consistencia interna alta. Estos hallazgos apoyan el uso de puntuaciones para representar estas dos dimensiones o factores al describir la percepción de los estudiantes sobre la labor del profesor.

Por otra parte, no se observaron factores distintos sobre el proceso de enseñanza, como ha sido evidenciado en algunos estudios (Cashin, 1989; Marsh y Dunkin, 1992; Parayitam, Desai y Phelps, 2007). Aparentemente, los estudiantes respondieron a los ítems relacionados con la forma en que enseñaba el profesor(a) fundamentados en una visión integral de lo que pasó en ese proceso y no a base de dimensiones distintas del mismo. La unidimensionalidad de los juicios de los estudiantes acerca de la calidad de la enseñanza ha sido evidenciada en algunos estudios (Beran y Violato, 2005; Griffin, 1999; Bierer y Hull, 2007). Por lo tanto, los hallazgos del presente estudio no apoyan el uso de puntuaciones para representar las dimensiones que se establecieron a priori para construir el COE.

En el análisis factorial se identificaron ocho ítems que no correspondieron a alguno de los dos factores extraídos. Estos ítems, sin embargo, no se deben descartar automáticamente por esta razón ya que los mismos no evidenciaron mala calidad. Debe examinarse la posible importancia que puedan tener éstos como fuente de información para que los profesores mejoren su ejecución o para que los comités de personal evalúen los profesores, antes de tomar una decisión final al respecto.

El análisis de las relaciones entre las puntuaciones del COE que describen las ejecuciones del profesor(a) y las medidas de otras variables arrojó hallazgos esperanzadores sobre la calidad de las medidas obtenidas con el instrumento. Por una parte, se evidenció la asociación entre la puntuaciones indicadoras de la labor del profesor(a) con una medida indirecta, aunque imperfecta, del aprovechamiento logrado por los estudiantes: la nota esperada en el curso. Se presume que si la enseñanza es efectiva, el aprovechamiento del estudiante sea alto. Se recomienda que en el futuro se

incorporaren algunos ítems para medir el aprovechamiento percibido por los estudiantes en el curso. Esta es la tendencia en la construcción de instrumentos que se utilizan en la actualidad con fines similares al del COE (Por ejemplo, el instrumento conocido como IDEA, que se utiliza en muchas instituciones de educación superior en Estados Unidos; Hoyt y Lee, 2002). La inclusión de otras medidas del aprovechamiento, en estudios futuros, podría ser una fuente importante de evidencias para validar las puntuaciones del instrumento objeto del presente estudio.

Entre las ocho variables consideradas en este estudio que no están en gran medida bajo el control del profesor y que podrían constituir amenazas a la validez de las puntuaciones del COE, se destacó la motivación para el curso. Esta fue la única variable que mostró una asociación importante, aunque débil, con las puntuaciones del COE. A pesar de que en este estudio se encontraron algunas ejecuciones del profesor asociadas con la motivación para el curso, los porcentajes de la variación común en estas asociaciones fueron bajos. Por lo tanto, se presume que la motivación para el curso depende en gran medida de factores ajenos al profesor. La importancia de esta variable como amenaza a la validez de las puntuaciones de los cuestionarios estudiantiles ha sido observada en estudios previos y la magnitud de su efecto ha sido similar al observado en este estudio (Hoyt y Lee, 2002).

Por otra parte, factores que usualmente se citan para argumentar en contra de la calidad de los datos obtenidos con instrumentos como el COE, no guardaron relaciones significativas con las puntuaciones de este instrumento. Estos factores fueron: (1) el nivel del curso, (2) si el curso era o no requerido, (3) el tamaño del grupo, (4) el año de estudios del estudiante, (5) la motivación para tomar el curso, (6) los hábitos de estudio, (7) el esfuerzo hecho en el curso y (8) la percepción de la dificultad del curso. Estos factores contribuyeron muy poco a explicar las puntuaciones del COE. En general, no se encontró evidencia para apoyar la necesidad de interpretar las puntuaciones en el contexto del nivel del curso, tamaño del grupo o si el curso es obligatorio o electivo.

Se podría cuestionar la calidad de la información relacionada con estos factores para explicar su débil asociación con las puntuaciones del COE. En efecto, todas estas variables, en particular las últimas cuatro, se midieron con un sólo ítem cada una. Esto aumenta el error de medición que, a su vez, afecta la magnitud de las correlaciones. No obstante, algunos hallazgos sugieren que esto no fue

del todo así. Algunas de dichas variables correlacionaron entre sí tal como como lógicamente se espera. Por ejemplo, se observó que a medida que la nota esperada aumentaba, mayor la frecuencia informada de las ejecuciones deseables al enseñar ($r = .31, p < .001$); a mayor la dificultad del curso percibida mayor el esfuerzo informado ($r = .58, p < .001$). Otro hallazgo que evidencia en cierto grado la validez de los datos obtenidos con el COE es la poca importancia que mostraron los ítems del factor cumplimiento de deberes para explicar la variación en la nota esperada, la motivación para el curso y la evaluación general del profesor(a). Sin embargo, los ítems sobre el proceso de enseñanza mostraron ser importantes para explicar dicha variación.

Las variables de invalidez potencial mostraron una relación mayor con la evaluación general del profesor que con la puntuación total del COE. Este hallazgo sugiere que, de utilizarse una sola puntuación para representar la percepción de los estudiantes sobre la labor del profesor en la enseñanza, sería preferible utilizar la puntuación total. Por otra parte, tal y como se esperaba, las variables de invalidez potencial mostraron una asociación más fuerte con el proceso de enseñar (Factor 1) que con el cumplimiento de los deberes y responsabilidades (Factor 2). Se supone que dichas variables tengan un efecto menor en este último factor.

En general, los hallazgos sostienen la práctica de informar los resultados del COE mediante: (1) la puntuación media de cada ítem; (2) la puntuación media obtenida en cada uno de los dos factores identificados: el proceso de enseñanza y el cumplimiento de deberes/responsabilidades, además de (3) la puntuación media de los ítems 14 al 39.

Llama la atención en este estudio la gran similitud entre los resultados obtenidos en los dos semestres considerados. Aunque se debe hacer un análisis para obtener evidencias acerca de la generalización de las inferencias derivadas de las puntuaciones del COE, la similitud entre dichos resultados apunta positivamente hacia la posibilidad de dicha generalización.

Este estudio de validez representa sólo una de varias posibles aproximaciones para validar las interpretaciones de las puntuaciones del COE, de acuerdo con su uso. La validación es un proceso complejo que incluye la recopilación de diversos tipos de evidencias mediante múltiples aproximaciones y métodos (APA, AERA y NCME, 1999). En este proceso, se acumulan las evidencias para apoyar o no las interpretaciones y los usos derivados de las puntuaciones del instrumento. Por lo

tanto, se validan las interpretaciones y el uso que se da a las puntuaciones del instrumento y no el instrumento como tal. En este estudio se obtuvieron evidencias relacionadas con algunos de los aspectos que Messick (1989, 1995) señala como importantes en la validación: (1) Sustantivo- razones teóricas de la consistencia observada en las repuestas, (2) estructural- configuración interna del instrumento y su dimensionalidad, y (3) externo – relaciones de las puntuaciones del instrumento con otras medidas. El aspecto de contenido – relevancia y representatividad de lo incluido en los ítems- se consideró anteriormente mediante los juicios de diferentes docentes del Recinto (Torres-Lugo, I y Bonilla Mújica, J., sin fecha). No se consideraron dos aspectos incluidos en la teoría de validez de Mesick (1989, 1995): la generalización- grado en que las inferencias derivadas de las puntuaciones se pueden generalizar a otras situaciones o tareas- y el consecuencial- consecuencias éticas y sociales del instrumento. Estos aspectos deben ser considerados en futuros estudios, aunque existe un gran debate respecto a este último aspecto (Crocker, L. ,1997).

Cabe señalar algunas limitaciones encontradas al realizar este estudio. Primero, la combinación de las respuestas omitidas con las respuestas No aplica evitaron su consideración apropiada en el análisis. Esto se debe evitar en el procesamiento de los datos del COE. En segundo lugar, las medidas de algunas variables consideradas como posibles amenazas a la validez (la motivación, los hábitos de estudio, el esfuerzo y la dificultad del curso) se basaron solo en un ítem cada una. No se sabe la magnitud del efecto que esto haya causado en los resultados. Hay ciertas indicaciones, señaladas anteriormente en este informe, que sugieren la posibilidad de que el efecto no sea de gran magnitud. En el futuro se deben considerar medidas que partan de múltiples ítems. La medida utilizada como indicadora de aprovechamiento (la nota esperada) es limitada. Por lo tanto, en estudios futuros, se deben añadir otras medidas tales como la nota obtenida en el curso, la percepción del estudiante sobre el aprovechamiento logrado en el curso, entre otras. La medida del aprovechamiento es indispensable para validar las puntuaciones de instrumentos sobre la calidad de la enseñanza.

Por último, es importante enumerar algunas recomendaciones adicionales que se derivaron de este trabajo.

1. Reducir el instrumento con la eliminación de los primeros 10 ítems. Los primeros cinco constituyen información que puede estar pre-codificada junto a las hojas de contestaciones. Los ítems 6 al 10 se incluyeron para estudios de validación y existen bancos de datos de varios años con la información obtenida con ellos para poder realizar los mismos.
2. Sustituir las preguntas de los ítems 41 y 42 por una o dos que evalúen mejor la calidad de la labor del profesor. En la literatura se ilustran varias de éstas.
3. Considerar la posibilidad de eliminar algunos otros ítems, por ejemplo: (a) aquellos cuya frecuencia de respuestas omitidas o *No aplica* fueron relativamente altas, (b) los que no se agruparon en uno de los dos factores. En particular se recomienda considerar la eliminación de los siguientes:

Ítem 19 – Demuestra conocimiento de los temas del curso. Este ítem no correlacionó al menos moderadamente con un factor. Esta información puede obtenerse de los pares.

Ítem 20 - Hace referencia a nuevos hallazgos, publicaciones u obras recientes en áreas relacionadas al curso. Este ítem tuvo la frecuencia mayor de omisiones y respuesta *No aplica*.

Ítem 30 – Hace referencia al libro de texto o lecturas durante el curso. Este ítem tuvo una frecuencia relativamente alta de omisiones y la respuesta *No aplica*. Además, no correlacionó al menos moderadamente con un factor.

Conclusión

Los resultados de este estudio sugieren que las puntuaciones obtenidas con los ítems del COE se comportaron en forma mayormente consistente con las expectativas sobre este tipo de instrumento. A pesar de que las medidas utilizadas como indicadoras de validez de dichos ítems no fueron los mejores posibles, los resultados fueron similares a los hallazgos de estudios publicados en revistas académicas especializadas en los que se han avaluado otros instrumentos con propósitos similares al COE. Los hallazgos presentados evidencian, en cierto grado, la validez de las puntuaciones derivadas con el instrumento que nos ocupa. Estudios futuros, mediante el uso de otros indicadores

de validez y otras estrategias y técnicas de análisis podrían ofrecer información adicional para validar estas puntuaciones.

Referencias

- American Psychological Association, American Educational Research Association, and National Council on Measurement in Education (APA, AERA y NCME; 1999). *Standards for educational and psychological testing*. Washington, DC: American Psychological Association.
- Arreola, R. A. (2007). *Developing a comprehensive faculty evaluation system*. Williston, VT: Anker Publishing.
- 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, 593-601.
- Bierer, S. B. y Hull, A. L. (2007). Examination of a clinical teaching effectiveness instrument used for summative faculty evaluation. *Evaluation & the Health Professions*, 30, 339-361.
- Cashin, W. E. (1989). Defining and evaluating college teaching. *IDEA Paper No. 21*. Manhattan, KS: Kansas State University, Center for Faculty Evaluation & Development.
- Cashin, W. E. (1995). Student ratings of teaching: The research revisited. *IDEA Paper No. 32*. Manhattan, KS: Kansas State University, Center for Faculty Evaluation and Development.
- Comité Institucional para la Elaboración de Instrumentos de Evaluación del Personal Docente. (2006, 12 de septiembre). *Informe sobre el "Cuestionario de Opinión Estudiantil" (COE)*. Informe enviado al Dr. Jorge I. Vélez Arocho, Rector y Presidente del Senado Académico del Recinto Universitario de Mayagüez.
- Costello, A. B. y Osborne, J. W. (2005). Best practices in exploratory factor analysis: Four recommendations for getting the most from your analysis. *Practical Assessment, Research & Evaluation*, 10(7). Disponible en línea en <http://paronline.net/getvn.asp?v=10&n=7>

- Crocker, L. (1997). Editorial: The great validity debate. *Educational Measurement: Issues and Practice*, 16, 4.
- Crocker, L. y Algina, J. (1986). *Introduction to classical and modern test theory*. NY: Holt, Rinehart and Winston.
- Garson, G. D. (2009). Factor analysis, en *Statnotes: Topics in multivariate analysis*. Obtenido el 15 de diciembre de 2009 en <http://faculty.chass.ncsu.edu/garson/pa765/statnote.html>
- Griffin, B. W. (1999). *Internal structure of scores obtained from the Student Ratings of Instruction instrument used at Georgia Southern University*. Obtenido el 23 de octubre de 2003 en www2.gasou.edu/acadaff/instruction/sturatings_a.html.
- Henson, R. K. y Roberts, J. K. (2006). Use of exploratory factor analysis in published research: Common errors and some comment on improved practice. *Educational and Psychological measurement*, 66, 393-416.
- Hoyt, D. P., & Lee, E. J. (2002). Basic data for the Revised IDEA System. *Technical Report No. 12*. The IDEA Center, Manhattan, KS.
- Marsh, H. W., & Dunkin, M. (1992). Students' evaluations of university teaching: A multidimensional perspective. *Higher Education: Handbook on theory and research* (Vol 8, pp 143-234). NY: Agathon.
- Menard, S. (2002). *Applied logistic regression analysis* (2nd ed.). Thousand Oaks, CA: Sage.
- Messick, S. (1989). Validity. En R.L. Linn (Ed.), *Educational Measurement* (3th. Ed.). New York: American Council on Education and Macmillan.
- Messick, S. (1995). Standards of validity and the validity of standards in performance assessment. *Educational measurement: Issues and Practice*, 14, 5-8
- Stevens, J. P. (2002). *Applied multivariate statistics fro the social sciences*. Mahwah, NJ: Lawrence Erlbaum.
- Toland, M. D., & De Ayala, R. J. (2005). A multilevel factor analysis of students evaluations of teaching. *Educational and Psychological Measurement*, 65, 272-296.

Torres-Lugo I. & Bonilla Mújica, J. (sf). *Resumen de estudios COE (2000-2004)*. Disponible en www.uprm.edu/evaluaciondocente/docs/EstCOE00-04.pdf

Wendorf, C. A., & Alexander, S. (2005). The influence of individual- and class-level fairness-related perceptions on student satisfaction. *Contemporary Educational Psychology, 30*, 190-206.

Apéndice
Histogramas de las Puntuaciones de los Ítems