

Evaluación del desempeño docente en cursos de especialización médica. Validación de un instrumento

Adrián
Martínez-González,¹
Alberto
Lifshitz-Guinzberg,²
Raúl
Ponce-Rosas,³
Víctor Aguilar⁴

¹Departamento de Salud Pública, Facultad de Medicina
²Facultad de Medicina
³Departamento de Medicina Familiar, Facultad de Medicina,
⁴Centro Médico Nacional La Raza, Instituto Mexicano del Seguro Social

Autores 1, 2 y 3, Universidad Nacional Autónoma de México

Distrito Federal, México

Comunicación con:
Adrián
Martínez-González.
Tel: (55) 5623 2476.
Correo electrónico:
aamg@servidor.unam.mx

RESUMEN

Objetivo: validar un instrumento para evaluar el desempeño docente en cursos de especialización médica.

Material y métodos: se procuró la validez de contenido por medio de la revisión de documentos, la opinión de expertos y de médicos estudiantes de varias especialidades. La validez de constructo se estimó con un análisis factorial exploratorio, conservando las preguntas de carga ≥ 0.50 y clasificadas en un solo factor; se calculó el coeficiente alfa de Cronbach.

Resultados: 157 residentes evaluaron a nueve profesores titulares. Se obtuvo una tasa de respuesta de 75.1 %. El cuestionario quedó constituido por 37 preguntas. Se identificaron cinco factores con valores propios ≥ 1 que explicaron en conjunto 52.5 % de la varianza total. El coeficiente alfa de Cronbach fue de 0.9596 para el cuestionario.

Conclusiones: el cuestionario alcanzó alta fiabilidad y puede ser utilizado para evaluar el desempeño docente dentro de las especialidades médicas.

SUMMARY

Objective: our purpose was to validate an instrument to assess the academic performance in medical specialization courses.

Methods: the design and content validity of the questionnaire had sustained in the literature review, as well as in the opinion of had been given by experts and residents. The construct validity was measure for exploratory factorial analysis, only questions with factor charge ≥ 0.50 and that classified as a single factor has been retained and Cronbach's alpha coefficient was calculated.

Results: residents (157) evaluated to 9 professors. The answer rate was of 75.1 %. The questionnaire has 37 items. Five factors were identified with eigen values ≥ 1 which explained 52.5 % of the total variance. The Cronbach's alpha coefficient was of 0.9596 for the questionnaire.

Conclusions: the instrument has a high grade of reliability and it can be used for evaluation and pursuit of the academic performance in medical specialties courses.

Recibido: 13 de febrero de 2007

Aceptado: 8 de mayo de 2007

Introducción

Evaluar la calidad del desempeño docente es una práctica cada vez más difundida en la educación superior, originada con el fin de dar respuesta a dos tipos de factores: los relativamente externos a las propias instituciones universitarias (escasez de recursos económicos, responsabilidad, satisfacción de necesidades sociales, competitividad y acreditación), que obligan a establecer patrones de excelencia y eficiencia; y los internos de las universidades que inciden en el avance de la educación ante la necesidad de la evaluación como elemento para el logro de un mejoramiento

continuo y una educación cada vez más eficaz y eficiente, a partir de criterios preestablecidos. La evaluación es el enjuiciamiento sistemático de la valía o el mérito de un objeto,¹ Esta definición implica siempre la realización de un juicio, cuyo papel dentro del proceso evaluativo es identificar puntos débiles y fuertes.

La mayoría de las instituciones educativas registran, evalúan y reconocen sistemáticamente las actividades de investigación, costumbre que contrasta con el menor interés en las actividades docentes, que frecuentemente solo se aprecian con sentido administrativo y, en comparación, reciben menos créditos.²⁻⁶

Palabras clave

- ✓ educación médica de posgrado
- ✓ mediciones educacionales
- ✓ profesorado

Key words

- ✓ graduate medical education
- ✓ educational measurement
- ✓ faculty

Durante las últimas décadas se ha multiplicado la demanda de evaluar el desempeño docente y de otorgar el reconocimiento que merecen los académicos de calidad excepcional, y de formular medidas formativas y correctivas, académicas, pedagógicas y actitudinales de la planta académica.⁷ Es decir, existe la necesidad de llevar a cabo una evaluación del desempeño docente, válida, confiable, objetiva, factible, ética y útil. Sin embargo, se reconoce que el desempeño docente es de naturaleza compleja y multidimensional (implica conocimientos, habilidades, actitudes, valores, investigación, orientación, organización, supervisión, tutoría, etcétera).

Aunque se utilicen los mejores y más modernos planes de estudio, infraestructura y métodos de enseñanza, los sistemas educativos continúan dependiendo de la calidad y desempeño de sus profesores. La importancia de esta labor y el papel de la evaluación como punto de apoyo para la toma de decisiones racionales, positivas y de compromiso participativo, justifican la necesidad de valorarla.^{1,8}

Por lo anterior, no se puede hablar de desempeño docente eficiente de manera general y absoluta sino, en nuestro caso, en función de las especializaciones médicas dado que la evaluación de la calidad del desempeño docente se presenta en una situación determinada, bajo métodos y estrategias docentes que buscan un propósito común: la formación de médicos especialistas de alto nivel.^{9,10}

Para evaluar la calidad del desempeño docente se han utilizado distintas estrategias complementarias, entre las que destacan la evaluación de los docentes mediante la opinión de los estudiantes, de los exalumnos, por logros del aprendizaje de los estudiantes, autoevaluación, evaluación por pares, mediante portafolio, entre otras.¹¹⁻¹⁴

La estrategia más ampliamente investigada y empleada, incluso en nuestro medio, es la evaluación que hacen los alumnos desde su particular punto de vista, de las actividades docentes de sus profesores, la cual ha mostrado ser una fuente de información consistente, válida y útil para retroalimentar y mejorar la docencia.^{6,15-18}

Se trata de un procedimiento centrado en la opinión de los alumnos sobre las estrategias de enseñanza, la relación profesor-estudiante, el desarrollo de habilidades para solucionar problemas, los tipos de evaluación del aprendizaje, la capacidad de motivación del docente, entre otras dimensiones. Para ello se utilizan instrumentos

que en su mayoría están basados en la noción de las dimensiones mencionadas. Lo anterior permite formular juicios de valor acerca de los puntos débiles y fuertes del desempeño docente.¹⁹⁻²⁶

Conviene resaltar la importancia de realizar la evaluación del desempeño de los profesores, ya que permite brindarles retroalimentación y la oportunidad de hacer modificaciones que redunden en beneficio de los estudiantes.^{27,28} Cuando se tiene información del desempeño de los profesores, los cuerpos colegiados de la institución pueden realizar acciones que beneficien a la planta docente en su conjunto, ya que a partir del diagnóstico desprendido de la evaluación es posible identificar los puntos débiles y tratar de fortalecerlos, poniendo en marcha los mecanismos y las estrategias necesarias para cubrir estas situaciones. Organismos nacionales (Consejo Mexicano para la Acreditación de la Educación Médica) e internacionales (*World Federation for Medical Education*)²⁹ destacan la evaluación del profesorado entre los criterios para la acreditación de las escuelas y facultades.

En la educación médica de nuestro país se han desarrollado estrategias y construido instrumentos para evaluar el desempeño docente a nivel de pregrado.^{11,30-35} El propósito de este estudio fue diseñar y validar un instrumento que permita evaluar el desempeño docente en cursos de especializaciones médicas con el fin de perfeccionar y reconocer esta labor.

Material y métodos

Estudio observacional, descriptivo, transversal, exploratorio; el proceso de diseño y validez del cuestionario se desarrolló en cuatro fases, las tres primeras para la validez de contenido y la cuarta para la validez de constructo y confiabilidad.

■ *Primera fase:* con base en la revisión de la literatura sobre evaluación del desempeño docente y en el Plan Único de Especializaciones Médicas de la Universidad Nacional Autónoma de México, realizada por un grupo interdisciplinario de expertos con posgrado y una larga trayectoria académica en relación con la educación médica, se identificaron las dimensiones y contenidos que definieron el “desempeño docente”. Se construyó un primer instrumento que consideró 40 preguntas distribuidas en cinco cate-

gorías. Se utilizó una escala de respuesta tipo Likert de cuatro opciones ordinales (nunca-siempre).

■ *Segunda fase:* el primer instrumento se presentó a 90 residentes de la especialidad de Pediatría Médica de la Secretaría de Salud del Distrito Federal, para que en su opinión determinaran si las categorías y contenidos propuestos definían el desempeño docente, y modificaran, agregaran o eliminaran categorías y preguntas, con lo cual se obtuvo un segundo instrumento con 51 preguntas.

■ *Tercera fase:* el segundo instrumento fue expuesto a cinco expertos en educación (con grado de maestría o doctorado en educación o áreas afines y con 10 o más años de experiencia docente) para que ubicaran las preguntas en las cinco categorías propuestas y que además valoraran su claridad con una escala de cuatro opciones, de “muy claro” a “no se entiende”. Las preguntas se les mostraron en forma aleatoria para evitar que los expertos clasificaran de manera sesgada los reactivos. Con los resultados se constituyó un tercer instrumento que se utilizó para la validez de constructo, con 51 preguntas distribuidas en cinco categorías.

■ *Cuarta fase:* la última versión se aplicó a la población blanco constituida por 209 residentes de las especializaciones de pediatría, cirugía pediátrica, cirugía general, otorrinolaringología, oftalmología, radiodiagnóstico, anestesiología, urgencias médico-quirúrgicas y medicina interna del Hospital General, Centro Médico Nacional La Raza, Instituto Mexicano del Seguro Social. La relación residente-docente en estas especialidades

fue de 17:1. Los criterios de inclusión fueron: todos los médicos residentes sin distinción de sexo o edad, inscritos en el curso universitario 2001 y con más de seis meses en el proceso educativo.

Fueron excluidos los residentes que no desearon participar. Los cuestionarios con respuestas incompletas fueron eliminados. El cuestionario final con 51 reactivos se aplicó en las aulas con un tiempo máximo de 20 minutos, por un médico especialista y un profesor de posgrado, a un total de 157 médicos residentes (pediatría 35, radiodiagnóstico 29, otorrinolaringología 27, cirugía pediátrica y general 20, urgencias médico-quirúrgicas 18, oftalmología 14, medicina interna 7, anestesiología 7). Se les explicó el objetivo del estudio y se les solicitó que evaluaran a nueve profesores titulares de la especialidad.

Para determinar la validez de constructo, la información obtenida se capturó en una base de datos en el programa SPSS versión 12; se realizó análisis multivariado utilizando la estrategia de análisis factorial exploratorio para identificar factores subyacentes. Respecto a la confiabilidad, se aplicó el análisis de consistencia interna de alfa de Cronbach.

Los supuestos mínimos y sus valores de referencia a cumplir en la validez de constructo y confiabilidad se muestran en el cuadro I, en la primera y segunda columnas. Se determinaron siete criterios y sus respectivos valores de referencia para considerar que el modelo de validación fuera adecuado. *A priori* se determinó que los siete criterios fueran superados para considerar al modelo útil y adecuado para evaluar el desempeño docente.

Cuadro I
Supuestos del modelo de desempeño docente para validar el constructo

Supuestos del modelo	Valor de referencia	
	Esperado	Obtenido
1. Determinante de la matriz de correlación múltiple	$p \leq 0.01$	$p = 0.0001$
2. Prueba de esfericidad de Bartlett	$p \leq 0.01$	$p = 0.0001$
3. Adecuación del muestreo (Kaiser-Meyer-Olkin)	≥ 0.60	0.89
4. Componentes principales y factores por obtener o confirmar (dimensiones del constructo)	≤ 5 a priori	11
5. Varianza explicada por el modelo	$\geq 50.0\%$	52.5%
6. Cargas mínimas de los reactivos para ser retenidos en los factores	≥ 0.50	0.51 a 0.80 para todas las variables
7. Coeficiente de confiabilidad alfa de Cronbach	≥ 0.70	0.95

Se utilizó análisis factorial para determinar los valores obtenidos. La rotación de los factores a la solución inicial fue varimax para optimar las cargas de las variables en los factores. Los eigen values (valores propios) de los 5 factores determinados a priori fueron: Factor 1 = 17.4, Factor 2 = 3.26, Factor 3 = 2.25, factor 4 = 2.14 y factor 5 = 1.69.

El modelo resultó confiable con alfa de Cronbach = 0.95 y válido con 6 criterios rebasados según los valores de referencia.

Resultados

El modelo final con sus valores de referencia para cada criterio se presenta en la tercera columna del cuadro I; seis criterios fueron superados. Los resultados permitieron el análisis complementario final del instrumento definitivo (anexo).

A través del análisis factorial solo se retuvieron las preguntas con carga ≥ 0.50 y que cargaron en un solo factor, por lo que se eliminaron 14 y el cuestionario quedó constituido por 37 preguntas. Se identificaron cinco factores con valores propios > 1 que explicaron en conjunto 52.5 % de la varianza total; la varianza de cada uno de los factores de mayor a menor y de acuerdo a su contenido fueron denominados “relación profesor-estudiante y motivación” (34.2 %), “metodología” (6.4 %), “evaluación” (4.4 %), “capacidad de solución de problemas” (4.2 %) y “conocimiento de la materia” (3.3 %). El coeficiente alfa de Cronbach fue de 0.9596 para el cuestionario global y para cada uno de los factores fue de 0.958, 0.959, 0.958, 0.958 y 0.959.

Discusión

La aplicación del análisis factorial exploratorio permitió reducir el número de preguntas del cuestionario e identificar cinco dimensiones: “relación profesor-residente y motivación”, “metodología”, “evaluación”, “capacidad de solución de problemas” y “conocimiento de la materia”.

Los resultados del análisis arrojaron que el factor de mayor relevancia en la evaluación del desempeño docente en las especializaciones médicas fue “la relación profesor-residente y motivación”, con una varianza explicada de 34.2 %. Este resultado es semejante al del estudio de Feletti³⁶ con una varianza de 48 %; en contraste, Batista³⁷ y De Grave³⁸ lo refieren en último lugar, con una varianza de 5 y 10.9 %, respectivamente. El factor en cuestión revela que el proceso educativo no solo implica el aprendizaje de conocimientos, habilidades y destrezas, sino también fomentar una adecuada relación médico-paciente. Para lo anterior es preciso que se desarrolle una adecuada relación profesor-residente en un ambiente de respeto y confianza recíprocos.

El factor I, “relación profesor-alumno y motivación”, fue el de mayor carga factorial. Para el

logro del aprendizaje es imprescindible el respeto, la comunicación y la creación de un ambiente adecuado. El respeto permite valorar las posibilidades de progreso y perfeccionamiento de cada persona y atender las diferencias individuales.³⁹

El conjunto de preguntas de este factor también implica los contenidos del aprendizaje. La función docente consiste en actuar como intermediario entre los contenidos del aprendizaje y el proceso de elaboración de cada uno de los residentes para lograr el conocimiento. Es el docente quien determina que la actividad del residente sea constructiva a través de la estimulación de la participación y de conferir importancia a los comentarios y puntos de vista de éste.

El aprendizaje se interrelaciona con lo motivacional. El docente deberá estimular la participación, dar credibilidad, respetar las ideas de los miembros del grupo y considerar los aspectos personales de los estudiantes. Esto favorece la disposición y actitud para el aprendizaje significativo.

Los preguntas relativas a los factores II, III y IV (“metodología”, “evaluación” y “solución de problemas”) versaron en torno a valorar la competencia del docente para facilitar el aprendizaje, evaluar y realimentar al estudiante. Así mismo, valúan las habilidades del docente para estimular el razonamiento clínico del residente para la evaluación y solución efectiva de los problemas de salud de los pacientes.

La evaluación debe ser congruente con el programa, los contenidos y los objetivos del curso y no reducirse a la medición del rendimiento de los estudiantes, sino proporcionar información significativa al profesor y a los estudiantes sobre lo que está sucediendo en el proceso de aprendizaje. La evaluación en este nivel es esencialmente formativa para los residentes al utilizar la realimentación para la autorregulación del aprendizaje. Durante la realimentación se les debe mostrar sus logros, potencialidades y deficiencias en el proceso educativo y la manera de superarlas.

La evaluación así entendida sirve como estímulo para aprender y cumple un papel fundamental en la medida en que a través de ella se autoriza al profesional para ejercer y se garantiza a la sociedad que quien quede autorizado posee la formación adecuada. Si la evaluación cumple con estos requisitos, el docente alcanzará el desarrollo personal y profesional del estudiante.

La evaluación de la solución de problemas sondea las habilidades del docente encaminadas a que el estudiante desarrolle su proceso de razonamiento. Éste es estimulado a lo largo de los cursos a través del análisis de los casos, que inicia desde que el estudiante identifica los síntomas y signos más importantes, formula preguntas y plantea problemas con sus posibles soluciones. Para solucionar el problema, el estudiante tiene que buscar nueva información, analizarla, cuestionarla, delimitar sus alcances y limitaciones y confrontarla con la práctica.

En este proceso de elaboración del conocimiento, la discusión y contrastación de los puntos de vista vertidos en el análisis y solución de los casos adquieren especial importancia, ya que de esta manera cada estudiante ejercita continuamente el proceso de razonamiento.

El estudiante, al desempeñar un papel activo en la elaboración del conocimiento, desarrolla una actitud crítica para analizar la información. Esta búsqueda y actitud favorecerá hábitos de estudio independiente y del pensamiento crítico y reflexivo para la toma de resoluciones inteligentes fundamentadas en la evidencia científica.

La finalidad última de la adquisición de conocimientos es su utilidad para entender determinadas situaciones y resolver problemas relacionados con la profesión médica y la salud de los pacientes o de la población.

En contraste, el factor V ("conocimiento de la materia") quedó en último lugar con una varianza explicada de 3.3 %, resultado semejante a lo encontrado por Irby⁴⁰ con una varianza de 3.6 %. Esto contrasta con el estudio de Stritter,⁴¹ quien lo identificó como el principal, con una varianza de 27 %; Irby⁴² en cambio, lo refiere en segundo lugar con una varianza de 11.7 %, así como Feletti,³⁶ con una varianza de 9 %. Las interrogantes de este factor se relacionan con el conocimiento del profesor, dominio, actualización y su vinculación con la práctica y hacen referencia a que no se puede enseñar lo que no se sabe.

Los contrastantes resultados probablemente se expliquen debido a que el Hospital General del Centro Médico Nacional La Raza es de alto nivel académico y asistencial, donde los profesores son expertos en su especialidad. Por lo anterior, es conveniente aplicar este instrumento de evaluación en diferentes hospitales y niveles de atención a la salud y observar los resultados.

Así mismo, es necesario considerar que los resultados obtenidos son de médicos formados que se encuentran en un proceso de especialización, en comparación con los otros estudios en estudiantes de pregrado. La metodología educativa utilizada en las especializaciones está centrada en la solución de problemas. Esto confiere una enorme importancia a la elaboración del conocimiento, al desarrollo de hábitos de estudio independiente y a la estimulación del pensamiento crítico y reflexivo por los residentes.

Una limitante de este estudio fue que la muestra de residentes no es representativa de toda la población de los mismos, por lo que los resultados no son generalizables al resto de los alumnos de las especializaciones médicas en esta institución.

Se debe considerar que el modelo no es puro, ya que en el análisis factorial hubo 11 factores donde se distribuyeron los reactivos eliminados. Esta situación obliga a pensar que es necesario probar los reactivos retenidos y los eliminados en contextos educativos semejantes para probar la consistencia de los primeros e identificar si los eliminados vuelven a distribuirse en factores no considerados *a priori* y con ello eliminarlos en definitiva.

El instrumento de evaluación muestra un alto grado de fiabilidad en las cinco dimensiones del desempeño docente, avalado por el resultado de consistencia interna alfa de Cronbach de 0.9596. Se debe tener en cuenta que en la práctica no existen coeficientes perfectos (1.0), y que la mayoría de los autores consideran coeficientes muy altos a los mayores de 0.81.^{43,44} Así, los estudiantes de diferentes especialidades pueden ser considerados evaluadores fiables de la actuación del docente.

Respecto a los criterios mínimos, que en este trabajo consideramos para emitir un juicio final acerca de la validez de constructo, es preciso mencionar que en los últimos 40 años diversos autores en el área médica han considerado solo algunos de los criterios que aquí proponemos: Valle,³³ Feletti,³⁶ Batista,³⁷ De Grave,³⁸ Irby,^{40,42} Stritter,^{41,46} Cotsonas,⁴⁵ Hayward,⁴⁷ Litzelman,⁴⁸ Chaves.⁴⁹

A lo largo del tiempo es posible observar cada vez más autores que consideran criterios adicionales para emitir un juicio final respecto al proceso de validación. Entre las décadas de 1970 y 1980, solo referían el tipo de rotación del análisis factorial sin proporcionar detalles técnicos y metodológicos del mismo, como lo hacen actualmente los autores.

Consideramos que para emitir un juicio de valor de los procesos de validez, es necesario y muy útil que se proporcione información adicional del análisis factorial para que los lectores analicen si con ello la validez de constructo tiene un apoyo objetivo y no simplemente por utilizar esta técnica se dé por hecho que se valida un instrumento de medición.⁵⁰

Es inadecuado considerar con carácter definitivo la validez de un instrumento de evaluación. Solo se puede señalar la validez de factores e indicadores particulares en relación con un criterio específico. Desde la perspectiva de la validez de constructo, no se valida de manera absoluta sino que se interpretan los datos obtenidos en contextos específicos y con un propósito determinado.

Cabe señalar que este trabajo constituye la fase inicial de la evaluación psicométrica del cuestionario de evaluación docente, por lo que es preciso continuar con otras investigaciones en diversas instituciones para evaluar con más precisión la validez de los resultados obtenidos en el Hospital General del Centro Médico Nacional La Raza.

Dada la variedad y complejidad del desempeño docente efectivo, además de la opinión de los residentes es importante tomar en cuenta otros tipos de evaluación como la de pares y la autoevaluación, a fin de elevar la calidad del proceso enseñanza-aprendizaje. Los instrumentos que se utilicen deberán ser confiables y válidos. Es susceptible de evaluación toda la planta académica en general, como el equipo docente de una especialidad o de un hospital.

Puesto que los profesores de medicina de las especializaciones médicas excepcionalmente cuentan con una verdadera formación docente, se han realizado esfuerzos importantes por profesionalizar la docencia, con resultados favorables mediante metodologías participativas.³¹ La evaluación es necesaria para realimentar y mejorar el desempeño docente en las especializaciones médicas, otorgar reconocimiento y apoyo a las labores docentes de buena calidad y orientar los cursos de inducción y formación al respecto, con el propósito de que el residente sea un médico especialista altamente competente con orientación humano-profesional (el ser), formación intelectual (el saber) y desempeño operativo eficiente (saber hacer y hacer).

Referencias

1. Stufflebeam D, Shinkfield. Evaluación sistemática. Barcelona, España: Paidós; 1993.
2. Blank R. Faculty support for evaluation of teaching: a test of two hypothesis. *J Higher Educ* 1978; 49:163-176.
3. Murray HG. The impact of formative and summative evaluation of teaching in North American universities. *Assessment and Evaluation in Higher Education* 1984;9:117-132.
4. McKeachie WJ, Lin YG, Dangherty M, Moffett MM, Meigler C, Nork J, Walz M, Baldwin R. Using student ratings and consultation to improve instruction. *Br J Educ Psychol* 1980;50:168-174.
5. Perry RP. Instruction in higher education. *J Educ Psychol* 1990;82(2):183-188.
6. Cohen PA. Effectiveness of student-rating feedback for improving college instruction: a meta-analysis of findings. *Res Higher Educ* 1990;13: 321-341.
7. McKeachie WJ. Student ratings. The validity of use. *Am Psychol* 1997;52:1218-1225.
8. Escudero-Escorza T. ¿Se pueden evaluar los centros educativos y sus profesores? Zaragoza, España: Instituto de Ciencias de la Educación, Universidad de Zaragoza; 1980.
9. Landshere G. Seleccionar a los educadores. *Perspectivas* 1980;3:349-356.
10. Universidad Nacional Autónoma de México. Plan Único de Especializaciones Médicas. México: UNAM; 1993.
11. Cruz-Iturrabarria SL. Consideraciones para la construcción de un modelo de evaluación de profesores. Tesis de Licenciatura en Pedagogía, Facultad de Filosofía y Letras, Colegio de Pedagogía, Universidad Nacional Autónoma de México, México, 1981.
12. Howard GS, Conway CG, Maxwell SE. Construct validity of measures of college teaching effectiveness. *J Educ Psychol* 1985;77(2):187-196.
13. Feldman KA. Instructional effectiveness of college teachers as judged by teachers themselves, current and former students, colleagues, administrators and external (neutral) observers. *Res Higher Educ* 1989;30:113-135.
14. Kremer JF. Construct validity of multiple measures in teaching, research, and service and reliability of peer ratings. *J Educ Psychol* 1990; 82(2):213-218.
15. Marsh HW. Students' evaluations of university teaching: dimensionality, reliability, validity, potential biases and utility. *J Educ Psychol* 1984; 76(5):707-754.
16. Marsh HW, Fleiner H, Thomas CS. Validity and usefulness of student evaluations of instructional quality. *J Educ Psychol* 1975;67(6):883-889.

17. Wilson RC. Improving faculty teaching. *J Higher Educ* 1986;57:196-211.
18. Abrami PC, D'Apollonia S, Cohen PA. Validity of student ratings of instruction: what we know and what we do not. *J Educ Psychol* 1990;82:219-231.
19. Lastra SA. Una evaluación de la tarea docente, a partir de las percepciones estudiantiles. *Rev Educ Sup ANUIES* 1983;12:96-109.
20. Slotnick BH, Rusell G, Durkovich G. Dimensions of medical students' perceptions of instruction. *J Med Educ* 1975;50:662-666.
21. Gómez-Junco H. La evaluación de la enseñanza por la opinión de los alumnos. *Rev Educ Sup ANUIES* 1976;17:30-47.
22. Rippey RM. *The evaluation of teaching in medical schools*. New York, NY: Springer Publishing; 1981.
23. Irby DM. Evaluating instruction in medical education. *J Med Educ* 1983;58:844-849.
24. Blanco SL, Sanz-Paz JJ. Evaluación de la docencia mediante cuestionario de alumnos: Universidad de Cantabria (curso 86-87). *Studia Paedagógica* 1988; 20:41-71.
25. Tejedor FJ, García-Valcárcel A. La evaluación de la calidad de la docencia universitaria (en el marco de la evaluación institucional) desde la perspectiva del alumno. En: Tejedor FJ, Rodríguez-Diéguez JL, editores. *Evaluación educativa II. Evaluación institucional. Fundamentos teóricos y aplicaciones prácticas. Documentos didácticos 157*. Salamanca, España: IUCE, Universidad de Salamanca; 1996. p. 5-83.
26. Luna SE. Jerarquía de las dimensiones de evaluación de la actividad docente en estudios de postgrado. En: Rueda-Beltrán M, Nieto-Gutiérrez J, editores. *La evaluación de la docencia universitaria*. México: Facultad de Psicología, División de Estudios de Postgrado, Centro de Investigaciones y Servicios Educativos, UNAM; 1996. p. 99-123.
27. Sánchez-Sosa JJ, Martínez-Guerrero JI. Diagnóstico y realimentación del desempeño docente mediante evaluaciones de alumnos. *Rev Mex Psicol* 1993;10(2):153-173.
28. Marsh HW, Roche LA. Making students evaluations of teaching effectiveness effective. *Am Psychol* 1997; 52:1187-1197.
29. World Summit on Medical Education. *Proceedings*. Edinburgh, UK, 8-12 August 1993. *Med Educ* 1994;28(Suppl 1):1-171.
30. Cruz-Talonia D. Desempeño académico del profesor del curso de histología de la Facultad de Medicina UNAM. *Rev Fac Med UNAM* 1985;28: 116-123.
31. Pérez-Rodríguez BA, Viniestra-Velázquez L. La formación de profesores de medicina. Comparación de dos estrategias educativas en el aprendizaje de la crítica de la información. *Rev Invest Clin* 2003;55(3):281-288.
32. Varela-Rueda CE. Evaluación estudiantil de la actuación del profesor y su efecto sobre la efectividad docente. *Gac Med Mex* 1989;125(5-6):151-159.
33. Valle RM, Petra I, Martínez-González A, Rojas-Ramírez JA, Morales-López S, Piña-Garza B. Assessment of student performance in problem-based learning tutorial sessions. *Med Educ* 1999; 33: 818-822.
34. Rodríguez-Carranza R, Martínez-González A, Ponce-Rosas R, Contreras E, Colina-Ramírez C, Cerritos A. Una nueva estrategia para evaluar la calidad del desempeño docente en las instituciones de educación superior. Resultados de su aplicación en la Facultad de Medicina de la UNAM. *Rev Educ Sup ANUIES* 2000;29: 129-141.
35. Valle RM, Alaminos I, Contreras E, Salas LE, Tomasini P, Varela M. Student questionnaire to evaluate basic medical science teaching (METEQ-B). *Rev Med IMSS* 2004;42(5):405-411.
36. Feletti GI, Doyle E, Petrovic A, Sanson FR. Medical student's evaluation of tutors in a group — learning curriculum. *Med Educ* 1982;16:319-325.
37. Batista JE. *La evaluación del profesor universitario*. Colombia: Universidad de Antioquia; 1987.
38. De Grave WS, Dolmans DHJM, van der Vleuten C. Tutor intervention profile: reliability and validity. *Med Educ* 1998;32:262-268.
39. Pasten-Núñez J. Una reflexión en torno a un profesor comprometido. *Academia Superior de Ciencias Pedagógicas de Valparaíso, Chile* 1983; No. 3-4, Año III: 15-16.
40. Irby DM. Clinical teacher effectiveness in medicine. *J Med Educ* 1978;53:808-815.
41. Stritter F, Baker RM. Resident preferences for the clinical teaching of ambulatory care. *J Med Educ* 1982;57:33-41.
42. Irby D, Rakestraw P. Evaluating clinical teaching in medicine. *J Med Educ* 1981; 56:181-186.
43. Ponce-Rosas ER, Gómez-Clavelina FJF, Irigoyen-Coria AE, Terán-Trillo M, Landgrave-Ibáñez S, Hernández-Gómez A, et al. Análisis de la confiabilidad de FACES III (versión en español). *Aten Primaria* 1999;23:479-484.
44. Muñoz J. *Psicometría*. Madrid, España: Universitas; 1996.
45. Cotsonas NJ. Student evaluation of clinical teaching. *J Med Educ* 1963;38:742-745.
46. Stritter F, Hain J, Grimes D. Clinical teacher reexamined. *J Med Educ* 1975;50:876-882.

