

Evaluación en educación médica

Martha Delgado Ramírez¹

Carlos Gómez-Restrepo²

Resumen

Introducción: La evaluación en educación médica es parte fundamental para la adecuada retroalimentación y valoración de los estudiantes que formamos. *Desarrollo:* Se revisan los propósitos generales de la evaluación en su vertiente sumativa y formativa. Se describen los procesos de evaluación sugeridos, teniendo en cuenta la pirámide de competencia de Miller; se presenta una herramienta para el diseño de evaluaciones y se postulan los conceptos de validez y reproducibilidad dentro de las evaluaciones. Se plantea la utilidad de la evaluación para el alumno y el docente. *Conclusiones:* El tema de evaluación debe salir de los parámetros de conocimiento que habitualmente hemos medido.

Palabras clave: Evaluación educacional, educación médica, psiquiatría, instituciones de enseñanza superior

Title: Assessment in Medical Education

Abstract

Introduction: The assessment of medical education is fundamental for proper feedback and evaluation of students. *Development:* The overall purpose of the evaluation is numerically and formatively revised. The suggested evaluation processes is described taking into account the Miller competence pyramid as a tool for designing evaluations and concepts of validity and reproducibility in assessments. The utility of student and the teacher is also raised. *Conclusions:* The issue of assessment should leave the knowledge parameters that are usually used.

Key words: Educational measurement; education, medical; psychiatry; higher education institutions

¹ Médica anestesióloga, epidemióloga clínica, profesora asociada y directora de carrera, Facultad de Medicina, Pontificia Universidad Javeriana. Bogotá, Colombia.

² Médico psiquiatra, psicoanalista, psiquiatra de enlace, epidemiólogo clínico y profesor titular, Departamento de Psiquiatría y Salud Mental, y director, Departamento de Epidemiología Clínica y Bioestadística, Facultad de Medicina, Pontificia Universidad Javeriana. Bogotá, Colombia.

Introducción

La formación de médicos y de especialistas exige procesos formativos de alta calidad que garanticen una sólida preparación disciplinar y el logro de las competencias profesionales, tan en boga hoy día. Esto supone hacer evidente en los procesos de evaluación la adquisición de las competencias establecidas por el currículo y articuladas de manera que respondan integralmente al perfil del egresado del programa (1). Dentro de los dominios de competencias nucleares, el ACGME (Consejo para la Acreditación de Educación para los Médicos Graduados) ha descrito en psiquiatría los siguientes (2):

- a. Cuidado del paciente
- b. Conocimiento médico
- c. Destrezas interpersonales y de comunicación
- d. Aprendizaje basado en la práctica
- e. Profesionalismo
- f. Práctica basada en sistemas

La evaluación efectiva requiere, por tanto, estar alineada con resultados de aprendizaje claramente definidos a priori y que puedan ser medidos; también, que el conjunto de métodos de evaluación seleccionados constituya una muestra representativa del desempeño del estudiante (3).

Por otra parte, es necesaria la utilización de pruebas estandarizadas aplicables a todos los alumnos,

que garanticen la equidad en el proceso y que los criterios para calificar y juzgar el nivel del desempeño sean explícitos, claramente definidos desde el comienzo y conocidos por los estudiantes de pregrado o de posgrado (3,4).

Para que la evaluación se constituya, efectivamente, en el motor del aprendizaje, tal como lo declaró George E. Miller hace varias décadas, es necesario que se establezcan mecanismos de retroalimentación oportuna y positiva para los estudiantes. La evaluación adecuadamente planeada e implementada da dirección al proceso de enseñanza-aprendizaje y es el elemento que genera mayor motivación para el mejoramiento continuo del alumno (4,5). En este aspecto, es fundamental que todos los profesores, los supervisores de práctica y los coordinadores de rotación, entre otros involucrados en el proceso de enseñanza y evaluación, sean conscientes de su gran influencia.

La evaluación puede servir para varios propósitos desde la vertiente del alumno; a saber:

1. Mejorar: Logrando identificar a estudiantes con conocimientos insuficientes para su nivel, de modo que busquen remediar las falencias; también, para motivar los alumnos de modo que busquen mejorar permanentemente y adquieran un hábito para el aprendizaje continuo a largo plazo.

2. Probar: Si los objetivos de aprendizaje se alcanzaron; también, para certificar el cumplimiento de niveles de logro de las competencias.

Y desde la vertiente del docente:

1. Mejorar: Se puede referir a los docentes, por cuanto ellos pueden evaluar si sus alumnos adquirieron los conocimientos o las competencias básicas que se les han enseñado; o, por otra parte, servir de insumo para dilucidar si la evaluación es justa y midió lo que se deseaba medir. Ejemplo de esto último podría ser una prueba demasiado difícil o mal realizada, y en la cual la mayoría de los estudiantes no obtiene una nota mínima de aprobación.
2. Retroalimentar: Se refiere a la posibilidad que se abre con una evaluación, de obtener retroalimentación y cuestionamientos por parte de los alumnos acerca de la amplitud, la precisión, la pertinencia y la justicia de la prueba realizada. Esta posibilidad puede o no abrirla el docente mismo; no obstante, abrirla es una oportunidad de gran importancia para mejorar en la carrera docente.

La evaluación puede ser de dos grandes tipos:

1. Evaluación sumativa: Indica si los estudiantes han logrado los

objetivos, y se usa para tomar decisiones de aprobar/reprobar y certificar al alumno. Es una evaluación de resultado, usualmente llevada a cabo al final de un curso o después de ciertos momentos de este (1,5).

2. Evaluación formativa: Orienta a los estudiantes sobre el proceso de aprendizaje que están teniendo. Se usa para mejorar el desempeño y se lleva a cabo a lo largo de todo el curso o de la rotación; ofrece una retroalimentación que le permite al alumno mejorar o ajustar la forma como ejecuta las diversas actividades donde está involucrado (1,5).

Existen numerosos métodos o estrategias de evaluación que han sido diseñados para evaluar diferentes dominios de la competencia. Ningún método por sí mismo tiene la posibilidad de evaluar todos los dominios, por lo cual es necesario diseñar evaluaciones que involucren diversas estrategias.

El conocimiento y la adquisición de competencias tienen diferentes niveles, por lo cual para el diseño de evaluaciones resulta útil tener en cuenta la taxonomía de Bloom y la Pirámide de la competencia de Miller.

La taxonomía de Bloom hace referencia específicamente a niveles de conocimiento, y puede ser utilizada en cualquier área. Esta clasificación indica qué tanto conoce sobre un tema la persona identificando cómo puede utilizar la información sobre

ese tema en particular; los niveles van de menor a mayor complejidad, según las siguientes categorías: conocer, comprender, aplicar, analizar, sintetizar y evaluar (6).

Por otra parte, la pirámide de Miller identifica las facetas o niveles de la competencia clínica, y permite, adicionalmente, el análisis de la validez de los procesos de evaluación articulado con estos niveles (7) (figura 1).

Figura 1. Pirámide de la competencia de Miller



La base de la pirámide representa el componente de conocimiento de la competencia; *sabe* (conoce información básica), seguido del *sabe cómo* (aplica ese conocimiento básico). Sin embargo, la evaluación de la competencia clínica supone considerar niveles más complejos, para observar el proceso de toma de decisiones y el actuar profesional del estudiante (7-9).

La selección de las estrategias o métodos de evaluación puede ser alineado con los objetivos según la pirámide de Miller; a continuación, un inventario de estrategias según los diferentes niveles:

Sabe y sabe cómo: Ambos aspectos pueden ser evaluados a través de exámenes escritos, con preguntas de selección múltiple, de correspondencia, de correspondencia extendida, de falso y verdadero, ensayos, respuestas cortas y simulación escrita; también, a través de exámenes orales donde se evalúen los conocimientos, el análisis y la aplicación, teniendo como base el contexto clínico (7-9).

Muestra cómo y hace: Pueden ser evaluados a través de pruebas de desempeño con paciente simulado, en laboratorio de destrezas, ECOES (Examen Clínico Objetivo y Estructurado —en inglés, OSCE— estaciones en centro de simulación clínica), gran caso clínico, miniexamen clínico, portafolio, observación directa, rúbricas y evaluación de 360° (7-9).

Tradicionalmente, la evaluación de los estudiantes de pregrado se ha centrado en examinar la base de la pirámide. Los métodos orientados a evaluar específicamente el conocimiento podrían ser apropiados en fases tempranas del proceso de formación; sin embargo, en la medida en que se diseñen currículos integrados que buscan desarrollar competencias, es necesaria una planeación cuidadosa de la evaluación como

elemento central en la formación, que considere diferentes estrategias capaces de poner de manifiesto diversos niveles en la adquisición de las competencias.

La evaluación debe tener una serie de atributos para que se la considere adecuada; debe ser válida, reproducible (confiable) y factible, e impactar de manera positiva el aprendizaje.

La validez se entiende como el grado en el que la evaluación o el examen miden adecuadamente el desenlace que pretende medir; es así como a partir de los resultados de la evaluación se busca hacer inferencias sobre el nivel de logro del estudiante. La validez se refiere, adicionalmente, a los datos y la información recolectados a través del examen, y que permiten una adecuada interpretación, para generar la calificación o el resultado que corresponda con el proceso de aprendizaje del estudiante, y son diseñados con un propósito específico, durante un momento particular en el tiempo (10).

A continuación, algunos ejemplos que podrían llevar a construir evaluaciones no válidas:

- El instrumento de evaluación tiene instrucciones poco claras, o carece de instrucciones.
- Las preguntas están mal redactadas; son ambiguas o ininteligibles.
- El método utilizado no es apropiado para medir la competencia de interés (por ejemplo, un exa-

men de múltiple escogencia para medir actitudes o habilidades).

- La persona encargada de diseñar la evaluación desconoce los objetivos que van a ser evaluados (11).

La validez es un elemento fundamental de la evaluación; sin embargo, no es posible hacer un análisis cuantitativo de ella, y es necesario llegar a conclusiones sobre su existencia mediante inferencias obtenidas a partir de los resultados o las evidencias de evaluación con diversos instrumentos (10,11).

La confiabilidad de una evaluación, o reproducibilidad, indica el grado de consistencia de ese proceso de medición, como una correlación teórica de puntajes de evaluaciones repetidas, estrechamente relacionado con el concepto de muestreo.

Se esperaría que si una evaluación es reproducible tenderá a producir resultados similares en aplicaciones repetidas en el tiempo. Algunos aspectos relacionados con una mayor confiabilidad de un examen incluyen: Mayor duración de la prueba, mayor extensión del instrumento (mayor número de preguntas), mayor número de examinadores llevando a cabo la evaluación, mayor número de casos a los que se expone un estudiante para ser examinado (12).

El diseño de cualquier evaluación es un reto que debe abordarse de manera sistemática y reproducible, mediante un plan replicable; por ello, se ha introducido la herramienta de *blueprint* o matriz de evaluación (13,14), como un elemento previo a

la construcción de la prueba y a la selección de las preguntas, y el cual asegura que el contenido del examen está alineado cuidadosamente con los objetivos de aprendizaje o las competencias que van a ser evaluadas. El uso de esta herramienta permite garantizar, en cierta medida, la validez de la evaluación que se está diseñando.

La matriz de evaluación, o *blueprint*, describe en detalle los dominios/tópicos/competencias, tal como están definidas en el currículo, que serán medidos o examinados en una evaluación en particular. Además, provee una guía para tener una muestra representativa de los contenidos por evaluar (13,14).

Los pasos para diseñar una matriz de evaluación son:

- Definir qué se va a evaluar —materia/tópico—.
- Ponderar: Asignar pesos o porcentajes a las unidades de contenido cubiertas —de acuerdo con la intensidad o el tiempo dedicado—.
- Definir las facetas/niveles que van a ser evaluados en cada una de las unidades.
- Definir el nivel de maestría por ser evaluado.
- Seleccionar el método de evaluación.

Como ejemplo de lo planteado, en una evaluación teórica de Introducción a la Psiquiatría (tabla 1) que ha sido abordada por dominios, es necesario asignar primero un peso a cada

uno de ellos; dicho peso va en relación con la intensidad horaria dedicada a cada dominio. Posteriormente se establece la proporción de preguntas dedicadas a evaluar objetivos de conocimiento (sabe y sabe cómo), de razonamiento clínico y de actitud, en concordancia con los objetivos y las competencias que intentó cubrir el curso. Una vez construida la matriz, se define que se hará un examen escrito con preguntas de escogencia múltiple y única respuesta, y se procederá a elaborar el número de preguntas o problemas indicados por dicha matriz.

La gran mayoría de los médicos que trabajan en universidades y hospitales universitarios supervisando y acompañando a estudiantes en proceso de formación (algunos, en fases preclínicas, y otros, en las rotaciones clínicas propiamente dichas) necesariamente tendrán que participar en procesos de evaluación de estos alumnos. Por ello, es necesario recordar siempre los objetivos de la evaluación, que van más allá de verificar en un momento particular el conocimiento o la habilidad del estudiante; el resultado de la evaluación bien entendida, y en la que se hace retroalimentación oportuna, constructiva y positiva al estudiante, orienta y motiva para el aprendizaje futuro (15,16).

Es importante tener en cuenta que se debe evaluar no solo el conocimiento, sino las actitudes, las habilidades comunicativas, el profesionalismo, el razonamiento clínico y las habilidades para trabajo en equipo, entre otras competencias (16).

Tabla 1. Matriz de evaluación. Examen de Introducción a la Psiquiatría-temáticas

Tema	Ponderación (N° de preguntas)	Conocimientos	Razonamiento clínico	Actitud
Neurociencias	30	15	9	6
Relación médico-paciente	15	3	4	8
Diagnóstico	30	15	10	5
Intervenciones- psicofarmacología	15	9	3	2
Intervenciones- psicoterapia	10	6	2	2

En el momento de decidir sobre cómo hacer la evaluación, es importante incluir múltiples métodos y diversos escenarios y contextos que logren evaluar realmente el desempeño del alumno (16,17). Finalmente, se debe ser cauteloso, para no convertir la evaluación en una herramienta punitiva, sino en un elemento fundamental del proceso de retroalimentación y consejería de los estudiantes, que sirva al cometido de proveer más aprendizajes.

Conclusión

Este artículo revisó de manera general los propósitos de la evaluación y los tipos de evaluación utilizados en procesos formativos, mencionó los métodos de evaluación sugeridos según el nivel en la pirámide de competencia de Miller y, finalmente, abordó los conceptos de validez y reproducibilidad en la evaluación y presentó una herramienta para el diseño de evaluaciones. Es necesario profun-

dicar en cada uno de dichos tópicos, para diseñar sistemas de evaluación robustos para los estudiantes y participar en ellos.

Agradecimiento

Al grant 0014559 ENRICH (Enhancing Research and Informatics Capacity for Health Information in Colombia). Pontificia Universidad Javeriana y University of Pittsburgh.

Referencias

1. Epstein RM, Hundert EM. Defining and assessing professional competence. *JAMA* 2002; 287:226-35.
2. ACGME Program Requirements for Graduate Medical Education in Psychiatry [internet]. 2007 [citado: 10 de mayo de 2012]. Disponible en: http://www.acgme.org/acWebsite/downloads/RRC_progReq/400_psychiatry_07012007_u04122008.pdf.
3. Epstein RM. Assessment in medical education. *N Engl J Med*. 2007;356:387-96.
4. Kromann CB, Jensen ML, Ringsted C. The effect of testing on skills learning. *Medical Education*. 2009;43:21-7.

5. Flores F, Contreras N, Martínez A. Evaluación del aprendizaje en la educación médica. *Revista de la Facultad de Medicina de la UNAM*. 2012;55:42-8.
6. Bloom B, Krathwohl D. Major categories in the taxonomy of educational objectives [internet]. 2012 [citado: 10 de mayo de 2012]. Disponible en: <http://krummeffamily.org/guides/bloom.html>.
7. Miller G. The assesment of clinical skills/competence/performance. *Acad Med*. 1990;65(suppl):S63-7.
8. Schuwirth L, van der Vleuten C. Merging views on assessment. *Med Educ*. 2004;38:1208-10.
9. Wass V, Van der Vleuten C, Shatzer J, et al. Assessment of clinical competence. *Lancet*. 2004;357:945-9.
10. Downing SM. Validity: on meaningful interpretation of assessment data. *Med Educ*. 2003;37:830-7.
11. Downing SM, Haladyna TM. Validity threats: overcoming interference with proposed interpretations of assessment data. *Med Educ*. 2004;38:327-33.
12. Downing SM. Reliability: on the reproducibility of assessment data. *Med Educ*. 2004;38:1006-12.
13. Banu S, Chacko TV. Blueprinting: a framework for health education programme evaluation. *Med Educ*. 2011;45:533.
14. Hamdy H. Blueprinting for the assessment of health care professionals. *Clin Teach*. 2006;3:175-9.
15. Wiliam D, Lee C, Harrison C, et al. Teachers developing assessment for learning: impact on student achievement. *Assess Educ*. 2004;11:49-65.
16. Murphy DJ, Bruce D, Eva KW. Workplace-based assessment for general practitioners: using stakeholder perception to aid blueprinting of an assessment battery. *Med Educ*. 2008;42:96-103.
17. Van der Vleuten C. Validity of final examinations in undergraduate medical training. *BMJ*. 2000;321:1217-9.

Conflictos de interés: Los autores manifiestan que no tienen conflictos de interés en este artículo.

Recibido para evaluación: 23 de junio de 2012

Aceptado para publicación: 27 de julio de 2012

Correspondencia
Martha B. Delgado Ramírez
Facultad de Medicina
Pontificia Universidad Javeriana
Carrera 7ª No. 40-62 Piso 8
Bogotá, Colombia
mdelgadoramirez@gmail.com