

La evaluación en educación médica

Expositor: José María Fornells Vallés⁽¹⁾

Filiación: ⁽¹⁾Institut d'Estudis de la Salut, Generalitat de Catalunya, Cataluña, España

Resumen

Este texto completo es la transcripción editada y revisada de una conferencia dictada en reunión clínica del Departamento de Medicina, Hospital Clínico Universidad de Chile. La publicación de estas actas ha sido posible gracias a una colaboración editorial entre Medwave y el Departamento de Medicina del Hospital Clínico Universidad de Chile. El director del Departamento de Medicina es el Dr. Alejandro Cotera.

Introducción

Para planificar en forma adecuada la evaluación de la educación médica es fundamental establecer desde el principio y en forma clara los objetivos del aprendizaje, es decir, las competencias que el estudiante debe adquirir al final del proceso, ya que según esto se escogerá los instrumentos que se va a utilizar.

Competencia profesional

El autor clásico Kane señala que *un profesional es competente en la medida en que utiliza los conocimientos, las habilidades, las actitudes y el buen juicio asociados a su profesión con la finalidad de desarrollarla en forma eficaz en aquellas situaciones que corresponden al campo de su práctica. La competencia, pues, pertenece al área del saber hacer* (1). Esta definición implica dos hechos: que la competencia profesional es un problema de integración de distintos aspectos en la formación de alguien; y que dicha competencia se demuestra a través de la actuación en un terreno práctico.

Otra definición interesante es la de Ronald M Epstein, que describe a la competencia profesional como *el uso cotidiano y juicioso del conocimiento, habilidades técnicas y de comunicación, juicio clínico, emociones, valores y reflexión en la práctica diaria para el beneficio de las personas y de la comunidad a la que se sirve* (2). En esta definición se introduce una dimensión social y también de integración.

En la Fig.1 se muestra la pirámide de Miller, que establece cuatro niveles de aprendizaje con sus respectivos niveles de evaluación de competencias. En el nivel más básico, el *conocer*, se aplica la evaluación de los conocimientos, es decir, la capacidad para recordar cosas; en el segundo nivel, el *cómo conocer*, se evalúa la capacidad para entender cómo ocurren los fenómenos, es decir, para buscar información, analizarla e interpretar resultados; en el tercer nivel, el *cómo mostrar*, se evalúa la ejecución de lo que se ha aprendido, es decir, implica una acción, pero en una práctica *in vitro*, no en condiciones de práctica real; finalmente en el vértice de la pirámide está el *hacer*, que se evalúa es la práctica real, *in vivo*.

Figura 1. Pirámide de Miller: niveles de aprendizaje y evaluación.

Cada elemento de la pirámide de Miller se asocia a una serie de instrumentos de evaluación. En la base están las pruebas de conocimientos, en especial las preguntas de respuestas múltiples; en el segundo nivel se aplican pruebas en un contexto clínico, por tanto para responderlas se requiere razonar, no solamente recordar información; en el tercer nivel se evalúa mediante simulaciones; y en el vértice de la pirámide se puede aplicar múltiples instrumentos, como portafolios, videos, etc. Cabe señalar que los sistemas de evaluación modernos se introducen en España en la década de los 60, con los exámenes tipo *test*, mientras que las pruebas basadas en simulaciones para evaluaciones objetivas estructuradas, surgidas a mediados de los años 70 en el Reino Unido, sólo se empezaron a desarrollar en los años 90. La época actual se caracteriza por el surgimiento de un nuevo paradigma, proveniente de la literatura y experiencia anglosajonas, que se enfoca en la evaluación de la práctica real.

Tipos de evaluación

Existen dos niveles o tipos de evaluación: formativa y sumativa. La evaluación formativa pretende determinar el grado de avance de un estudiante en sus objetivos de aprendizaje, con el objetivo de subsanar los déficits que se detecten en ese momento; no tiene como objetivo calificar para aprobar o reprobar, por lo tanto, no es coercitiva. La evaluación sumativa, sancionadora o acreditativa tiene como objetivo determinar si el estudiante adquirió cierto cuerpo de conocimientos o competencias y por lo tanto va asociada a una nota o calificación que le entrega o no el derecho a pasar al siguiente ciclo. Lo que hace que una evaluación sea sumativa o formativa no es el tipo de instrumento que se utiliza, sino el propósito. Por ejemplo, el examen con base en preguntas de respuesta múltiple se asocia tradicionalmente a una evaluación sumativa, que se mide mediante una nota, pero también se puede utilizar en un objetivo formativo si se permite a los alumnos corregir el examen si falló.

Los instrumentos de evaluación deben cumplir con ciertas características:

- Validez y fiabilidad: significa que el instrumento debe medir exactamente lo que se quiere medir y que los resultados sean reproducibles.
- Aceptación y transparencia, es decir, que las reglas del juego estén bien definidas, que los estudiantes que están en proceso formativo sepan realmente que van a ser evaluados y qué trascendencia tiene la evaluación.
- Factibilidad, es decir, cuánto cuesta, qué recursos se va a necesitar para poner en marcha una evaluación. Se puede tener un sistema fantástico, pero habrá que ver con qué facilidad se puede implementar este elemento fundamental a la hora de hacer un análisis y determinar qué instrumentos se van a utilizar para un determinado tipo de estudiante, en un determinado perfil.
- Impacto educativo: no se puede dejar de tener en cuenta este elemento. Todo proceso, todo instrumento de evaluación tiene implicaciones o repercusiones sobre la persona que se está formando y sobre cómo se va a formar.

Un ejemplo de instrumento de evaluación es el examen MIR, que se utiliza en España como requisito de ingreso a una especialidad médica. Este examen consiste en un *test* de 250 preguntas de respuesta múltiple, en el que se evalúa fundamentalmente los dos primeros niveles de competencia, es decir, el conocimiento y el conocer cómo. En este contexto es importante recordar lo que Miller señalaba en el sentido de que *la evaluación de alguna manera condiciona el sistema de aprendizaje*, por lo tanto ningún sistema de evaluación es inocuo.

Cuando se plantea una evaluación se debe tener claros los siguientes aspectos:

1. ¿A quiénes se va a evaluar?
2. ¿Para qué se les va a evaluar?
3. ¿Cómo se les va a evaluar? ¿Qué métodos vamos a utilizar, que instrumentos?
4. ¿Con qué nivel? Nivel de excelencia, nivel de mínimos, todos estos son elementos importantes.
5. ¿Quién evalúa? Va a ser una evaluación externa o interna.

Evaluación en la formación de pregrado

El 13 de febrero de 2008 el Estado español dictó un orden que fija los objetivos y competencias que los estudiantes deben alcanzar durante su formación de seis años en la Facultad de Medicina, agrupándolas en siete grandes bloques:

- Valores profesionales, actitudes y comportamientos éticos
- Fundamentos científicos de la medicina
- Habilidades clínicas
- Habilidades de comunicación
- Salud pública y sistemas de salud
- Manejo de la información

- Análisis crítico e investigación

Esto permite determinar desde el principio el sistema de evaluación que se va a poner en marcha para ver si se consiguen estos objetivos y se adquieren estas competencias.

La evaluación clínica objetiva estructurada (ECOЕ)

La ECOE es una utilización combinada, sistemática y estandarizada de distintos instrumentos de evaluación, principalmente para evaluar las competencias previamente definidas a través de un circuito de estaciones. En la Fig. 2 se representa la ECOE de un alumno va pasando por una serie de estaciones en las que encuentra pacientes simulados, preguntas de respuesta múltiple, en función de objetivos que se han definido previamente (Fig. 2).

Figura 2. Esquema general de la evaluación clínica objetiva estructurada (ECOЕ)

Las competencias tradicionales que se evalúan en la ECOE son: anamnesis, exploración física, procedimientos, diagnóstico, manejo del paciente y habilidades comunicativas. Además, en los últimos años se ha incorporado a otras competencias, como el profesionalismo, valoración crítica, investigación, etc.

La ECOE tiene varias ventajas:

1. Tiene validez de superficie, ya que reproduce un amplio abanico de situaciones de la práctica.
2. Permite evaluar un amplio espectro de habilidades en un espacio relativamente pequeño, gracias al circuito que se utiliza.
3. Permite establecer estándares de competencia mediante una lista de chequeo.
4. Reduce la variabilidad que depende del paciente y del examinador, porque los pacientes son los mismos y su actuación con cada alumno y los evaluadores u observadores están entrenados para calificar ante las mismas respuestas y comportamientos.
5. Permite evaluar el producto de un período de formación.
6. Tiene un formato flexible.
7. Permite pesquisar déficits en el programa de formación.

La ECOE tiene las siguientes desventajas:

1. Es un proceso difícil de organizar desde el punto de vista logístico y en general consume más recursos.
2. La consistencia es un aspecto crítico en la ECOE, especialmente si se utiliza para una evaluación sumativa para decidir si alguien es apto o no, ya que la actuación de un alumno ante un paciente simulado puede ser igual frente a pacientes reales, lo que obliga a trabajar con cierta cantidad de casos clínicos.
3. La gestión de los recursos humanos y de los aspectos organizativos es compleja.
4. El desarrollo de las distintas fases puede tomar mucho tiempo, entre otras causas porque la evaluación se realiza en un espacio que normalmente no trabaja allí todo el tiempo.

Se ha determinado que para una evaluación formativa la ECOE debe tener entre ocho y quince estaciones y hasta tres horas. Para una evaluación sumativa se requiere como mínimo veinte estaciones, con una duración de cuatro a seis horas. En resumen, la ECOE para una evaluación sumativa que para una evaluación formativa.

Evaluación de competencias clínicas en las facultades de Medicina

La ECOE se aplica en Cataluña para evaluar las competencias clínicas en licenciados, es decir, en estudiantes de sexto año

Medicina, desde hace varios años, ya que los métodos tradicionales de evaluación no permiten evaluar en forma adecuada este aspecto. La experiencia obtenida en el período 1994-2006 está publicada (3).

Entre 1994 y 1997 se trabajó con diez estaciones y entre 1998 y 2006, con veinte estaciones y catorce casos clínicos, para obtener un mínimo margen de seguridad. Además, a partir de 2001 se incorporó una fase escrita con 150 preguntas de selección múltiple, 20 de ellas asociadas a imágenes, lo que permite ampliar mucho más el ámbito de las competencias a evaluar. Entre 1994 y 2006 participaron 4.331 alumnos de sexto año y en 2006 con 565 alumnos de las cuatro facultades. En ese lapso el coeficiente de fiabilidad alfa de Cronbach se mantuvo sobre 0,80, lo que es considerado como recomendable, tanto en la fase de multiestación como en la fase escrita de la prueba (Tabla I).

Tabla I. Coeficiente de fiabilidad alfa de Cronbach para la ECOE. FME= fase de multiestación, FE= fase escrita

Un estudiante de sexto año en España no tiene un internado como el de Chile, donde los estudiantes tienen relación con pacientes y el contacto con los pacientes es limitado y la responsabilidad frente a un paciente real también es limitada, por lo tanto es difícil. Por eso se ha elevado una propuesta al Ministerio español para que considere a la ECOE como un elemento más dentro de la formación de las especialidades, que actualmente solamente evalúa conocimientos.

Evaluación de la formación especializada en España

Las comisiones nacionales designadas para ello establecen la definición de competencias y elaboran el Programa Nacional de Formación Especializada, que se convierte en Ley cuando el Ministerio lo aprueba. En el último tiempo se ha ido definiendo en forma cada vez más precisa el camino que debe alcanzar el futuro especialista o médico en su proceso de especialización, así como el proceso de evaluación formativa y sumativa. En febrero de 2008 se emitió un Decreto que modifica sustancialmente los sistemas de evaluación y formación especializada, dando más énfasis a la evaluación formativa, sin dejar de considerar la evaluación sumativa; sin embargo el tema es complejo y se ha convertido en un objeto de exenta de controversia.

En el curso de la formación especializada hay tres elementos fundamentales a evaluar:

- Preguntar, de palabra o por escrito.
- Observar lo que hace, en el quirófano, en la consulta.
- Mirar lo que escribe. Un médico genera documentos que son importantes, que tienen repercusión y que hay que ver cómo lo está haciendo.

El tutor en la evaluación formativa de especialidades médicas

En España ha surgido con especial fuerza en los últimos años la figura del tutor, que es el profesional encargado de seguir el proceso de formación del residente. El tutor debe: determinar los objetivos específicos del aprendizaje; evaluar en qué medida se van alcanzando; utilizar los recursos de *preguntar, observar y mirar*, mediante distintos métodos; y finalmente debe establecer la triangulación, que es la comunicación entre el residente y el adjunto. La triangulación, que es la comunicación entre el residente y el adjunto, es fundamental para que el seguimiento de ese residente sea adecuado (Figura 3).

Figura 3. Triangulación tutor-residente-adjunto.

Un elemento importante de la evaluación sumativa es el *feedback o retorno de información*. Jack Ende, que ha escrito artículos brillantes sobre el tema, señala que *sin feedback los errores no se corrigen, los aspectos positivos no se refuerzan y la competencia se desarrolla de forma empírica o no se llega a conseguir* (4).

En España la evaluación sumativa depende del Ministerio, que la aplica través de Comisiones Nacionales en centros sanitarios de formación. Los hospitales, los tutores y los adjuntos siguen siendo ejecutores, pero no tienen competencias en el diseño de

fundamental tener claro las competencias que se van a evaluar; a quién se va a evaluar; con qué nivel; y cómo será la evaluación final, que no es lo mismo que un examen final. En la evaluación final se deberá aplicar una serie de métodos con un peso ponderado predeterminado; por ejemplo: preguntas de respuesta múltiple 20%, evaluación continuada 20%, observación estructurada 30%, auditorías de historia 15% y portafolio, 15%; porcentajes todos ellos con un significado meramente ilustrativo.

Observación estructurada de la práctica clínica OEPC (mini-CEX)

Actualmente se está impulsando en Cataluña la aplicación de la Observación estructurada de la práctica clínica (OEPC), también denominada mini-CEX, en su versión original inglesa. Se trata de un instrumento para evaluar habilidades clínicas que tiene las siguientes características: se realiza con casos clínicos reales, es decir, en el vértice de la pirámide; hay un observador cada vez; se presentan distintos casos en distintas sesiones; se utiliza una ficha estructurada para recoger la información; entrega retroalimentación al evaluado y se lleva a cabo en un tiempo fijo de 5-10 minutos. El mini-Cex fue creado en los Estados Unidos por el *American Board of Internal Medicine* a partir del antiguo *Case-Based Examination*, el examen tradicional, que consumía mucho tiempo, una a dos horas y se hacía con un solo paciente y un solo observador, lo que afectaba la validez y fiabilidad y ocasionaba mucho estrés al examinado.

El mini-CEX evalúa una serie de competencias, al igual que la ECOE y entrega los mismos resultados con menor costo y mayor rapidez. Las competencias valoradas están: habilidades de entrevista clínica, habilidades de exploración física, cualidades humanas del médico, *counseling skills* y organización y eficiencia; esto sin perder de vista un objetivo muy importante, que es elaborar un informe satisfactorio que sirvan de base para la evaluación global final del evaluador. Un ejemplo de perfil satisfactorio es: *demuestra habilidades de entrevista clínica, capacidad de síntesis y de resolución y tiene en cuenta aspectos de eficiencia, valorando riesgos y beneficios en el diagnóstico y tratamiento.*

El mini-CEX recoge los siguientes datos descriptivos: evaluador, residente, año de residencia, fecha, entorno clínico (urgencias, consulta, etc.), edad del paciente, motivo principal de la consulta, complejidad del caso, categoría del observador (adjunto, tutor), perfil de paciente. Los observadores prácticos se recomienda al observador presentarse al paciente si es necesario; cuidar su posición de modo que tener buen acceso a la información sobre el accionar del evaluado; tener conocimiento del nivel esperado y exigible al observado; estar familiarizado con la ficha de evaluación (descriptores de las competencias) y mantener la concentración y atención. Los estudios demuestran que el mini-CEX proporciona resultados fiables sobre la base de seis encuentros clínicos al año valorados por distintos observadores.

La observación estructurada tiene varias ventajas:

- Es representativa de la realidad, ya que se evalúa con distintos casos y en entornos distintos.
- La validez es buena.
- Hay varios observadores, lo que aumenta la fiabilidad.
- Hay mejor *feedback*.
- Es factible en cuanto a ejecución.
- Favorece la auto-responsabilidad del residente, que dentro de ciertos márgenes de tiempo participa en la evaluación.
- Puede ser valorada por los mismos residentes observados.
- Es eficiente y fácilmente aplicable.
- El nivel de satisfacción de evaluador y evaluado es alto.
- Es flexible.
- Es medible, de modo que tiene un valor estadístico.

Entre las limitaciones de la observación estructurada se cuentan las siguientes:

- Se necesita la colaboración de otros docentes o tutores.
- Se necesita hacer la recopilación y seguimiento de las distintas sesiones, es decir, se debe registrar la información que se generan.

- Se requiere consenso entre los observadores en cuanto a los descriptores de las distintas competencias, de modo que se les debe entregar cierto entrenamiento para que no haya diversidad de criterios a la hora de valorar actuaciones del residente.
- Los observadores deben ser personas activas.

En Cataluña se estableció que los encargados de aplicar el mini-CEX son los tutores, de modo que éstos deben asistir a una formación de cierta duración para explicarles cómo se utiliza este instrumento. En una encuesta que se envió por correo electrónico a 120 tutores de un taller, se les preguntó si estaban aplicando el mini-CEX: 50% de ellos contestaron la encuesta y de éstos, alrededor de 50% demuestra que la formación es un elemento necesario, pero no suficiente para lograr que el uso de este instrumento se masifique. Se están diseñando estrategias para superar los posibles obstáculos, como la falta de voluntad y de recursos. Se está planificando la formación para hacer un seguimiento de las personas que han recibido formación para aplicar el mini-CEX.

Evaluación de la práctica profesional

La práctica profesional es el tercer eslabón del proceso educativo, después del pregrado y de la formación especializada. La evaluación de la práctica profesional, como describió Norcini en 2003, con base en:

- Resultados de la atención médica: morbilidad, mortalidad, reingresos, etc.
- Procesos de atención: cumplimiento de protocolos, actividades preventivas.
- Volumen de actividad: juntas médicas, atenciones, etc.

Entre los instrumentos disponibles para valorar la práctica profesional, en pleno vértice de la pirámide de Miller, están: los diarios de aprendizaje, *logbooks* o portafolios; la observación de la práctica en forma directa o a través de grabación de video de situaciones relevantes. En el caso de la observación, el video tiene la ventaja de que permite al evaluado verse a sí mismo en acción, lo que ayuda a corregir errores. Los instrumentos descritos tienen distinto grado de validez, fiabilidad y factibilidad (Tabla II).

Tabla II. Propiedades de algunos instrumentos de evaluación de la práctica profesional.

Se denomina *portafolio* a la recogida, más o menos sistemática, de elementos que evidencian el aprendizaje o nivel de cumplimiento de los objetivos establecidos con anterioridad, y que va acompañado de un proceso de reflexión. Actualmente en España se está implementando el portafolio y ya se está implementando algunas experiencias concretas con este instrumento; de hecho la Sociedad Española de Medicina Interna está llevando a cabo un proyecto piloto de ejecución de un portafolio de objetivos y alcance moderados que permite al observador observar las dificultades que se encuentren para llevarlo a cabo. En Medicina Familiar y Comunitaria también existe un proyecto del Ministerio, que es mucho más ambicioso, pero aún no se ha puesto en marcha.

Cees van der Vleuten, Director del Departamento de Educación Médica de la Universidad de Maastricht, Holanda, una de las universidades de educación médica en Europa junto a la de Dundee, en Escocia, plantea que en la formación médica lo principal no es aplicar un instrumento de evaluación más o menos objetivo, sino elaborar un programa que defina los niveles de evaluación a través de todo el desarrollo de la formación que es un proceso continuo, de modo que los instrumentos que se apliquen en cada nivel se definan con base no sólo en su validez, sino también considerando su impacto educativo, aceptabilidad, recursos que consuma y finalidad de la evaluación (5). Es distinto el caso de un residente que tiene derecho a ejercer como especialista, a realizar una evaluación formativa (Fig. 4).

Figura 4. Desarrollo profesional continuo individual: niveles de evaluación.

La implementación de un portafolio contempla las siguientes etapas:

1. Definir el propósito.
2. Identificar las competencias a desarrollar o a evaluar.
3. Seleccionar el material o evidencias del *portafolio*.
4. Desarrollar un sistema de puntuación y planificación del proceso de evaluación.
5. Seleccionar y formar a los evaluadores.
6. Orientar a los profesionales-estudiantes.
7. Realizar pilotaje del instrumento.
8. Establecer evidencias de fiabilidad y validez.

En el desarrollo de un portafolio hay tres aspectos clave: el diseño del instrumento, la estrategia de aplicación y la evaluación.

En resumen, existen distintos tipos de portafolio, en distintos ámbitos y con distintos propósitos, pero en todos los casos se requieren evidencias de aprendizaje y competencias que facilite la reflexión sobre el propio proceso de aprendizaje. El diseño, aplicación y evaluación de portafolios son diferentes según se utilice en el ámbito de la formación pregraduada, especializada o en el desarrollo profesional. Los usuarios de la acción pueden ser estudiantes, docentes, residentes o profesionales en ejercicio. Igualmente condicionante es el propósito del portafolio en cada uno de los ámbitos, sumativo o formativo. El portafolio, además de permitir valorar los resultados del aprendizaje, también permite el desarrollo de estrategias, actividades, habilidades y procesos cognitivos esenciales para el aprendizaje a lo largo de toda la vida.

Cuando se plantea un portafolio de ejecución obligada es importante evitar que ocurran *efectos perversos*, como que se pierda espontaneidad y se induzca a la persona a hacer más bien lo políticamente correcto en función de lo que se exige, lo que limita el valor del dicho portafolio. Por todo ello, no se debe usar el portafolio para medir aspectos de la competencia que se pueden medir normalmente. No obstante, sí se puede incluir estos resultados en el portafolio. No hay ningún instrumento que pueda evaluar por sí solo, pero en su conjunto; para esto se requiere un sistema integrado de evaluación que combine diferentes métodos.

Consideraciones finales

- La evaluación no es sólo un problema de medición o de bondad inherente a un determinado problema de diseño de un Programa apropiado.
- Un Programa de Evaluación debe considerar los siguientes elementos clave: fiabilidad, validez, aceptabilidad y recursos.
- La fiabilidad es un problema más bien de muestreo que de estandarización u objetividad.
- La validez es un tema de integración de competencias.
- El principal problema de la fiabilidad es la especificidad de contenido, o de caso, lo que constituye un aspecto clave, dentro del cual se debe considerar qué se va a evaluar (contenidos, competencias) y en qué contextos evaluar (contextos o entornos clínicos) y quién va a evaluar.

Referencias

1. Kane MT. The assessment of professional competence. *Eval Health Prof.* 1992 Jun;15(2):163-82. [Volver arriba](#) | [CrossRef](#) | [PubMed](#) |
2. Epstein RM, Hundert EM. Defining and assessing professional competence. *JAMA.* 2002 Jan 9;287(2):226-35. [Volver arriba](#) | [CrossRef](#) | [PubMed](#) |
3. Kronfly Rubiano E, Ricarte Díez JI, Juncosa Font S, Martínez Carretero JM. Evaluación de la competencia clínica en medicina de cataluña, 1994-2006: evolución de los formatos de examen hasta la evaluación clínica objetiva es (Barc). 2007 Dec 1;129(20):777-84. [Volver arriba](#) | [CrossRef](#) | [PubMed](#) |
4. Ende J. Feedback in clinical medical education. *JAMA.* 1983 Aug 12;250(6):777-81. [Volver arriba](#) | [CrossRef](#) | [PubMed](#) |
5. Wass V, Van der Vleuten C, Shatzer J, Jones R. Assessment of clinical competence. *Lancet.* 2001 Mar 24;357(9231):1033-7. [Volver arriba](#) | [CrossRef](#) | [PubMed](#) |

