

Sistemas de evaluación

Margarita Barón Maldonado

Universidad de Alcalá de Henares. Presidenta Association for Medical Education in Europe (AMEE)

Se acepta que la profesión y la práctica médicas sean frecuentemente calificadas como de “alto riesgo” por la propia naturaleza de su objetivo profesional: el cuidado y la promoción de la salud del ser humano. Parece evidente que una alta calidad de los profesionales se manifestará en una práctica médica excelente, lo cual mitigaría el riesgo y la incertidumbre hasta los límites de lo aceptable. Esta calidad se busca a través de un proceso de formación extenso, riguroso, específico y continuo que no termina sino con el cese del ejercicio profesional.

Sin embargo, no es suficiente diseñar programas y procesos de formación excelentes, sino que es preciso demostrar que su aplicación produce el impacto positivo que se desea. Es imprescindible pues, la evaluación continua, rigurosa y específica del profesional de la Medicina en sus distintas etapas de desarrollo: desde el profesional en proyecto como estudiante de la titulación de Medicina en la Facultad, hasta el profesional en el ejercicio pleno de su práctica médica, pasando por el residente en fase de especialización.

Como *marco conceptual para la evaluación* es preciso considerar lo siguiente:

a) la evaluación forma parte intrínseca del *currículum*; b) el impacto retroactivo que los sistemas de evaluación tienen sobre el aprendizaje o, lo que es lo mismo, los objetivos de aprendizaje, podrán facilitarse aplicando un sistema de evaluación coherente y consistente con ellos, o en caso contrario constituirán un obstáculo; c) el valor formativo de la evaluación y el impacto positivo que la información transparente de los resultados de la evaluación tiene sobre el aprendizaje; d) la valoración de la operatividad de la evaluación y... otros aspectos (fig. 1 y fig.2).

Por lo tanto, los sistemas y la técnica de evaluación deberán seleccionarse en función de los objetivos de aprendizaje, orientarlos a su consecución y formarán parte inseparable del diseño curricular y de las estrategias para el desarrollo personal de los profesionales.

La selección de un tipo u otro de evaluación, ya sea sumativa o formativa, debe ser congruente con el objetivo de la evaluación y capaz de medir los parámetros que desean valorarse. Es imprescindible pues, previamente al diseño de la prueba, tener en cuenta las características que deben presidir su selección y diseño. Consecuentemente, la *selección y el diseño* de la prueba debe basarse en determinadas características¹: 1) Enfocada al ejercicio profesional específico 2) Centrada en el propósito. 3) Válida y reproducible. 4) Sus resultados deben poderse comparar con estándares determinados previamente y sus consecuencias deben haberse preestablecido con transparencia. 5) Ser aplicable

La selección de un tipo u otro de evaluación, ya sea sumativa o formativa, debe ser congruente con el objetivo de la evaluación y capaz de medir los parámetros que desean valorarse. Es imprescindible pues, previamente al diseño de la prueba, tener en cuenta las características que deben presidir su selección y diseño. Consecuentemente, la *selección y el diseño* de la prueba debe basarse en determinadas características¹: 1) Enfocada al ejercicio profesional específico 2) Centrada en el propósito. 3) Válida y reproducible. 4) Sus resultados deben poderse comparar con estándares determinados previamente y sus consecuencias deben haberse preestablecido con transparencia. 5) Ser aplicable

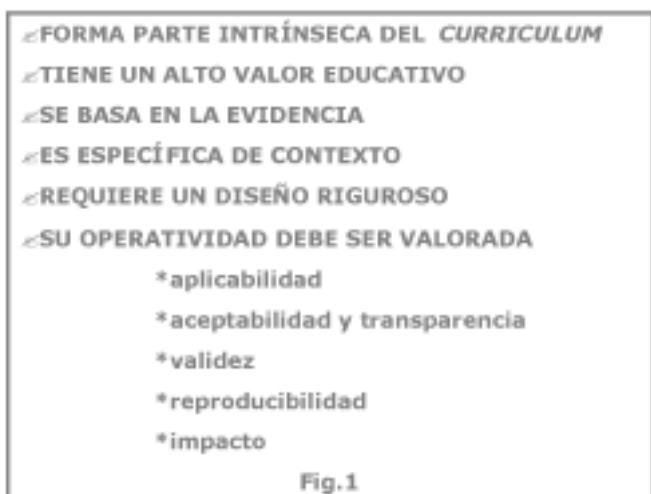




Fig. 3

(fig. 3). La complejidad de la ciencia médica y la de la formación de sus profesionales es ampliamente conocida por todos los educadores y por ello la característica primera es fundamental y constituye el punto de partida de la selección y diseño de la prueba; la evaluación pretende determinar el valor de un individuo o población como poseedor/a de un conjunto de conocimientos y capacidades que se corresponden con su nivel de desarrollo profesional, su orientación específica a una determinada práctica médica, el ámbito en el que tiene lugar su actividad etc.; coherentemente pues, la prueba debe enfocarse específicamente en la/s actividad/es concreta/s que ocurren en las circunstancias identificadas como definidoras de la misma.

Todavía más, es necesario *categorizar los componentes del ejercicio profesional* que pueden ser evaluados y que relacionan las esferas cognitiva, emocional y de la conducta. Para unos, la categorización desagregada, cuyos componentes forman el todo a evaluar, está constituida por los conocimientos, las habilidades y las actitudes. Miller, 1990, expresa en su muy conocida, afamada y citada pirámide su categorización en cuatro estratos que constituyen los componentes a evaluar. La base amplia y extensa compuesta por los conocimientos, esfera cognitiva, que evolucionan hacia la esfera conductual mediante la superposición más sofisticada del manejo de los conocimientos: saber lo que se puede hacer con ellos, demostrar que se hace y hacer. Todo ello referido específicamente al profesional de la ciencia médica y a su ejercicio. El concepto de Miller, ha resultado ser extraordinariamente productivo en cuanto que ha generado un buen número de pruebas específicas cuyo propósito se centra en evaluar los dos estratos superiores de su categorización, esto es, el demostrar que se sabe hacer y el hacerlo. Pocas veces un aporte conceptual ha tenido como consecuencia mayor aplicación práctica.

El nuevo paradigma del aprendizaje basado en *competencias de salida*, traducción libre de *outcomes-learning*, considera que el demostrar que se sabe hacer y el hacerlo son la expresión máxima de la competencia del profesional en su ejercicio en el nivel que corresponda. La planificación curricular en este tipo de aprendizaje sigue un camino inverso al tradicional; comienza con la determinación de las competencias de salida que ha de exhibir el graduado, el especialista o determinado profesional y termina en la evaluación del mismo, (fig. 4) mediante instrumentos diseñados específicamente para detectar que sabe hacer y hace determinadas prácticas utilizando y aplicando adecuadamente sus conocimientos. A este nivel pues, se está evaluando e intentando cuantificar la conducta. Diversas instituciones académicas han establecido un listado de competencias de salida acorde con el graduado, por ejemplo, que se proponen obtener. Gran aceptación ha tenido el modelo diseñado por el Institute for International Medical Education (IIME)² que propone siete dominios amplios conteniendo cada uno de ellos una serie de competencias específicas que deben ser evaluadas: en total 60. Este proyecto ya ha superado la fase de “estudio piloto” mediante el diseño de pruebas adecuadas e implementadas en determinadas Facultades de diversos países. Otro modelo en fase de ejecución corresponde al proyecto ideado por las Facultades de Medicina de Escocia y conocido con el nombre de el “*Scottish doctor*”³ que se basa en una propuesta de Harden⁴ de la enseñanza basada en competencias de salidas. En general, las instituciones que se ocupan de la formación de los médicos y de su desarrollo profesional proponen listados de competencias aparentemente diferentes por su denominación; sin embargo todas ellas pueden incluirse dentro de una u otra dimensión



Fig. 4

amplia, o dominio, que se encuentran implícita o explícitamente en todos los listados.

Las pruebas orientadas a la evaluación de competencias son de diversa naturaleza y, debido a la dificultad de la cuantificación de la conducta, han sido sometidos a riguroso escrutinio e investigación, habiendo dado lugar a una literatura científica muy extensa. Su refinamiento y sofisticación persigue eliminar el sesgo y la subjetividad como rasgos principales de la medición de conducta, por lo que su validez y reproducibilidad han sido objeto de estudio intenso. La determinación de estándares previos, tarea de dificultad y complejidad enormes, pero necesaria, ha sido también exhaustivamente trabajada⁵.

Como pruebas de *evaluación de competencias* se pueden citar desde la Prueba de Progreso al Portafolios, pasando por las Escalas Globales o la OSCE. Cada una de ellas presenta caracteres diferenciales en cuanto a propósito de aplicación, validez y fiabilidad⁶. Muy utilizada en el momento actual es la OSCE, "*Objective Structured Clinical Examination*" propuesta por Harden y Gleeson en 1979. La elaboración en su diseño y forma de aplicación, representa lo más cercano a la verdadera y fiable valoración de los estratos superiores de la

pirámide de Miller. Este instrumento será ampliamente tratado en el capítulo siguiente.

BIBLIOGRAFÍA

1. Crossley J, Humphris G & Jolly B. Assessing health professionals. *Medical Education* 2002; 36:800-804.
2. Core Committee Institute for International Medical Education. Global minimum essential requirements in medical education. *Medical Teacher* 2002; 24:130-135.
3. Simpson JC, Furnace J, Crosby J et al. 2002. The scottish doctor –learning outcomes for the medical undergraduate in Scotland: a foundation for competent and reflective practitioners. *Medical Teacher*. 2002; 24: 136-143.
4. Harden RM, Crosby JR, Davis MH & Friedman M. AMEE guide N°1: Outcome-based education: Part- From competency to meta-competency: a model for the specification of learning outcomes. *Medical Teacher* 1999; 21:546-552.
5. Norcini JJ. Setting standards on education tests. *Medical Education* 2003; 37:464-469.
6. Friedman Ben-david M. Principles of Assessment. In: A Practical Guide for Medical Teachers. John A. Dent & Ronald M. Harden eds. Elsevier Churchill Livingstone, 2nd ed. 2005, pp. 282-292.