

La prueba de 2016 fue la más sencilla desde 2009

Examen MIR: cada vez más fácil y con menos capacidad de discriminación

Un estudio psicométrico demuestra que el examen MIR de 2016 fue el más fácil y con menos capacidad discriminativa desde 2009. El estudio, que analiza las respuestas de 23.809 candidatos, concluye que la homogeneidad y alta preparación de los aspirantes dificulta cada vez más diseñar una prueba con un alto valor de discriminación. Aun así, concluye que la prueba tiene "una calidad media adecuada".



Varios alumnos esperan en el aula de la facultad a que empiece una de las pruebas MIR.
DM

DIARIO MEDICO. Francisco Goiri. Madrid
18 septiembre, 2018

A pesar del alto grado de consenso que parece existir sobre la calidad del examen MIR, **la dificultad y capacidad de discriminación de la prueba ha disminuido en las últimas convocatorias**, alcanzando en 2016 su nivel más bajo de complejidad y eficacia para discriminar entre los mejores y los peores candidatos. Ésta es una de las principales conclusiones de un estudio conjunto publicado en la revista de la *Fundación Educación Médica* en el que se hace un **análisis psicométrico** del examen MIR entre las convocatorias de los años 2009 y 2017.

Aunque los autores (médicos, ingenieros y matemáticos) insisten en que sus conclusiones deben enmarcarse en el contexto de un estudio psicométrico y matemático, instan a los responsables del Ministerio de Sanidad a “**augmentar la capacidad discriminativa de la prueba**, al menos hasta alcanzar los valores previos al año 2016, porque, de lo contrario, el examen podría perder calidad”.

El estudio se basa en las respuestas que **23.809 candidatos** que se examinaron en los nueve años analizados introdujeron voluntariamente en una aplicación web expresamente desarrollada para el estudio por el Curso Intensivo MIR Asturias. La muestra supera los 1.600 candidatos en cada una de las convocatorias estudiadas, y en las dos últimas está por encima del 30 por ciento del total de examinados. Aun así, los autores admiten que la **limitación fundamental del trabajo** ha sido, precisamente, no poder acceder a los resultados de todos los candidatos que se presentaron al examen en el periodo de análisis.

Los autores instan a aumentar la capacidad discriminativa de la prueba, al menos hasta los valores previos a 2016

Entre otros indicadores, las preguntas fueron analizadas utilizando su **índice de dificultad** (cociente entre quienes aciertan una pregunta y el número total de examinados), su **índice de dificultad corregido por el azar** (introduce una variable que corrige el posible efecto de acertar una pregunta por casualidad) y su **índice de discriminación** (capacidad de una pregunta para distinguir entre quienes consiguen notas altas y bajas en una prueba). Cuanto mayores sean los dos primeros índices, más fácil es una pregunta, y cuanto más alto sea el tercero más discriminativa resulta.

Datos objetivos

Con estos datos en la mano, el examen MIR de 2016 fue el más fácil de los nueve años estudiados, con **un índice de dificultad de 0,6909, frente a un promedio de 0,6252** (más bajo y, por tanto, más difícil) en los exámenes de las

convocatorias 2009 a 2015. El índice de dificultad corregido por el azar ofrece resultados similares: el examen de 2016 tiene un valor de 0,5971, frente a un promedio de 0,5462 en los siete años previos.

Curiosamente, la prueba del año 2017 vuelve a ser más difícil (con un índice de dificultad de 0,6750 y de dificultad corregido de 0,5740), pero “no sabemos si esto responde a una voluntad expresa del ministerio de volver a hacerla más compleja o a otros factores”, matiza Jaime Baladrón, director del Curso MIR Asturias y uno de los autores del estudio.

En cuanto al índice de discriminación, el del año 2016 también fue el examen menos discriminativo de toda la serie, con una diferencia de 0,0910 sobre el promedio de los exámenes de las convocatorias anteriores.

El número de aspirantes españoles ha crecido año a año: en 2009 eran el 55,82% del total y en 2017 suponían ya un 74,44%

Fernando Sánchez Lasheras, del Departamento de Matemáticas de la Facultad de Ciencias de la Universidad de Oviedo y otro de los autores del estudio, afirma a DM que “desde el punto de vista psicométrico, **la mejor forma de saber si una pregunta es discriminativa es probarla previamente sobre un grupo determinado de personas**, lo que nos permitiría calibrar su capacidad de discriminación. Por motivos evidentes, esto no se puede hacer con el diseño actual de la prueba MIR”.

La pérdida de calidad psicométrica del examen tiene que ver, según los autores, con el contenido y estructura de las propias preguntas, pero también con factores extrínsecos a la prueba, como la preparación, la edad y la nacionalidad de los examinados.



Curvas con la probabilidad de acierto en el examen en los 9 años analizados. En 2016, la probabilidad es más alta que en ningún otro año.

En cuanto a la prueba en sí, desde el año 2015 las preguntas tienen cuatro alternativas de respuesta (frente a las cinco anteriores), pero la penalización por fallar no ha variado: se pierde una pregunta acertada por cada tres erróneas. En suma, **la penalización estadística por contestar preguntas al azar se ha reducido desde 2015**, “y esto quizás haya favorecido un aumento en las puntuaciones de los examinados que tenían un menor nivel de conocimientos”, aventura el estudio. En un examen con poca capacidad de discriminación, muchas de las puntuaciones de los que se presentan se concentran en un rango estrecho y pequeñas variaciones, fruto del azar, en el resultado de uno de los aspirantes se traducen en grandes cambios en el número de orden final. **“En el caso del examen MIR, eso afecta, claro, a la posibilidad de acceder o no a la especialidad médica deseada”, dice Baladrón.**

Según Sánchez Lasheras, en puridad “no se puede hablar de calidad de la prueba, sino de discriminación. Lo que ocurre es que un examen que resulte menos discriminativo cumple peor su propósito de ordenar a los examinados en función de su nivel de conocimiento para elegir plaza”.

Pese a todo, el estudio concluye que el examen MIR tiene una “fiabilidad global muy elevada y una calidad media adecuada”

Pero la disminución de la dificultad y capacidad de discriminación de la prueba también tiene que ver, según el director del Curso MIR Asturias, “con un cambio en la composición de los médicos que se presentan al examen”.

Los estudios previos del Ministerio de Sanidad demuestran que los aspirantes de nacionalidad española y los jóvenes recién graduados en las facultades de Medicina son los que mejores resultados promedio obtienen tradicionalmente en la prueba MIR, y ambos grupos han crecido mucho en los últimos años, “lo que explicaría la aparente mayor facilidad del examen en los últimos años”, dice Sánchez Lasheras. En concreto, el porcentaje de examinados españoles ha pasado de poco más de la mitad en 2009 (un 55,82 por ciento del total) hasta el 74,44 por ciento en 2017.

Además, el aumento del *numerus clausus* en las facultades de Medicina y la apertura de nuevos centros en varias comunidades autónomas en los 9 años objeto del estudio han hecho que también aumente el número de aspirantes recién graduados, que estadísticamente también logran mejores notas medias que los no recién graduados de mayor edad. Por resumirlo con una frase del propio estudio, “**el MIR es cada año más español y más joven**”.

Fiabilidad elevada

Sánchez Lasheras concluye que el MIR es “**un examen objetivo, con una fiabilidad global muy elevada, estructuralmente válido y con una calidad media adecuada**, pero la progresiva mejora del nivel medio de los que se presentan cada año a la prueba hace que cada vez sea más complejo para los examinadores mantener una estabilidad en los valores promedio de dificultad y discriminación”.

En la misma línea, el director del Curso Intensivo MIR Asturias dice que “**la mejora de la calidad psicométrica global de la**

prueba contribuiría a incrementar su capacidad discriminativa y, por tanto, a mejorar su eficacia clasificadora del conjunto de los médicos que se examinan en función de su nivel de conocimiento”.

Pérdida de oportunidad

Las 225 preguntas del examen MIR se completan con **diez de reserva**, que se activan si la comisión calificadora anula alguna de las iniciales. Una de las propuestas del estudio para mejorar la calidad psicométrica de la prueba es que este mecanismo se aproveche para prescindir de las preguntas con menos capacidad discriminativa. **En 2016 y 2017 se anularon sólo 3 y 4 preguntas**, respectivamente (frente a 7 en 2014 y 8 en 2013), y no precisamente las menos discriminativas. Según los autores, no aprovechar este mecanismo supone “una pérdida de oportunidad” para mejorar la calidad psicométrica del examen.