

Original

# Pruebas de acceso a la formación sanitaria especializada para médicos y otros profesionales sanitarios en España: examinando el examen y los examinados

Albert Bonillo

Departamento de Psicobiología y Metodología de Ciencias de la Salud, Universitat Autònoma de Barcelona, Bellaterra, Barcelona, España

## INFORMACIÓN DEL ARTÍCULO

### Historia del artículo:

Recibido el 21 de junio de 2011

Aceptado el 14 de septiembre de 2011

On-line el 22 de diciembre de 2011

### Palabras clave:

Internado y residencia  
 Criterios de admisión escolar  
 Médicos hospitalarios  
 Psicometría

### Keywords:

Internship and residency  
 School admission criteria  
 Hospitalists  
 Psychometrics

## R E S U M E N

**Objetivos:** Estudiar las pruebas de acceso a la Formación Sanitaria Especializada de las convocatorias 2005 y 2006. Se pretende evaluar la calidad de los exámenes y explorar las variables de los aspirantes que permiten predecir la puntuación final.

**Métodos:** El Ministerio de Sanidad y Consumo proporcionó las respuestas de los 23.136 aspirantes de ambas convocatorias, así como variables demográficas y el valor baremado de su expediente académico. **Resultados:** Se realiza un análisis de ítems a partir de las respuestas de los aspirantes para así evaluar la fiabilidad de las pruebas. Además, se calculan modelos de regresión lineal para estudiar qué variables permiten predecir la puntuación final de un aspirante.

**Conclusiones:** Las pruebas de acceso a la Formación Sanitaria Especializada tienen una excelente calidad psicométrica. Serían optimizables reduciendo el número de alternativas y eliminando algunos ítems más a posteriori. Por último, los alumnos españoles son los que mejor nota media ajustada logran.

© 2011 SESPAS. Publicado por Elsevier España, S.L. Todos los derechos reservados.

## Access tests to specialized health training for doctors and other healthcare professionals in Spain: examining the exam and the examined candidates

### A B S T R A C T

**Objectives:** This study examines the accessing tests to Specialised Health Training for 2005 and 2006 calls. It aims to assess the quality of exams and explore candidates' variables that allow predicting the final score.

**Methods:** The Ministry of Health provided the 23,136 candidates' answer from both calls, plus demographic variables and the normalised value of their student record.

**Results:** Item's analysis is made from candidates' answers for evaluating the test reliability. In addition, it's been calculated linear regression models for studying which variables allow predicting a candidate's final score.

**Conclusions:** the accessing tests to Specialised Health Training have excellent psychometric quality. It would be improved by reducing the number of choices and eliminating some items more afterwards. Finally, Spanish students are the ones who achieved best adjusted mean score.

© 2011 SESPAS. Published by Elsevier España, S.L. All rights reserved.

## Introducción

La nota final que da acceso a la formación sanitaria especializada se obtiene a partir del expediente académico y del resultado de una prueba específica de cada programa. Ésta consiste en un examen tipo test con cinco opciones de respuesta<sup>1</sup>. El aspirante puede escoger especialidad y centro en función del orden de su nota final. La plaza a la cual se aspira ha dado lugar al nombre «popular» de la prueba: MIR para medicina, FIR para farmacia y EIR para enfermería, entre otros.

La formación sanitaria especializada no ha recibido demasiada atención por parte de la literatura, y las pruebas que rigen su acceso aún menos, con la salvedad del MIR<sup>2</sup>. Se han realizado propuestas de reforma del acceso<sup>3</sup>, así como análisis de la validez del formato de respuesta múltiple respecto a otros también habituales

en la docencia médica<sup>4</sup>. En otros países, exámenes análogos son estudiados y analizados profusamente; a modo de ejemplo, una búsqueda no exhaustiva en Medline arroja 270 referencias sobre el USMLE (*United States Medical Licensing Examination*).

Existen varias directrices con recomendaciones para redactar óptimamente preguntas de tipo test<sup>5,6</sup>, y se ha estudiado profundamente el efecto que tiene el no respetarlas, también en pruebas médicas<sup>7-9</sup>. Las guías indican que un aspecto importante a estudiar son las propiedades psicométricas de un examen, tanto más si éste se repite a lo largo del tiempo y su resultado tiene efectos sociales importantes. En una prueba de rendimiento, como es un examen, esto implica analizar la dificultad y la discriminación de cada una de las alternativas<sup>10</sup>. Además, es de interés estudiar cómo se relacionan las características de los individuos con su rendimiento. Este análisis permite no sólo conocer a los participantes sino corregir posibles sesgos de las pruebas.

En ninguna de las búsquedas realizadas en bases bibliográficas especializadas (Medline, Scielo, Cochrane Plus y Medes) hemos

Correo electrónico: albert.bonillo@uab.cat

**Tabla 1**  
Exámenes de acceso a la formación sanitaria especializada (España 2005-2006): descripción de variables por programa y año de convocatoria

	EIR-05 N = 4.653	FIR-05 N = 829	FIR-06 N = 849	MIR-05 N = 7.908	MIR-06 N = 8.897
Edad, años, media (DT)	27,1 (5,0)	26,7 (3,9)	26,0 (4,1)	29,2 (6,2)	28,5 (6,4)
Años fin estudios, media (DT)	4,8 (4,0)	2,8 (2,9)	3,0 (3,0)	3,5 (4,8)	3,7 (5,0)
Expediente, media (DT)	1,76 (.41)	1,79 (.55)	1,75 (.52)	1,89 (.54)	1,88 (.55)
Examen, media (DT)	85,3 (35,3)	331,6 (125,1)	348,8 (130,0)	297,0 (136,4)	306,9 (142,0)

DT: desviación típica.

encontrado un análisis de las propiedades de las pruebas de acceso a la formación sanitaria especializada, y tampoco ningún trabajo que aporte datos sobre qué características de los aspirantes son predictoras de un buen resultado en las pruebas.

Este trabajo pretende cubrir este vacío, cuantificando dichos aspectos. Considerando la larga historia de las pruebas de acceso a la formación sanitaria especializada, esperamos encontrar que sus propiedades psicométricas sean adecuadas. En cuanto a las características de los participantes, no tenemos ninguna hipótesis, ya que no hay estudios previos sobre ellas.

## Método

Tras efectuar una petición, la Subdirección General de Ordenación Profesional del Ministerio de Sanidad, que es la encargada de gestionar las pruebas de formación sanitaria especializada, aceptó proporcionarnos datos relativos a las convocatorias de los años 2005 y 2006, que se realizaron en enero de 2006 y de 2007, respectivamente.

Las matrices contenían, para las convocatorias de MIR, FIR y EIR, datos individuales de cada aspirante presentado, obviamente, sin identificador personal alguno. De la convocatoria de EIR se trabajó sólo con la de 2005 porque los datos de la de 2006 estaban dañados y no podían recuperarse.

Las variables registradas de cada individuo pueden agruparse en tres bloques: por un lado, demográficos (sexo, edad y nacionalidad), por otro, relativas a su currículum académico (nota baremada de su licenciatura o diplomatura [rango teórico entre 1 y 4] y año en que la finalizó); y por último las relativas a la prueba que completó, es decir, la respuesta dada a cada ítem/pregunta del examen y su nota final. La **tabla 1** muestra la descripción de estas variables para los 23.136 aspirantes estudiados.

Las pruebas de MIR y de FIR tienen 250 preguntas más 10 de reserva, que se usan para, por orden, sustituir a las que son eliminadas por su mal rendimiento o ambigüedad tras un proceso de análisis interno y de impugnación externa. El examen de EIR sigue el mismo esquema, con la salvedad de que usa 100 preguntas y 10 de reserva. Cada pregunta acertada suma tres puntos, cada una fallada resta uno y se obvian las preguntas no contestadas. El rango teórico de puntuaciones es entre -250 y 750 para MIR y FIR, y de -100 a 300 para EIR.

Para nuestro análisis hemos eliminado las preguntas que fueron anuladas y aquellos aspirantes que no contestaron ninguna. Dejaron su examen completamente en blanco, en las convocatorias de 2005 y de 2006, respectivamente, 1085 (12,1%) y 1359 (14,1%) en el MIR, 119 (12,5%) y 126 (12,9%) en el FIR y 1246 (21%) en el EIR.

En la primera parte de los resultados se presenta el análisis de ítems de cada convocatoria, que se basa en dos índices. Por un lado, la dificultad de una pregunta se entiende como la proporción de personas que la aciertan de entre los que la han contestado. En ítems de elección múltiple (esto es, no binarios) es habitual corregir esta proporción restándole la de errores (dividida entre el número de opciones de respuesta menos uno) para así sustraer del acierto aquella parte que se explica por puro azar<sup>11</sup>. Paradójicamente, un valor alto del índice indica baja dificultad. Por otro lado, se entiende como discriminación de un ítem su capacidad para distinguir entre

los participantes con un alto frente a un bajo rendimiento<sup>12</sup>. Puede calcularse mediante la correlación entre la puntuación del ítem y la total de la prueba, o bien utilizando el índice D. Éste, que ha sido el utilizado en el presente trabajo, se calcula mediante la diferencia de proporciones de acertantes entre el grupo de personas con un rendimiento en la prueba superior al 73% del resto frente al grupo que ha obtenido un rendimiento inferior al 23%. Ebel<sup>13</sup> propuso puntos de corte para evaluar el índice D, siendo menor de 0,2 inaceptable, menor de 0,3 revisable, menor de 0,4 aceptable pero mejorable, y mayor de 0,4 óptimo. Además, es deseable que la D sea muy semejante entre distractores de un mismo ítem, ya que sería indicador de que son iguales en atractivo<sup>14</sup>.

En la segunda parte del apartado de resultados se utilizan modelos de regresión lineal para explicar la puntuación total de la prueba a partir de las variables demográficas (sexo, edad, nacionalidad) y relativas a la licenciatura/diplomatura (años desde su finalización y puntuación del expediente baremada). Puesto que no se trata de datos muestrales sino poblacionales, no tiene sentido realizar pruebas de significación ni mostrar los errores estándar. Se utiliza esta técnica multivariada para valorar la contribución propia de cada predictor, ajustando su efecto por el resto, y evitar la confusión inherente a todo estudio no experimental<sup>15</sup>. Para poder comparar variables con distintas métricas, o la misma entre diferentes convocatorias, se muestran los coeficientes  $b$  y  $\beta$ .

Se han incluido en las ecuaciones todas las variables predictoras, tanto demográficas como relativas al pasado académico del candidato, que estaban registradas. La variable «nacionalidad» sólo se ha incluido en las convocatorias de MIR por ser éstas las que presentan proporciones no anecdóticas de extranjeros presentados.

Para calcular los índices del análisis de ítems se han creado algoritmos *ad hoc*. Todos los cálculos se han efectuado con SPSS v.19.

## Resultados

### Análisis de ítems

La **figura 1** muestra las curvas de dificultad (ordenando por ésta los ítems en abscisas) de cada una de las pruebas. En este tipo de gráfico, una pendiente de 45° indicaría un escalamiento óptimo de los ítems, ya que habría preguntas para todos los grados de dificultad posibles.

En primer lugar, se constata que hay dificultades negativas y que sólo pueden explicarse por aplicar la corrección del azar a ítems muy difíciles. En segundo lugar, cabe tener en cuenta que la prueba de enfermería cuenta con 100 ítems (frente a los 250 del resto de las convocatorias), cosa que explica la evidente diferencia en las pendientes. En tercer lugar, las curvas de las pruebas de farmacia son superiores a las de medicina, lo que indica que son pruebas más fáciles. En cuarto lugar, las curvas son muy semejantes intraprogramas, es decir, la dificultad de las pruebas es muy semejante año tras año.

La **figura 2** muestra, mediante un diagrama de cajas, la discriminación de la opción correcta y de cada uno de los distractores, ordenados éstos de mayor a menor, de las cinco convocatorias analizadas (esto es, [250 ítems  $\times$  2 programas  $\times$  2 años + 100 ítem de

Download English Version:

<https://daneshyari.com/en/article/1073701>

Download Persian Version:

<https://daneshyari.com/article/1073701>

[Daneshyari.com](https://daneshyari.com)