

Selección de residentes en un hospital de comunidad: resultados obtenidos por los egresados de distintas universidades y análisis de la calidad del examen administrado

AMANDA GALLI*¹, DIEGO HÄBICH², HORACIO YULITTA^{1,2}

¹Saidem (Sociedad Argentina de Investigación y Desarrollo en Educación Médica); ²Hospital Alemán

* ✉ amandagalli@ciudad.com.ar / tel-fax: 4827-5120

RESUMEN

Introducción: Las distintas jurisdicciones e instituciones privadas en Argentina desarrollan sus propios sistemas de selección de residentes. El Hospital Alemán realiza anualmente un concurso para seleccionar 13 residentes para 8 especialidades; se postulan al mismo egresados de universidades públicas y privadas. El propósito de este trabajo es presentar los resultados obtenidos en el concurso de 2006 y analizar la calidad del examen implementado. *Material y métodos:* A partir de un examen escrito de 100 preguntas de selección múltiple se compararon los puntajes obtenidos según sexo, universidad de egreso y especialidad elegida. Asimismo se realizó la revisión de la calidad de las preguntas del examen según la escala de Galofré, el índice de dificultad y un análisis cuali-cuantitativo. *Resultados:* La media de los puntajes fue 62,1 (40-83 DS 7,66). No hubo diferencia significativa por sexo. Los promedios de puntajes más altos correspondieron a egresados de tres universidades públicas. Nueve universidades privadas no presentaron diferencias significativas con las universidades públicas; y seis universidades privadas no tuvieron diferencias significativas con la Universidad de Buenos Aires. Los mejores promedios de puntajes corresponden a los postulantes a Pediatría y a Clínica Médica. Al evaluar las preguntas del examen con la escala de Galofré, el índice de dificultad y el análisis cuali-cuantitativo se concluye que el examen resultó de una calidad muy buena.

Palabras clave: selección de residentes - seguimiento de egresados - preguntas de selección múltiple - banco de preguntas - índice de dificultad.

ABSTRACT

Selection of residents in a community hospital: Results obtained by graduates from different universities and evaluation of the test quality. The different jurisdictions and private health institutions are using their own systems for the selection of residents. The «Hospital Alemán» annually selects 13 residents for 8 specialties among graduates of public and private universities using a 100 multiple choice questions test. The results of the 2006 evaluation have been stratified according sex, university and selected specialty. Mean of marks was 62,1 (40-83 SD 7.66) without significant difference by sex. The best results were of three public universities graduates. Nine private universities average scores do not present significant differences with public universities; six private universities do not have significant differences with the Universidad de Buenos Aires. The best averages correspond to Pediatric and Internal Medicine candidates. When evaluating the MCQ using the Galofré scale, the difficulty index and a quali-quantitative analysis, it was concluded that the test was of very good quality.

Key words: postgraduate medical training- residency selection - application process - admissions test - educational measurement - improve measurement instruments - MCQ items - scoring test data - item analysis.

INTRODUCCIÓN

En Argentina, como en otras partes del mundo, las residencias constituyen la modalidad de formación y especialización más importante para médicos de re-

ciente graduación. Cada jurisdicción –por ejemplo la Ciudad Autónoma de Buenos Aires, o la Provincia de Buenos Aires– y las distintas instituciones privadas, desarrollan y aplican sus propios sistemas de selección. No existe, como en otros países, un examen

nacional unificado para el ingreso al sistema de residencias. Por ejemplo, en Panamá es responsabilidad de la Facultad de Medicina de la Universidad de Panamá, organizar y aplicar la prueba única de Conocimientos Generales Básicos para seleccionar a todos los aspirantes de plazas de residencias médicas en el territorio nacional. Durante un lapso de cinco años (2000-2006) el examen ha sido confeccionado con la participación de la Junta Nacional Examinadora de los Estados Unidos de América de acuerdo a convenio firmado entre el Ministerio de Salud, la Universidad de Panamá y la Junta Nacional Examinadora de los Estados Unidos de América¹. En México el *Examen nacional de aspirantes a residencias médicas* se implementa desde 1977 como mecanismo de ingreso para cursar alguna especialidad médica en las instituciones de salud, públicas y privadas². En España, en 1977 se estableció el examen nacional de ingreso a las residencias (MIR)³.

En nuestro país existen 28 facultades de medicina, 11 de gestión estatal y 17 de gestión privada (Cuadro 1), en las que ingresan anualmente unos 14.000 alumnos y se gradúan alrededor de 6.000 médicos por año⁴.

En el área metropolitana, egresados de las distintas universidades se postulan a las residencias del Gobierno de la Ciudad de Buenos Aires y de los establecimientos privados. Un médico recién graduado se presenta a varias pruebas de selección en su intento de ingresar a un programa de residencia.

Generalmente la selección de residentes se basa en un examen escrito y en una entrevista personal. En algunas instituciones se exige, como requisito para la inscripción, un promedio en la carrera de grado de 7 puntos como mínimo.

«Los directores o responsables de la admisión de candidatos a los programas de residencias médicas habitualmente seleccionan un conjunto de factores como la edad, los antecedentes personales y de la escuela secundaria del aspirante, el promedio de la ca-

rrera de grado, el puntaje obtenido en el examen de ingreso a la residencia, los antecedentes del internado, las entrevistas y las cartas de referencia o recomendación, a fin de generar un ranking final de selección en base al mérito»⁵.

En términos generales se aplican los exámenes y se establece el ranking en función de los puntajes obtenidos^{6, 7, 8, 9, 10, 11} pero no se hace una revisión de la calidad del examen utilizado.

... «los evaluadores tienen la obligación profesional de que las evaluaciones propuestas o finalizadas estén sujetas a una evaluación competente». Se entiende por meta-evaluación la tarea de evaluar la calidad del proceso de evaluación¹².

El Hospital Alemán realiza anualmente un concurso para cubrir 13 cargos de residentes en 8 especialidades. La prueba de selección es un examen escrito con 100 preguntas de selección múltiple. Pueden postularse médicos egresados de facultades públicas y privadas, siempre que se hayan graduado con un promedio de carrera de 7 puntos como mínimo. Los egresados que obtienen los mejores puntajes en el examen pasan a una entrevista personal. Los resultados de la entrevista son los que definen qué postulantes se incorporan al Hospital como residentes¹⁰.

Resulta de interés comparar los resultados obtenidos en este concurso con los presentados por otros autores^{6, 8} e identificar diferencias y similitudes con otros concursos de selección de residentes.

Los propósitos de esta publicación son: informar los resultados obtenidos por los egresados de distintas universidades en el concurso realizado en el Hospital Alemán en el año 2006, compararlos según sexo, universidad de egreso y especialidad elegida, y presentar la metodología utilizada para evaluar la calidad de las preguntas del examen.

MATERIALES Y MÉTODOS

Se inscribieron 440 postulantes a través de internet en el concurso del año 2006, quienes respondieron a un examen escrito de 100 preguntas de selección múltiple con 5 opciones. Se exploraron conocimientos de las 4 disciplinas clínicas básicas y conceptos de salud pública.

Las preguntas fueron redactadas, en primera instancia, por los jefes de los servicios del hospital y luego revisadas por un grupo técnico. Finalmente dos profesores, de Clínica Médica y de Cirugía, seleccionaron las preguntas que se incluyeron en el examen.

Todos los inscriptos rindieron el mismo examen el mismo día, a la misma hora y en el mismo lugar. Las 100 preguntas se presentaron en un cuadernillo impreso y las respuestas debían ser registradas en un formulario adecuado para su lectura electrónica.

Cuadro 1
Carreras de Medicina de Gestión Estatal
y de Gestión Privada

Gestión estatal	UBA, Córdoba, Tucumán, La Plata, UNNE, Cuyo, del Sur, Litoral, Rosario, del Comahue, La Rioja.
Gestión privada	Abierta Interamericana, Aconcagua, Adventista del Plata, Austral, Barceló, Católica de Córdoba, Católica de Cuyo, Cemic, del Salvador, Favaloro, Hospital Italiano de Buenos Aires, Hospital Italiano de Rosario, Maimónides, Mendoza, Morón, UCA, UCES.

La corrección de los exámenes se realizó con un lector óptico, con generación automática de puntajes. Cada respuesta correcta valía un punto. De este modo se obtuvo un registro computarizado de los resultados, según sexo, universidad de egreso, puntaje obtenido y especialidad solicitada.

Los resultados obtenidos se describen por medidas de tendencia central y de dispersión según corresponda (promedio, mediana, rango y desvío estándar). Se comparan universidades públicas y privadas y cada una de las universidades privadas y el promedio de las públicas y con el promedio de la UBA. El bajo número de postulantes de algunas universidades hace que en algunos casos la comparación pierda significancia estadística.

Los resultados fueron analizados con el programa *Primer of Biostatistics*, Glantz-McGraw-Hill, mediante diferencia de proporciones, t-test y test de ANOVA.

La evaluación de la calidad de las preguntas se realiza una vez aplicado el examen. Cada una de las 100 preguntas se analiza utilizando la escala de Galofré, el índice de dificultad y el análisis cuali-cuantitativo. Los resultados obtenidos con esta revisión son utilizados para el mejoramiento de las preguntas que serán archivadas en un *banco de ítem* con el propósito de utilizarlas en otro examen.

La **escala de Galofré**¹³ es una sistematización de las recomendaciones publicadas para la redacción de buenas preguntas de selección múltiple^{14, 15}.

Se toman en cuenta 10 requisitos o condiciones y los resultados se resumen en una escala de 5 puntos (Tabla 1). Permite identificar los defectos de construcción de una pregunta y es un recurso práctico para orientar el proceso de mejoramiento de las preguntas elaboradas por especialistas en el contenido y/o docentes de una asignatura.

La calidad de las preguntas se clasifica con una puntuación de 1 a 5.

Una puntuación de 5 expresa la mejor calidad posible, corresponde a una pregunta que tiene viñeta y no presenta ningún defecto de construcción; un 4 corresponde a una pregunta sin viñeta y sin defectos de construcción.

Si una pregunta, sin importar si tiene o no viñeta, tiene uno, dos, tres o más defectos, recibe un puntaje de 3, 2, 1 respectivamente. Preguntas con 3 puntos o menos no deben ser incluidas en los exámenes.

El **índice de dificultad** es la cantidad de respuestas correctas que acumula una pregunta. Las preguntas son muy fáciles si son respondidas correctamente por más del 85% de los alumnos; relativamente fáciles entre 84 y 69%; dificultad media entre 32 a 68%; relativamente difícil entre 31 a 16% y muy difíciles si son respondidas correctamente sólo por el 15% o menos de los alumnos^{15, 16}. Se considera que los exámenes deben tener preguntas con una dificultad me-

Tabla 1
Escala de Galofré

Los 10 criterios o requisitos a tener en cuenta para la realización de buenas preguntas de selección múltiple son:

- **Presencia de viñeta:** se refiere a la presencia de un caso clínico o procedimiento de laboratorio o problema.
- **Nivel taxonómico de aplicación de conocimiento o superior:** que la pregunta requiera usar un proceso mental superior al simple recuerdo de un dato específico.
- **Enunciado completo:** el lector puede comprender la pregunta y pensar una respuesta antes de leer las opciones.
- **Concordancia gramatical entre el enunciado y las opciones:** respetar la concordancia de género y número
- **Distractores verosímiles:** todas las opciones deben parecer verdaderas, igualmente atractivas.
- **Opciones homogéneas:** similares en cuanto al contenido.
- **Extensión similar entre las opciones:** todas las opciones deben tener una longitud similar, más o menos el mismo número de palabras.
- **Opciones ordenadas:** lógica o numéricamente. Opciones con números en orden creciente o decreciente.
- **Evitar el uso de negaciones:** como Excepto – No es
- **Evitar las opciones Ninguna y Todas las anteriores:** bastaría saber que dos opciones son correctas para saber que si existe una opción de «todas las anteriores» ésta tiene que ser la respuesta.

dia para poder identificar los sujetos de mejor nivel de conocimientos^{14, 15, 16, 17, 18}.

El **análisis cuali-cuantitativo** de una pregunta se basa en la cantidad de veces que cada una de las opciones es seleccionada como respuesta correcta. Una buena pregunta de selección múltiple se caracteriza porque todas las opciones ofrecidas –clave y distractores– resultan atractivas y son elegidas por algunos de los examinandos. Si un distractor no resulta elegido por nadie significa que no funcionó como tal^{14, 15, 16, 17}.

El análisis cuali-cuantitativo de cada una de las preguntas del examen debe ser un paso previo a la comunicación de los resultados del examen.

La revisión de las preguntas teniendo en cuenta cómo fueron respondidas por el total de los examinandos es una buena práctica que permite detectar las preguntas «que no han funcionado bien» y decidir si todas las preguntas son válidas o si, por error técnico, alguna pregunta debe ser anulada. Si se anulara alguna pregunta se modificaría el puntaje máximo posible de la prueba y podría variar el puntaje obtenido por cada sujeto. En la actualidad se dispone de tecnología informática-electrónica que permite una rápida lectura de los resultados y la elaboración de una estadística suficiente como para tomar decisiones como la de anular alguna pregunta por razones técnicas^{14, 16, 17, 18, 19}.

RESULTADOS

Rindieron el examen 440 médicos todos con nota promedio de la carrera universitaria de 7 puntos o más. El puntaje promedio en el examen fue de 62,1 respuestas correctas; variaron entre 83-40 y el desvío estándar fue de 7,66. La mediana fue de 65.

El 55% de los postulantes fueron mujeres. No se evidenciaron diferencias significativas al analizar el promedio de respuestas correctas según género. ($p=0,254$). Los resultados discriminados por género se presentan en la Tabla 2.

En cuanto a la universidad de egreso, el 64% de los postulantes fueron egresados de ocho universida-

Tabla 2
Resultados obtenidos en el examen discriminados por género

Género	Promedio	Rango	DS	p
Mujeres	64,44	41 – 80	7,40	0,254 NS
Hombres	63,36	40 - 81	7,95	

Tabla 3
Cantidad de postulantes y resultados obtenidos en el examen según Universidad de egreso. (El nombre de cada Universidad ha sido reemplazado por un número antecedido por la sigla PUB –pública– o PRIV –privada–)

Universidad	Cantidad de postulantes	Rango	Promedio de puntajes	DS
Públicas				
PUB1	195	44- 83	66,57	6,66
PUB2	30	40- 76	60,43	8,20
PUB3	22	42- 69	61,36	6,81
PUB4	20	52- 81	65,50	6,70
PUB5	14	55- 72	64,50	5,22
PUB6	9	54- 69	61,55	5,05
PUB7	3	57- 74	67	8,88
PUB8	1	-	62	-
Privadas				
PRIV1	40	46- 74	63,53	7,88
PRIV2	40	51-78	64,80	6,10
PRIV3	20	41- 72	57,35	9,59
PRIV4	10	44-71	58,10	9,78
PRIV5	10	51-66	60	5,16
PRIV6	9	54- 67	61,67	5,38
PRIV7	9	48-70	58,40	6,42
PRIV8	7	49- 73	61,71	9,49
PRIV9	7	51- 68	56,14	6,89
PRIV10	4	48-73	61,25	12,14
PRIV11	5	56-69	64	5
PRIV12	3	59-70	64,33	5,50
PRIV13	1	-	64	-

des públicas y el 36% de trece universidades privadas. Los tres promedios de puntajes más altos correspondieron a tres universidades públicas (67; 66,57; 65,50); los más bajos a tres privadas (58,1; 57,35; 56,14) (Tabla 3).

Existió una diferencia significativa ($p=0,0001$, t-test) entre el promedio de puntajes obtenidos por los egresados de universidades de gestión estatal (65,22) y los de las universidades de gestión privada (61,77).

Al comparar el promedio de puntajes obtenidos por los egresados de cada una de las universidades privadas versus el promedio de las públicas (ANOVA), ocho privadas no presentaron diferencias significativas mientras que otras cuatro privadas obtuvieron puntajes significativamente más bajos (Tabla 4).

Al comparar el promedio de puntajes obtenidos por los egresados de cada una de las privadas versus el promedio de la UBA (PUB1), se observó que los resultados de cinco universidades privadas no presentan diferencias significativas con ella, mientras que otras siete arrojan resultados significativamente más bajos (Tabla 5).

Los puntajes promedio obtenidos por los aspirantes a las especialidades de Pediatría (76,9) y a Clínica Médica (76,1) fueron significativamente superiores a los de otras especialidades ($p=0.007$).

Cada servicio de las distintas especialidades concursadas, entrevistó a los 10 postulantes con puntajes más altos y en función de los resultados de la entrevista se hizo la selección final. Todos los ingresantes

Tabla 4
Comparación del promedio de puntajes de graduados de Universidades Privadas y el promedio de todas las Universidades Públicas

Universidad	Cantidad de postulantes	Promedio de puntajes en el examen	Desvío estándar	Valor de p
Públicas				
(todas)	293	65,22	7,06	
PRIV2	40	64,80	6,10	0,72 NS
PRIV1	40	63,53	7,88	0,165 NS
PRIV3	20	57,35	9,59	<0,001
PRIV4	10	58,10	9,78	0,002
PRIV5	10	60	5,16	0,021 NS
PRIV6	9	61,67	5,38	0,136 NS
PRIV7	9	58,40	6,42	0,005
PRIV	7	56,14	6,89	<0,001
PRIV8	7	61,71	9,49	0,198 NS
PRIV11	5	64	5	0,7 NS
PRIV10	4	61,25	12,14	0,27NS
PRIV12	3	64,33	5,50	0,82 NS

En la tabla no se incluyen datos de la Universidad PRIV13 porque sólo se presentó un postulante.

Tabla 5

Comparación del promedio de puntajes de los graduados de cada una de las Universidades Privadas con el promedio de la Universidad de Buenos Aires (PUB1)

Universidad	Cantidad de postulantes	Promedio de puntajes en el examen	Desvío estándar	Valor de p
PUB1	194	66,58	6,66	
PRIV2	40	64,80	6,10	0,12 NS
PRIV1	40	63,53	7,88	0,01
PRIV3	20	57,35	9,59	<0,001
PRIV4	10	58,10	9,78	<0,001
PRIV5	10	60	5,16	0,002
PRIV6	9	61,67	5,38	0,03
PRIV7	9	58,40	6,42	<0,001
PRIV9	7	56,14	6,89	<0,001
PRIV8	7	61,71	9,49	0,06 NS
PRIV11	5	64	5	0,39 NS
PRIV10	4	61,25	12,14	0,12 NS
PRIV12	3	64,33	5,50	0,56 NS

En la tabla no se incluyen datos de la Universidad PRIV13 porque sólo se presentó un postulante.

Tabla 6

Ingresantes al programa de residencias médicas del Hospital Alemán según especialidad, universidad de egreso y puntaje obtenido en el examen

Especialidad (Cargos)	Universidad de egreso (Puntaje obtenido en el examen)
Pediatría (2)	Tucumán (68) – UNNE (76)
Clínica Médica (2)	Austral (73) – UN Córdoba (66)
Cirugía (2)	UBA (73) – Austral (70)
Imágenes (2)	Del Salvador (69) – UBA (68)
Ginecología y Obstetricia (1)	Austral (67)
Oftalmología (1)	Austral (73)
Cardiología (1)	UBA (67)
Urología (1)	UNNE (76)
Traumatología (1)	Austral (73)

tienen puntajes más altos que el promedio del grupo total de postulantes (Tabla 6).

Calidad del examen

No se anuló ninguna pregunta por razones técnicas.

Según la escala de Galofré, se identificaron 83 preguntas sin defectos de construcción, 11 preguntas con un defecto de construcción y 6 preguntas con dos defectos de construcción.

El defecto de construcción más frecuente fue un bajo nivel taxonómico, es decir que la pregunta solía sólo el recuerdo de datos.

En esta oportunidad la revisión de las preguntas con la escala Galofré se hizo a posteriori, una vez que ya se había aplicado el examen a todos los postulantes. La única decisión en relación a las seis preguntas con dos defectos de construcción fue no incluirlas en el banco de ítems²⁰⁻²¹.

Índice de dificultad. El examen resultó una prueba de dificultad media ya que 83 preguntas resultaron con dificultad media, 12 muy fáciles y 5 muy difíciles.

Análisis cuali-cuantitativo: 87 preguntas con las cinco opciones igualmente atractivas. En 13 preguntas se encontraron distractores que no funcionaron como tales. Se presentan dos ejemplos del análisis cuali-cuantitativo.

Cuadro 2

Análisis cuali-cuantitativo. Porcentaje de sujetos que elige cada una de las opciones. Ejemplo 1

¿Cuál es el tratamiento de elección en la angina estreptocócica?

opciones	elegida por el
a) Penicilina VO x 7 días.	20,85%
b) Amoxicilina VO x 7 días.	28%
c) Penicilina VO x 10 días.	19% clave
d) Amoxicilina EV x 7 días	15,53%
e) Azitromicina EV x 5 días.	16,60%

Las 5 opciones ofrecidas resultaron igualmente atractivas, las elecciones se distribuyen en forma pareja. Pregunta relativamente difícil para este grupo de postulantes ya que sólo el 19% la respondió correctamente

Cuadro 3

Análisis cuali-cuantitativo. Porcentaje de sujetos que elige cada una de las opciones. Ejemplo 2

¿En cuál de las siguientes patologías NO está involucrado el Estreptococo B hemolítico grupo B?

opciones	elegida por el
a) Trabajo de parto pretérmino	35,32%
b) Rotura prematura de membranas	29,57%
c) Corioamnionitis	9,36%
d) Aborto	25,11% clave
e) Sepsis neonatal.	0,64%

La opción e) no resultó atractiva, no es un buen distractor, se debe buscar otra opción para mejorar la pregunta y archivarla en el banco de ítems para un próximo examen.

DISCUSIÓN

Las características del grupo de egresados que postuló en el Hospital Alemán son similares a las de otros grupos que participaron en otros concursos^{6,8}. Resulta difícil establecer comparaciones con los datos de la Dirección de Capacitación Profesional y Técnica del Gobierno de la Ciudad de Buenos Aires por la gran diferencia en la cantidad de postulantes (en el año 2006 se presentaron 3603 médicos) y fundamentalmente porque es un concurso para cargos de residentes y de concurrentes en el que los egresados se inscriben sin límite de edad ni de años de graduación^{22,23}.

No son muchos los datos publicados sobre resultados de los exámenes de selección de residentes en nuestro medio pero con la información disponible se advierte que, en el área metropolitana, tanto en el sector público como en el privado, alrededor del 45% de los postulantes son egresados de la UBA y que entre un 20-30% son egresados de otras universidades públicas. En los dos últimos años se observa una mayor presencia de los egresados de universidades privadas. En los hospitales de comunidad, donde se concursan sólo cargos de residentes, el porcentaje de egresados de universidades privadas es más alto que en el Gobierno de la Ciudad de Buenos Aires^{6,8,22,23}.

La tendencia a una mayor proporción de mujeres que se observa en la matrícula de las carreras de medicina ya es también evidente entre los postulantes a residencias. No se observaron diferencias en el rendimiento del examen entre varones y mujeres. Los resultados encontrados al comparar el rendimiento de los egresados de las distintas universidades son coincidentes, en gran medida, con los datos publicados por Neuman y cols.⁶ quienes tampoco encontraron diferencia significativa entre UBA y las universidades Austral, del Salvador, Nacional de Cuyo y Favaloro.

En el Hospital Alemán, como en otros establecimientos privados, el factor decisivo en la selección de los finalistas es la entrevista personal en la que se toman en cuenta condiciones y rasgos personales de los candidatos que los evaluadores consideran importantes y que son apreciados de un modo subjetivo. Los candidatos que ingresaron –que cubrieron los 13 cargos concursados– fueron seleccionados entre los que se presentaron a las entrevistas. Cuando son citados los postulantes que lograron los más altos puntajes en el examen se producen «corridas». Esto quiere decir que los egresados que participaron en varios concursos y que en todos los exámenes obtuvieron buenos puntajes, pueden optar por otra residencia y no presentarse a la entrevista en el Hospital Alemán. De esta forma el cargo queda disponible para postulantes que obtuvieron puntajes un poco más bajos. En este concurso el puntaje máximo obtenido por un postulante fue 83 puntos y el puntaje máximo

de un ingresante fue 76 puntos. En la práctica los establecimientos asistenciales se «autoevalúan» analizando la brecha entre los mejores puntajes obtenidos en el examen y los puntajes de los que finalmente se incorporan a la residencia.

En algunos ámbitos, académicos y no académicos, existe una creencia o preconcepción acerca de las universidades privadas: «aprueban a todos los alumnos porque pagan»; este supuesto no se confirma en este estudio dado que los egresados de algunas privadas tienen tan buen rendimiento como los de UBA y algunos egresados de universidades de gestión estatal tienen tan mal rendimiento como egresados de algunas universidades privadas. Resultados similares ya fueron reportados por otros autores^{6,8,9}.

La mayoría de las publicaciones sobre selección de residentes analizan las características de los postulantes y los resultados del proceso de selección^{6,8,9,11,23,24,25} pero no es frecuente el análisis de la calidad del examen utilizado para la selección.

En general se considera que un buen examen es aquel cuyas preguntas tienen un índice de dificultad del 50%, no hay que olvidar que el índice de dificultad de una pregunta siempre es relativo al grupo de sujetos que rinde el examen; la misma pregunta puede ser muy fácil si es respondida por expertos y muy difícil si es respondida por un grupo de novatos¹⁵. Si el propósito del examen es la selección de un pequeño grupo de candidatos es recomendable que las preguntas tengan un mayor grado de dificultad. La validez y confiabilidad de un examen depende de la cantidad y de la calidad de las preguntas²⁶. La construcción de exámenes de buen nivel es una ardua tarea y es conveniente ir archivando las preguntas utilizadas y evaluadas en un *banco de ítems*. Un banco de ítems es un archivo en el que las preguntas están clasificadas por temas y según sus propiedades psicométricas^{20,21}. En la literatura especializada se recomienda construir los exámenes con preguntas ya «probadas» y evaluadas^{14,15,19}. En nuestro medio no es muy frecuente este trabajo de mejorar y archivar preguntas, predomina más el criterio de que cada vez las preguntas deben ser totalmente nuevas.

En esta oportunidad, la escala de Galofré fue utilizada cuando el examen ya había sido aplicado pero es recomendable tener presente los diez criterios considerados en la escala en el momento de redactar y corregir preguntas.

Los datos que se obtienen aplicando la escala de Galofré, el índice de dificultad y el análisis cuali-cuantitativo son muy útiles en el proceso de redacción y corrección de las preguntas elaboradas por los distintos profesores, especialistas, jefes de servicio o jefes de residentes que son quienes habitualmente redactan preguntas para los exámenes de ingreso a la residencia.

En distintas oportunidades se ha observado el interés y motivación que despierta el recibir *feedback* sobre cómo funcionaron «mis preguntas»; el análisis de los resultados de un examen también es un recurso muy motivador para ser utilizado en el entrenamiento de docentes y médicos de planta involucrados en la generación de preguntas.

Al observar los resultados obtenidos por los médicos recién graduados en los exámenes de selección de residentes llama la atención la dispersión de los puntajes obtenidos. Todos los postulantes han recibido su título de médico, con un promedio general de carrera de 7 puntos y más pero un gran número de ellos pudieron responder correctamente apenas un poco más de la mitad de las preguntas (mediana del examen: 65 puntos) de un examen de dificultad media.

Las carreras de medicina deberían revisar sus sistemas de evaluación y afinar sus instrumentos de evaluación porque han examinado y aprobado alumnos que en los exámenes de ingreso a la residencia demostraron serias deficiencias de información. Habría que analizar la validez y confiabilidad de los exámenes instrumentados por los profesores universitarios y/o habría que plantear la necesidad de instalar, como tienen muchos otros países, una instancia de evaluación integral nacional para otorgar la licencia para el ejercicio profesional; una especie de control de calidad de los egresados.

Agradecimientos: A Daniel Silva por apoyo técnico en procesamiento de los datos.

REFERENCIAS BIBLIOGRÁFICAS

1. Panamá en <http://www.docencia-css.org.pa/PDF/reglaresidencia.pdf> Acceso 10/07/07.
2. Méjico en. <http://www.facmed.unam.mx/informe/2003/alumnos.html> Acceso 10/07/07.
3. España en <http://ortopedia.rediris.es/tribuna/art9.html> Acceso 10/07/07.
4. Ministerio de Educación. SPU. Anuario de Estadísticas Universitarias. 2005.
5. Borracci R, Arribalzagala E. Aplicación de análisis de conglomerados y redes neuronales artificiales para la clasificación y selección de candidatos a residencias médicas. *Educación Médica* 2005; 8 (1): 22-30.
6. Neuman M, Questa U, Kaufman R. Concurso de residencias médicas en la ciudad de Buenos Aires: importancia del género y universidad. Dirección de Capacitación y Desarrollo, Secretaría de Salud, Gobierno de la Ciudad de Buenos Aires. *Educación Médica* 2004; 7 (2): 90-96.
7. Alves de Lima A, Botto F, Azzari F, Galli A, Guzman L, Trivi M, Albertal J, Belardi, J. Exámenes de selección para postulantes a residencias en cardiología. Variables que se asocian con los resultados. *Rev Argent Cardiol* 2001; 69: 327-331.
8. Eymann A, Busaniche J, Carrió S, Durante E. Características de los aspirantes a la residencia de un hospital universitario: 2003-2006 CAEM 2006. Libro de resúmenes. pág. 163.
9. Rodríguez P, Listovsky G, Galli A, Lloveras S, Argente H. Rendimiento de postulantes e ingresantes al sistema de residencias médicas 2006. Libro de resúmenes. CAEM 2007 pág. 94.
10. Galli A, Habich D, Rotholz N. Resultados de la evaluación de aspirantes a residencias en un hospital de comunidad según sexo, universidad de egreso y especialidad elegida. CAEM 2006. Libro de resúmenes pág. 141.
11. Slobayen A, D'Angelo W, Markowsky C. Valor predictivo del promedio en ingresantes al sistema de residencias médicas. Libro de resúmenes CAEM 2007. pág. 90.
12. Scriven 1968 citado por Stufflebeam D, Shinkfield A. Evaluación sistemática Paidós. Buenos Aires 1987.
13. Galofré A. Construcción del índice de calidad de las preguntas. 2005. <http://www.saidem.org.ar/docs/mbts.php> Acceso 12 de julio 2007.
14. Case S, Swanson D. Nacional Board of Medical Examiners. Cómo elaborar preguntas para evaluaciones escritas en el área de ciencias básicas y clínicas. Tercera edición. Philadelphia. 2006.
15. Guilbert JJ. Guía pedagógica para el personal de salud. OPS/OMS. Ginebra. 1981.
16. Castro C, Galli A. Evaluación educacional. Módulos 1-2-3. Afacimera. Buenos Aires. 2002. <http://www.saidem.org.ar/docs/mbts.php> Acceso 15 de julio 2007.
17. Lafourcade P. Evaluación de los aprendizajes. Kapelusz. Buenos Aires. 1969.
18. Thorndike R, Hagen E. Medición y evaluación en Psicología y educación. Trillas. México. 1996.
19. Atkins Word D Elaboración de tests. Desarrollo e interpretación de los test de aprovechamiento. Trillas. México. 1969.
20. Barbero, M. I. Bancos de ítems. En J. Muñiz (ed.). *Psicometría*. Madrid. Universitat. 1996.
21. Martínez Cervantes R, Moreno Rodríguez R. Construcción de un banco de ítems informatizado para la evaluación de conocimientos sobre una materia universitaria. Publicado en: J. M. de Mesa, R. Castañeda y L. M. Villar (eds). Asegurar la calidad en las universidades (vol. II). Actas de las III Jornadas Andaluzas de Calidad en la Enseñanza Universitaria. Universidad de Sevilla, 2002. 329-344. <http://www.aloj.us.es/gmbytdlc/fevaluacionenfnt.pdf> Acceso 15/09/07.
22. Listovsky G, Galli A, Rodríguez P, Lloveras S, Giorgi MJ, Horszowski A. Postulantes a residencias médicas en la Ciudad de Buenos Aires. Libro de resúmenes. CAEM 2007 pág. 94.
23. Listovsky G, Lloveras S, Rodríguez P, Argente H, Giorgi J. Residencias y concurrencias médicas - Gobierno de la Ciudad de Buenos Aires: perfil de concursantes e ingresantes 2005. Libro de resúmenes. CAEM 2006 pág. 180.
24. Markowsky C, D'Angelo W, Slobayen. Ingreso a residencias de salud de mayores de 35 años. Libro de resúmenes. CAEM 2007 pág. 87.
25. Carestía G, Pasquare C, Carestía P. El sistema de selección de residentes de un hospital municipal descentralizado: 20 años de trayectoria. Libro de resúmenes. CAEM 2006 pág. 30.
26. Juárez Díaz González N, García Carrasco F. Taller: examen escrito. <http://www.certificacionespecialistas.org.mx/pdfs/11.pdf> Acceso 15/09/07.