



Calidad del examen estatal escrito en la especialidad de Medicina General Integral. 1ra convocatoria. Facultad de Ciencias Médicas de Holguín. Curso 2016-2017.

Quality of the state exam written in the specialty of Integral General Medicine. 1st call. Faculty of Medical Sciences of Holguín. Course 2016-2017.

Magalys Moreno Montañez,¹ Pedro A. Díaz Rojas,² Juana María González Marrero,³ Manuel Osorio Serrano.⁴

1 Especialista de II grado en Medicina General Integral. Auxiliar. Master en Educación Médica. Departamento de MGI. Universidad de Ciencias Médicas de Holguín. Holguín. Cuba.

2 Especialista de II Grado en Histología. Profesor Titular. Doctor en Ciencias en Educación Médica. Departamento de Ciencias Básicas. Universidad de Ciencias Médicas de Holguín. . Cuba.

3 Especialista de II grado en Medicina General Integral. Auxiliar. Master en Longevidad Satisfactoria. Departamento de MGI. Universidad de Ciencias Médicas de Holguín. Holguín. Cuba.

4 Especialista de II grado en Medicina General Integral. Auxiliar. Master en Atención Primaria de Salud y Educación Médica. Departamento de MGI. Universidad de Ciencias Médicas de Holguín. Holguín. Cuba.

Correspondencia: magalysm@infomed.sld.cu

RESUMEN

El presente estudio resulta un trabajo de investigación con el objetivo de caracterizar al examen estatal escrito en la especialidad de MGI, primera convocatoria, del curso 2016-2017 en la Facultad de Ciencias Médicas de Holguín. Se utilizó la metodología P Díaz para analizar la calidad del examen estatal que permitió su caracterización. Se realizó el análisis de los resultados de los indicadores de calidad, tales como; Alfa de Cronbach, índice de dificultad y coeficiente de correlación biserial. En el examen estatal escrito sobresale que es un examen mixto, con mayoría de preguntas test objetivo, predominio del tipo selección múltiple de verdadero o falso; su confiabilidad es aceptable aunque resaltaron las preguntas fáciles de carácter reproductivo. Solo dos preguntas discriminaron adecuadamente a los residentes de alto y bajo rendimiento por lo que la triangulación metodológica permitió inferir que ello estuvo relacionado de forma general con el diseño de las preguntas. Se hacen conclusiones.

Palabras Clave: instrumentos de evaluación, Medicina General Integral , indicadores de calidad.

ABSTRACT

The present research has as main objective to characterize the written state exam in the specialty of family medicine concerning it's first exam in the course 2016-2017 in the University of Medicine of Holguin province. The methodology used was that of Pedro Diaz to analyze the quality of the written exam state exam which let us characterize such exam .The analysis of quality indicators such as : Cronbach Alfa, biserial correlation coefficient and difficulty of degree. Concerning the written state exam it points out the fact that it comprises a mixture of set of questions being predominant test- objective questions as well as a selection of true or false multiple choice exercises where it's liability is important to be mentioned ,though easy questions of reproductive character were present , just two questions were able of characterizing in an adequate way the high and low resident performance in the exam in which the methodological triangulation gave us the possibility to infer that it was closely related to the design of the set of questions employed in the exam. Conclusions were made.

Key words: Evaluation instruments, Family medicine , quality indicators

INTRODUCCIÓN

La formación de profesionales con una alta preparación científica, cultural y humanista, constituye una prioridad para todos los Ministerios de Salud del mundo y de Cuba, desde esta perspectiva la Educación Médica tiene el reto de formar un profesional de perfil amplio con profundos conocimientos que den respuesta a los problemas de salud que se presenten una vez graduado.

Es imprescindible que a lo largo de todo su proceso formativo de grado y especialidad se les vaya verificando el progreso del aprendizaje para lograr que al concluir cada etapa y finalmente al egresar, estén preparado de forma adecuada para ejercer un desempeño laboral eficiente.¹

El sistema de formación del Especialista de Primer Grado en Medicina General Integral parte del continuo proceso de perfeccionamiento que caracteriza este régimen de estudio, que desarrolla una reingeniería del programa en el año 2005 a partir de los cambios que acontecen en el internado ,con el propósito de producir las transformaciones en los modos de actuación del Médico General para convertirlo en especialistas de perfil amplio, ajustado a las exigencias actuales y perspectivas del desarrollo económico y social de Cuba.²

El Programa de la Residencia consta de dos años. Cada año académico se planifica para 48 semanas, de ellas 45 se dedican a desarrollar el programa docente y tres a la preparación y realización de los exámenes de promoción. Tiene una estructura modular, con un total de 29 módulos, de los cuales 25 se desarrollan en la A. P. S. y cuatro en los hospitales, e incluye además tres cursos, tres rotaciones y una estancia en MNT.

La evaluación en el régimen de la residencia implica la medición y valoración de las transformaciones logradas en los conocimientos, hábitos, habilidades, conductas y valores del residente, definidas en el modelo del médico especialista³.

La evaluación educativa es un proceso que involucra la elaboración, aplicación y análisis de los instrumentos de medición. La función principal de un instrumento de medición, cuando se crea como medida para inferir las capacidades de las personas, es ofrecer información para la correcta toma de decisiones⁴.

A partir del curso 2011-2012 en la Universidad de Ciencias Médicas de Holguín se inician los primeros estudios sobre medición de la calidad en los exámenes certificativos de las diferentes disciplinas en la enseñanza del pregrado, no así en la formación del posgrado, particularmente en la especialidad de MGI, por lo que los instrumentos de evaluación son contruidos de forma empírica.

Por lo anterior expuesto, el problema científico que sustenta la presente investigación es que no existe una caracterización sobre bases científicas de la calidad del examen estatal escrito de la especialidad de MGI, 1ra convocatoria en la Facultad de Ciencias Médicas de Holguín durante el curso 2016-2017.

El presente trabajo tuvo como **objetivo**: Caracterizar la calidad del examen estatal escrito de la especialidad de MGI, 1ra convocatoria en la Facultad de Ciencias Médicas de Holguín durante el curso 2016-2017.

DISEÑO METODOLÓGICO

Para la caracterización del examen estatal escrito de la especialidad en Medicina General Integral, el universo estuvo definido por los 250 exámenes estatales escritos de los residentes de segundo año de la especialidad de MGI que se presentaron a la primera convocatoria.

Se utilizaron métodos teóricos y procedimientos estadísticos que garantizaron la triangulación metodológica efectuada.

Los **métodos teóricos**, garantizaron la revisión documental de investigaciones científicas en educación médica sobre calidad de instrumentos de evaluación, programa de la residencia

en MGI, documentos normativos del sistema de evaluación en el posgrado, informes y orientaciones metodológicas del área, entre otras fuentes.

Se utilizó la guía metodológica para el análisis de calidad de instrumentos de evaluación propuesta por P Díaz⁵, con el objetivo de caracterizar el examen estatal escrito; para ello se utilizaron las siguientes variables: Tipo de examen, cantidad de preguntas y temáticas que aborda, cantidad de ítem por preguntas y áreas temáticas, formato de preguntas, calificaciones por preguntas y del examen, correspondencia entre el fondo de tiempo asignado en la planificación docente y cantidad de ítem que se explora en el examen.

Se utilizaron **procedimientos estadísticos** como el análisis números enteros, frecuencia simple y la distribución porcentual para presentar los resultados en tablas. Para el análisis de los resultados de los indicadores de calidad se utilizó el coeficiente alfa de Cronbach,^{6,7} trabajado por González Sapsin y Ortíz Romero, proporciona una medida de la consistencia interna de los exámenes y puede calcularse aún en el caso de que se empleen preguntas de diferentes formatos. El Índice de Dificultad del examen, corresponde a la proporción de personas que responden correctamente una pregunta de la prueba. , mientras mayor sea esta proporción menor será el grado de dificultad. Para su cálculo se dividió el número de residentes que contestaron correctamente el ítem, entre el número total de examinados. Usualmente a esta proporción se le denota con la letra p, (dificultad del ítem).^{6,7-8}

Según autores consultados ⁸ los resultados de este indicador pueden oscilar en: 5 % para preguntas fáciles, 20 % para preguntas medianamente fáciles, 50 % para preguntas de dificultad media, 20 % para preguntas medianamente difíciles, 5 % difíciles.

Como el examen corresponde a una evaluación estatal cuyo por ciento de aprobado es 70 puntos, la autora consideró oportuno realizar un ajuste a la distribución de los valores de p, de manera tal que 70 % del contenido estuviese incluido en preguntas de dificultad media y difícil, quedando de la siguiente manera: 10 % para preguntas fáciles ($p > 0.86$), 20 % para preguntas medianamente fáciles (p entre 0.74 - 0.86), 40 % para preguntas de dificultad media (p entre 0.53 - 0.73), 20 % para preguntas medianamente difíciles (p entre 0.32 - 0.52), 10 % para preguntas difíciles ($p < 0.32$).⁹

El Coeficiente de Discriminación punto de correlación biserial se utiliza para saber si los estudiantes de más alto rendimiento obtienen las respuestas correctas.^{6, 7, 8, 9}

Según estándares internacionales⁶ la distribución de los resultados del coeficiente punto de correlación biserial es: $r_{pbis} < 0$ preguntas que discriminan negativamente, $r_{pbis} = 0 - 0.14$ preguntas que discriminan pobremente, $r_{pbis} = 0.15 - 0.25$ preguntas que discriminan regular, $r_{pbis} = 0.26 - 0.35$ preguntas con buen poder discriminativo, $r_{pbis} > 0.35$ preguntas con excelente poder de discriminación.

RESULTADOS

El examen estatal escrito se estructuró de 10 preguntas con 55 ítems, según el formato de ellas se clasificó como un examen mixto, las preguntas en su mayoría fueron del tipo test objetivo.

Tabla 1. Distribución de ítems según formato de las preguntas de la primera convocatoria del examen estatal escrito de la especialidad en MGI.

Preguntas	Cantidad de ítem	Ensayo R Cortas	Verdadero o falso	o Enlace	Selección Múltiple CS	Selección Múltiple CA
1	2	2				
2	8				8	
3	3	1				2
4	6			6		
5	9		9			
6	6			6		
7	10		10			
8	3		2		1	
9	2	1		1		
10	6	6				
Total	55	10	21	13	9	2

Fuente: Examen estatal

Tabla 2. Correspondencia entre fondo de tiempo y cantidad de ítem que se explora en el examen estatal escrito de la especialidad en MGI.

Preguntas	Módulos	Contenidos	Fondo de tiempo	%	Cantidad de ítem explorados	%
1,2,6	M 4.Salud Familiar	Familia, crecimiento y desarrollo, gestorragias de la 1ra y 2da mitad del embarazo	8h	25	1- 2 2- 8 6-6	36,36
3	M 14. Afecciones respiratorias y ORL	Tuberculosis Pulmonar	2h	6,25	3	6,82
4,10	M 21.Enfermedades infecciosas	Enfermedades infecciosas	8h	25	4-6 10-6	27,27
5	M3.Salud Comunitaria	ASIS	2h	6,25	9	13,64
7	M 24. Afecciones dermatológicas	Enfermedades dermatológicas	4h	12,5	10	20,45
8	M 18. Afecciones endocrino metabólicas	Diabetes Mellitus	2h	6,25	3	6,82
9	M 20. Afecciones neurológicas	Enfermedades cerebrovasculares	6h	18,75	4	9,09

Fuente: Examen estatal

Tabla 3. Índice de dificultad por preguntas en la primera convocatoria del examen estatal escrito de la especialidad en MGI.

Preguntas	Aprobados	Desaprobados	Dificultad	Tipo de dificultad
1	160	90	0.64	media
2	203	47	0.81	medianamente fácil
3	249	1	0.99	fácil
4	239	11	0.95	fácil
5	228	22	0.91	fácil
6	246	4	0.98	fácil
7	198	52	0.79	medianamente fácil
8	229	21	0.91	fácil
9	244	6	0.97	fácil
10	170	80	0.68	media
Examen	249	1	0.99	fácil

Fuente: Procesamiento estadístico

De forma general la dificultad del instrumento fue de 0.99, por lo que se considera un examen fácil.

Tabla 4. Coeficiente de discriminación, punto de correlación biserial en la primera convocatoria del examen estatal escrito de la especialidad en MGI.

Preguntas	Aprobados		Desaprobados		r_{pbis}	Discriminación
	n	x	n	x		
1	160	85.94	90	78.38	0.47	buena
2	203	84.92	47	76.58	0.42	regular
3	249	83.9	1	79	0.04	pobremente
4	239	84.3	11	74.63	0.26	pobremente
5	228	84.45	22	77.93	0.24	regular
6	246	84.02	4	75.12	0.14	pobremente
7	198	85.96	52	75.95	0.52	regular
8	229	84.3	21	79.28	0.18	regular
9	244	84.14	6	73.16	0.22	regular
10	170	86.24	80	78.86	0.44	buena

Fuente: Procesamiento estadístico

El coeficiente de discriminación punto de correlación biserial demostró que las preguntas 3, 4 y 6 son clasificadas como fáciles y discriminan pobremente (r_{pbis} 0 - 0.14). Las preguntas 2, 5, 7, 8 y 9 a pesar de ser fáciles discriminan de forma regular (r_{pbis} 0.15 - 0.25); las preguntas 1 y 10 con dificultad media tienen buen poder de discriminación .

Tabla 4. Dificultad por contenido en la primera convocatoria del examen estatal escrito de la especialidad en MGI.

Contenidos	Módulos	Apro- bados	Desa- probados	Dificultad	Tipo de dificultad
Familia	4	160	90	0.64	media
Crecimiento y desarrollo	4	203	47	0.81	medianament e fácil
Tuberculosis Pulmonar	14	249	1	0.99	fácil
Enfermedades infecciosas	21	239	11	0.95	fácil
ASIS	3	228	22	0.91	fácil
Gestorragias del embarazo	4	246	4	0.98	fácil
Afecciones dermatológicas	24	198	52	0.79	medianament e fácil
Diabetes Mellitus	18	229	21	0.91	fácil
Enfermedades cerebrovasculares	20	244	6	0.97	fácil
Dengue	21	170	80	0.68	media

Fuente: Procesamiento estadístico.

En los resultados del índice de dificultad por contenidos, las preguntas con mayor dificultad fueron la 1 con 0.64 % de aprobados y la 10 con 0.68% ambas con dificultad media.

Tabla 5. Coeficiente Alfa de Cronbach en la primera convocatoria del examen estatal escrito de la especialidad en MGI.

	1	2	3	4	5	6	7	8	9	10	
1	0.69 17										
2		0,69 63									
3			0.67 87								
4				0.71 51							
5					0.71 08						
6						0.70 59					
7							0.69 58				
8								0.68 52			
9									0.70 00		
10										0.708 3	
TOTAL											0,717 8

Fuente: Procesamiento estadístico.

El resultado del coeficiente α de Cronbach, por preguntas y total. El valor total 0.71 demuestra que el instrumento tiene una confiabilidad aceptable.

DISCUSIÓN

La metodología aplicada permitió caracterizar el examen estatal escrito. De los 55 ítems que fueron utilizados en el instrumento, existen diferentes formatos en la confección de las preguntas: ítem de respuesta alternativa verdadero o falso, selección múltiple complemento

simple, de enlace. Las preguntas de selección múltiple verdadero o falso fueron las más utilizadas con una frecuencia total de 21 ítems.

Las preguntas de verdadero - falso constituyen una variedad de pregunta de selección múltiple, con la diferencia de que las alternativas que se emplean son sólo dos, por lo que también se conocen como preguntas de respuesta alternativa. Están compuestas por un conjunto de afirmaciones de preparación relativamente fácil. Su utilización debe reducirse a proposiciones que sean inequívocamente verdaderas o demostrablemente falsas¹.

Basándose en la literatura consultada la autora asumió que se deben tener en cuenta a la hora de la confección de las preguntas de verdadero - falso los aspectos siguientes:

- Redactar la información de modo que tenga una idea solamente.
- Asegurarse de que los ítems estén escritos de manera que no haya tendencia a la equivocación.
- No introducir elementos falsos en una afirmación que es correcta.
- No utilizar palabras textuales de libros.
- Las afirmaciones verdaderas deben ser aproximadamente de igual número que las falsas, sin usar la alternancia regular.
- Eliminar adverbios como: "todo", "siempre", "ninguno", así como sinónimos que expresan generalizaciones tan amplias que tienen mucha posibilidad de ser falsas.
- Eliminar también el empleo de otros términos menos definidos como "usualmente", "a veces" y "alguna vez", que suelen formar parte de proposiciones verdaderas.
- Se debe emplear un lenguaje preciso, incluyendo términos numéricos cuando sea necesario.
- Evitar la utilización de proposiciones negativas, especialmente las de negación doble.

Las preguntas de verdadero - falso han sido estudiadas por varios pedagogos y en algunos casos se han realizado propuestas para mejorarlas por la susceptibilidad de su ambigüedad, los efectos del azar y que por tanto brinden mayor información acerca de los conocimientos del examinado. Las preguntas utilizadas fundamentalmente en este examen fueron reproductivas más que de pensamiento y análisis médico.

La correspondencia entre el fondo de tiempo planificado y su exploración en los ítems del examen escrito no tuvo un comportamiento adecuado. Las preguntas 1, 2, 6 correspondieron a contenidos del módulo 4 de Salud Familiar que se imparte durante un mes, distribuidas en dos horas semanales en diferentes formas organizativas durante la actividad docente presencial que representan un 25%, sin embargo se exploran en 16 ítem del examen para 36,3%. La pregunta 5 correspondiente al módulo de Salud Comunitaria está planificada para dos horas, 6,25% de la planificación, se exploró en 9 ítem del

instrumento para un 13,64%. La pregunta 7 correspondiente a contenidos del módulo 24 de Afecciones dermatológicas tiene 4 horas planificadas para 12,5% y se exploró en 10 ítem, 20,45%. La pregunta 9 exploró contenidos del módulo 20 de afecciones cerebrovasculares que se imparte según calendario docente en tres semanas, 18,75% sin embargo quedó por debajo su exploración en el examen escrito con solo cuatro ítem 9,09%. Estos resultados de forma general coinciden con los obtenidos por González Sapsin⁶, Ortiz Romero⁷, Quiñones Gutiérrez⁸ en sus trabajos de terminación de maestría en el área de formación del pregrado. La autora de este trabajo considera que la falta de correspondencia está dada por el formato de preguntas empleados y carácter reproductivo.

Coincide además con Díaz Rojas P⁵ que plantea debe encontrarse una proporción adecuada de ítems en función del tiempo que se le asignó al tratamiento del contenido.

En el índice de dificultad del instrumento analizado predominaron en más del 50% las preguntas fáciles lo que explica el alto grado de aprobados. Según Backhoff E.⁹ y otros autores que lo citan^{6,7-10} la dificultad del examen debe oscilar entre 0.5 y 0.6. De las 10 preguntas del examen solo la 1 y la 10 tuvieron una dificultad media. A juicio de la autora el predominio de las preguntas fáciles coincide con el carácter reproductivo del examen.

Las preguntas 3, 4 y 6 son clasificadas como fáciles y discriminan pobremente a los residentes de alto y bajo rendimiento (r_{pbis} 0 - 0.14). Las preguntas 2, 5, 7, 8 y 9 a pesar de ser fáciles discriminan de forma regular (r_{pbis} 0.15 - 0.25); las preguntas 1 y 10 con dificultad media tienen buen poder de discriminación. Las preguntas 3, 4 y 6 son clasificadas como fáciles y discriminan pobremente (r_{pbis} 0 - 0.14). La autora considera que las preguntas fáciles que no discriminan de manera adecuada deben ser rediseñadas con carácter productivo de manera tal que el residente pueda evidenciar los conocimientos adquiridos durante su proceso de aprendizaje.

Las preguntas 1 y 10 sobre las áreas temáticas de familia y dengue fueron las de mayor dificultad y las que discriminan adecuadamente a los estudiantes de mayor y menor rendimiento. La autora considera que éstas preguntas a pesar de ser las de mayor dificultad en el examen escrito, son contenidos que se imparten desde el segundo año hasta la culminación de la carrera, se profundiza y consolida en la formación posgradual, por lo que sus resultados pudieran estar relacionados con la dirección del proceso docente por el tutor en la educación en el trabajo, el aprovechamiento del estudio individual por parte del residente en formación.

Según la autora en la pregunta 1 sobre familia, tradicionalmente sus resultados en los exámenes estatales escritos es controvertida, relacionado al diseño de la pregunta, los contenidos que incorpora y la clave de calificación

El análisis de la consistencia interna del examen aplicado fue satisfactorio con un valor de 0,71, por lo que resulta confiable. Este resultado concuerda con el coeficiente que se propone en las Indicaciones Metodológicas para el Planeamiento, Diseño y Control de la Calidad de los Exámenes Escritos para la Evaluación de la Competencia y el Desempeño Profesional,¹² del MINSAP, los coeficientes de confiabilidad descritos por Anastasi y Weiner¹¹ citados por otros autores¹¹ y a los estándares propuestos por el Colegio de médicos de Canadá^{10,11} donde consideran aceptables coeficientes entre 0.69 y 0.71. Los resultados de los coeficientes por preguntas son estables, es decir no están por encima del coeficiente total del examen, por tanto no afectan este valor obtenido.

CONCLUSIONES

- Se caracterizó el examen estatal escrito de la especialidad en MGI en su primera convocatoria. Según el formato de las preguntas empleadas se clasificó como un examen mixto, con predominio de las preguntas de selección múltiple tipo verdadero o falso. No existió adecuada correspondencia entre el fondo de tiempo y la cantidad de ítem explorados en el examen.
- el análisis de los indicadores de calidad demostró que el examen tuvo una confiabilidad aceptable, predominio de preguntas fáciles y bajo poder discriminativo.
- Sistematizar con el claustro y directivos del proceso, la planeación de los instrumentos de evaluación escritos

REFERENCIAS BIBLIOGRÁFICAS

1. Salas Perea RS, Salas Mainegra A. Educación Médica Contemporánea. Retos, Procesos y Metodologías. Bucaramanga. Colombia: Universidad de Santander; 2014.
2. MINSAP. Instituto Superior de Ciencias Médicas de la Habana. Vicerrectoría de desarrollo. Programa de la residencia en Medicina General Integral. Reingeniería del plan de estudio. 2005
3. MINSAP. Reglamento del régimen de residencias. Capítulo VI De la evaluación. 2004
4. Salas Perea RS, Salas MI. Evaluación para el aprendizaje en ciencias de la salud. EDUMECENTRO. [Internet] 2017;9(1) 208-227. Disponible en: <http://www.revedumecentro.sld.cu/index.php/edumc/article/view/354/html>
5. Díaz Rojas PA, Leyva Sánchez E. Metodología para determinar la calidad de los instrumentos de evaluación. Educ Med Super [Internet]. 2013 [citado 10 Jun 2013];27(2):[aprox. 18 p.]. Disponible en: <http://scielo.sld.cu>

6. Glez Sapsin K. Calidad del examen final de la asignatura Morfofisiología Humana II. Curso 2009-2010. [Tesis para optar por el título de Master en Educación Médica. Holguín: UCM; 2012]
7. Ortiz Romero GM. Calidad del examen final de la asignatura Metodología de la Investigación y Estadística. Carrera de Medicina. Curso 2010-2011. (Tesis para optar por el título de Master en Educación Médica para optar por el título de Master en Educación Médica . Holguín: UCM; 2012.
8. Quiñones Gutiérrez D. Calidad de los exámenes estatales del curso 2011-2012, 2012-2013 en la carrera de licenciatura en Enfermería. [Tesis para optar por el título de Master en Educación Médica. Holguín: UCM; 2014]
9. Backhoff E, Larrazolo N, Rosas M. Nivel de dificultad y poder de discriminación del Examen de Habilidades y Conocimientos Básicos (EXHCOBA). Rev Electrón Invest Educ. 2000. [citado 2 oct 2013]; 2(1). Disponible en: <http://redie.uabc.mx/ojs/index.php/redie/article/view/15>
10. Moreno Montañez M. Caracterización del examen estatal escrito de la especialidad en Medicina General Integral. 1ra convocatoria. República Bolivariana de Venezuela. Mayo 2006. [Tesis para optar por el título de Máster en Educación Médica]. ENSAP; 2007.
11. Salas Perea R. Calidad de los instrumentos de evaluación. Módulo Evaluación Material de estudio Maestría en Educación Médica. En CD-ROM Maestría en Educación Médica. ENSAP: 2005.
12. Cuba. Ministerio de Salud Pública. Proyecto de Indicaciones Metodológicas para el Planeamiento, Diseño y Control de la Calidad de los Exámenes Escritos para
13. Evaluación de la Competencia y el Desempeño Profesionales. La Habana: Escuela Nacional de Salud Pública; 2001.