

Valoración metodológica de la confección de temarios de exámenes finales de Medicina y Estomatología

Methodological assessment of the preparation of lists of topics in final exams of medicine and dentistry studies

MsC. Margarita Rosa Roméu Escobar, Dr. Cs. José Aurelio Díaz Quiñones

Facultad de Ciencias Médicas "Dr. Raúl Dorticós Torrado". Cienfuegos, Cuba.

RESUMEN

Interesaba precisar cómo se cumplían las indicaciones y recomendaciones en la confección de temarios de exámenes finales. Se revisaron 99 temarios de los cursos 2012/2013 y 2013/ 2014, de la Facultad de Ciencias Médicas "Dr. Raúl Dorticós Torrado", Cienfuegos, con el objetivo de valorar aspectos metodológicos. El formato de pregunta más utilizado fue el ensayo de respuestas cortas. El 87,8 % tenía las raíces bien confeccionadas, el 60,6 % los ítems y el 61,6 % las escalas de calificación. Los errores más frecuentes fueron: tarea imprecisa en preguntas de ensayo modificado, preguntas tipo complemento simple con tres opciones de respuestas y escalas con aprobado en menos del 60 %. El 65 % de las preguntas eran reproductivas. En general, los temarios están confeccionados adecuadamente, pero se observan desaciertos en la elaboración de preguntas y escalas de calificación. El predominio de preguntas reproductivas limita la validez conceptual.

Palabras clave: evaluación del aprendizaje, exámenes test objetivos.

ABSTRACT

This article was aimed at finding out how the indications and recommendations for the preparation of the list of topics for the final exams. Ninety nine of them from 2012-2013 and 2013-2014 academic years in "Dr Raul Dorticos Torrado" medical sciences school, Cienfuegos province were revised in order to assess the methodological aspects. The most used format was the short-answer test. In the studied group, 87.9 % had well-constructed stems, 60.6 % had good items and 61.6 % showed adequate evaluation scales. The most frequent errors were imprecise tasks in modified essay questions, single complement-type questions and evaluation scales where the pass mark is lower than 60%. Sixty five percent of questions were "reproduction-type" questions. Generally speaking, the list of items was adequately constructed, but there were observed some mistakes in the construction of questions and in the evaluation scales. The predominance of reproduction-type questions restricts the conceptual validity of tests.

Keywords: learning evaluation, objective exams.

INTRODUCCIÓN

La evaluación del aprendizaje es, probablemente, el componente del proceso enseñanza-aprendizaje (PEA) que más debates origina y que más variaciones conceptuales ha experimentado. En varios documentos rectores están declaradas las normativas de los sistemas de evaluación de las asignaturas que componen los planes de estudio de la Educación Médica Superior en Cuba.¹⁻³

El examen escrito es la modalidad utilizada en la mayoría de los exámenes finales. Lo precede generalmente un examen práctico donde se evalúan habilidades específicas, técnicas y procedimientos inherentes a cada asignatura, según se trate de asignaturas de formación básica, básico-clínica o clínica.³

Los programas de las asignaturas declaran los objetivos instructivos a cumplir³ y durante años se ha sistematizado la orientación metodológica a los profesores sobre la confección de exámenes escritos, utilizando diferentes vías, por lo que los autores consideran que el claustro está apto para esta tarea.

Diversos métodos para evaluar calidad en la confección de temarios, su poder discriminativo y validez se han desarrollado en diferentes facultades del país.⁴⁻⁶

En esta ocasión se trata de una visión general de los aciertos y desaciertos, como experiencia compartida del claustro de profesores, por supuesto, con las limitaciones que impone el no poder analizar todos los componentes del proceso.

Por tanto, se planteó como problema de investigación el siguiente: no se conoce con precisión cómo se cumplen las indicaciones y recomendaciones para la elaboración de los temarios escritos de los exámenes finales en la Facultad de Ciencias Médicas de Cienfuegos.

Consecuentemente el objetivo ha sido valorar aspectos metodológicos de la confección de temarios de exámenes finales de las asignaturas de las carreras de Medicina y Estomatología en la facultad de Ciencias Médicas "Dr. Raúl Dorticós Torrado", de la provincia de Cienfuegos.

MÉTODOS

Se realizó una investigación descriptiva y transversal, cuyo universo estuvo formado por los temarios de exámenes finales ordinarios de las asignaturas propias de las Ciencias de la Salud, de ambas carreras, en los cursos 2012/2013 y 2013/2014. Se excluyeron las asignaturas que no tienen examen final escrito y el internado de Medicina.

Se pudieron incluir 50 asignaturas: 29 asignaturas de la carrera de Medicina y 21 de Estomatología que representaron 96,6 y 67,8 % respectivamente de las posibles a incluir según criterios antes declarados. Se dejaron de incluir 11 asignaturas porque no se pudieron localizar los temarios.

Los autores seleccionaron al azar 2 temarios por asignatura, aplicados en exámenes ordinarios en ambos cursos, excepto en una asignatura de Medicina en la que solo se tuvo acceso a un temario, por tanto, la muestra incluyó 99 temarios de exámenes finales: 57 correspondientes a la carrera de Medicina y 42 a Estomatología.

Se contrastaron los objetivos y habilidades declaradas en los programa de las asignaturas con los contenidos de los exámenes.

Se incluyeron los siguientes aspectos:

- Número de preguntas por temario.
- Tipos de preguntas según formato.
- Errores en la elaboración de las preguntas: raíces e ítems.
- Errores en la elaboración de claves de respuestas y en las escalas de calificación.
- Cumplimiento de los objetivos del programa y niveles de asimilación de las preguntas.

Incorrecto significa que al menos un elemento estaba errado en el temario.

RESULTADOS

De los 57 temarios de Medicina, 36 tenían 7 preguntas y 21 temarios 5 preguntas. De los 42 de Estomatología, 31 de 5 preguntas y 11 de 7 preguntas.

En ambas carreras el tipo de examen fue mixto: con preguntas de ensayo modificado y de test objetivo. El patrón más generalizado fue abordar un tema por pregunta.

De 327 preguntas revisadas en Medicina, se observó la siguiente distribución de los formatos de preguntas:

- Ensayo de respuesta corta (ERC) 38,6 %.
- Respuesta alternativa (VF) 19,3 %.
- Completar oraciones 17,5 %.
- Relacionar columnas 12,3 %.
- Complemento simple (CS) 7,6 %.
- Complemento agrupado (CA) 4,6 %.

De 232 preguntas revisadas en Estomatología se precisó la siguiente frecuencia:

- Ensayo de respuestas cortas (ERC) 24 %.
- Respuesta alternativa verdadero-falso (VF) 23 %.
- De relacionar columnas 20,3 %.
- Completar oraciones 11,7 %.
- Complemento agrupado 6,9 %.
- Complemento simple (CS) 3,9 %.

En ambas carreras en conjunto, de 559 preguntas, se destaca que las preguntas tipo ERC (37 %) fueron las más utilizadas, más en Medicina que en Estomatología y que las preguntas con formato CS fueron muy pocas (6 %). Las preguntas VF son habitualmente muy empleadas (21 %). Por otro lado, se incluyeron pocas del tipo complemento agrupado (5,5 %). El resto de los formatos fueron a discreción: 15,5 % tipo relacionar columna y 15 % tipo completar oraciones.

De los 57 temarios revisados en Medicina, en 49 (86%) las raíces de las preguntas estaban correctamente elaboradas. En 30 temarios (53 %) estaban correctos los ítems y formatos de las preguntas y en igual número de temarios las escalas de calificación eran adecuadas.

De los 42 temarios de la carrera de Estomatología, 38 (90,5 %) tenían las raíces de las preguntas correctas. En 30 temarios (71,4 %) estaban correctos los ítems. En 31 temarios (73,8 %) las escalas de calificación eran adecuadas.

En general, el 87,8 % de los temarios tenía las raíces bien confeccionadas, el 60,6 % los ítems y el 61,6 % las escalas de calificación.

Los desaciertos fundamentales fueron los siguientes:

- En las raíces de las preguntas se plantea la tarea de forma directa, sin previamente introducir el asunto o contenido de que se trata.
- En la elaboración de ítems:

Tareas ambiguas, particularmente en el formato de ERC no se precisa el alcance o el número de elementos a aportar en la respuesta cuando es necesario.

Formato de complemento simple con tres opciones de respuesta.

Preguntas de relacionar con distractores en ambas columnas o heterogeneidad de los elementos comparados o sin distracto con igual número de elementos en ambas columnas.

Completar donde es plausible varias respuestas o con múltiples espacios sucesivos o intermedios o falta de coherencia gramatical.

Modificaciones parciales de formatos o mezcla de ítems.

- En las escalas de calificación los sesgos más frecuentes fueron que se fija un punto de corte de la escala en menos del 60 % para el aprobado y escalas incompletas.

En cuanto al cumplimiento de objetivos, el desacuerdo fundamental fue que hubo en Medicina 194 preguntas reproductivas (59 %) y en Estomatología 173 (74 %). En general, de las 559 preguntas, 367 eran de tipo reproductivo que representaron el 65 %.

DISCUSIÓN

Predominaron los temarios que abordan un tema por pregunta, siendo escasos los que integran contenidos de varios temas en una pregunta o aplican el análisis progresivo, aún cuando sería posible siguiendo la lógica de la asignatura.

Los formatos utilizados en la elaboración de temarios de exámenes finales son de tipo mixto, con preguntas de ERC y de test objetivo, lo cual es muy favorable ya que las preguntas de ERC muestran las capacidades de los estudiantes para expresar de forma lógica ideas esenciales y además permite evaluar aspectos ortográficos y de redacción, de manera que se obtiene un balance adecuado con las de test objetivo. Pero, a veces fueron utilizadas para citar, enumerar o identificar pocos elementos asociados o no a situaciones-problemas, por lo cual no se aprovechan todas sus ventajas.

Los autores han observado que el formato VF es muy utilizado debido a que aparentemente es de fácil elaboración, sin embargo, puede tener poca validez si no se cumplen las recomendaciones para su elaboración.

De igual forma, los autores consideran que las preguntas con formato de CS permiten discriminar adecuadamente los saberes de los estudiantes y abarcar mayor cantidad de contenidos de uno o varios temas, por lo que es lamentable que sea limitado su uso.

Por otro lado, se reconoce la poca pertinencia del uso de pocas o ninguna preguntas tipo complemento agrupado en exámenes cortos, por su calificación extrema: 2 o 5. Por supuesto sería posible utilizarla si se elaboraran varios incisos tipo CA en la pregunta.

En la elaboración de las raíces de las preguntas el desacuerdo más importante fue omitir la introducción al asunto o temas que se incluirían en la pregunta. Los autores desean subrayar que la introducción orienta al estudiante sobre la tarea a cumplir y el temario gana en calidad y armonía, en contraposición a un temario donde solamente se reflejen órdenes a ejecutar en abstracto.

Acerca de los desaciertos más importantes por formato de pregunta es conveniente comentar lo siguiente:

Existen recomendaciones, ventajas y desventajas para el uso de preguntas test objetivos las cuales han sido socializadas desde hace varios años a través de diferentes vías y que se originaron de las pautas ofrecidas por el Dr. C. *Salas Perea*, durante el diplomado impartido en esta institución y de otras fuente^{7,8} que constituyen las referencias en las que se basan las opiniones y comentarios que se hacen a continuación, además de la experiencia profesional acumulada por los autores.

Es un requisito de las preguntas de ensayo que se explique exactamente el alcance, lo que se quiere que responda el estudiante tanto cualitativamente como en términos de cantidad de elementos que se solicitan, si el número no es estrictamente exacto por la propia naturaleza de lo que se pregunta. La ambigüedad de la tarea indicada no aumenta el rigor de la pregunta, pero si dificulta la respuesta.⁹

Existen estudios que demuestran que la capacidad para comprender enunciados de preguntas y problemas en los exámenes es un factor clave, por lo que elaborarlos con precisión es imprescindible.^{9,10}

Los autores consideran que el formato de complemento simple con tres opciones de respuestas por inciso resta validez a la pregunta al facilitar respuestas que pueden decidirse al azar sin que estén avaladas por real conocimiento.

Por otra parte, el formato de relacionar columnas con heterogeneidad de los elementos a comparar conduce al análisis compartimentado guiándose el estudiante por elementos gramaticales o aproximaciones lógicas. La ausencia de distractores resta rigor y los distractores en ambas columnas no aumentan el rigor intrínseco de la pregunta, si no que aportan confusión al estudiante.

En las preguntas para completar oraciones elaboradas con ambigüedad sería posible más de una respuesta lógica. Cuando en este tipo de pregunta el tema no está orientado en la raíz, las posibles respuestas lógicas pueden ser muchas más. A veces la defensa de los profesores es plantear «así se lo di en clase», pero no es convincente en tanto no se está evaluando memoria. El espacio debe destinarse para un elemento de alto valor científico, e idealmente que sea posible uno y nada más que uno. Varios espacios sucesivos confunden al estudiante y evidencian que quizás se debió adoptar otro formato.

Las mezclas de formatos e ítems no se aconsejan, aunque su uso está bien extendido. Quizás se redujeran si los objetivos de la pregunta estuvieran bien precisados por el profesor que la elabora.

En la Universidad de Villa Clara se realizó una investigación en la que se identificaron los siguientes desaciertos: instrumentos con falta de datos en el encabezamiento, predominio de preguntas de respuestas cortas o restringidas en la mayoría de los exámenes, ítems con errores de redacción y diseño, uso indiscriminado de términos como: menciona, diga, escasas preguntas de desarrollo que permitan evaluar la ortografía y redacción, instrumentos evaluativos sin adecuado balance de las preguntas de tipo desarrollo y test objetivos con sus respectivas variantes, a pesar de la preparación pedagógica de los profesores.⁵

Como se observa hay coincidencia en varios desaciertos.

En cuanto a las claves de repuestas, los desaciertos fundamentales se observaron en la determinación del punto de corte de las escalas numéricas que significa aprobado o 3 para la escala cualitativa. Este aspecto puede ser fácilmente consensuado por directivos y metodólogos.

Un aspecto positivo es haberse logrado claves de respuestas bien explícitas en la mayoría de las preguntas de ERC, pero se contrapone a lo impreciso de las tareas solicitadas a los estudiantes, como ya se expresó.

Las escalas habitualmente se establecen porque los profesores necesitan operacionalizar las calificaciones en las preguntas test objetivos -creadas para calificaciones cuantitativas- para poder decidir una calificación cualitativa, equilibrada y justa para cada pregunta. Las combinaciones de calificaciones de preguntas para obtener la nota general del temario sí está sustentada en documento normativo. Estos aspectos deberían analizarse en el futuro pues son fuentes de sesgos.

El aspecto más relevante en cuanto a validez conceptual fue la desproporción de preguntas de reproducción, por lo que se puede afirmar que algunos objetivos no son evaluados con el nivel de asimilación de los contenidos que se declaran en los programas.

La reproducción del conocimiento siempre está presente en la base del aprendizaje, de hecho se observaron programas que lo declaran para algún objetivo, pero en exámenes finales de asignaturas cuyos objetivos tributan directamente al perfil profesional del egresado, es inexplicable que aún predominen preguntas cuyas tareas se planteen a nivel reproductivo, varias de ellas francamente memorísticas y otras con un nivel de generalización demasiado alto, según criterio de los autores, ya que deben tratarse con mayor concreción. También fue notable la escasa cantidad de preguntas donde se solicite argumentar, fundamentar o simplemente explicar y otras habilidades lógico-formales con las que los estudiantes deben operar en diversas áreas de conocimientos.

En las ciencias médicas el objeto de estudio es el proceso salud/enfermedad y su atención en el ser humano, por tanto, el objeto es un sujeto a escala individual —clínica— o grupos de sujetos —salud pública— con toda sus complejidades, por ello, el proceso enseñanza aprendizaje y la evaluación del aprendizaje deben acercarse a la realidad misma de la práctica profesional, según los niveles de sistematicidad, asimilación y complejidad declarados.^{11,12}

"En la medida en que las actividades de evaluación y el diseño del instrumento se aproximan a la realidad, o sea, están dirigidas a valorar o medir la aplicación social e independiente del sistema de conocimientos, habilidades y modos de actuación del estudiante, mayor será su validez...".⁶

En muy pocas asignaturas se observó que se declararan objetivos por pregunta en la clave y los autores infieren que no es habitual trabajar con tablas de especificaciones. Este es un ejercicio muy útil para evaluar contenidos con niveles de asimilación y complejidad adecuados. Por supuesto, es imprescindible haberlos tratado a similares niveles durante el PEA.

La calidad en el diseño y confección de temarios de exámenes se gana en el colectivo de asignatura. El diseño y elaboración de un temario de examen final requiere dominio de los contenidos, conocimiento del programa de la asignatura, suficiente dedicación de tiempo, consenso entre los profesores que desarrollan el proceso y comprensión del alcance, recomendaciones, ventajas y desventajas de los diferentes formatos de preguntas.

En el colectivo de asignatura hay que definir *qué evaluar*: contenidos, objetivos con niveles de asimilación, complejidad e integración definidos; *cómo evaluar*: cuáles técnicas o procedimientos emplear, tipo de examen, formatos y números de preguntas; *cómo calificar*: elaborar claves precisas y completas y establecer las pautas y escalas de calificación. Deben lograr la integración de los contenidos siguiendo la lógica de la profesión y el enfoque de sistema.^{13,14}

Por último, y aunque no son objetos ni resultados de esta investigación, convendría mencionar dos modificaciones necesarias: revalorizar el papel activo del estudiante en el proceso evaluativo incluyendo sus opiniones sobre instrumentos de evaluación y reconocer que más tarde o más temprano será necesario el rediseño curricular para las evaluaciones por competencias en Educación Médica Superior.¹⁵⁻¹⁷

Los autores consideran que esta investigación tiene el mérito de alertar sobre sesgos que probablemente no han sido identificados de forma consciente por los colectivos de asignatura.

En términos generales, la mayoría de los temarios de exámenes están confeccionados adecuadamente, pero se observaron desaciertos en la elaboración de preguntas y escalas de calificación.

El predominio de preguntas reproductivas limita la validez conceptual de los exámenes finales.

REFERENCIAS BILIOGRÁFICAS

1. Ministerio de Educación Superior. Reglamento de Trabajo Docente Metodológico. Resolución 210/2007. La Habana: Gaceta Oficial de la República de Cuba; 2007.
2. Ministerio de Salud Pública. Reglamento para la organización del proceso docente educativo en los centros de Enseñanza Médica Superior. La Habana: Editorial Ciencias Médicas; 2009.
3. Instituto Superior de Ciencias Médicas. Plan de estudios de la carrera de Medicina La Habana: Editorial Ciencias Médicas; 2009.
4. Carrazana LA, Salas Perea RS, Ruiz Salvador AK. Nivel de dificultad y poder de discriminación del examen diagnóstico de la asignatura Morfofisiología Humana I. Educ Med Super [Internet]. 2011 [citado 07 Jul 2014];25(1):103-14. Disponible en: http://scielo.sld.cu/scielo.php?script=sci_arttext&pid=S0864-21412011000100010&lng=es

5. López Espinosa GJ, Quintana Mujica R, Rodríguez Cruz O, Gómez López L, Pérez De Armas A, Aparicio Manresa G. El profesor principal y su preparación para diseñar instrumentos de evaluación escritos. *Rev Edu Med Centro* [Internet] 2014 [citado 07 Jul 2014];6(2):9 p. Disponible en:
6. Díaz Rojas PA, Leyva Sánchez E. Metodología para determinar la calidad de los instrumentos. *de evaluación Educ Med Super* [Internet]. 2013 [citado 17 Jul 2014];27(2):269-86. Disponible en:
http://scielo.sld.cu/scielo.php?script=sci_arttext&pid=S0864-21412013000200014&lng=es
7. Cortada N. Manual para la construcción de test objetivos de rendimiento. Buenos Aires: Paidós; 1968.
8. Moreno R, Martínez RJ, Muniz J. Directrices para la construcción de ítems de elección múltiple. *Psicothema* 2004;16(3):490-7.
9. Ayala-Valenzuela R, Messing-Grube H. Comprender los enunciados en un examen escrito: ¿dónde está el problema? *Educ Med Super* [Internet]. 2013 [citado 17 Jun 2014];27(2):211-9:8p. Disponible en:
http://scielo.sld.cu/scielo.php?script=sci_arttext&pid=S0864-21412013000200008&lng=es
10. López López M. ¿Sabes enseñar a describir, definir y argumentar? La Habana: Editorial Pueblo y Educación; 1989.
11. Fernández Sacasas JA. Educación médica: el objeto, el sujeto desdoblado y el contexto. *Educ Med Super* [Internet]. 2012 [citado 12 Jul 2013];26(3):361-3. Disponible en: http://scielo.sld.cu/scielo.php?script=sci_arttext&pid=S0864-21412012000300001&lng=es
12. Espinosa Brito A. La clínica y la Medicina Interna. *Rev cubana med* [revista en la Internet]. 2013 [citado 02 Ene 2014];52(3):214-27. Disponible en:
http://scielo.sld.cu/scielo.php?script=sci_arttext&pid=S0034-75232013000300008&lng=es.1
13. Díaz Velis Martínez E, Ramos Ramírez R, Wong Orfilalí T. Reflexiones acerca de la labor metodológica de la asignatura en el currículo de Medicina. *Rev Edu Med Centro* [Internet]. 2010 [citado 07 Jul 2014];2(2):10. Disponible en:
<http://www.revedumecentro.sld.cu/index.php/edumc/article/view/76>
14. Díaz Quiñones JA. ¿Se corresponde el proceso evaluativo actual con el proceso de enseñanza aprendizaje asumido? *Rev Gaceta Médica Espirituana* [Internet] 2012 [citado 07 Jul 2014];14(1):8. Disponible en:
<http://bvs.sld.cu/revistas/gme/pub/sum.14.%281%29/sumario.html>
15. Gual A. La universidad que queremos (II). En búsqueda de una mejora contrastada. *Educ Med* [Internet]. 2012 [citado 12 Jun 2013];15(1):1-3 Disponible en: <http://scielo.isciii.es/pdf/edu/v15n1/editorial.pdf>

16. Gil-Flores J. La evaluación del aprendizaje en la universidad según la experiencia de los estudiantes. Estudios sobre educación [revista en la Internet]. 2012 [citado 12 Jun 2013];(22):133-53. Disponible en: <http://dspace.unav.es/dspace/bitstream/10171/22636/2/ESE22-07-Gil-Flores.pdf>
17. Ion G, Cano E. La formación del profesorado universitario para la implementación de la evaluación por competencias. Educación XXI. 2012; 15(2):249-70.

Recibido: 13 de octubre de 2014.

Aprobado: 21 de octubre de 2014.

Margarita Rosa Roméu Escobar. Facultad de Ciencias Médicas "Dr. Raúl Dorticós Torrado". Cienfuegos, Cuba. Correo electrónico: salud5005@ucm.cfg.sld.cu