



# Revista Médica de Trujillo

Publicación oficial de la Facultad de Medicina de la Universidad Nacional de Trujillo - Perú

## Artículo Original

### Eficiencia de preguntas de opción múltiple con 3 alternativas

#### Efficiency of three-option multiple choice question

Luis Arturo Camacho-Saavedra<sup>1,2,a,b</sup>, Juan Jorge Huamán-Saavedra<sup>3,c,d</sup>, Jorge Omar Plasencia-Alvarez<sup>4,a,e</sup>.

1 Profesor Asociado Sección Farmacología. Departamento de Ciencias Básicas Médicas. Facultad de Medicina. Universidad Nacional de Trujillo.  
2 Médico especialista en Medicina Interna. Hospital II Florencia de Mora – Essalud 3 Profesor Principal Sección Bioquímica. Departamento de Ciencia Básicas Médicas. 4 Profesor Auxiliar Sección Bioquímica. Departamento de Ciencia Básicas Médicas. Facultad de Medicina. Universidad Nacional de Trujillo.  
a Magister b Orcid: <https://orcid.org/0000-0002-9680-2010> c Doctor. d Orcid: <https://orcid.org/0000-0003-3842-555X> e Orcid: <https://orcid.org/0000-0003-3979-8382>

#### Correspondencia.

Luis Arturo Camacho Saavedra.  
Correo: [lcamacho@unitru.edu.pe](mailto:lcamacho@unitru.edu.pe)  
Teléfono: 051-949937353.  
Dirección: Av. Los Angeles 242-  
AIOB Víctor Larco

Recibido: 12/10/20

Aceptado: 20/11/20

#### RESUMEN

**Introducción:** En evaluación se ha venido utilizando las preguntas de opción múltiple (POM) con 5 alternativas. Sin embargo, se ha encontrado que con 5 alternativas, más de 2/3 de ítems tenían 2 o más distractores no funcionales (DNF), con un promedio de 2,12 DNF; por ello se sugiere que se podrían utilizar menos opciones (3 o 4) sin afectar la calidad de las preguntas.

**Objetivo:** Determinar la eficiencia de las POM con 3 opciones en pruebas de bioquímica médica.

**Métodos:** Experimental, en pruebas de 260 examinandos de bioquímica médica divididos en dos grupos: A con 75 ítems de 5 opciones y B con 63 ítems de 3 opciones. Se comparó el promedio de notas, número de aprobados, cantidad de DNF, índices de dificultad y de discriminación.

**Resultados:** 9600 respuestas en el grupo A y 8316 respuestas en el grupo B. El promedio de notas fue de 12,33 ± 2,8 en A y 13,73 ± 2,5 en B.

La proporción de alumnos aprobados es de 83,59% en el grupo A y 92,42% en el grupo B.

El promedio de DNF fue de 1,71 en A y 0,68 en B ( $p < 0,002$ ).

El 42,67 % de las preguntas tiene índice de dificultad entre 33 y 66 en A y 33,33 % en B (NS).

El índice de discriminación (ID)  $> 0,20$  se halló en 66,7% en A y en 54% en B (NS).

**Conclusión:** La eficiencia de las preguntas con 3 opciones es similar a las preguntas con 5 opciones

**PALABRAS CLAVES:** Preguntas de examen. Educación médica. Evaluación educacional. Preguntas de opción múltiple.

#### SUMMARY

**Introduction:** it is of current use the five-option multiple choice questions (MCQ) as a part of evaluation. However it was found that with five options there were more than 2/3 of items with 2 or more non-functional distractor (NFD), with an average of 2,12 NFD; thus it has been suggested that 3 or 4 options could be used without impacting the quality of questions.

**Objective:** to determine the efficiency of three option multiple choice question in medical biochemistry course.

**Methods:** Experimental, in 260 tests of examinee on medical biochemistry divided in two groups: A with 75 items 5 options y B with 63 items of three options. It was compared the average exam scores, number of approved students, quantity of NFD, difficulty index and discrimination index.

**Results:** 9600 answers in A group and 8316 answers in B group. The average exam score was 12,33 ± 2,8 in A and 13,73 ± 2,5 in B.

The proportion of approved students was 83,59% in A group and 92,42% in B group.

The average NFD was 1,71 in A and 0,68 in B ( $p < 0,002$ ).

The difficulty index at 33-66 range was found in 42,67 % of questions in A and 33,33 % in B (NS).

The discrimination index at  $> 0,2$  range was found in 66,7% of questions in A and 54% in B (NS).

**Conclusion:** the efficiency of three option multiple choice questions was similar to five options questions

**KEY WORDS:** Examination Questions. Medical education. Educational assessment. Multiple-choice item.

## INTRODUCCION

La evaluación constituye una parte importante del proceso educativo; con ella, utilizando criterios previamente establecidos y mutuamente convenidos con los estudiantes, se obtienen evidencias necesarias para juzgar, retroalimentar y calificar el aprendizaje de cada estudiante. Mientras más oportunidades de evaluación existan, con su correspondiente retroalimentación, más se estará utilizando la evaluación como estrategia de enseñanza<sup>1</sup>. Es caracterizada por tres constantes, a saber, obtención de la información, formulación de juicios y toma de decisiones<sup>2</sup>.

Un objetivo importante de las pruebas o exámenes es comunicar qué es lo que se considera importante, así, cuando se elabora preguntas debe decidirse primero qué incluir en una prueba. Es necesario elegir una muestra de tópicos; no se puede preguntar todo. La naturaleza de la muestra determina hasta qué punto la estimación de la habilidad real es reproducible (confiable, generalizable) y exacta (válida). Si la muestra no representa adecuadamente el dominio más amplio que nos interesa (por ejemplo, si se incluye solamente contenido cardiovascular en una prueba de medicina general), los resultados de la prueba serán sesgados y no permitirán estimar el desempeño en el dominio que nos interesa. Si la muestra es muy pequeña, los resultados del examen podrán no tener la confiabilidad necesaria (reproducibles, confiables) para asegurar que reflejen la habilidad real del examinando<sup>3</sup>.

En 1990 Miller describió mediante una pirámide cuatro niveles de evaluación: en la base se sitúan los conocimientos, el segundo nivel corresponde a la competencia, el tercer escalón corresponde a la actuación y, en el vértice, la práctica que evalúa lo que el médico realiza con sus propios pacientes y en situaciones del todo reales. En la actualidad se usan los métodos indirectos que se sitúan en la base de la pirámide e incluyen a los exámenes escritos, orales estructurados, pacientes estandarizados, simulaciones por ordenador y el Examen Clínico Objetivo Estructurado (ECO). Los métodos directos se acercan al vértice de la pirámide e incluyen auditoría de historias clínicas, videograbaciones, indicadores de la práctica clínica<sup>4,5</sup>.

En la Facultad de Medicina de la Universidad Nacional de Trujillo (FMUNT) se lleva el curso de

Bioquímica Médica, en el Tercer Año, pertenece al Departamento de Ciencias Básicas y se ubica en el primer semestre. La asignatura se divide en cinco unidades y entre los parámetros de evaluación se incluye las pruebas con preguntas de opción múltiple al final de cada unidad que constan de 30 preguntas con 5 opciones<sup>6</sup>. Sin embargo, Jurado-Núñez y Col, encontraron que en preguntas de 5 alternativas, más de 2/3 del total de ítems tenían 2 o más distractores no funcionales (DNF), con un promedio de 2,12 DNF<sup>7</sup>. Se ha sugerido que se podrían utilizar menos opciones (3 o 4) sin afectar la calidad de las preguntas, así, Rodríguez en un metaanálisis de estudios de 80 años concluyó que el número óptimo de opciones parece ser 3, basado en criterios de dificultad, discriminación, confiabilidad y evidencia de validez<sup>8</sup>.

En la FMUNT se utiliza las preguntas de opción múltiple con 05 alternativas como parte de la evaluación del estudiante, similar a lo utilizado por la asociación peruana de facultades de medicina (ASPEFAM) en el examen nacional de medicina (ENAM) desde 2003<sup>9</sup>.

Por lo antes descrito, el objetivo de la presente investigación es determinar la eficiencia de las preguntas de opción múltiple con 3 frente a 5 opciones en exámenes de Bioquímica Médica de la Universidad Nacional de Trujillo, en base a la comparación de promedios de notas, proporción de aprobados, DNFs, índices de dificultad e índices de discriminación.

## MÉTODOS

El presente estudio se desarrolló en la Facultad de medicina de la Universidad Nacional de Trujillo, se elaboró inicialmente la prueba para el grupo A con preguntas de 5 opciones, correspondiente a las unidades 2, 4 y 5 del curso de Bioquímica Año 2019, luego se suprimieron dos opciones según criterio de experto (02 profesores) considerando a priori la menor factibilidad de dichas opciones, quedando la prueba B con tres alternativas. Se aplicó el examen de 30 preguntas en las 3 unidades distribuyéndose aleatoriamente en dos grupos de 50 y 47 alumnos cada uno.

Luego se construyó una base de datos en Excel con las opciones elegidas por cada alumno, a partir de la cual se determinó la proporción de ítems contestados, el promedio de notas y proporción de aprobados; se determinó los distractores no

funcionales, usando para el presente estudio, el criterio del “distractor usado por menos del 5% de evaluados”; se determinó el grupo fuerte (27% de alumnos con notas más altas) y el grupo débil (27% de alumnos con notas más bajas), a partir de los cuales se determinó los índices de dificultad y de discriminación. Luego se clasificó los ítems según criterios usados por ENAM, tomando como óptimos el índice de dificultad situado entre 33 y 66 y el índice de discriminación mayor a 0,20<sup>10</sup>.

## ANÁLISIS ESTADÍSTICO.

Los datos obtenidos se presentan en cuadros y gráficas y el análisis se realiza en base a proporciones, promedios y desviaciones estándar, para comparar los Distractores no funcionales, índices de dificultad y discriminación se usó la prueba t de Student para establecer diferencias.

## DEFINICIONES OPERACIONALES:

**Pregunta de selección múltiple:** O Item, Consiste en preguntas (tipo test), eligiendo la respuesta correcta de entre las cinco opciones propuestas<sup>11</sup>.

**Opción, alternativa o Distractor:** es cada una de las respuestas alternativas que se propone para que el alumno escoja<sup>11</sup>.

**Eficiencia:** nivel óptimo de funcionalidad y congruencia para posibilitar la concreción de metas con calidad<sup>12</sup>.

**Distractor no funcional (DNF):** es aquel distractor que fue elegido por menos del 5% de evaluados<sup>7</sup>.

**Índice de dificultad:** Proporción de alumnos que responden correctamente el ítem. Permite determinar en qué medida una pregunta es fácil o difícil. En realidad es un índice de facilidad, puesto que cuanto más alto es, más fácil es la pregunta. Este índice va de 0 a 100, siendo 100 muy fácil y 0 muy difícil. El índice de dificultad recomendado se sitúa entre 33 y 66 (10,13). Para obtenerlo se utilizó la fórmula 1:

**Índice de discriminación:** Este índice determina la selectividad de la pregunta para distinguir el grupo fuerte del débil. Este índice va de -1 a +1, considerándose un índice de discriminación bueno a excelente el mayor a 0,20 (10,13). Para obtenerlo se utilizó la fórmula 2:

### Fórmula 1

$$IDif = \frac{\text{Número de aciertos Grupo Fuerte} + \text{Número de aciertos Grupo Débil}}{\text{Número tota de estudiantes Grupo Fuerte} + \text{Grupo Débil}}$$

### Fórmula 2

$$IDis = \frac{\text{Número de aciertos Grupo Fuerte} - \text{Número de aciertos Grupo Débil}}{(\text{Número tota de estudiantes Grupo Fuerte} + \text{Grupo Débil}) \times 0,5}$$

Índice de Discriminación	Valoración ASPEFAM
Negativo	Malo
0,0 - 0,1	Deficiente
0,11 - 0,2	Aceptable
0,21 - 0,40	Bueno
0,41 - 0,68	Excelente

## CONSIDERACIONES ÉTICAS.

La confidencialidad de la información obtenida se mantuvo de la siguiente manera: La data se tomó de las hojas de respuesta de los exámenes y registradas en el instrumento de Excel sin nombres ni códigos de los alumnos, solo número de orden asignado; por lo tanto no existen sujetos vulnerables que se vean afectados por el procesamiento de la información.

## RESULTADOS

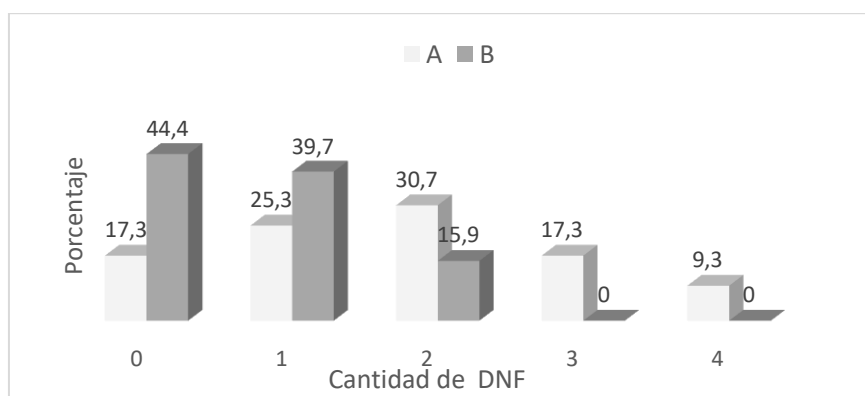
La muestra de estudio estuvo dada por los distractores elegidos en todos los ítems respondidos por los alumnos de medicina del tercer año en su tabla de respuestas, llegando a un total de 9825 respuestas en el grupo A y 8316 respuestas en el grupo B, debido a que se incluyeron 75 preguntas de 5 opciones, contando con las respuestas de 131 alumnos en el grupo A y 63 preguntas de 3 opciones de 132 alumnos en el grupo B, después de haber descartado hojas de respuesta incompletas o con marcado dudoso. El enunciado fue el mismo en las preguntas de las pruebas A y B.

El promedio de notas fue de  $12,33 \pm 2,8$  en el grupo A y  $13,73 \pm 2,5$  en el grupo B. La proporción de alumnos aprobados es de 81,68% en el grupo A y 92,42% en el grupo B. Se encontró 17,33% de distractores funcionales en el grupo A y 44,44% en el grupo B ( $p < 0,02$ ). La cantidad total de distractores no funcionales (DNFs) fue 131 en grupo A y 45 en grupo B; el promedio de DNFs fue de 1,71 en A y 0,68 en B ( $p < 0,002$ ). Se halló hasta 4 distractores no funcionales en el grupo A (Figura N° 1).

El 42,7 % de las preguntas tiene índice de dificultad entre 33 y 66 en el grupo A y 33,33 % en el grupo B. Se verificó por medio de la prueba de Kolmogorov-Smirnov que tiene distribución normal, luego se aplicó la prueba t de Student, resultando que no hay diferencia (Figura N° 2).

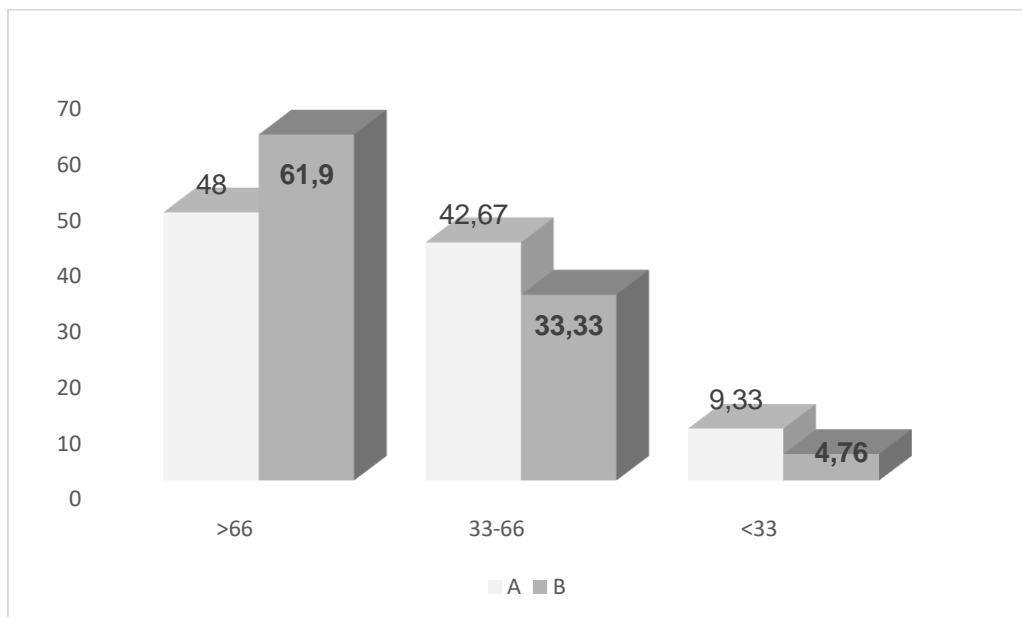
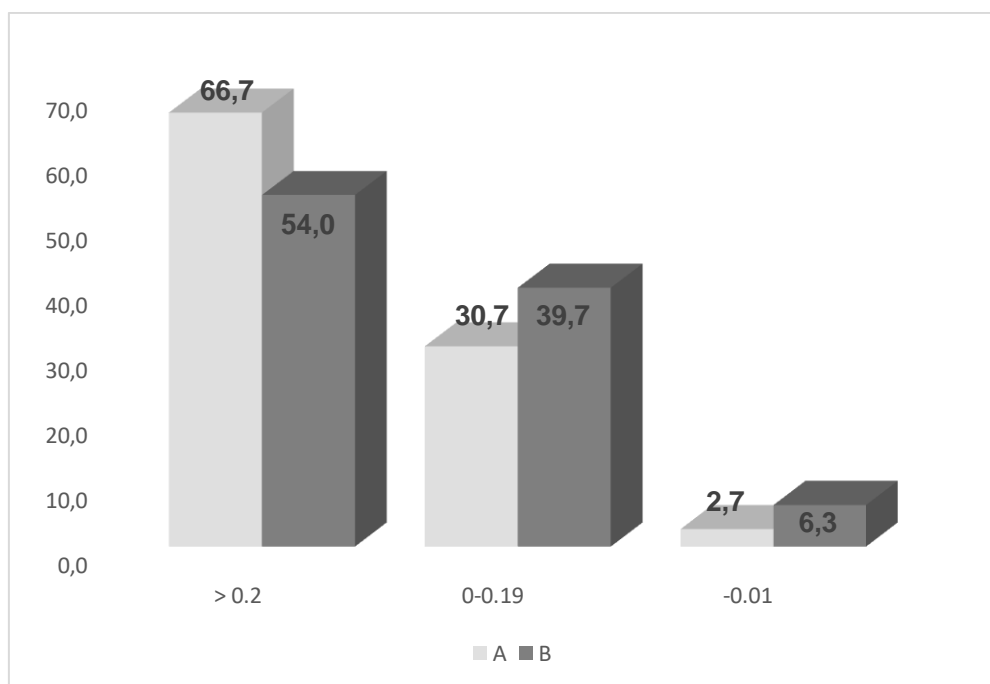
El índice de discriminación (ID) calificado como bueno y excelente ( $> 0,20$ ) se halló en 66,7% en el grupo A y en 54% de los ítems del grupo B. Se comprobó que tiene distribución normal por medio de la prueba de Kolmogorov-Smirnov, luego se aplicó la prueba t de Student, resultando que no hay diferencia (Figura N° 3).

**Gráfico 1. Comparación de cantidad de distractores no funcionales (DNF) por ítem según porcentaje.**



**Leyenda: A:** con 5 opciones

**B:** con 3 opciones

**Gráfico 2. Comparación de ítems con 5 y 3 opciones según Índice de dificultad**N.S.  $p > 0.13$ **Gráfico 3. Comparación de ítems con 5 y 3 opciones según Índice de discriminación**N.S.  $p > 0.24$ 

## DISCUSIÓN

El promedio de notas no se modifica significativamente al reducir el número de opciones. La proporción de aprobados se incrementa de 83.6%

a 92.4% cuando se reduce la cantidad de alternativas de 5 a 3.

La cantidad de distractores no funcionales se reduce a menos de la tercera parte cuando se reduce a 3 opciones, observándose un aumento significativo de

distractores funcionales. En general, el promedio de DNFs se reduce muy significativamente con 3 opciones, lo cual conlleva mejor calidad de la pregunta.

Se ha sugerido que la importancia de contar con el menor número de DNFs estriba en que mejora el índice de discriminación y mantiene en rango aceptable el índice de dificultad <sup>7</sup>. Si bien es cierto que las guías para elaborar preguntas, hechas por expertos internacionales, recomiendan la creación de tantos distractores como sea posible, también recomiendan establecer 3 opciones como base, pudiéndose crear más opciones pero con calidad <sup>7, 14-16</sup>.

El índice de dificultad en rango óptimo (de 33 a 66), si bien baja de 43 a 33% al reducir a 3 opciones, esta diferencia no es significativa. El porcentaje de preguntas fáciles aumenta de manera no significativa y el de preguntas difíciles se reduce de manera no significativa.

El índice de discriminación bueno a excelente (> 0.2) disminuye con 3 opciones pero de manera no significativa. Se encontró en total 6 ítems con índice de discriminación negativa en ambos grupos, los cuales se mantienen en porcentajes bajos. Del mismo modo se encontró 3 y 5 ítems respectivamente con índice de discriminación muy alto (>0.6), el cual, si bien puede ser considerado como DNF según Jurado-Nuñez y colaboradores, no fue usado como criterio para el presente estudio <sup>7</sup>. En un estudio inicial, evaluando la 1ra unidad del curso en 79 exámenes se halló resultados similares en cuanto a la no significancia en estos índices <sup>17</sup>.

Se destaca el hecho de que la supresión de opciones para pasar de 5 a 3 fue realizada con criterio de experto por dos docentes de la asignatura, considerando a priori la menor plausibilidad de dichas alternativas, ya que las preguntas se ejecutaban por primera vez; esto generó la incertidumbre de si la supresión impactaría en la calidad de las preguntas, encontrándose finalmente que no hubo diferencias en los índices de dificultad y discriminación; al respecto, Rodríguez encontró en una revisión de 49 estudios, que no hubo diferencia en el método de supresión de opciones, después de analizar 4 métodos: supresión aleatoria, supresión de distractores inefectivos, supresión de distractores más efectivos y adición de distractores y ninguno resultó en modificación importante de los resultados <sup>8</sup>. Un estudio realizado en exámenes

de enfermería utilizó datos de análisis de ítems para eliminar los distractores no funcionales en ítems de 4 distractores y dejarlo con 3, resultando en cambios mínimos en los índices de dificultad y discriminación <sup>18</sup>.

Una ventaja adicional que se ha descrito para preguntas con 3 opciones es que permite ahorrar tiempo en su elaboración para cubrir mayor contenido, incrementando la validez y fiabilidad, pudiendo ser más fáciles para docentes inexpertos, quienes estarían propensos a elaborar distractores no plausibles <sup>19</sup>.

Se espera contribuir con esta investigación correspondiente al análisis psicométrico de reactivos de opción múltiple, agregando una pequeña evidencia de que este procedimiento existe y es necesario para ejecutar una adecuada evaluación sumativa del aprendizaje, tal como lo señala Sánchez-Mendiola en su editorial de la revista de investigación en educación médica <sup>20</sup>. Es pertinente mencionar finalmente, que para el año 2020 la asociación de facultades de Medicina del Perú (ASPEFAM), señala en las bases del examen nacional de medicina (ENAM) que será con 4 opciones <sup>21</sup>.

### Limitaciones del estudio:

Se parte de ítems nuevos. De haberse partido de preguntas de un banco se podría haber contado con los índices de dificultad y discriminación, los que habrían permitido seleccionar preguntas óptimas.

No se ha incluido otros criterios para definir a los distractores no funcionales

No se ha obtenido el índice de confiabilidad.

Como fortaleza, el presente estudio constituye un punto de partida para estudios similares en otras áreas de la carrera de Medicina.

### CONCLUSIONES

La eficiencia de las preguntas con 3 opciones es similar a las preguntas con 5 opciones

### REFERENCIAS BIBLIOGRÁFICAS

1. Comisión de Modernización Pedagógica. Pontificia Universidad Católica del Perú: Evaluación de los aprendizajes en la universidad. [www.pucp.edu.pe/cmp/docs/la\\_evaluacion.pdf](http://www.pucp.edu.pe/cmp/docs/la_evaluacion.pdf)

2. Blázquez F. La evaluación como proceso de diálogo. En: Rodríguez rojo M: Didáctica general. Edit Biblioteca Nueva. Madrid 2002 Cap X pp. 341-368.
3. Case SM, Swanson DB. Cómo construir preguntas de selección múltiple para ciencias básicas y ciencias clínicas. 3ra ed. 2005. National Board of Medical Examiners. 3750. market Street. Philadelphia PA 19106. [www.nbme.org](http://www.nbme.org)
4. [www.ecoetitulo.net](http://www.ecoetitulo.net) Proyecto ECOE TITULO MF 03: Instrumentos evaluativos. Documento sem FYC).
5. Romero S. ECOE: Evaluación clínica objetiva estructurada. I. Competencias y su evaluación. Medicina de Familia (And) 2002;1:49-52).
6. Silabo del Curso Bioquímica Médica. 3er Año. Facultad de Medicina. Universidad Nacional de Trujillo. Año 2019.
7. Jurado-Núñez A, Flores-Hernández F, Delgado-Maldonado L, Sommer-Cervantes H, Martínez-González A, Sánchez-Mendiola M. Distractores en preguntas de opción múltiple para estudiantes de medicina: ¿cuál es su comportamiento en un examen sumativo de altas consecuencias?. Inv Ed Med 2013;2(8):202-210
8. Rodríguez MC. Three options are optimal for multiple-choice: a meta-analysis of 80 years of research. Educational Measurement: Issues and Practice 2005;24(2):3-13.
9. Asociación Peruana de facultades de Medicina. Examen nacional de medicina (ENAM) 2003- experiencia piloto. Informe final. Lima 2003. Disponible en: <http://www.aspefam.org.pe/documentos/Informe%20Final%20ENAM.pdf>
10. Asociación Peruana de Facultades de Medicina. ENAM experiencia de 3 años: 2003-2005. Disponible en: <http://www.aspefam.org.pe/cpem/ponencias/PRESENTACIONE/S/03%20DICIEMBRE/Mesa%20Redonda1.pdf>
11. Davis BG. Quizzes, Tests and exams. En: Davis BG. Tools for teaching. Jossey-Bass Publisher. San Francisco 1993. Disponible en: <https://www.indiana.edu/~istr695/readingsfall2013/Tools%20For%20Teaching.pdf>
12. Hernández D, Vargas A, Almuiñas J, García J. Los indicadores actuales de la eficiencia académica. Pedagogía Universitaria 2015;20(3):53-62. Disponible en: <https://www.semanticscholar.org/paper/Los-indicadores-actuales-de-la-eficiencia-necesidad-Falc%C3%B3n-Jim%C3%A9nez/77b6847d5708456a140a1f8e5bd1f5aa24f43679>
13. Camacho-Saavedra L. análisis de las preguntas de opción múltiple en exámenes aplicados en exámenes de la asignatura Farmacología Básica Facultad de Medicina Universidad Nacional de Trujillo. Libro de resúmenes II Congreso Peruano de Educación Médica 2007. ASPEFAM T30. Disponible en [www.aspefam.org.pe/cpem2007](http://www.aspefam.org.pe/cpem2007)
14. Haladyna T, Downing S, Rodríguez M. A Review of Multiple-Choice Item-Writing Guidelines for Classroom Assessment. Applied Measurement in Education 2002;15(3), 309-334
15. Moreno R, Martínez R, Muñiz J. Directrices para la construcción de ítems de elección múltiple. Psicothema 2004; 16(3):490-497
16. Rivera-Jiménez J, Flores-Hernández F, Alpuche-Hernández A, Martínez-González A. Evaluación de reactivos de opción múltiple en medicina. Evidencia de validez de un instrumento. Inv Ed Med 2017;6(21):8-15
17. Camacho-Saavedra L, Huaman-Saavedra Jorge, Plasencia-Avarez Jorge. Eficiencia de preguntas de opción múltiple con 3 opciones en la Universidad Nacional de Trujillo. Libro de resúmenes V Congreso Internacional de Educación Médica CIEM2019. ASPEFAM T77. Lima 30,31 Mayo y 1 de Junio 2019. Disponible en: [http://www.aspefam.org.pe/ciem2019/libro\\_resumenes2019.pdf](http://www.aspefam.org.pe/ciem2019/libro_resumenes2019.pdf)
18. Tarrant M, Ware J. A comparison of the psychometric properties of three- and four-option multiple-choice questions in nursing assessments. Nurse education today 2010;30(6):539-543
19. Dehna A, Nasser H, Hosseini A. A Comparison between Three-and Four-Option Multiple Choice Questions. Procedia - Social and Behavioral Sciences. 2014;98:398-403
20. Sánchez-Mendiola M. La investigación en la educación médica moderna: ¿local o global?. Inv Ed Med 2014;3(12):175-176
21. ASPEFAM. Bases del Examen Nacional de Medicina – ENAM (Aprobadas por la Asamblea General de ASPEFAM, 18 de enero de 2020) <http://www.aspefam.org.pe/enam/enam2020e/bases.pdf>

Citar como: Camacho-Saavedra LA, Huamán-Saavedra JJ, Plasencia-Alvarez JO. Eficiencia de preguntas de opción múltiple con 3 alternativas. Rev méd Trujillo 2020;15(4):163-9