



## Calidad del examen parcial de la asignatura Sangre y Sistema Nervioso en la carrera Estomatología

Quality of the partial test from Blood and Nervous System subject in Dentistry career

Oscar Rodríguez Reyes<sup>1\*</sup> <https://orcid.org/0000-0002-3042-9153>

Martha Fajardo Puig<sup>1</sup> <https://orcid.org/0000-0003-2437-6996>

Norys Esther Pérez Sánchez<sup>1</sup> <https://orcid.org/0000-0001-5408-8181>

Margarita Hernández Cunill<sup>1</sup> <https://orcid.org/0000-0002-4959-946X>

Ileana González Palomo<sup>1</sup> <https://orcid.org/0000-0002-8433-1821>

Kirenia Montoya González<sup>1</sup> <https://orcid.org/0000-0002-9406-5811>

<sup>1</sup>Universidad de Ciencias Médicas de Santiago de Cuba. Facultad de Estomatología. Departamento de Ciencias Básicas Biomédicas. Santiago de Cuba, Cuba.

\*Autor para la correspondencia. Correo electrónico: [orreyes7807@infomed.sld.cu](mailto:orreyes7807@infomed.sld.cu)

### RESUMEN

**Introducción:** Estimar la calidad de cualquier instrumento evaluativo escrito, es requisito esencial no solo en su proceso de construcción, sino también en la obtención de información para la toma correcta de decisiones.

**Objetivo:** Evaluar la calidad del examen parcial de la asignatura Sangre y Sistema Nervioso, disciplina Bases Biológicas de la Estomatología, carrera Estomatología.

**Métodos:** Estudio observacional analítico y transversal, en 27 exámenes correspondientes al examen parcial de la asignatura Sangre y Sistema Nervioso, disciplina Bases Biológicas de la Estomatología, carrera de Estomatología, Universidad de Ciencias Médicas de Santiago de Cuba, curso 2022. Se analizó



la correspondencia del fondo de tiempo y la cantidad de ítems explorados en el examen, se calcularon índices de dificultad y de discriminación de incisos y preguntas, coeficiente de discriminación biserial puntual y de consistencia interna.

**Resultados:** Existió correspondencia adecuada entre el fondo de tiempo asignado a los contenidos de las temáticas y la cantidad de ítems explorados en el examen. Predominaron los reactivos con índice de dificultad media y buen poder de discriminación. La consistencia interna de las preguntas y del examen general se evaluó de excelente.

**Conclusiones:** La calidad del examen parcial aplicado en la asignatura Sangre y Sistema Nervioso de la disciplina Bases Biológicas de la Estomatología, carrera de Estomatología, Universidad de Ciencias Médicas de Santiago de Cuba, curso 2022, es adecuada.

**Palabras clave:** carrera; sistema nervioso; tiempo; dificultad; discriminación; confiabilidad; reactivos.

## ABSTRACT

**Introduction:** To estimate the quality of any written evaluative instrument is an essential requirement not only in its construction process, but also in obtaining information for correct decision making.

**Objective:** To evaluate the quality of the partial test from Blood and Nervous System subject in Biological Bases of Dentistry discipline, Dentistry career.

**Methods:** Analytical and cross-sectional observational study, in 27 tests corresponding to the partial test of Blood and Nervous System subject, Biological Bases of Dentistry discipline, Dentistry career, Santiago de Cuba Medical Sciences University, course 2022. The correspondence from the background of time and the number of items explored in the test were analyzed, indices of difficulty and discrimination of items and questions, punctual biserial discrimination coefficient and internal consistency were calculated.

**Results:** There was an adequate correspondence between the amount of time assigned to the contents of the themes and the number of items explored in the test. Items with a medium difficulty index and good discrimination power prevailed. The internal consistency of the questions and general test was evaluated as excellent.



**Conclusions:** The quality of the partial test from Blood and Nervous System subject, Biological Bases of Dentistry discipline, Dentistry career, Santiago de Cuba Medical Sciences University, course 2022, is adequate.

**Keywords:** career; nervous system; time; difficulty; discrimination; reliability; reagents.

Recibido: 26/10/2022

Aprobado: 24/03/2023

## INTRODUCCIÓN

La evaluación constituye el mecanismo regulador del proceso enseñanza-aprendizaje; posibilita su retroalimentación, dirección y perfeccionamiento. Como uno de los aspectos más polémicos y de mayor interés en la práctica pedagógica, permite en la educación médica superior cubana, el control y la valoración de los conocimientos, habilidades, hábitos y modos de actuación que los estudiantes adquieren sobre los objetivos de cada disciplina, asignatura y del plan de estudio. Según *Blanco Pereira* y otros,<sup>(1)</sup> para que la evaluación sea eficaz debe de estar vinculada con la competencia de los profesores, los métodos y medios empleados, así como de la calidad del currículo.

Al concluir la confección de cualquier instrumento evaluativo escrito, sea frecuente, parcial o final, se presenta el problema de determinar en qué medida la información obtenida a través de su aplicación, refleja el nivel de competencia del examinado. Desde esta perspectiva, estimar la calidad del instrumento es requisito esencial, no solo en su proceso de construcción, sino también en la obtención de información para la toma de decisiones correctas.<sup>(2)</sup>

Múltiples investigadores, como *Villamarín Guevara*,<sup>(3)</sup> *Díaz Rojas* y otros,<sup>(4)</sup> coinciden al afirmar que, dentro de los indicadores de calidad de los exámenes escritos que con mayor frecuencia se estudian, están la validez, confiabilidad, discriminación y dificultad. Aseguran que el análisis en el colectivo de profesores, durante el proceso de planeación y de elaboración de las preguntas e ítems, la forma de definir



la clave de calificación, de determinar el éxito y el fracaso; así como, la experiencia del claustro docente en la temática evaluación, es fundamental para el logro de un instrumento válido y confiable.

La carrera de Estomatología en Cuba, incluye en su currículo base del plan de estudio E, la disciplina Bases Biológicas de la Estomatología (BBE); es integradora y sigue la tendencia actual de las ciencias biomédicas en la unificación interdisciplinaria. Su sustento se materializa en la identificación de los núcleos básicos de los contenidos, esencialidades, relación con los problemas de salud bucal que los futuros egresados deben ser capaces de enfrentar y resolver; así como las tareas que en interacción con la familia, la comunidad y el medio ambiente, deben desarrollar para darles solución y definen la futura competencia.<sup>(5)</sup>

Sangre y Sistema Nervioso es una de las asignaturas de mayor complejidad dentro de las BBE. Se imparte en el primer año (segundo período) del currículo base, e incluye el estudio de los sistemas hemolinfopoyético y nervioso, temas de vital importancia para el futuro estomatólogo general. Su evaluación parcial abarca los contenidos del tema hemolinfopoyético y los resultados docentes, de modo habitual no siempre han sido los esperados por el claustro de profesores de las ciencias básicas biomédicas. La carencia de una estimación científica sobre la calidad del examen parcial de la asignatura Sangre y Sistema Nervioso, constituye el problema científico de la presente investigación.

El objetivo es evaluar la calidad de este instrumento evaluativo, durante el curso académico 2022.

## MÉTODOS

Se realizó un estudio observacional analítico y transversal, en la facultad de Estomatología, Universidad de Ciencias Médicas de Santiago de Cuba, entre los meses de septiembre a octubre de 2022.

El universo de estudio estuvo constituido por 27 exámenes, correspondientes a la prueba parcial de la asignatura Sangre y Sistema Nervioso, disciplina Bases Biológicas de la Estomatología, del plan de estudio E, carrera Estomatología, aplicada los estudiantes del primer año del curso académico 2022 (las calificaciones se incluyen como archivo complementario al artículo).

El instrumento evaluativo abarcó el tema sistema hemolinfopoyético, al cual según plan de estudio le corresponde un fondo de tiempo de 16 horas. Quedó estructurado en 5 preguntas con un total de 40 ítems.



Según el formato se clasificó como un examen mixto, con predominio de ítems test objetivo. Las temáticas evaluadas incluyeron: generalidades de sangre, serie roja, defensa del organismo, hemostasia y coagulación sanguínea, tejido hemolinfopoyético.

Para la clasificación de los ítems según dificultad esperada, se asumió la escala considerada por *Díaz Rojas* y otros<sup>(4)</sup> en: 5 % para ítems fáciles, 20 % para ítems medianamente fáciles, 50 % para ítems de dificultad media, 20 % para ítems medianamente difíciles y 5 % para los ítems difíciles.

En una hoja de cálculo de Microsoft Excel, se tabularon las calificaciones de cada pregunta y la general del examen objeto de estudio, se incluyó además el resultado de cada ítem del instrumento, categorizado en aprobado (1) y desaprobado (0), lo que garantizó su posterior procesamiento estadístico.

Para dar salida al objetivo propuesto, se definieron y operacionalizaron las siguientes variables:

Índice de dificultad del ítem (Di): número de estudiantes examinados que respondió correctamente el ítem, dividido entre el total de examinados.

Índice de dificultad de la pregunta (Dp): suma del total de estudiantes que respondieron correctamente cada uno de los ítems que corresponden a la pregunta, dividido entre el número de ítems de esta. Luego se calculó la media de respuestas correctas para la pregunta y se dividió por el total de alumnos examinados.

Índice de dificultad del examen (De): suma del total de estudiantes que respondieron correctamente cada ítem del examen, dividido entre el número de ítems y a continuación entre el número de examinados. El índice de dificultad para ítems, preguntas y examen se clasificó en las categorías:<sup>(7,8)</sup>

- Difícil: ( $< 0,32$ )
- Medianamente difícil: ( $0,32 - 0,52$ )
- Dificultad media: ( $0,53 - 0,73$ )
- Medianamente fácil: ( $0,74 - 0,86$ )
- Fácil: ( $> 0,86$ )



Índice de discriminación (Id): se seleccionó el grupo de mejores calificaciones (en sentido descendente, a partir de la más alta) y el grupo de notas más bajas (en sentido ascendente, a partir de la más baja). Para el cálculo del Id se empleó la siguiente fórmula:

$$Id = \frac{GA(\text{aciertos}) - GB(\text{aciertos})}{N (\text{grupo mayor})}$$

Id: índice de discriminación del reactivo.

GA: número de aciertos en el reactivo i del 33 % de personas con las puntuaciones más altas en el test.

GB: número de aciertos en el reactivo i del 33 % de personas con las puntuaciones más bajas en el test.

N: número de personas en el grupo más numeroso (GA o GB).<sup>(4)</sup>

Para la evaluación de la calidad de los ítems y preguntas en términos de índice de discriminación, se empleó la siguiente escala, utilizada por *Martínez Pérez* y otros:<sup>(7)</sup>

- Excelente: mayor o igual de 0,40 (conservar)
- Buena: entre 0,30 – 0,39 (posibilidades de mejorar)
- Regular: entre 0,20 – 0,29 (necesidad de revisar)
- Pobre: entre 0,00 – 0,19 (descartar o revisar a profundidad)
- Pésimo: < 0,00 (descartar definitivamente)

Relación discriminativa (Rd): para determinar si un reactivo o pregunta es aceptable o no a partir de su índice de discriminación. Está dado por el cociente entre el índice de discriminación del ítem -Id (i)- y su norma discriminativa (ND).<sup>(4)</sup>

$$Rd = \frac{Id(i)}{ND}$$

El valor de la norma discriminativa se obtuvo a partir de la siguiente fórmula:

<http://scielo.sld.cu>

<http://www.revmedmilitar.sld.cu>





$$ND = 0,3 * (Di) \text{ si } 33 \% \leq (Di) \leq 67 \%$$

o

$$ND = 1,00 - (Di) \text{ si } 67 \% \leq (Di) \leq 100 \%$$

Si  $R_d$  mayor que 1: se considera el ítem aceptable

Si  $R_d$  (0,60-0,99): se recomienda analizar el ítem

Si  $R_d$  menor de 0,6: se recomienda rechazar el ítem

Coefficiente de discriminación del punto biserial puntual ( $r_{pbis}$ ): para calcular la efectividad discriminativa de un reactivo o pregunta. Se utilizó para conocer si los estudiantes de calificaciones más altas son los que obtuvieron las respuestas correctas, qué tanto poder predictivo tuvo el reactivo, y cómo puede contribuir a las predicciones.<sup>(4)</sup> Para el cálculo se utilizó la siguiente fórmula:

$$r_{pbis} = \frac{\bar{X}_1 - \bar{X}_0}{S_x} \times \sqrt{\frac{n_1 n_0}{n(n-1)}}$$

$\bar{X}_1$  = media de las puntuaciones totales de aquellos que respondieron correctamente el ítem.

$\bar{X}_0$  = media de las puntuaciones totales de aquellos que respondieron incorrectamente el ítem.

$S_x$  = desviación estándar de las puntuaciones totales.

$n_1$  = número de estudiantes que respondieron correctamente el ítem.

$n_0$  = número de estudiantes que respondieron incorrectamente el ítem.

$n = n_1 + n_0$

La distribución de los resultados del  $r_{pbis}$  que se tuvo en cuenta fue la siguiente:

- $< 0$  – discriminan negativamente
- $0 - 0,14$  – discriminan pobremente



- 0,15 – 0,25 – discriminan regular
- 0,26 – 0,35 – buen poder discriminativo
- > 0,35 – excelente poder de discriminación

Consistencia interna de reactivos (Cir):<sup>(3)</sup> se determinó a través del cálculo del coeficiente de Küder-Richardson (KR20) mediante la siguiente fórmula:

$$KR20 = \frac{k}{k-1} * \left[ 1 - \frac{\sum p * q}{Vt} \right]$$

KR20 = coeficiente de confiabilidad

k = número total de ítems en el instrumento

Vt = varianza total

$\sum p*q$  = sumatoria de la varianza de los ítems

p = TRC/N: total de respuestas correctas (TRC) entre el número de sujetos participantes (N)

q = 1-p

Consistencia interna del examen (Cie):<sup>(4)</sup> se determinó a través del cálculo del coeficiente de confiabilidad alfa de Cronbach, mediante el paquete estadístico SPSS v.22.

La consistencia interna de reactivos y del examen en general se evaluó de la siguiente forma:

- $\leq 0,53$  – confiabilidad nula
- 0,54 – 0,59 – confiabilidad baja
- 0,60 – 0,65 – confiable
- 0,66 – 0,71 – muy confiable
- 0,72 – 0,99 – excelente confiabilidad
- 1,00 – confiabilidad perfecta





La existencia o no de correspondencia entre la cantidad de horas por temáticas y el número de ítems explorados en la prueba parcial, se determinó a través del coeficiente de correlación por rangos de Spearman.<sup>(6)</sup>

El nivel de significación empleado para todas las pruebas estadísticas fue de  $\alpha = 0,05$ .

Conforme a los aspectos bioéticos y legales, se solicitó a las autoridades académicas correspondientes de la facultad de Estomatología, Universidad de Ciencias Médicas de Santiago de Cuba, la autorización para el desarrollo de la investigación. Se mantuvo durante todo el estudio, la confidencialidad e integridad de los datos obtenidos de los exámenes.

## RESULTADOS

La tabla 1 muestra el fondo de tiempo de los contenidos y su presencia en el examen parcial. Como se aprecia, existió un balance equilibrado de horas según temáticas abordadas. La relación entre fondo de tiempo e ítems del instrumento fue positiva y considerable, con un coeficiente de correlación de Spearman de 0,82.

**Tabla 1** - Fondo de tiempo de los contenidos del tema según programa de la asignatura y presencia de los incisos en las temáticas

Contenidos del tema hemolinfopoyético	Fondo de tiempo		Debían planearse		Presencia en el examen	
	Horas	%	Incisos	%	Incisos	%
Generalidades de sangre y serie roja	4	25	20	25	15	30
Defensa del organismo	4	25	10	25	5	10
Hemostasia y coagulación	4	25	10	25	10	20
Tejido Hemolinfopoyético	4	25	10	25	10	20

Rho de Spearman: 0,82 (p = 0,184, p > 0,05).

La correspondencia entre la cantidad de incisos según su índice de dificultad esperada con los reales del examen se muestra en la tabla 2.



**Tabla 2** - Incisos de la prueba parcial de Sangre y Sistema Nervioso según índice de dificultad esperado y real

Incisos según índice de dificultad	Incisos esperados	% de incisos esperados	Incisos reales	% de incisos reales
Fáciles	2	5	6	15
Medianamente fáciles	8	20	9	22,5
De dificultad media	20	50	16	40
Medianamente difíciles	8	20	9	22,5
Difíciles	2	5	-	-
Totales	40	100	40	100

Rho de Spearman: 0,97 (p= 0,005, p< 0,05).

En el examen existió un incremento en un 10 % de los incisos fáciles con respecto a lo esperado. El resto de los reactivos estuvo bastante cercano a lo deseado. El coeficiente de correlación de Spearman, evidenció una correlación positiva con significación estadística entre los incisos esperados y los reales obtenidos.

El índice de dificultad por preguntas, su tipología y por temáticas, se resume en la tabla 3. Predominaron en un 60 % del instrumento evaluativo las preguntas con índice de dificultad media. Las temáticas serie roja y defensa del organismo, cuyas preguntas son de tipología problemática, fueron las de mayor complejidad para los estudiantes, al mostrar menor índice de dificultad. El examen en sentido general se catalogó de dificultad media.

**Tabla 3** - Índice de dificultad de las preguntas, su tipología, temáticas y el examen

Preguntas	Temáticas	Tipo de pregunta	Índice de dificultad
1	Generalidades de sangre	Selección múltiple complemento simple	0,77
2	Serie roja	Problemática de respuestas cortas	0,56
3	Defensa del organismo	Problemática de respuestas cortas	0,53
4	Hemostasia y coagulación sanguínea	Selección múltiple complemento agrupado	0,61
5	Tejido Hemolinfopoyético	Test de verdadero o falso	0,78
Examen			0,66



Los resultados en relación a la discriminación de los ítems de las preguntas, se agrupan en la tabla 4.

**Tabla 4 - Índice de discriminación de los incisos según preguntas**

Preguntas	Temáticas	< - 0,01	0,00 a 0,19	0,20 a 0,29	0,30 a 0,39	> 0,39	Total
1	Generalidades de sangre	1	2	2	2	-	7
2	Serie roja	-	-	-	-	8	8
3	Defensa del organismo	-	-	-	-	5	5
4	Hemostasia y coagulación sanguínea	-	-	1	1	8	10
5	Tejido hemolinfopoyético	1	1	3	2	3	10
Total		2	3	6	5	24	40

Más de la mitad de los reactivos del examen (24) discriminan de forma adecuada, para un 60 % del total. Las preguntas 1 y 5 que corresponden a las temáticas de generalidades de sangre y tejido hemolinfopoyético, incluyeron entre ambas, 2 ítems de escasa discriminación (5 %); 3 que lo hacen de modo pobre (7,5 %) y 5 que discriminan regular (12,5 %).

La tabla 5 proporciona una visión general sobre el índice de dificultad, de discriminación y la relación discriminativa de las preguntas.

**Tabla 5 - Índice de dificultad, índice de discriminación y relación discriminativa por preguntas**

Preguntas	Tipo de pregunta	Índice de dificultad	Índice de discriminación	Relación discriminativa
1	Selección múltiple complemento simple	0,77	0,10	1,17
2	Problémica de respuestas cortas	0,56	0,72	4,63
3	Problémica de respuestas cortas	0,53	0,71	3,93
4	Selección múltiple complemento agrupado	0,61	0,53	3,19
5	Test de verdadero o falso	0,78	0,32	1,53

La pregunta de selección múltiple complemento simple exhibió un pobre índice de discriminación. En las restantes se aprecia buena correspondencia entre las 3 variables y el tipo de pregunta. La relación discriminativa habla a favor de las preguntas 2 y 3.



El  $r_{pbis}$  calculado mostró un valor de 0,19 para la temática de generalidades de sangre (pregunta 1), lo que confirma también que discrimina regular y necesita revisión profunda. El resto de las temáticas poseen un excelente poder de discriminación ( $r_{pbis} > 0,35$ ).

La consistencia interna de los reactivos (Cir) y del examen en general (Cie) mostró una excelente confiabilidad ( $KR20 = 0,91$ ,  $\alpha$  de Cronbach = 0,78).

## DISCUSIÓN

A pesar de que existió una correspondencia adecuada entre el fondo de tiempo asignado en la planificación docente a cada temática abordada, con la cantidad de incisos explorados en el examen, autores como *Martínez Pérez* y otros,<sup>(7)</sup> manifiestan que cuando esta correspondencia falla, se resta validez al instrumento, pues el contenido debe ser una muestra representativa del constructo o dominio definido en el programa de la asignatura. Añaden, además, que la selección de temas a evaluar en el examen, se centra en la relevancia y la representatividad del contenido objeto de la evaluación, por tanto, ha de reflejar fielmente los núcleos básicos y sus esencialidades.

En relación con la correspondencia entre la cantidad de incisos, según su índice de dificultad esperada, con los del examen; los resultados coinciden con los reportados por *Moreno Montañez* y otros,<sup>(9)</sup> quienes en el análisis de calidad de un examen final escrito de Medicina Comunitaria, con 41 ítems, obtuvieron un predominio de incisos con nivel de dificultad media y ausencia de ítems difíciles. Hallazgos diferentes son reportados por *Gómez López* y otros,<sup>(2)</sup> en los que se evalúa el índice de dificultad y la discriminación de un examen parcial de Bioquímica; se encontró que 4 % de los reactivos fueron difíciles, 17 % medianamente difíciles, 43 % de dificultad media y 20 % fáciles.

Las temáticas relacionadas con serie roja y defensa del organismo, muestran los más bajos índices de dificultad (0,56 y 0,53 respectivamente). A juicio de los autores, resultaron las de mayor complejidad, por ser de tipo problémica, y exigir un nivel de asimilación aplicativo por parte de los estudiantes. En este sentido *Minte Münzenmayer* y otros,<sup>(10)</sup> expresan que los procedimientos evaluativos aplicados en las universidades promueven de manera general, niveles de razonamiento y dimensiones del proceso cognitivo básicos. En otras palabras, los alumnos no son evaluados para lograr aprendizajes, son



evaluados bajo parámetros tradicionales, centrados en la repetición, en la memoria, no en el desarrollo de competencias intelectuales de orden superior.

Rodríguez Graña y otros,<sup>(11)</sup> Camacho Saavedra y otros<sup>(12)</sup> y González Machado y otros,<sup>(13)</sup> exponen que el nivel medio de dificultad de un examen escrito debe oscilar entre 0,5 y 0,6, lo que concuerda con los resultados de la presente investigación.

El Id y el  $r_{pbis}$  permitieron identificar que solo la pregunta 1, de tipología selección múltiple complemento simple, que abarcó los contenidos relacionados con la temática generalidades de sangre, discriminó pobremente entre los estudiantes con altas y bajas calificaciones, al obtener un índice de dificultad catalogado como medianamente fácil (0,77). Este resultado sugiere, como proponen Villamarín Guevara<sup>(14)</sup> y Menacho Chiok y otros,<sup>(15)</sup> que debe ser revisada a profundidad por el claustro de profesores de la asignatura o descartarla de forma definitiva para futuros exámenes parciales. Se sugiere, además, que la pregunta 5, aunque posee buena calidad, tiene posibilidades de mejorarse.

Los valores del coeficiente  $\alpha$  de Cronbach y el KR20 demuestran que el instrumento tiene una excelente confiabilidad o consistencia interna, lo que concuerda con los estándares propuestos por el Colegio de Médicos de Canadá, el cual considera excelente entre 0,72 y 0,99; y los establecidos por la Comisión Nacional de Evaluación de la Competencia y el Desempeño del Ministerio de Salud Pública, con valores superiores a 0,60.<sup>(16)</sup>

Se concluye que la calidad del examen parcial aplicado en la asignatura Sangre y Sistema Nervioso de la disciplina Bases Biológicas de la Estomatología, carrera de Estomatología, en la Universidad de Ciencias Médicas de Santiago de Cuba, en el curso académico 2022 es adecuada, al presentar un índice de dificultad media, buena discriminación de sus preguntas y una excelente consistencia interna.

## REFERENCIAS BIBLIOGRÁFICAS

1. Blanco Pereira ME, Martínez L, González Gil A, Jordán Padrón M. Calidad del examen final teórico de Morfofisiología I en la facultad de Ciencias Médicas de Matanzas. Cursos 2012-2013 y 2013-2014. Revista Médica Electrónica. 2015 [acceso: 11/10/2022]; 37(4):323-32. Disponible en: <http://scielo.sld.cu/pdf/rme/v37n4/rme030415.pdf>

<http://scielo.sld.cu>

<http://www.revmedmilitar.sld.cu>



2. Gómez López VM, Rosales Gracia S, García Galaviz JL, Berrones Sánchez KI, Berrones Sánchez CM. Índice de dificultad y discriminación de ítems para la evaluación en las asignaturas básicas de medicina. *Revista Educación Médica Superior*. 2020 [acceso: 11/10/2022]; 34(1):e1727. Disponible en: <http://www.ems.sld.cu/index.php/ems/article/view/1727/1015>
3. Villamarín Guevara RS. Indicadores de calidad de los instrumentos de evaluación y el rendimiento académico [Tesis doctoral]. Perú: Universidad Nacional de San Marcos; 2017 [acceso: 11/10/2022]. Disponible en: [https://cybertesis.unmsm.edu.pe/bitstream/handle/20.500.12672/6636/Villamarin\\_gr.pdf?sequence=3&isAllowed=y](https://cybertesis.unmsm.edu.pe/bitstream/handle/20.500.12672/6636/Villamarin_gr.pdf?sequence=3&isAllowed=y)
4. Díaz Rojas PA, Leyva Sánchez E. Metodología para determinar la calidad de los instrumentos de evaluación. *Educ Med Super*. 2013 [acceso: 11/10/2022]; 27(2):269-86. Disponible en: <http://scielo.sld.cu/pdf/ems/v27n2/ems14213.pdf>
5. Sánchez Fernández NV, Sánchez Clemente A, Thomas Ferrer F, Tabera García ME. Programa de la Disciplina Estomatología Integral. Análisis del plan de estudio "D". *Revista Información Científica (RIC)*. 2017 [acceso: 11/10/2022]; 96(2):280-7. Disponible en: <https://www.redalyc.org/journal/5517/551764114014/html/>
6. Ortíz Romero GM, Díaz Rojas PA, Llanos Domínguez OR, Pérez Pérez SM, González Sapsin K. Dificultad y discriminación de los ítems del examen de Metodología de la Investigación y Estadística. *EDUMECENTRO*. 2015 [acceso: 09/09/2022]; 7(2):19-35. Disponible en: [http://scielo.sld.cu/scielo.php?script=sci\\_arttext&pid=S2077-28742015000200003](http://scielo.sld.cu/scielo.php?script=sci_arttext&pid=S2077-28742015000200003)
7. Martínez Pérez R, Ríos Rodríguez M, Roque Marrero Y, Caballero Pedrón KY. Evidencias de validez del examen final de la asignatura Biología molecular, carrera de Medicina. *Revista Electrónica Medimay*. 2019 [acceso: 12/9/2021]; 26(3):338-48. Disponible en: <http://www.medimay.sld.cu/index.php/rcmh/article/view/1483/1677>
8. Baladrón J, Sánchez-Lasheras F, Romero Ladrero JM, Curbelo J, Villacampa Menéndez P, Jiménez-Fonseca P. Evolución de los parámetros dificultad y discriminación en el ejercicio de examen MIR. Análisis de las convocatorias de 2009 a 2017. *FEM: Revista de la Fundación de Educación Médica*.





2018 [acceso: 16/10/2022]; 21(4):181-93. Disponible en:

[https://scielo.isciii.es/scielo.php?script=sci\\_arttext&pid=S2014-98322018000400003](https://scielo.isciii.es/scielo.php?script=sci_arttext&pid=S2014-98322018000400003)

9. Moreno Montañez M, Lardoeyt Ferrer M, Ramos Arencibia AL. Calidad del examen final escrito de Medicina Comunitaria en la facultad de Medicina Cabinda, Angola. Revista Conrado. 2020 [acceso: 13/10/2022]; 16(73):214-20. Disponible en:

[http://scielo.sld.cu/scielo.php?script=sci\\_arttext&pid=S1990-86442020000200214](http://scielo.sld.cu/scielo.php?script=sci_arttext&pid=S1990-86442020000200214)

10. Minte Münzenmayera A, Sepúlveda Obreque H, Jaramillo Alvarado R, Díaz-Levicoy D.

Evaluación en Educación Superior: características y demandas cognitivas de preguntas escritas. Revista de Estudios y Experiencias en Educación (REXE). 2021 [acceso: 11/10/2022]; 20(44):43-52.

Disponible en: <https://revistas.ucsc.cl/index.php/rexe/article/view/1002/752>

11. Rodríguez Graña T, Sondón Fernández ME, Peña González M, Hernández Tamayo M, Robles García R. La calidad del examen final de la asignatura Morfofisiología IV. CCM. 2018 [acceso: 09/10/2022]; 22(3):446-62. Disponible en:

[http://scielo.sld.cu/scielo.php?script=sci\\_arttext&pid=S156043812018000300009&lng=es](http://scielo.sld.cu/scielo.php?script=sci_arttext&pid=S156043812018000300009&lng=es)

12. Camacho-Saavedra LA, Huamán-Saavedra JJ, Plasencia Alvarez JO. Eficiencia de preguntas de opción múltiple con tres alternativas. Rev méd Trujillo. 2020 [acceso: 16/10/2022]; 15(4):163-9.

Disponible en: <https://revistas.unitru.edu.pe/index.php/RMT/article/view/3214>

13. González Machado EC, Reyes Piñuelas EP, López Ortega M. Construcción de una prueba para evaluar aprendizajes en educación superior. En: Congreso Nacional de Investigación Educativa San Luis Potosí 2017; 2017. San Luis Potosí: Universidad Autónoma de Baja California. [acceso:

16/10/2022]. Disponible en: <http://www.comie.org.mx/congreso/memoriaelectronica/v14/doc/2300.pdf>

14. Villamarín Guevara RS. Índices de la calidad de los instrumentos de evaluación. Revista Redipe. 2017 [acceso: 14/10/2022]; 6(5):238-45. Disponible en:

<https://revista.redipe.org/index.php/1/article/view/274>

15. Menacho Chiok CH, Cano Alva Trinidad JM. Evaluación de pruebas informatizadas aplicando la teoría clásica de los test y la teoría de respuesta al ítem. Anales Científicos. 2020 [acceso: 15/10/2022]; 81(2):278-88. Disponible en: [https://revistas.lamolina.edu.pe/index.php/acu/article/view/1638/pdf\\_270](https://revistas.lamolina.edu.pe/index.php/acu/article/view/1638/pdf_270)





16. Ríos Hidalgo N, Rivero Machado IC, Ordoñez DJ, Sitchao Suárez NM, López Vergara JM, Garcés Sigas JR. La retención de contenidos biomédicos relacionados con la enseñanza desarrolladora en Anatomía Patológica. Escuela Latinoamericana de Medicina. Panorama Cuba y Salud. 2017 [acceso: 13/10/2022]; 12(2):15-23. Disponible en:  
<http://www.revpanorama.sld.cu/index.php/panorama/article/view/643>

## Conflictos de interés

Los autores del presente artículo declaran no tener conflicto de interés.

## Contribuciones de los autores

Conceptualización: *Dr. C. Oscar Rodríguez Reyes, M. Sc. Margarita Hernández Cunill, Dra. Martha Fajardo Puig*

Análisis formal: *Dr. C. Oscar Rodríguez Reyes*

Investigación: *Dr. C. Oscar Rodríguez Reyes, M. Sc. Margarita Hernández Cunill, Dra. Martha Fajardo Puig,*

Metodología: *Dr. C. Oscar Rodríguez Reyes*

Administración del proyecto: *Dr. C. Oscar Rodríguez Reyes*

Recursos: *M. Sc. Ileana González Palomo, Lic. Norys Esther Pérez Sánchez, Dra. Kirenia Montoya González*

Supervisión: *Dr. C. Oscar Rodríguez Reyes*

Visualización: *Dr. C. Oscar Rodríguez Reyes*

Redacción -borrador original: *Dr. C. Oscar Rodríguez Reyes, M. Sc. Margarita Hernández Cunill, Lic. Norys Esther Pérez Sánchez, Dra. Martha Fajardo Puig*

Redacción -revisión y edición: *Dr. C. Oscar Rodríguez Reyes, M. Sc. Ileana González Palomo, Dra. Kirenia Montoya González*