



CALIDAD DEL INSTRUMENTO DE EVALUACIÓN FINAL DE LA ASIGNATURA CÉLULA TEJIDOS Y SISTEMA TEGUMENTARIO

Autores: Dra. Marisela Resco Zequeira¹, Dra. Ibis Ofelia Sánchez Rodríguez², Lic. Yeumicelis González Verdecia³.

1- Especialista de Primer Grado en Histología. Especialista de Primer Grado en MGI. Profesor Asistente.

2- Especialista de Primer Grado en Histología. Especialista de Primer Grado en MGI. Profesor Asistente.

3-. Especialista de Primer Grado en Histología. Profesor Asistente.

Facultad de Ciencias Médicas de Manzanillo. Granma. Cuba

RESUMEN

Introducción: La evaluación del aprendizaje es un componente esencial del proceso de enseñanza-aprendizaje como vía de retroalimentación para su propia dirección y perfeccionamiento. **Objetivo:** Evaluar la calidad del Examen Ordinario de la Asignatura Célula Tejidos y Sistema Tegumentario a través del nivel de dificultad y el poder de discriminación del instrumento. **Material y Métodos:** Se realizó un estudio descriptivo en el campo de la evaluación. El universo estuvo integrado por 407 exámenes que corresponden al total de estudiantes presentados al examen. Las variables de estudio fueron el nivel de dificultad y poder de discriminación (índice de discriminación y coeficiente de discriminación). Para el procesamiento estadístico se realizó el cálculo de índices, coeficiente y por cientos. **Resultados:** El examen estuvo constituido por 7 preguntas, predominaron las de tipo Test Objetivo con un total de 5 que representa el 71,42%. La pregunta 2 tuvo mayor número de aciertos mientras que la pregunta 5 tuvo el mayor número de errores o menor número de aciertos. De las preguntas, 3 (42,85%) fueron de dificultad media, 2 (28,57%) altamente fácil, 1 (14,28%) medianamente fácil y 1 (14,28%) altamente difícil. El examen de forma general calificó de dificultad media con un índice de 0,69. El 71,42% de las preguntas presentaron una discriminación excelente. El examen de forma general tuvo una discriminación excelente con un índice de 0,69. El coeficiente de discriminación para todas las preguntas siempre fue positivo por lo que se catalogó como adecuado oscilando entre 0,62 y 0,69. **Conclusiones:** La calidad del examen es adecuada a partir de la coherencia entre las preguntas, el nivel de dificultad y el poder de discriminación.

Palabras claves: Evaluación del Aprendizaje, Índice de Dificultad, Poder de Discriminación, Índice de Discriminación, Coeficiente de Discriminación.



INTRODUCCIÓN

La enseñanza constituye el proceso de organización y dirección de la actividad cognoscitiva en el cual interactúan el profesor y el estudiante y por ello se denomina proceso de enseñanza-aprendizaje. El mismo constituye un sistema en el que cada uno de sus componentes se interrelacionan, siendo la evaluación el mecanismo regulador del sistema y puede considerarse como el instrumento de control de la calidad del producto resultante del proceso. En este sentido la evaluación del aprendizaje es un componente esencial del proceso de enseñanza-aprendizaje como vía de retroalimentación para su propia dirección y perfeccionamiento.^{1, 2, 3}

En el sistema de educación médica superior cubano, la evaluación del aprendizaje implica el control y la valoración de los conocimientos, habilidades y hábitos, así como de los modos de actuación que los estudiantes van adquiriendo a través del proceso docente educativo, de acuerdo con los objetivos de cada asignatura, estancia o rotación en particular y del plan de estudio en general.

Por ello es imprescindible que a lo largo de toda la carrera comprobemos en qué medida estamos preparando al estudiante para enfrentarse exitosamente a su responsabilidad social y si su aprendizaje se corresponde con los objetivos de cada asignatura, disciplina, año de estudio y la carrera en general.^{1, 2}

Cuando se utilizan instrumentos de alto impacto, como es el caso de los exámenes finales de las asignaturas, es necesario conocer los indicadores técnicos que definen la calidad del instrumento evaluativo que se emplea. Por su dimensión y por el poderoso impacto social que tienen, la elaboración de exámenes debe ajustarse a rigurosos estándares de calidad, pues los defectos en la elaboración de los mismos tienen efectos negativos en el currículo, el proceso docente y los estudiantes.⁴

Reviste gran importancia que la calidad de los exámenes aplicados sea verificada, pues sirven como instrumento de medición educativa, para inferir capacidades de las personas, ofrecer información correcta para la adopción de decisiones y permitir evaluar la marcha de procesos docentes.^{1, 5}

Diversos estudios realizados tanto en nuestro país como en otros, reportan defectos en la elaboración de instrumentos evaluativos escritos,⁶ los cuales repercuten de manera perjudicial en las características psicométricas y, por tanto, en la validez de dichos instrumentos.



Es necesario que el sistema de evaluación sea lo más sólido posible, se impone que los instrumentos que se empleen tengan la calidad requerida y se pongan a prueba.¹

Dentro de los indicadores de calidad de los exámenes escritos que con mayor frecuencia se trabajan está la validez, confiabilidad, dificultad y discriminación, los dos últimos fueron los estudiados en esta investigación.¹

Estos indicadores aportan información útil a estudiantes y profesores sobre el grado de dificultad que presenta un examen o contenido evaluado para trabajar en la superación de los problemas encontrados, además permite diferenciar los estudiantes que dominan un conocimiento o habilidad de los que no lo han logrado aún, valorar el nivel de autopreparación alcanzado y orientar mejor el trabajo diferenciado; referencias necesarias para la planificación, organización y ejecución de las actividades para mejorar lo relacionado con la evaluación y con ello el proceso enseñanza-aprendizaje.¹

Desde la creación del Ministerio de Educación Superior, se ha mantenido como una de sus funciones principales el perfeccionamiento continuo de los planes de estudio, que en determinados momentos adquirió tal significación que condujo a transformaciones curriculares.

La implementación del Plan E en la carrera de Medicina comienza a partir del curso 2019-2020 e incluye en su currículo como una de sus disciplinas a Bases Biológicas de la Medicina, la misma es integradora y sigue la tendencia actual de las ciencias básicas biomédicas en la unificación interdisciplinaria para tratar los problemas comunes o afines de varias ramas científicas. La asignatura Célula, Tejidos y Sistema Tegumentario se imparte en el primer semestre de primer año e incluye los contenidos referentes a la célula eucariota, tejidos básicos y sistema tegumentario.⁷

La valoración sobre bases científicas de la calidad de la evaluación del aprendizaje de esta asignatura permitirá identificar las principales dificultades presentes en el proceso evaluativo de la misma y así trazar estrategias que permitan corregir los problemas encontrados, y realizar actividades de capacitación dirigidas a los profesores principales de las asignaturas responsables de la elaboración de los bancos de preguntas para la confección de los exámenes, y de esta forma aportar resultados que permitan realizar acciones que garanticen elevar la calidad en la formación de los estudiantes para el perfeccionamiento del proceso enseñanza-aprendizaje principalmente el sistema evaluativo.



Problema Científico: Desconocimiento de la calidad (dificultad y discriminación) del examen ordinario de la asignatura Célula Tejidos y Sistema Tegumentario aplicado al primer año de la carrera de Medicina en la Facultad de Ciencias Médicas de Granma.

OBJETIVOS

General:

Evaluar la calidad del Examen Ordinario de la Asignatura Célula Tejidos y Sistema Tegumentario a través del nivel de dificultad y el poder de discriminación del instrumento.

Específicos

1. Describir la estructura del instrumento evaluativo utilizado teniendo en cuenta el número de preguntas, procedimientos y técnicas empleados en la elaboración de cada pregunta.
2. Determinar la frecuencia de errores, el nivel de dificultad y el poder de discriminación por preguntas y del instrumento evaluativo en general.

MATERIAL Y MÉTODO

Se realizó un estudio descriptivo en el marco de la educación médica (en el campo de la evaluación), para evaluar la calidad del Examen Ordinario de la Asignatura Célula Tejidos y Sistema Tegumentario a través del nivel de dificultad y el poder de discriminación del instrumento, aplicado al primer año de la Carrera de Medicina en la Facultad de Ciencias Médicas de Granma, Curso 2019-2020.

El universo de estudio estuvo integrado por 407 exámenes que corresponden al total de estudiantes presentados al examen de Célula Tejidos y Sistema Tegumentario en el curso 2019- 2020.

Se clasificaron las preguntas teniendo en cuenta procedimientos y técnicas empleados para su elaboración y se determinó la frecuencia de errores de cada una.

Los métodos estadísticos se utilizaron para la identificación del nivel de dificultad y del poder de discriminación de cada pregunta y del instrumento como un todo.

El índice de dificultad y el índice de discriminación se calcularon a través del software de hoja de cálculo Excel de Windows XP, junto con una calculadora



científica. Se trabajó con números enteros, por ciento e índices para la presentación de los resultados en tablas.

Nivel de dificultad: Proporción de personas que responden correctamente dentro del total de examinados, es inversamente proporcional a la dificultad, entre mayor sea esta proporción menor será la dificultad. Bajo una perspectiva estadística se expresa como índice de dificultad (p) y, como cualquier índice, debe estar cercano a una proporción de 0.5-0.6.^{1, 8} Su cálculo se realizó por la fórmula de Backhoff.⁸

Se distribuyó los valores de p en: altamente difícil menos de 0,32, medianamente difícil de 0,32-0,52, dificultad media de 0,53-0,73, medianamente fácil de 0,74-0,86 y altamente fácil más de 0,86 de acuerdo a los parámetros propuestos por Backhoff.⁸

Se interpretaron los resultados acorde con la curva de distribución de frecuencias donde se espera que, según las respuestas correctas, dadas por los estudiantes, un 5% considera fácil el examen, 20% medianamente fácil, 50% con dificultad media, 20% medianamente difícil y 5% difícil.⁸

Poder de discriminación: Permite distinguir a los estudiantes de altos y bajos rendimientos, además aporta criterio de predicción sobre los resultados del estudiante, pues si el examen y una pregunta miden la misma habilidad o competencia, se puede esperar que quien tuviera una puntuación alta en todo el examen deberá tener altas probabilidades de contestar correctamente los temas evaluados en sus preguntas. Para identificarlo se utilizan dos indicadores: índice de discriminación y coeficiente de discriminación.⁸

Índice de discriminación: permite distinguir los estudiantes de altos y bajos rendimientos, solo se valora 54% (27% más alto y 27% más bajo) de los estudiantes evaluados, por tanto, se dividió el grupo de estudiantes en tres subgrupos, de acuerdo con calificación que obtuvieron en el examen, quedó así un tercio superior, un tercio medio y un tercio inferior.⁸

Entre más alto es el índice de discriminación (D_i), la prueba diferenciará mejor a los estudiantes con altas y bajas calificaciones. Si todos los estudiantes con altas calificaciones contestan correctamente una pregunta y todos los estudiantes con bajas calificaciones contestan incorrectamente, entonces el $D_i = 1$, (valor máximo de este indicador); si sucede lo contrario $D_i = -1$ (valor máximo negativo), si ambos grupos contestan por igual $D_i = 0$ (valor mínimo de discriminación).

Un índice de discriminación negativo refleja error en la determinación de la clave o que es una pregunta muy difícil, lo cual impediría discriminar correctamente



un grupo de estudiantes para valorar si dominan o no el contenido evaluado. Cuando el índice es altamente positivo se considera una pregunta excelente, ajustada a la clave, pero si disminuye el índice refleja que la pregunta se podrá modificar, debido a que muy pocos estudiantes seleccionaron la respuesta correcta, o un gran número de estudiantes seleccionó otra respuesta como correcta, lo cual orienta a la revisión para valorar si esa otra respuesta predominante sería igualmente correcta, si hay problemas de redacción que hagan que el estudiante interprete diferente a la respuesta correcta, o si consta de preguntas con respuestas muy elementales que disminuyan el poder de discriminación que se persigue.⁹

Para el cálculo del indicador índice de discriminación (Di) se aplicó la fórmula:^{3, 7, 8}

Pasos seguidos para el cálculo de Di: se organizaron los resultados de los exámenes en orden decreciente, se calculó 27% de los resultados más altos y de los más bajos, se calculó el número que representa 46% restante, se aplicó la fórmula.^{3, 7, 8}

El Di se distribuyó según propuesta de Backhoff:⁸ mayor de 0,39 calidad excelente (Conservar), de 0,30 – 0,39 buena calidad (Posibilidades de mejorar), de 0,20 – 0,29 calidad regular (Necesidad de revisar), de 0,00 – 0,20 pobre calidad (Descartar o revisar a profundidad) y < -0.01 pésima calidad (Descartar definitivamente).

Coeficiente de discriminación: a diferencia del índice de discriminación, tiene en cuenta a todos los estudiantes evaluados, permite predecir sobre los resultados del estudiante, pues si el examen y una pregunta miden la misma habilidad o competencia, se puede esperar que quien tuviera una puntuación alta en todo el examen deberá tener altas probabilidades de contestar correctamente los temas evaluados en sus preguntas, como sucede en los coeficientes con cifras positivas. Los coeficientes de discriminación con cifras positivas tienen adecuado poder de discriminación y predicción, pero disminuye mientras más se acerque a cero. Un coeficiente cero o negativo, significa que la pregunta evalúa algo diferente a lo que evalúa el resto del examen. Cuando este coeficiente es muy bajo o negativo generalmente indica una pregunta con errores.^{1, 8, 10}

El cálculo del coeficiente de discriminación (r_{pbis}) se realizó por la fórmula de Backhoff.⁸

Según estándares internacionales la distribución de los resultados del coeficiente discriminación punto de correlación biserial (r_{pbis}) es: menor de 0 discriminan negativamente, de 0 - 0,14 discriminan pobremente, de 0,15- 0,25



discriminan regular, de 0,26 -0,35 buen poder discriminativo, mayor de 0,35 excelente poder de discriminación.

Los resultados del coeficiente de discriminación se distribuyeron en: ¹ adecuado (coeficiente positivo) e inadecuado (coeficiente negativo).

RESULTADOS

Tabla 1. Procedimientos y Técnicas utilizados en el Examen Ordinario de Células Tejidos y Sistema Tegumentario. Facultad de Ciencias Médicas de Granma. Curso 2019-2020.

Pregunta	Procedimiento	Técnica
1	Test Objetivo	Verdadero y Falso
2	Test Objetivo	Enlace o Apareamiento
3	Test Objetivo	Selección Múltiple
4	Test Objetivo	Complemento Simple
5	Ensayo	Respuesta Larga
6	Test Objetivo	Enlace o Apareamiento
7	Ensayo	Respuesta Corta

Fuente: Exámenes

En el presente estudio según se muestra en la Tabla 1 el examen estuvo constituido por 7 preguntas, predominaron las preguntas de tipo Test Objetivo o Respuesta Estructurada con un total de 5 que representa el 71,42% y dentro de estas las de enlace o apareamiento.

Tabla 2. Frecuencia de errores calculada al Examen Ordinario de Célula Tejido y Sistema Tegumentario. Facultad de Ciencias Médicas de Granma. Curso 2019-2020.

Pregunta	Frecuencia de Errores	%	Aprobados	%
1	38	9,33%	369	90,67%
2	22	5,40%	385	94,60%
3	64	15,72%	343	84,28%
4	70	17,19%	337	82,81%
5	152	37,34%	255	62,66%
6	112	27,51%	295	72,49%
7	31	7,61%	376	92,39%
Examen	67	16,46%	340	83,53%

Fuente: Procesamiento Estadístico



En la **Tabla 2** se aprecia que la pregunta 2 tuvo mayor número de aciertos con un total de 385 para un 94, 60% mientras que la pregunta 5 tuvo el mayor número de errores o menor número de aciertos con un total de 152 que representa el 37,34%. Resultaron aprobados en el examen 340 estudiantes que representa el 83,53% y desaprobados 67 estudiantes para un 16,46%.

Tabla 3. Dificultad calculada al examen de Célula Tejido y Sistema Tegumentario. Facultad de Ciencias Médicas de Granma. Curso 2019- 2020.

Pregunta	p	Dificultad
1	0,76	Medianamente Fácil
2	0,93	Altamente Fácil
3	0,54	Dificultad Media
4	0,53	Dificultad Media
5	0,30	Altamente Difícil
6	0,63	Dificultad Media
7	0,88	Altamente Fácil
Examen	0,69	Dificultad Media

Fuente: Procesamiento Estadístico **P:** Índice de Dificultad

En la **Tabla 3** según los resultados obtenidos del examen de acuerdo a su índice de dificultad se observó que de las 7 preguntas 3 (42,85%) fueron de dificultad media, 2 (28,57%) altamente fácil, 1 (14,28%) medianamente fácil y 1 (14,28%) altamente difícil. El examen de forma general calificó de dificultad media con un índice de 0,69.

Tabla 4. Poder de discriminación calculado a cada pregunta y al el examen general del curso 2019- 2020.

Pregunta	Poder de Discriminación			
	Di	Discriminación	rpbis	Discriminación
1	0,51	Excelente	0,69	Excelente
2	0,86	Excelente	0,60	Excelente
3	0,10	Pobre	0,66	Excelente
4	0,10	Pobre	0,66	Excelente
5	0,41	Excelente	0,64	Excelente
6	0,86	Excelente	0,63	Excelente
7	0,89	Excelente	0,63	Excelente
Examen	0,69	Excelente	0,62	Excelente

Fuente: Procesamiento Estadístico

Di: Índice de Discriminación

rpbis: Coeficiente de Discriminación



En la **Tabla 4** teniendo en cuenta el índice de discriminación se aprecia que 5 de las preguntas que representa el 71, 42% presentaron una discriminación excelente. El examen de forma general tuvo una discriminación excelente con un índice de 0,69.

Los resultados obtenidos al determinar el coeficiente de discriminación del punto biserial muestran que las 7 preguntas y el examen como un todo discriminaron de forma excelente, con valores entre 0,62 a 0,69. Este siempre fue positivo por lo que se catalogó como adecuado.

DISCUSIÓN

El examen ordinario estuvo compuesto por siete preguntas. Predominaron las preguntas de tipo test objetivo. En estas preguntas el educando elige su respuesta dentro de varias, no se puede negar que son fáciles y rápidas en su calificación, permiten evaluar un número mayor de estudiantes y un número mayor de contenido, pero deben ser elaboradas con rigurosidad.

En las revisiones realizadas se pudo encontrar autores como Carrazana A et al¹, que analizaron en sus investigaciones instrumentos evaluativos con formato mixto, resultados que evidencian la tendencia de elaboración de exámenes en este formato en ciencias médicas, Zayas¹¹, plantea que las preguntas más empleadas son las de test objetivo, pero contrariamente las preguntas de mayor preferencia por los estudiantes fueron las de ensayo.

Las autoras consideran que lo correcto sería utilizar exámenes de formato mixto, utilizando ambos tipos de preguntas y de esta forma aprovechar las ventajas de cada una y minimizar sus desventajas.

Al determinar la frecuencia de errores se observó que la pregunta 2 tuvo mayor número de aciertos con un 94, 60% o sea fue respondida correctamente por más examinados mientras que la pregunta 5 tuvo el mayor número de errores o menor número de aciertos con el 37,34%. Para el grupo de examinados la pregunta 5 es más difícil que la pregunta 2. Coincidiendo esto con el índice de dificultad calculado a estas preguntas. Resultaron aprobados en el examen 340 estudiantes que representa el 83,53% y desaprobados 67 estudiantes para un 16,46%.

En todas las preguntas se determinó el índice de dificultad siendo de dificultad media el 42, 85% de estas. El resultado difiere con lo que se desea obtener para un examen, pues como promedio deben oscilar en un rango de dificultad media a medianamente fácil.^{1, 8, 12, 13}

Acorde con la curva de distribución de frecuencias según las respuestas correctas dadas por los estudiantes, un 28, 57% considera fácil el examen, un



14, 28% medianamente fácil, el 42, 85% lo considera de dificultad media, y difícil lo considera el 14, 28%. El resultado difiere con lo que se desea obtener para un examen un 5% considera fácil el examen, 20% medianamente fácil, 50% con dificultad media, 20% medianamente difícil y 5% difícil.⁸

A consideración de las autoras no se realizó con la profundidad requerida la planificación del examen y esto pudo ser la causa de la diferencia de los valores porcentuales esperados y observados de los incisos fáciles y difíciles.

El promedio general de dificultad del examen analizado fue de 0,69, clasificándose de dificultad media lo cual provocó que más de 80% de los estudiantes lograran responder el examen satisfactoriamente. Según Carrazana,¹ el nivel medio de dificultad del examen debe oscilar entre 0.5 y 0.6, aunque Backhoff⁸ extiende el límite superior de este rango hasta 0,73, criterio asumido por los autores de este trabajo.

Es consenso para diferentes autores que han publicado sobre el tema que un examen que posea dificultad media, es superior, en confiabilidad de resultados, a aquellos que tengan mucha dificultad o que resulten muy fáciles, debido a que aumenta el criterio para diferenciar los estudiantes que dominan un contenido de los que no aún lo hacen. Es decir, exámenes muy fáciles o muy difíciles dan poco margen de discriminación, debido a que los estudiantes son igualados, en un examen muy fácil la mayoría tiene altos resultados, en un examen muy difícil la mayoría tiene muy bajos resultados, estas preguntas no contribuyen a la fiabilidad.^{1, 8, 10, 14}

El índice de discriminación osciló entre 0,41 y 0,89, así la mayoría de las preguntas presentaron excelente discriminación y un mínimo con pobre discriminación. Estas cifras muestran que el instrumento presenta buena eficacia para diferenciar los estudiantes que tienen buenos resultados de aquellos con resultados bajos, define sin dificultad aquellos estudiantes que tienen el conocimiento correcto de aquellos que no lo tienen, resultados diferentes se encontró en otros estudios afines.^{1, 8, 15, 16}

La excelente discriminación general del instrumento concuerda con la frecuencia de errores del examen lo que se corresponde con lo que se espera de los exámenes con una dificultad media. Estos resultados coinciden con investigaciones realizadas por Blanco Pereira y colaboradores.¹⁴

Las dos preguntas con índice discriminativo pobre (28, 27%) coincidieron en ser clasificadas de dificultad media según su índice de dificultad, fueron elaborados de tipo test objetivo, con poca profundidad y sus resultados no permitieron predecir que los estudiantes que obtuvieron altas puntuaciones en el examen deben haber obtenido altas calificaciones en las preguntas.



Es criterio de las autoras que las preguntas de dificultad media que no discriminaron adecuadamente, debe valorarse la necesidad de descartar o revisar con profundidad para lograr reelaborarlas con mayor complejidad, y puedan predecir que el estudiante que salió bien en el examen obtuvo buenos resultados en las preguntas.

Se debe tener presente que cuando se hace necesario la revisión de un examen se verifique: si las preguntas responden a los objetivos generales a evaluar, si otras respuestas predominantes serían igualmente correctas, si existen problemas de redacción que hagan que el estudiante interprete diferente a la respuesta correcta, o si constan preguntas con respuestas muy elementales o muy difíciles que disminuyan el poder de discriminación que se persigue.^{1, 8}

El coeficiente de discriminación, el cual tiene en cuenta la totalidad de los estudiantes, mostró cifras positivas en todas las preguntas y en el examen de forma general, lo cual revela que todas las preguntas fueron adecuadas, estos resultados coinciden con las cifras promedio que se buscan en un examen cuando se quiere valorar la capacidad de predicción de un instrumento.^{1, 8} Así, se puede inferir que las preguntas miden el mismo conocimiento o habilidad que está midiendo el examen total, es decir, las preguntas permitieron predecir que el estudiante que salió bien en el examen debe haber salido bien en cada una de las preguntas, lo cual coincide con diferentes estudios publicados,¹ pero se diferencia de Backhoff,⁸ que reportó al menos una pregunta que no discrimina ni permite predecir.

CONCLUSIONES

La calidad del examen Célula Tejidos y Sistema Tegumentario es adecuada a partir del análisis de la coherencia entre las preguntas, el nivel de dificultad y el poder de discriminación. Sugerimos incorporar a estas investigaciones métodos cualitativos que permitan valorar la calidad de los instrumentos, así como trabajar la línea de la evaluación del aprendizaje en las actividades metodológicas de los colectivos de asignaturas que integran la disciplina Bases Bilógicas de la Medicina y de esta forma profundizar en el diseño de los instrumentos de evaluación, que favorecen la eficiencia del proceso docente educativo.



REFERENCIAS BIBLIOGRÁFICAS

1. Carrazana Lee Armando, Salas Perea Ramón Syr, Ruiz Salvador Ana Karelia. Nivel de dificultad y poder de discriminación del examen diagnóstico de la asignatura Morfofisiología Humana I. Educ Med Super [Internet]. 2011 Mar [consultado 4 de Junio de 2021]; 25(1): 103-114. Disponible en: http://scielo.sld.cu/scielo.php?script=sci_arttext&pid=S086421412011000100010&lng=es
2. Chaviano Herrera Orlando, Baldomir Mesa Tamara, Coca Meneses Olga, Gutiérrez Maydata Alfredo. La evaluación del aprendizaje: nuevas tendencias y retos para el profesor. EDUMECENTRO [Internet]. 2016 Dic [consultado 6 de Junio de 2021]; 8(4): 191-205. Disponible en: http://scielo.sld.cu/scielo.php?script=sci_arttext&pid=S207728742016000700014&lng=es
3. Salas Perea, Ramón Syr, Salas Mainegra, Arlene (2017). Modelo formativo del médico cubano. Bases teóricas y metodológicas. La Habana: Editorial Ciencias Médicas.
4. 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 [Internet]. 2017 [consultado 10 de Junio de 2021]; 6(21). Disponible en: <https://reader.elsevier.com/reader/sd/pii/S2007505716300060?token=51734986279B99F4E310778549CB0CCF24ABF493ACE983DF0F23A795665B6EF8CA4924C2EDEA23E4BED3FFEFEC81EF2E>
5. Sánchez Hernández E, Medina Pavón M, Rodríguez García M, Vega Van Der Meer L, de la Torre Vega G. Indicadores de calidad para un examen teórico de la especialidad de medicina general integral. MEDISAN [Internet] 2015 [consultado 18 de Junio de 2021]; 19 (2). Disponible en: http://scieloprueba.sld.cu/scielo.php?script=sci_arttext&pid=S1029-30192015000200002&lng=es&nrm=iso
6. 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. [Internet]. 2018 Sep [consultado 24 de Junio de 2021]; 22(3). Disponible en: http://scielo.sld.cu/scielo.php?script=sci_arttext&pid=S156043812018000300009&lng=es



7. Fernández Jiménez M, González Jardines M, Dovale Borjas A, Rodríguez Pérez I, Iglesias Ramírez B, González Aguilar V. Programa de Célula Tejidos y Sistema Tegumentario. Comisión Nacional de Carrera de Medicina. Julio 2019.

8. Backhoff E, Larrazolo N, Rosas M. Nivel de dificultad y poder de discriminación del Examen de Habilidades y conocimientos Básicos (EXHCOBA). Revista Electrónica de investigación Educativa [Internet]. Mayo 2000. [consultado 4 de Junio de 2021]; 2(1). Disponible en: <http://redie.uabc.mx/vol2no1/contenido-backhoff.html>

9. Carrazana Lee A, Álvarez Bustamante G, Quesada Rodríguez M, Hidalgo Cerito Y. Dificultad y discriminación de exámenes ordinarios de la asignatura Célula, Tejidos y Sistema Tegumentario en Ciencias Básicas. Rev haban cienc méd [Internet]. 2018 [consultado 1 de Julio de 2021]; 17(2): [278-289]. Disponible en: <http://www.revhabanera.sld.cu/index.php/rhab/article/view/2149>

10. Morales P. Análisis de ítems en las pruebas objetivas. Universidad Pontificia Comillas. Madrid. [Internet]. 2009 [consultado 10 de Julio de 2021]. Disponible en: <https://educrea.cl/wp-content/uploads/2014/11/19-nov-analisis-de-items-en-las-pruebas-objetivas.pdf>

11. Zayas González M. Estudio descriptivo de los instrumentos evaluativos en la asignatura Farmacología. [Trabajo de terminación de Diplomado en Educación Médica Superior]. Villa Clara: Instituto Superior de Ciencia Médica "Serafín Ruíz de Zarate Ruiz"; 2002.

12. Argudín Somonte E, Díaz Rojas P, Leyva Sánchez E. Índice de Dificultad del examen de Morfofisiología Humana I. Super [Internet]. 2011 Jun [consultado 20 de Julio de 2021]; 25 (2): Educ Med 97-106. Disponible en: http://scielo.sld.cu/scielo.php?script=sci_arttext&pid=S08642141201100020007&lng=es

13. García Reyes XM, Damiani Cavero JS, Hernández González BL. Caracterización del examen final ordinario de Morfofisiología III, por el índice de dificultad. Polo docente Julio Trigo-Enrique Cabrera 2009-2011. Rev. Cong Univ. [Internet]. 2016 [consultado 25 de Julio de 2021]; 5(3). Disponible en: <http://www.congresouniversidad.cu/revista/index.php/rcu/article/view/746>

14. Blanco Pereira M.E, Martínez L, González Gil A, Jordán Padrón M. Calidad del examen final teórico de Morfofisiología Humana I en la Facultad de Ciencias Médicas de Matanzas. Cursos 2012-2013 y 2013-2014 Rev. Med. Electrón. [Internet] 2015 [consultado 4 de Junio de 2021]; 37(4). Disponible en: http://scielo.sld.cu/scielo.php?script=sci_arttext&pid=S16841824201500040003&lng=es



15. Ortiz Romero GM, Díaz Rojas PD, Llanos Domínguez OR, Pérez Pérez SM, González Sapsin K. Dificultad y discriminación de los ítems del examen de Metodología de la Investigación y Estadística. Rev EDUMECENTRO [Internet]. 2015 [consultado 3 de Agosto de 2021]; 7(2). Disponible en: http://scielo.sld.cu/scielo.php?script=sci_abstract&pid=S2077-28742015000200003

16. Martínez Pérez R, Ríos Rodríguez M, Roque Marrero Y, Caballero Padrón KY. Evidencias de validez del examen final de la asignatura Biología molecular, carrera de Medicina. Medimay [Internet]. 2019 [consultado 18 de agosto de 2021]; Sep-Dic; 26(3):338-48. Disponible en: <http://www.medimay.sld.cu/index.php/rcmh/article/view/1483>