

**Universidad de Ciencias Médicas de Holguín
Facultad de Ciencias Médicas “Mariana Grajales Coello”**

Maestría en Educación Médica.

Título: Calidad del examen final de la asignatura
Morfofisiología Humana II. Curso 2010-2011.

Autora: Dra. Kariné González Sapsin

Especialista de Segundo Grado en Medicina General Integral

Tutora: Lic. María Elisa Sondón Fernández

Prof. Auxiliar. MSc. en Educación Médica

Asesor: Dr. Pedro Augusto Díaz Rojas

Profesor Titular. MSc. en Educación Médica

**TRABAJO PARA OPTAR POR EL TÍTULO DE MASTER EN EDUCACIÓN MÉDICA
Año 2012**

DEDICATORIA

A mi pequeño hijo Luís Ángel, por ser la luz que ilumina mi vida.

A mis padres, porque sin ustedes no sería quien soy.

A mí querido esposo, quien con su amor y ayuda en el plano profesional me ha dado la fuerza necesaria para lograr este propósito.

AGRADECIMIENTO

A mi tutora María Elisa, por su guía y orientación acertada para el logro de este proyecto.

Al profesor Pedro Díaz Rojas, por sus sabias y valiosas recomendaciones que permitieron la realización de este trabajo.

Al claustro de profesores de la Maestría en Educación Médica, por toda la experiencia y conocimiento aportado.

RESUMEN

El presente estudio es un trabajo de investigación en el campo del Proceso de Enseñanza Aprendizaje, cuyo objeto es el instrumento evaluativo de la asignatura Morfofisiología Humana II y el objetivo general de la investigación es caracterizar la calidad del examen de esta unidad curricular aplicado a estudiantes del primer año de la carrera de medicina de la Facultad de Ciencias Médicas de Holguín en el curso escolar 2010 – 2011. Se emplearon métodos teóricos, empíricos y procedimientos estadísticos que garantizaron la triangulación de la investigación. Se describieron las áreas de las ciencias básicas y tipo de preguntas utilizadas en el examen, se identificó la correspondencia entre el fondo de tiempo de los contenidos de la unidad curricular y su exploración en el examen y se realizó el análisis de los indicadores de calidad: Índice de dificultad, índice de discriminación D, correlación del punto biserial y Alfa de Cronbach; dentro de las características principales del instrumento escrito sobresalen que se exploraron contenidos básicos de Embriología, Anatomía Humana, Fisiología e Histología, con predominio de Anatomía Humana, el mayor por ciento fue de preguntas de ensayo y no existió correspondencia entre el fondo de tiempo dedicado a cada tema y la cantidad de ítems presentes en el examen. La valoración de los indicadores de calidad mostraron que el instrumento según el índice de dificultad fue clasificado de fácil, donde existió correspondencia entre el índice de dificultad por preguntas con su discriminación. El coeficiente alfa de Cronbach mostró confiabilidad en su consistencia interna.

INDICE

	Página
1. INTRODUCCIÓN	1
2. MARCO TEORICO	6
3. OBJETIVOS	30
4. DISEÑO METODOLÓGICO	31
5. ANÁLISIS Y DISCUSIÓN DE LOS RESULTADOS	36
6. CONCLUSIONES	51
7. RECOMENDACIONES	52
8. REFERENCIAS BIBLIOGRÁFICAS	53
9. ANEXOS	

INTRODUCCIÓN

Formar al hombre nuevo que asuma el dinamismo de la era del conocimiento ha provocado la implementación de reformas universitarias en algunas acciones, con la pretensión de adecuar los modelos educativos a las exigencias de competitividad que se requieren. Estos cambios, se enfocan hacia la transformación de la rigidez disciplinaria y hacia el desarrollo del proceso pedagógico centrado en el aprendizaje.¹

La sociedad contemporánea requiere de cambios cualitativos en el sistema de salud, a fin de incrementar la satisfacción de las necesidades de la población y de la comunidad, mediante el empleo óptimo de los recursos humanos y materiales existentes, esto nos plantea la necesidad de promover la formación de un nuevo profesional.^{2, 3} La Universidad está llamada a formar personas capaces de responder a los problemas de la sociedad, su misión es, sobre todo, formar hombres con capacidad de evolucionar a lo largo de toda la vida y de situarse frente a los problemas de la misma.^{4, 5}

La Educación Médica Superior no está exenta de esta misión, ya que ella es la encargada de formar un profesional que responda a las necesidades de salud de la sociedad, que trabaje en la solución de los problemas médicos del individuo, familia y la comunidad y a partir de la promoción de salud revierta los índices de enfermedades.⁴

Es por lo anteriormente planteado que surge la necesidad de formar un profesional con competencias diagnósticas y terapéuticas capaz de brindar atención médica integral, mediante las acciones básicas no solo al individuo sino también a la familia, la comunidad y el medio ambiente; portador además de valores éticos, humanísticos, solidarios y de actitud ciudadana; llamado a transformar la situación de salud existente en correspondencia con las exigencias de la sociedad actual.

La eficiencia del proceso docente educativo se expresa precisamente en graduados capaces de cumplir con la máxima calidad, el encargo que le plantea la sociedad. Por ello es imprescindible que a lo largo de toda la carrera se compruebe en qué medida se preparan los educandos para enfrentar exitosamente su responsabilidad social y si su aprendizaje se corresponde con los objetivos de cada asignatura,

disciplina, año de estudio y la carrera en general. 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, y como parte de ella, el control y la valoración de los modos de actuación que los estudiantes adquieren a través del desarrollo del proceso docente.⁶

Evaluar es un proceso difícil y complejo, cuya ausencia simplificaría el proceso enseñanza - aprendizaje, pero le cortarían el termómetro de análisis de su calidad y obligaría a los docentes a permanecer en la oscuridad, al ignorar los resultados de sus esfuerzos en lograr la formación del estudiante y la transmisión del conocimiento científico. La evaluación tiene una serie de sesgos que hacen la tarea difícil, sin embargo aporta retroalimentación constante, valoración formativa y favorece que el educando madure y desarrolle sus propias técnicas para la evaluación de su estudio, su trabajo y su aprendizaje.⁷

En esto radica la importancia de la evaluación del aprendizaje de los estudiantes, la dirección efectiva del proceso docente educativo y la determinación de su grado de eficiencia; pero a su vez también en ello radica su complejidad.⁸

La evaluación es un proceso socialmente necesario, algo inherente al hombre que como ser social, al realizar su valoración, refleja en su conciencia la significación que tienen objetos y fenómenos de la realidad. Es necesario que el sistema de evaluación sea lo más sólido posible y que se apliquen instrumentos evaluativos fiables, pues evaluar es una acción eminentemente humana como habíamos expresado anteriormente donde esta presente, en algún grado, el subjetivismo de los implicados.

El origen de la evaluación vinculada a la actividad docente se pierde en el tiempo, se hace referencia a exámenes escritos y orales en la antigua China, para seleccionar funcionarios en la administración imperial y reducir así la influencia de la burocracia.⁹

La evaluación aparece en la didáctica como una necesidad sentida en la Edad Media, ligada a la concesión de grados, del aspirante a artesano para obtener la calificación de enseñante y a partir del siglo XIII se formaron corporaciones de maestros que constituían los “colegios de doctores” que tenían el derecho de conferir títulos “*ius promovendi*”, constituyendo esto uno de los primeros pasos a la formación

de la universidad en el sentido en que hoy se comprende; así en 1231 se decretó que nadie ejercería la profesión médica sin antes haber rendido exámenes en la escuela de Salerno, que de ese modo adquiriría atributos de universidad.⁶

Al hablar de la historia de la evaluación en Cuba debemos referirnos a que existen datos de que la misma tuvo sus inicios en la época colonial, con la aplicación de formas de comprobación que se les aplicaban a los esclavos, aprendices y artesanos.¹⁰ En la etapa de neo-república no se produjeron cambios con relación a la evaluación, continúan los exámenes en las escuelas, las escalas de calificaciones apenas se transforman, aunque existieron preocupaciones por parte del profesorado en este sentido.

Durante la primera mitad del siglo XX y hasta la década de los 60, la función de la evaluación fue la de comprobar los resultados del aprendizaje, ya se tratase en términos del rendimiento académico o del cumplimiento de los objetivos propuestos.¹¹

En la década del 50 el panorama educacional era bastante pobre y carente de un sistema científico coherente, matizado por la casi absoluta despreocupación de los gobernantes de turnos. Esta situación cambia radicalmente con el triunfo de la Revolución Cubana el 1° de Enero de 1959 y la aplicación en los primeros años del proyecto de alfabetización, al cual le siguieron otros proyectos educacionales que continúan hasta nuestros días.

En la actualidad existen profesores e investigadores cubanos que han dedicado gran parte de su trabajo a la evaluación, entre los que se destacan: Salas, Arteaga, Hatin y otros.¹²

En Cuba y en otras partes del mundo, se han iniciado planes de estudio de medicina con un esquema "integrador" de las ciencias básicas biomédicas a través de una disciplina académica denominada Morfofisiología Humana, que agrupa contenidos correspondientes a las disciplinas científicas de Anatomía Humana, Embriología, Histología, Bioquímica, Fisiología y Genética que siguen, en lo esencial, un eje integrador por aparatos y sistemas.^{13- 15.}

Uno de los pilares fundamentales para conseguir un adecuado desarrollo de la capacidad integradora lo constituye la evaluación y es justamente en ella donde encontramos más dificultades.¹³

La asignatura Morfofisiología Humana II se desarrolla en el primer año de la carrera de medicina, se parte del sistema de objetivos generales establecidos, como necesario nivel de análisis para la concepción de los diferentes instrumentos de evaluación y control, según las características del contenido a asimilar. En la medida en que se logre el cumplimiento de los objetivos propuestos, la evaluación constituirá un elemento de retroalimentación y de dirección continua del proceso docente.¹⁶

El control final de una asignatura (o disciplina) comprueba el logro de sus objetivos más generales y esenciales, cumple las distintas funciones de la evaluación y su forma más usual es el examen final.¹⁷

El proceso educacional puede y debe ser mejorado regularmente, la presencia de los estudiantes en el proceso, en las fases de planeamiento, ejecución y evaluación, es la mejor forma de dar un marco de realidad y de asegurar la posibilidad de mejorar. La evaluación debe ser hecha en relación al perfil profesional y al marco de los objetivos y metas.¹⁸

Para lograr una buena calidad en los instrumentos evaluativos, estos no deben ser contruidos de forma empírica y de esta forma garantizamos que midan el cumplimiento de los objetivos educacionales.¹⁹

En esto radica la importancia del planeamiento de un examen que permite el balance de preguntas y problemas, de modo que los diferentes componentes del contenido y su nivel de asimilación, estén representados en proporción a la importancia que les confiere el grupo que construye el instrumento de evaluación.¹⁹

Toda investigación que aporte resultados que permitan realizar acciones que garanticen elevar la calidad de los educandos, adquiere relevancia y aún más si permiten perfeccionar el sistema evaluativo. La importancia teórica, práctica y social de este trabajo se sustenta en las ideas anteriores, el mismo se relaciona con el perfeccionamiento del proceso enseñanza-aprendizaje en el ámbito de la asignatura Morfofisiología Humana II.

Por todo lo expresado anteriormente nos planteamos el siguiente **problema científico: No existe una caracterización científicamente fundamentada sobre la calidad de los instrumentos evaluativos de la asignatura Morfofisiología Humana II.**

La caracterización sobre bases científicas de la calidad de la evaluación del aprendizaje de esta asignatura, permitió identificar las dificultades que se presentan en la misma y servirá de referente para trazar estrategias que propicien corregir las desviaciones encontradas. Sus resultados permitirán realizar acciones capacitantes dirigidas a los profesores y así lograr el perfeccionamiento de la evaluación de esta asignatura que resulta vital para el buen desarrollo docente de cualquier institución.

MARCO TEORICO

El proceso enseñanza-aprendizaje constituye un verdadero par dialéctico en el cual el primer componente debe organizarse y desarrollarse de manera tal que facilite la apropiación del conocimiento a partir de la realidad objetiva.

Como consecuencia del proceso de enseñanza, ocurren cambios sucesivos e ininterrumpidos en la actividad cognoscitiva del estudiante; con la ayuda del maestro o profesor, que dirige su actividad conductora u orientadora hacia el dominio de los conocimientos, así como a la formación de habilidades, hábitos y valores acordes con su concepción científica del mundo.⁹

La autora de este trabajo comparte el criterio con otros autores que este proceso constituye un sistema en el que cada uno de sus componentes: Objetivos, contenidos, métodos, medios y evaluación, se interrelacionan, pero donde los objetivos constituyen la categoría rectora.

La evaluación es el mecanismo regulador del sistema y puede considerarse como el instrumento de control de la calidad del producto resultante del proceso.²⁰

Los objetivos determinan los contenidos, sistema de conocimientos y habilidades que se concretan en determinados modos de actuación; los que se enseñan a los estudiantes mediante las diversas actividades docentes, al aplicar y desarrollar determinados métodos de enseñanza y auxiliados con el empleo de los diferentes medios de enseñanza.²¹

Todo ello conduce por tanto al establecimiento para cada nivel, carrera, disciplina, semestre y asignatura de un sistema de evaluación bien planificado, estructurado e interrelacionado, en cuya determinación participan activamente la facultad, los comités verticales y horizontales, los departamentos docentes y los colectivos de asignaturas (o disciplinas).¹⁷

La evaluación del aprendizaje es una parte esencial del proceso de enseñanza-aprendizaje que nos posibilita su propia dirección, así como el control y la valoración de los modos de actuación que los estudiantes adquieren a través del desarrollo del proceso docente, al comprobar el grado con el que se alcanzan los objetivos propuestos. La evaluación del aprendizaje, así definida, en su sentido amplio, abarca tanto el control como la valoración de sus resultados.¹⁷

Como se ha planteado anteriormente en la educación médica superior cubana, los objetivos educacionales constituyen la categoría rectora del proceso docente-educativo, pero no hay dudas de que la evaluación influye en lo que enseñamos y determina de manera directa lo que los educandos aprenden.¹⁷

Se denomina evaluación al juicio con que culmina el análisis del logro de los objetivos. Este juicio o evaluación se expresa en una calificación que lo define. Las calificaciones son formas convencionales establecidas que expresan en categorías el resultado de la evaluación, lo que permite ordenar o clasificar el rendimiento de los estudiantes.¹⁷

Para expresar las diferentes categorías se utilizan escalas, las cuales pueden ser simbólicas o numéricas, existen diversos criterios acerca de estas y lo cierto es que ambas tienen sus ventajas y desventajas. Lo esencial radica en que la calificación sea objetiva, que exprese una correcta correspondencia con la calidad de la asimilación y por lo tanto, la necesidad de establecer criterios unitarios, que disminuyan la influencia del factor subjetivo, y que frente a una actuación determinada, con iguales resultados, el grupo de profesores otorguen iguales calificaciones.¹⁷

Son muchos los autores que se han detenido en el estudio de la evaluación desde distintos enfoques y puntos de vistas en el proceso de enseñanza y también en el proceso pedagógico profesional, caracterizado este último, por la implicación que el sujeto que aprende debe tener en relación a su preparación y la condicionalidad que se presenta en relación con los valores políticos, sociales, morales, y laborales necesarios en él, que fortalecen el desarrollo de las cualidades de la personalidad.⁹

Salas Perea define la evaluación como la actividad que consiste en el análisis e interpretación de los resultados de diferentes instrumentos de medición sobre la actuación de un educando o candidato ante una actividad y/o tarea, en comparación con un patrón o modelo estandarizado, y que se efectúa con la finalidad de tomar una buena decisión.¹²

Según el colectivo de autores del Ministerio de Educación de Cuba (MINED) bajo la dirección del Instituto Central de Ciencias Pedagógicas, la evaluación se ha definido en su sentido más amplio, como un componente esencial del proceso de enseñanza,

que parte de la definición misma de los objetivos y concluye con la determinación del grado de eficiencia del proceso; dada por la medida en que la actividad del educador y el alumno haya logrado como resultado los objetivos propuestos.¹²

Otros autores conciben la evaluación como un proceso objetivo y continuo, que se desarrolla en espiral, y que consiste en comparar la realidad con un modelo, de manera que los juicios que se obtengan de esta comparación actúen como información retroalimentadora, que permita adecuar el plan de estudio a la realidad o cambiar parte de esta.¹²

González Pérez M, define la evaluación del aprendizaje, como “(...) actividad cuyo objetivo es la valoración del proceso y resultados del aprendizaje de los estudiantes a los efectos fundamentales de orientar y regular la enseñanza para el logro de las finalidades de la formación”, plantea además que la mayoría de las definiciones actuales coinciden en reconocer, como procesos básicos de la evaluación, la recogida de información y la emisión de un juicio valorativo.^{22, 23.}

Según Giovanis Lafrancesco “la evaluación es un proceso sistemático y permanente que comprende la búsqueda y obtención de información de diversas fuentes acerca de la realidad del desempeño, avance, rendimiento o logro del estudiante, de la calidad de los procesos empleados por los docentes, la organización y análisis de la información a manera de diagnóstico, la determinación de su importancia y pertenencia de conformidad con los objetivos que se esperan alcanzar, todos con el fin de tomar decisiones que orienten el aprendizaje y los esfuerzos de la gestión docente”.¹²

La resolución 210-207 expresa que la evaluación es un proceso consustancial al desarrollo del proceso docente educativo. Tiene como propósito comprobar el grado de cumplimiento de los objetivos formulados en los planes y programas de estudio de la educación superior, mediante la valoración de los conocimientos y habilidades que los estudiantes adquieren y desarrollan; así como, por la conducta que manifiestan en el proceso docente educativo. Constituye a su vez, una vía para la retroalimentación y la regulación de dicho proceso.²⁴

Este concepto planteado anteriormente se ajusta perfectamente a las intenciones de la autora. En resumen la evaluación es la valoración del logro de los objetivos

formativos y se refiere a la eficiencia del proceso de asimilación y a la calidad de sus resultados.

La evaluación del aprendizaje en la educación superior tiene un carácter continuo, cualitativo e integrador; y debe estar basada fundamentalmente, en el desempeño del estudiante durante el proceso de aprendizaje. Se debe desarrollar de manera dinámica, donde no solo evalúe el profesor, sino que se propicie la participación de los estudiantes mediante la evaluación grupal y autoevaluación, para lograr un ambiente comunicativo en este proceso.^{24, 25}

A nivel mundial las funciones de la evaluación del aprendizaje ocupan un lugar importante en la literatura pedagógica. En los diferentes trabajos revisados^{12, 21, 26- 28} no se estructuran de la misma manera, aunque las ideas que expresan los diferentes autores son similares.

Parra Pineda MO y González Pérez M^{20, 22} las describen de la siguiente forma:

Social: Tienen que ver con la certificación del saber, la acreditación, la selección y la promoción. Los títulos que otorgan las instituciones educativas, a partir de resultados de la evaluación, se les atribuye socialmente la cualidad de simbolizar la posesión del saber y la competencia, en función de los valores dominantes en cada sociedad y momento.

De control: Observación y análisis de la realidad educativa. Por su significación social y sus implicaciones en la vida de los educandos, los resultados de la evaluación constituyen un instrumento potente de poder del evaluador sobre los evaluados.

Pedagógicas: Ellas incluyen las funciones de diagnóstico, pronóstico, de creación de ambiente escolar, de afianzamiento del aprendizaje, de recurso para la individualización, de retroalimentación, de motivación y de preparación de los estudiantes para la vida.

Salas Perea RS y Pérez Espinosa LM las reagrupan de forma sencilla y didáctica, así tenemos que la evaluación del aprendizaje cumple las siguientes funciones: Retroalimentación, instructiva, comprobación y control y educativa.^{17, 20}

Función de retroalimentación: Está directamente vinculada con la dirección del aprendizaje. En diferentes momentos del proceso docente (muestras, controles) establecemos cómo se encuentra el aprendizaje de los estudiantes (conocimientos,

habilidades, hábitos y modos de actuación) en correspondencia con los objetivos establecidos en la asignatura, semestre, año académico, ciclo y esta información retroalimenta al profesor, le señala como está el proceso de asimilación del contenido de la enseñanza. En dependencia de los resultados alcanzados, el profesor determinará las correcciones que es necesario introducir en la estrategia docente (contenido, métodos, ritmos, consultas, etcétera) para acercar cada vez más los resultados a las exigencias de los objetivos.

Por otra parte, el estudiante también recibe información que lo retroalimenta, acerca de su propio progreso en las transformaciones que debe haber logrado en el sistema de conocimientos, habilidades y en sus modos de actuación; le propicia el desarrollo de su autoevaluación y lo orienta hacia donde tiene que dirigir sus esfuerzos en el trabajo independiente y el auto estudio. Algunos autores designan esta función de retroalimentación como de diagnóstico y desarrollo del proceso de enseñanza-aprendizaje.

Función instructiva: La evaluación continua contribuye a incrementar, sistematizar, profundizar y generalizar los conocimientos y las habilidades, o sea, la actividad cognoscitiva de los estudiantes; propicia el trabajo independiente y el auto estudio, con lo cual crea las condiciones para la consolidación de los modos de actuación y hacer mas efectivo el proceso del aprendizaje. Algunos autores la denominan función lógico-cognoscitiva.

Función de comprobación y control: Nos informa sobre el grado en que se cumplen los objetivos de la enseñanza (formativos), o sea, nos permite controlar el estado del proceso docente y la efectividad y calidad de la enseñanza y del aprendizaje. La evaluación nos posibilita verificar si los estudiantes han adquirido el nivel de preparación requerido conforme con los objetivos establecidos y a su vez valorar la calidad de los planes y programas de estudio, así como de la estrategia empleada, lo que nos asegura el continuo perfeccionamiento del proceso docente educativo.

Función educativa: La evaluación correctamente organizada y aplicada, donde los estudiantes demuestran los resultados de su esfuerzo ante el profesor y su grupo de estudio, contribuye a que incremente su responsabilidad en el aprendizaje, se

plantee mayores exigencias y constituya un motor impulsor, consciente y voluntario en el aumento de la calidad del auto estudio y del trabajo independiente; todas sus acciones las convierte de un deber en un placer.

La autora opina que las funciones de la evaluación del aprendizaje se relacionan e interactúan dialécticamente entre sí, por lo que todas se consideran importantes.

La evaluación del aprendizaje en la educación superior puede producirse en tres momentos o etapas del proceso: Preliminar; durante el proceso de enseñanza-aprendizaje y diferido.⁸

El preliminar se realiza con el objetivo de evaluar el nivel de partida del estudiante, pueden ser exámenes de ingreso, donde se evalúa la posibilidad de un individuo de ser aceptado en un centro de educación superior o en determinada carrera y exámenes de diagnóstico previo al inicio de una asignatura o materia, en los que se evalúa el nivel de desarrollo inicial que tiene el estudiante.

El control durante el proceso de enseñanza-aprendizaje se realiza durante, o al final de una disciplina o asignatura, así como también al concluir la carrera, puede ser formativa y acumulativa. La evaluación formativa se desarrolla durante el proceso docente, la acumulativa o sumativa se desarrolla al concluir una asignatura, ciclo, carrera o especialidad.

El control diferido es el que se realiza después de concluida la enseñanza de una disciplina, asignatura, ciclo o carrera, tiene como objetivo el comprobar la retención de los conocimientos, habilidades y modos de actuación adquiridos en ella.

La aplicación de las diferentes formas de evaluación se puede realizar mediante múltiples métodos, procedimientos y técnicas. Los principales métodos empleados en la realización de los controles son: Teóricos, prácticos y teórico-prácticos.

Los métodos teóricos se emplean para evaluar los conocimientos y habilidades intelectuales, mientras que los prácticos se utilizan para evaluar las habilidades manuales o destrezas y los modos de actuación. En los controles prácticos se debe tener bien definido el objetivo de la observación y auxiliarse de técnicas como las listas de observación o comprobación, las que contienen la secuencia de las acciones que deben ser observadas y las escalas de calificaciones que contienen

series de formulaciones sobre las que se debe pronunciar el estudiante (escalas de Thurstone y las de Likert).²⁹

Numerosos autores (Bloom, Hastings y Madaus; 1975; Aray, 1983; Villarroel, 1990; Zabalza, 1991; Fernández Pérez, 1994; Casanova, 1999), han hecho importantes aportes a la clasificación, elaboración y aplicación de variadas formas de evaluación, en particular de los exámenes escritos.³⁰ Sin embargo, la autora comparte el criterio que este es un tema en desarrollo que obliga a todos los maestros, a replantearse nuevas concepciones para mejorar el diseño y la elaboración de los exámenes.

Los instrumentos evaluativos escritos pueden dividirse en dos grandes familias según la literatura:^{6,31,32} Los exámenes de desarrollo, de respuesta libre, tradicionales o de ensayo y los exámenes de respuesta estructurada o test objetivos.

Puede existir combinación de preguntas de uno u otro tipo y entonces se denominaría examen mixto.

Esta clasificación, según varios autores se basa en la categoría de las preguntas o "ítems", que puedan ser de respuestas libres o de ensayo y de respuestas estructuradas o test objetivos.

En el primer tipo, el sujeto aporta su propia respuesta, como por ejemplo, la pregunta de ensayo, que requiere de una respuesta extensa; la pregunta de respuesta corta, que no necesita más que una o dos oraciones para responderla y la de completamiento, que sólo requiere de una palabra o frase para responderse.

Estas preguntas, llamadas también de desarrollo o tradicionales, son muy útiles ya que miden la capacidad del individuo para relacionar, seleccionar y organizar el material, así como su habilidad para expresar las ideas con calidad y precisión.

La autora está de acuerdo en que las pruebas de desarrollo permiten conocer la profundidad en el dominio de una temática; pero tienen la dificultad que es escasa la muestra del contenido a evaluar y la subjetividad del evaluador influye en gran medida en la puntuación otorgada al sujeto, ya que es mucho el tiempo requerido para responderla y no permiten la estandarización de los resultados al ser prácticamente imposible elaborar claves de calificación.

Como resumen podemos señalar que dentro de las ventajas que presentan este tipo de preguntas tenemos: Posibilita que el educando exponga sus conocimientos,

ordene sus ideas y demuestre su capacidad para expresarse en su idioma, y dentro de las desventajas: Limita sensiblemente el número de áreas a explorar, la objetividad es difícil de conseguir, proporciona escasa retroalimentación al educando y al claustro, y su calificación consume demasiado tiempo, es subjetiva y sus resultados poco confiables.

En la literatura revisada se describen indicaciones generales para su uso como: Que el grupo de estudiantes sea pequeño, los objetivos del control contemplen evaluar las habilidades de autoexpresión escrita, que los examinadores se sientan satisfechos con evaluar a través de pocas preguntas una muestra de los contenidos que deben ser evaluados, que los examinadores tengan experiencia y se aplica cuando es más importante evaluar la libre expresión del educando o candidato que la objetividad y confiabilidad de una calificación.

Se recomienda: No emplear muchas preguntas o preguntas que sean demasiado extensas, si se usa varias preguntas, establecer una gama de complejidad o de dificultad entre ellas, preparar con antelación el esquema de calificación con su clave bien detallada para cada posible respuesta y redactar un conjunto de orientaciones generales para la prueba con enunciados que abarquen el plan general por el que han de regirse los examinados al resolver la prueba, la forma en que debe escribirse la respuesta (en prosa coherente o esquema) y el tiempo que dispone para resolver el examen.

Las preguntas de ensayo de respuestas cortas algunos autores las consideran como preguntas puramente de ensayo y otros como un tipo de pregunta situada entre las preguntas de ensayo y las de test objetivo, de ahí que se denominan también semiobjetivas. Éstas reducen las desventajas de las preguntas de tipo de ensayo, incrementan la muestra de contenido a evaluar y se limitan las preguntas sólo a aquellas que conllevan respuestas libres, pero cortas.

Las ventajas descritas para las preguntas de ensayo de respuestas cortas son:^{6, 31, 32.}

Fáciles de confeccionar, se incrementan las áreas de materias a examinar al aumentar el número de preguntas, los examinadores pueden ponerse de acuerdo sobre las posibles respuestas a aceptar como correctas y en cómo calificarlas y la objetividad y confiabilidad de las calificaciones se incrementa considerablemente.

En la literatura revisada se describen desventajas:^{6, 31, 32.} La amplitud de las áreas/materias a examinar es aún limitada, la calificación consume mucho tiempo y persiste el juicio subjetivo alrededor del límite de lo correcto/incorrecto de las respuestas emitidas por los educandos y las respuestas solicitadas son frecuentemente reproductivas más que de razonamiento y discusión, las preguntas pueden ser triviales o insignificantes de acuerdo con los objetivos propuestos.

La redacción de este tipo de pregunta es de gran importancia y se sugieren una serie de recomendaciones para la confección de las mismas, entre las que se destacan:^{6, 31, 32.}

- Ampliar el número de preguntas a realizar con el fin de incrementar el número de áreas o materias a evaluar.
- Cada ítem debe tratar contenidos importantes y no triviales.
- Usar un lenguaje preciso en relación con la materia a explorar, para evitar las proposiciones indefinidas o abiertas.
- Si el problema requiere de una respuesta numérica, deben señalarse las unidades en que debe expresarse.
- En el caso de preguntas de completamiento deben omitirse sólo las palabras claves, sin dejar muchos espacios en blanco.
- Los espacios vacíos deben ponerse al final del enunciado de modo que parezca una pregunta común.
- No deben sacarse oraciones de un texto para realizar la pregunta o hacer énfasis sólo en la memoria literal del texto.
- No deben ponerse oraciones textuales de los libros.
- Elaborar las claves de respuestas bien detalladas con sus correspondientes calificaciones.

Las preguntas de ensayo modificado constituyen otra variante de preguntas de ensayo cuyo formato se basa en un problema o situación donde los elementos son presentados en etapas sucesivas. Se utilizan como contenidos fundamentales problemas o situaciones de salud.

A partir de la presentación de la situación problema inicial se formulan una serie sucesiva de etapas / problemas que constan de una batería de preguntas de ensayo

tradicional o de ensayo de respuestas breves acerca del desarrollo de la situación, donde las respuestas en cada etapa son relativamente libres, ya que no son sugeridas al evaluado.³²

Las preguntas de respuestas estructuradas o test objetivos, son aquellas en que el educando elige su respuesta dentro de varias, proporcionadas por el elaborador del examen.

Este tipo de preguntas es útil cuando se examina un grupo grande de estudiantes, amplio contenido muestral a evaluar, necesidad de obtener calificaciones muy confiables, evaluación imparcial, justa y libre del “efecto de halo” con necesidad de divulgar los resultados con prontitud.

La autora coincide con muchos autores que el examen tipo test objetivo, es útil y fácil de calificar, pero no debe abusarse del mismo.

Existen sugerencias generales para la elaboración de este tipo de preguntas dentro de las que tenemos:³² Desarrollar el ítem a partir de ideas relevantes, seleccionar temas o ideas que permitan la elaboración de ítems con el máximo poder de discriminación, elaborar ítems que midan objetivos importantes, presentar el ítem con la máxima claridad posible, evitar la inclusión de elementos no funcionales, evitar la construcción de ítems basados en elementos demasiado específicos, adaptar la dificultad del ítem al nivel del grupo, eliminar la fraseología estereotipada, seleccionar una muestra de los aspectos que constituirán el examen y definir cuales técnicas o tipos de test, cómo escribir y presentar cada ítem, cómo ordenar los ítems, cómo administrar el examen, cómo evaluar e interpretar la calificación y cómo determinar la validez y confiabilidad.

Dentro de las variantes de este tipo de preguntas tenemos las de selección múltiple, de respuestas alternativa (verdadero-falso) y el de igualamiento o pareación.

Salas Perea refiere, que las preguntas de selección múltiple son las más flexibles y eficaces y pueden ser utilizadas para evaluar conocimientos, vocabularios, comprensión, aplicación de principios, o habilidades para interpretar datos.³¹

La estructura de las mismas está compuesta por dos partes fundamentales: La guía o núcleo que presenta el problema (también llamada base, cuerpo, raíz o tallo) y la lista de posibles respuestas alternativas (también se le denomina complemento).

Se describen tres variantes de este tipo de preguntas:

Selección múltiple tipo complemento simple: Aquí hay una sola alternativa sencilla para seleccionar. Esta puede ser complemento simple positivo y complemento simple negativo.

Selección múltiple tipo complemento agrupado: En este tipo de preguntas se ofrecen un grupo de aseveraciones y las alternativas son combinaciones de ellas, son más elaboradas y requieren de procesos mentales a la hora de seleccionar una alternativa.

Selección múltiple tipo simulación y análisis progresivo: Es el tipo más complejo y requiere de un meticuloso esfuerzo en su elaboración.

A las preguntas de análisis progresivo se les critica que si la respuesta a la primera pregunta es errónea, incrementan las posibilidades subsecuentes de error, pero debemos tener en cuenta que si el problema es relevante, este “peso extra” puede ser justificado.

Las consideraciones generales a tener en cuenta para confeccionar ítems de selección múltiple son:^{31, 32.}

- La guía puede aparecer en forma de proposición incompleta o en forma de pregunta.
- Debe poseer de 4 a 6 ítems o alternativas en cada pregunta para reducir el acierto por azar.
- El núcleo debe incluir todo lo que se pueda del texto de la pregunta, de modo que sea relativamente largo y las diversas alternativas lo más corta posible.
- Debe garantizarse que exista una sola respuesta correcta o que sea con toda la claridad.
- Cuando se requiere medir comprensión o capacidad para aplicar principios, los términos en que se enuncia deben ser novedosos, de modo que no sean muy semejantes a los textos o a notas de clase y así se disminuye la posibilidad que la respuesta correcta se deba a la memorización mecánica de lo leído o escuchado.
- No puede haber incongruencia gramatical en las preguntas, cada alternativa debe completar el núcleo de forma gramaticalmente correcta.

- Ser cuidadosos con las bases “negativas” y evitar dobles negaciones, ya sea en la base como en las alternativas de respuestas.
- Elaborar alternativas que no provoquen controversias en cuanto a cuál es la respuesta correcta.
- Todas las alternativas deben ser plausibles, al emplear los errores comunes de los educandos.
- Utilizar un lenguaje preciso, con una extensión y complejidad similar, se deben usarse expresiones frecuentes en los estudiantes.
- Evitar en lo posible el empleo de “todas las anteriores sean ciertas” o “todas las anteriores son falsas”.
- Evitar situar la respuesta correcta en posición preferente y no ubicarla siempre en una misma posición.
- El contenido de una oración incompleta en la base debe ser equivalente a una pregunta.

Las preguntas de verdadero-falso constituyen una variedad de pregunta de selección múltiple, con la diferencia que las alternativas que se emplean son sólo dos, por lo que también se conocen como preguntas de respuesta alternativa; están compuestas por un conjunto de afirmaciones de preparación relativamente fácil. Su utilización debe reducirse a proposiciones que sean inequívocamente verdaderas o demostrablemente falsas.^{6, 31, 32.}

Según la literatura revisada al confeccionar las preguntas de verdadero - falso se deben tener en cuenta los aspectos siguientes:^{31, 32} Redactar la información de modo que tenga una idea solamente, asegurarse que los ítems estén escritos de manera que no haya tendencia a la equivocación, no introducir elementos falsos en una afirmación que es correcta, no utilizar palabras textuales de libros, evitar frases entre paréntesis o incidencias que tengan poco que ver con la idea principal de la afirmación, las afirmaciones verdaderas deben ser aproximadamente de igual número que las falsas, sin usar la alternancia regular, eliminar adverbios como: “todo”, “siempre”, “ninguno”, así como sinónimos que expresan generalizaciones tan amplias que tienen mucha posibilidad de ser falsas, eliminar también el empleo de otros términos menos definidos como “usualmente”, “a veces” y “alguna vez”, que

suelen formar parte de proposiciones verdaderas, se debe emplear un lenguaje preciso, incluyendo términos numéricos cuando sea necesario, evitar la utilización de proposiciones negativas, especialmente las de negación doble y tratar que la extensión de las aseveraciones verdaderas sea similar a las de las falsas.

Este tipo de preguntas han sido estudiadas por varios pedagogos y en algunos casos se han realizado propuestas para mejorarlas y así reducir su ambigüedad, los efectos del azar y que por tanto brinden mayor información acerca de los conocimientos del examinado. En estos casos se propone subrayar la parte más importante para que el examinado concentre su atención en esa parte y utilizar variantes de estímulos previos a partir de un gráfico, un mapa, una tabla, etc.³¹

Las preguntas de pareamiento o asociación son formas especiales de las preguntas de selección múltiple; se distinguen de las corrientes de selección múltiple en que en lugar de un solo problema con un grupo de respuestas a escoger, hay varios problemas, cuyas respuestas deben sacarse de una sola lista de alternativas posibles.

Con esta pregunta se mide el aprovechamiento con eficacia y autonomía de espacio ya que pueden utilizarse láminas, gráficos e imágenes.

Dentro de las consideraciones y recomendaciones que debemos tener en cuenta para redactar preguntas de pareamiento encontramos: Las preguntas de una misma serie deben ser homogéneas, el número de respuestas a elegir debe ser mayor que el de los problemas presentados. Si fueran iguales, alguno de los problemas no se asocia y otro se repite, las alternativas deben presentarse en un orden lógico. Debe procurarse que el conjunto de ítems oscile entre 5 y 10, usar para cada columna un encabezado que describa adecuadamente su contenido; si las columnas difieren en cuanto a la amplitud del enunciado, los examinados deben elegir la respuesta en la columna con una menor carga de lectura y las instrucciones deben especificar cuáles son las bases del apareamiento y si las opciones de respuesta pueden utilizarse más de una vez.

Dentro de las variantes de preguntas de asociación encontramos: Asociación simple, asociación compuesta, relación causa efecto, asociación del término excluido, variación relativa, asociación extendida y comparación cuantitativa.

A opinión de la autora la variante más utilizada es la asociación simple, se plantea en la literatura recomendaciones para su elaboración: Debe existir coincidencia gramatical entre ambas columnas, cada columna debe tener contenidos del mismo tipo, si la columna de opciones es una lista de palabras o números es conveniente su ordenamiento, incluir más enunciados que respuestas, más respuestas que enunciados y si igual número, es recomendable que una respuesta sirva a más de un enunciado, número mínimo de ítems entre 8-10 y todas las respuestas deben ser plausibles.

La autora comparte el criterio que los diferentes tipos de preguntas que se enunciaron anteriormente presentan ventajas y desventajas, el éxito radica en la selección adecuada de las mismas y esto depende del conocimiento que se tenga de cada una de ellas, para así poder utilizarlas adecuadamente al elevar sus ventajas y reducir sus desventajas. Además considera que al elaborar un examen final de una asignatura se deben utilizar formatos mixtos y debe existir una adecuada correspondencia entre uno u otro formato a utilizar.

Muchas veces, se comienza a proyectar un examen días antes de ser aplicado cuando en realidad, es algo que debe comenzarse a prever desde las primeras semanas del curso, reajustándose las perspectivas periódicamente según sean los resultados de aprendizajes que van obteniéndose en las restantes tareas evaluativas (frecuentes).³⁰

Al construir un instrumento evaluativo, este no debe realizarse empíricamente sino que debe ser planificado, con el fin de contribuir a incrementar la efectividad y la eficiencia, en la literatura se plantean las etapas siguientes:^{19, 31, 33.}

1. Definir qué debe ser evaluado.
2. Seleccionar los métodos y procedimientos a emplear.
3. Cómo calificar un examen.
4. Cómo establecer un patrón de aprobado/desaprobado.

Para definir el contenido de una evaluación tenemos que desarrollar los pasos siguientes:¹⁹

- I. Identificar los problemas que los educandos deben ser capaces de enfrentar y/o resolver.

- II. Para cada problema definiremos las tareas en las que se espera que el educando sea competente.
- III. Preparar una tabla de especificaciones o de contenidos para determinar la selección de los problemas y tareas que deben ser incluidos en el instrumento evaluativo

Existen varios sistemas para categorizar las operaciones cognitivas que son usadas en el desarrollo de las especificaciones del instrumento. La más conocida internacionalmente en el ámbito de la evaluación es la taxonomía de Bloom (1956) que contiene las siguientes categorías:¹⁹

Conocimiento: Recuerda el material en la misma forma en que fue presentado durante la instrucción.

Comprensión: Permite el traslado, interpretación o extrapolación de un concepto dentro de alguna forma diferente de la práctica originaria o presentación en clase.

Aplicación: Es capaz de resolver nuevos problemas a través del uso de principios familiares o generalizaciones.

Análisis: Puede descomponer un problema o sistema en sus componentes elementales al usar un proceso que requiere del reconocimiento de los múltiples elementos de un todo y la relación que existe entre estos y/o los principios de esa relación u organización.

Síntesis: Combina elementos en un todo, usa una estructura original o la solución de un problema que requiera de la combinación de varios principios en una nueva situación.

Evaluación: Permite el empleo de criterios propios o ajenos para enunciar juicios críticos en términos de precisión, consistencia de la lógica de los puntos de vista del examinado de tipo artístico, filosófico, éticos o morales.

En Cuba empleamos cuatro niveles de asimilación a saber: Familiarizarse, reproducir, producir y crear.

El primer nivel o familiarizarse, implica que el educando tiene una idea de conjunto o noción del contenido que le es informado aunque no pueda reproducirlo ni utilizarlo, no es capaz de resolver ninguna situación que se le presente.

El segundo nivel o reproducir, implica que el educando conoce el contenido o la situación problema planteada y es capaz de repetirlo.

El tercer nivel o producir, implica que el educando está preparado para aplicar el contenido ya conocido ante una situación o problema nuevo, o sea está preparado para saber emplear el conocimiento o habilidad aprendida.

El cuarto nivel o crear, supone la capacidad del educando para enfrentar y resolver situaciones nuevas para las que no dispone de todos los elementos.

Thorndike y Hagen (1977) proponen categorías como las de recordar, reconocer, identificar, definir, aplicar, analizar, sintetizar, generalizar, predecir o evaluar.¹⁹

La autora asume el criterio de que lo importante es, que el diseñador de los instrumentos reconozca que en una pregunta o problema pueden incidir diferentes niveles de procesos cognitivos y lo tenga en cuenta a la hora de elaborar el instrumento.

La estructura más usada para la selección de las preguntas o problemas del instrumento es la llamada tabla de especificaciones o de contenidos (Blueprint de los anglo-sajones) que incluye generalmente los contenidos y los procesos a evaluar.

En la selección de los métodos y procedimientos a emplear se debe precisar los objetivos y contenidos a evaluar. Para su selección debemos tener en cuenta: Emplear con preferencia aquellos que pongan a los educandos en situaciones reales o lo más parecido posible; las tareas clínicas a evaluar prácticamente indican los métodos, procedimientos y técnicas a emplear y valorar las restricciones que pueden condicionar su selección, como claustro docente, tiempo disponible para el examen, cantidad de educandos, recursos para el diseño, entre otros.

En la calificación del examen se emplean diferentes técnicas, las cuales deben estar adecuadamente diseñadas con suficiente anterioridad y validadas inicialmente por consenso de expertos y posteriormente por análisis matemáticos de sus resultados.

El establecimiento del patrón de aprobado / desaprobado, es un proceso complejo. Debemos partir del análisis de su contenido en correspondencia con los objetivos educacionales a evaluar.

Se han desarrollado diferentes enfoques a partir de la propuesta de Livingston y Ziesky en 1982, citadas por otros autores, que se pueden resumir en dos grandes grupos:^{31, 33.}

Enfoque normativo: Establece un patrón relativo que depende de la competencia mostrada por cada examinado. La calificación de aprobado se establece con posterioridad al examen.

Enfoque criteriado: Parte de un patrón absoluto que se establece con anterioridad a la realización del examen. La selección de la calificación del aprobado es arbitraria e independiente de los resultados que obtengan los estudiantes.

De forma general todos los instrumentos evaluativos deben reunir las características siguientes:^{31, 34.} Sus contenidos deben estar directamente relacionados con los objetivos educacionales, ser realista y práctico, tener validez y confiabilidad, ser objetivo y factible de realizar, ser completo, pero lo más breve posible y ser preciso, pero claro en su redacción.

Miriam González plantea que la evaluación, y el acto evaluativo como unidad, suponen operaciones o subprocesos que van desde el establecimiento de los objetivos o propósitos, la delimitación y caracterización del objeto de evaluación, la definición (selección, elaboración) y aplicación de los instrumentos hasta la recogida de información, el procesamiento y análisis de dicha información, su interpretación y expresión en un juicio evaluativo, la retroinformación y toma de decisiones derivadas del mismo, su aplicación y valoración de resultados. Para recomenzar en un ciclo ascendente, progresivo, que permite, en su dinámica, imprimir el auténtico significado de esta actividad. La evaluación supone, además, la propia valoración de la misma.²²

Para poder evaluar la calidad del proceso enseñanza aprendizaje es necesario que se realice un análisis sistemático e interpretación de los resultados obtenidos, mediante la aplicación de diferentes instrumentos de medición, con la finalidad de adoptar las mejores decisiones posibles en cada caso o situación.³⁵

En la literatura especializada aparecen conceptos para evaluar la calidad de los instrumentos evaluativos. En este sentido Guilbert recomienda tener en cuenta otros aspectos entre los que se destacan:³⁶

Objetividad: Grado de concordancia entre los juicios emitidos por evaluadores independientes y competentes sobre lo que constituye una buena respuesta para cada una de las preguntas de un instrumento evaluativo.

Pertinencia: Es el grado de respeto de los criterios establecidos en la selección de las preguntas para que sean conformes con los fines y objetivos de la evaluación.

Equilibrio: Grado de concordancia entre la proporción de preguntas reservadas a cada uno de los objetivos y la proporción ideal para un instrumento de este tipo.

Equidad: Grado de concordancia entre las preguntas planteadas en el instrumento evaluativo y el contenido que se pretende evaluar.

Discriminación: Calidad de cada pregunta del instrumento evaluativo que permite distinguir los estudiantes de altos y bajos rendimientos.

Eficacia: Calidad de un instrumento evaluativo que permite el mayor número posible de respuestas independientes por unidad de tiempo.

Se plantea que un instrumento evaluativo tiene calidad en la medida en que es válido, confiable y sus resultados son generalizables.

La correcta aplicación del sistema de evaluación del aprendizaje, que tiene un carácter cualitativo e integrador, se basa en un conjunto de condiciones que constituyen las **premisas básicas** que debe satisfacer, las cuales se encuentran estrechamente relacionadas entre sí y cuya inobservancia impide se cumplan efectivamente sus funciones; con fines didácticos exclusivamente las estructuramos en principios y exigencias.¹²

Los principios de la evaluación son la objetividad, la sistematización y continuidad, así como su carácter sistémico. Las exigencias que debe satisfacer son las de la validez y la confiabilidad.

Objetividad, para que una evaluación sea objetiva es imprescindible, que a partir de los objetivos definidos, determine el dominio del contenido de la enseñanza y las cualidades que lo caracterizan, de acuerdo con el nivel de profundidad precisado y su correspondencia con el nivel de asimilación exigido y que a su vez esta valoración se realice mediante la forma y tipo de evaluación que asegure la validez del control.

La sistematización y continuidad de la evaluación está dada porque el número de muestras (controles) del universo de las actividades docentes que realice el

estudiante y la regularidad con que se efectúen, posibilite la generalización del criterio del nivel de aprendizaje alcanzado. En este sentido es necesario puntualizar que el aprovechamiento del estudiante siempre hay que caracterizarlo con un criterio dinámico, evolutivo y activo de su tendencia.

La autora opina que la frecuencia con que se realicen las evaluaciones es importante definir las ya que debe existir una cantidad adecuada de estas, ni muchas que recarguen al estudiante ni muy pocas que resulten insuficientes.

Su frecuencia debe ser tal que permita ayudar oportunamente al estudiante y orientarlo al fin deseado y retroalimente a su vez al profesor, también debe estar en dependencia del éxito que logra el estudiante durante el desarrollo del proceso docente y de la necesidad que tenga el estudiante de ser sometido al control.

Su carácter sistémico está dado en la necesidad de concebirlo como un componente del proceso de enseñanza-aprendizaje y en estrecha interrelación e interdependencia con los restantes componentes del mismo objetivo, contenido, método, formas y medios de enseñanza, así como en el papel activo que tienen que desarrollar tanto el profesor como el estudiante para su efectividad.

La validez, en su acepción básica, es la correspondencia entre lo que se pretende verificar por el instrumento, procedimiento o método y lo que realmente se mide o evalúa, y ello abarca tanto a las exigencias propias del control como a la determinación de los índices valorativos que permiten la calificación.³⁷

Podemos preguntarnos: ¿Mide el instrumento evaluativo lo que realmente debe medir? ¿Cuál es el grado de correspondencia entre las preguntas del instrumento evaluativo y los contenidos a evaluar?³⁶

En la medida en que las actividades de la 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, hábitos, habilidades y modos de actuación por el estudiante, mayor será la validez de dicho control. Mientras más teórico, reproductivo y alejado de la práctica esté un control, mucho menor será su validez.

Usualmente se clasifica en cuatro grupos:³⁶

a) Validez de contenido.

b) Validez de criterio.

c) Validez de construcción.

d) Validez funcional.

a) La validez de contenido está en correspondencia con los conocimientos y las habilidades (tanto cualitativa como cuantitativamente) que evalúan o miden el instrumento de control sean lo suficiente para definir el grado de aprovechamiento del estudiante, sobre la base de los objetivos propuestos.¹⁷

La validez de contenido consiste en el examen sistemático del contenido del instrumento evaluativo para determinar, si constituye una muestra representativa del dominio de los conocimientos y/o habilidades que se pretende evaluar.³⁶

b) Validez de criterio: Cuando se realiza un estudio de validez de criterio, se comparan los resultados de la aplicación del instrumento evaluativo con otros criterios externos que se le llama "variable-criterio".

Se dice que un instrumento evaluativo tiene validez de criterio cuando, a partir de los resultados obtenidos, se pueden hacer inferencias válidas sobre una variable que no se puede medir directamente por el instrumento evaluativo.

Los estudios de validez de criterio se clasifican en dos tipos: Validez predictiva y validez concurrente.³⁶

La validez predictiva indica la efectividad del instrumento evaluativo en un resultado futuro. Para esto se comparan las calificaciones del instrumento evaluativo con una medida directa de la ejecución posterior de los sujetos llamado criterio.

La validez concurrente se refiere a la interrelación del instrumento evaluativo y un criterio de medición efectuado en el momento en que se aplica el instrumento evaluativo.

c) Validez de construcción: Los estudios de validez de construcción de un instrumento evaluativo se realizan cuando no se dispone de una medida directa del atributo a medir, de tal manera que hay que sustituirlo por un modelo hipotético.

d) La validez funcional está en correspondencia con la selección del tipo de actividad (acciones, tareas, problemas) que debemos controlar en el estudiante, para evaluar sus modos de actuación de acuerdo con los objetivos propuestos. La autora está de acuerdo que ambos tipos de validez se interrelacionan y complementan.¹⁷

Es una variante de la validez de contenido, pero que adquiere personalidad propia en la educación médica. Está en correspondencia con el tipo de actividad con que debemos verificar el nivel alcanzado por el educando en un grupo de habilidades que conforman su modo de actuación.³⁶

La confiabilidad es la estabilidad en los resultados de un control, ya sea al repetirlo, o al ser calificado por distintos profesores. Significa que hay constancia en los resultados obtenidos y que, por tanto, es representativo del grado de aprovechamiento alcanzado por el estudiante en el tipo de control realizado.^{17, 36.}

La validez y la confiabilidad están estrechamente relacionadas. Un control que cumpla con las exigencias de la validez, tiene un alto grado de probabilidad de ser confiable; sin embargo, no necesariamente ocurre así a la inversa. Por ejemplo, un control puede ser confiable, pues al ser calificado por diferentes profesores se obtienen los mismos resultados y sin embargo, no tener validez ya que el instrumento no mide los objetivos establecidos. En primer orden hay que garantizar la validez de todo instrumento evaluativo y en segundo orden su confiabilidad.

En la literatura revisada se reportan varios procedimientos para cuantificar la confiabilidad de los instrumentos evaluativos:^{20, 36.}

- a) El coeficiente θ
- b) El coeficiente ω
- c) El coeficiente α_0
- d) El coeficiente de Kuder-Richarson
- e) El coeficiente alfa de Cronbach

La autora comparte el criterio con varios autores que el coeficiente alfa de Cronbach:^{20, 36- 39.} Es el más utilizado en los estudios contemporáneos de calidad de los instrumentos evaluativos, lo que nos permite comparar nuestros resultados con los que se reportan en la literatura mundial, nos proporciona una medida de la consistencia interna de los instrumentos evaluativos y puede calcularse aún en el caso de que empleemos preguntas de diferentes formatos en el instrumento evaluativo.

En la literatura contemporánea, la mayoría de los autores, siguiendo a Anastasi y a Weiner, aceptan como adecuados en los exámenes escritos basados en preguntas de tipo test objetivo, los coeficientes que oscilan entre 0,80 y 0,85.²²

El Colegio de Médicos de Familia de Canadá en sus exámenes de certificación de especialistas ha reportado que:

En los exámenes escritos basados en respuestas cortas (SAMPs: Short Answer Management Problems) considera aceptables coeficientes entre 0,69 y 0,7.

En los exámenes de simulación oral en el consultorio médico (SOOs: Simulated Office Orals) han obtenido coeficientes entre 0,54 y 0,63 y consideran como aceptables los superiores a 0,60.

En el examen clínico objetivamente estructurado (ECOE) con pacientes estandarizados y preguntas de ensayo de respuestas cortas han considerado como aceptables coeficientes entre 0,71 y 0,77. En este tipo de examen otros autores han reportado coeficientes entre 0,40 y 0,88.

En Cuba a partir de los exámenes que se diseñan y aplican desde el año 1996, en el marco de las inspecciones territoriales del Ministerio, se han realizado varios estudios y la Comisión Nacional de Evaluación de la Competencia y el Desempeño Profesional decidió establecer como aceptable, un coeficiente alfa de Cronbach de 0,6.²⁰

Existen otros indicadores que miden la calidad de un examen escrito tipo test objetivo como son el análisis del nivel de dificultad y el poder de discriminación.^{31, 40.}

Se entiende como nivel de dificultad de un ítem a la proporción de personas que responden correctamente una pregunta de una prueba. Entre mayor sea esta proporción, menor será su dificultad, por lo que se trata de una relación inversa; a ésta proporción se le denota con la letra *p*, e indica la dificultad del ítem.

El manual del Examen de Habilidades y Conocimientos Básicos (EXHCOBA) de la universidad de Baja California en México considera el nivel medio de dificultad del examen entre 0.5 y 0.6.³¹ Estos valores se tomarán para el análisis de los resultados.

El poder de discriminación del examen es la cualidad de cada pregunta del instrumento que permite distinguir los estudiantes con altos y bajos rendimientos.

Usualmente se utilizan dos formas para medir la capacidad discriminativa del ítem: El índice de discriminación y el coeficiente de discriminación.

Entre más alto es el índice de discriminación, el ítem diferenciará mejor a los estudiantes con altas y bajas calificaciones. Su valor se denomina con la letra D.

El coeficiente de discriminación se mide a través de dos formas: El punto de correlación biserial y el coeficiente de correlación biserial.³¹

El coeficiente de discriminación ofrece ventajas con respecto al índice de discriminación, el primer método toma en cuenta a todos y cada uno de los estudiantes evaluados, mientras que el segundo solo toma 54 % de ellos 27 % más alto y 27 % más bajo.³¹

El coeficiente punto de correlación biserial se mide para saber si los estudiantes de alto rendimiento son los que obtienen las respuestas adecuadas.

El coeficiente de correlación biserial se calcula para determinar el grado en que las competencias que mide el instrumento también las mide el ítem.

La autora coincide con otros autores que sugieren mejor el uso del coeficiente punto de correlación biserial ya que aporta más sobre la validez predictiva del instrumento.

En la asignatura Morfofisiología Humana II se realizan evaluaciones frecuentes, evaluaciones parciales y la evaluación final.

El examen final es un tipo de evaluación que comprueba los objetivos generales de las asignaturas o disciplinas. Evalúa fundamentalmente, los conocimientos y las habilidades adquiridos por el estudiante, y su capacidad para integrarlos, sistematizarlos, aplicarlos y generalizarlos.²⁴

En dependencia del modo en que se realice, el examen final puede contribuir a comprobar en qué medida los estudiantes han incorporado a su conducta los valores precisados en los objetivos generales de la asignatura o disciplina.²⁴

La tendencia que debe predominar en los exámenes finales es su carácter integrador, de modo que puedan evaluar objetivos generales de varias asignaturas o disciplinas que se imparten en un determinado período lectivo.²⁴

Se realiza mediante un examen práctico y uno teórico a través de un instrumento elaborado para ese fin. En este tipo de evaluación se consideran a todos los

estudiantes y es una importante referencia para la nota final de la asignatura en curso.

El instrumento que se aplica se elabora por los profesores principales de la asignatura y asesores de la misma y sus resultados forman parte del currículo individual de los estudiantes dentro de la carrera.

Por todo lo anterior expuesto la autora asume en este estudio el uso de los indicadores de calidad: Índice de dificultad, índice de discriminación, coeficiente de discriminación y Alfa de Cronbach que garantizarán la caracterización de la calidad del instrumento evaluativo escrito de la asignatura Morfofisiología Humana II.

OBJETIVOS

Objetivo General

Caracterizar la calidad del examen de la unidad curricular Morfofisiología Humana II, aplicado a estudiantes del primer año de la carrera de medicina de la Facultad de Ciencias Médicas de Holguín, curso escolar 2010 - 2011.

Objetivos específicos

1. Describir las áreas de las ciencias básicas y tipo de preguntas utilizadas en el examen.
2. Identificar la correspondencia entre el fondo de tiempo de los contenidos de la unidad curricular y su exploración en el examen.
3. Valorar la calidad del instrumento de evaluación escrito aplicado.

DISEÑO METODOLÓGICO

El estudio que se realizó responde a una tesis de investigación en el campo del proceso de Enseñanza-Aprendizaje, cuyo objeto de estudio es el instrumento evaluativo de la asignatura Morfofisiología Humana II y el objetivo general de la investigación es caracterizar la calidad del examen de esta unidad curricular aplicado a estudiantes del primer año de la carrera de medicina de la Facultad de Ciencias Médicas de Holguín en el curso escolar 2010 - 2011.

El universo estuvo constituido por el total de exámenes finales teóricos de la asignatura Morfofisiología Humana II del curso 2010-2011, la muestra fue representada por 182 exámenes finales teóricos de la primera convocatoria.

Se utilizaron métodos teóricos, empíricos y procedimientos estadísticos que garantizaron, la triangulación metodológica de la investigación y permitió la caracterización del examen.

Métodos teóricos:

Fueron empleados la revisión documental y bibliográfica en las temáticas siguientes: Programa analítico de la asignatura Morfofisiología Humana II, documentos normativos de la asignatura, trabajos investigativos, artículos sobre evaluación y otras fuentes de interés para la autora.

El análisis-síntesis y la inducción-deducción, se utilizaron en el procesamiento de la información teórica y empírica, lo que permitió la caracterización del objeto de la investigación, la determinación de los fundamentos teóricos y metodológicos y la elaboración de las conclusiones, así como el análisis de los resultados en sus diferentes momentos.

Métodos empíricos:

Se aplicó una entrevista a expertos (Anexo 1) para caracterizar el examen según tipos de preguntas, área de las ciencias exploradas en cada uno de los incisos y contenido que evalúa cada ítem; para esto fueron seleccionados 5 profesores con más de 10 años de experiencia docente en las ciencias básica biomédicas, con categoría docente y en este momento en función de profesores principales, jefe de los departamentos de ciencias morfológicas y fisiológicas y asesores de la asignatura

Morfofisiología Humana II de la Facultad de Ciencias médicas “Mariana Grajales Coello”.

Se revisaron los exámenes y se clasificaron las preguntas e incisos según las áreas de las ciencias que exploran, el contenido temático, las calificaciones por ítems y fondo de tiempo por temas en una tabla de cálculo Excel de Microsoft Office 2003 confeccionada para este propósito.

El **procesamiento estadístico** se realizó mediante el análisis de frecuencia simple y la distribución porcentual para presentar los resultados en tablas y gráficos. Se utilizó además el Test de Correlación de Spearman.

En el análisis de los resultados de los indicadores de calidad se tuvo en cuenta el índice de dificultad, proporción de personas que responden correctamente una pregunta de la prueba. Mientras mayor sea esta proporción menor será el grado de dificultad, fue calculado para cada ítem, pregunta, tema y el examen.

El cálculo del índice de dificultad de cada ítem se realizó a través de la siguiente fórmula: ^{4, 41, 42.}

$$p = \frac{A}{Nt}$$

Donde:

p= Índice de dificultad del inciso.

A= Número de aciertos del inciso.

Nt= Número total de presentados en el examen.

Se definieron los intervalos según la proporción de estudiantes aprobados en ellos, se estableció que los incisos que tuvieron una proporción de aprobados:

- Menor de 0,40 se consideran difíciles.
- Entre 0,41 y 0,60 se consideran medianamente difíciles.
- Entre 0,61 y 0,80 se consideran de dificultad media.
- Entre 0,81 y 0,90 se consideran medianamente fáciles.
- Entre 0,91 y 1 se consideran fáciles.

Esto permitió determinar la clasificación de los ítems atendiendo a la dificultad real presentada en el instrumento de evaluación.

Según los autores consultados los resultados de este indicador pueden oscilar en:

5 % para preguntas fáciles.

20 % para preguntas medianamente fáciles.

50 % para preguntas de dificultad media.

20 % para preguntas medianamente difíciles.

5 % difíciles.

Para calcular el índice de dificultad de las preguntas y de los temas explorados en el examen se utilizó la siguiente fórmula: ⁴

$$Dt = \frac{\sum (At1 + At2 + \dots + Att)}{nt} \div Nt$$

Donde:

Dt = Índice de dificultad de la temática.

At = Número de respuestas correctas de los incisos de la temática.

nt = Número de incisos de la temática.

Nt = Número total de presentados en el examen.

Se determinó el poder discriminativo para cada ítem, pregunta y la media del examen a través del índice de discriminación y el coeficiente de discriminación.

Aunque hay varias maneras equivalentes de calcular el índice de discriminación, en este trabajo utilizamos el índice de discriminación D. La fórmula que se tuvo en cuenta fue:^{42,43}

$$D_i = \frac{GA_{aciertos} - GB_{aciertos}}{N_{grupomayor}}$$

Donde:

Di: Índice de discriminación del reactivo i

GAaciertos: Número de aciertos en el reactivo i del 27% de personas con las puntuaciones más altas en el test.

GBaciertos: Número de aciertos en el reactivo i del 27% de personas con las puntuaciones más bajas en el test.

Ngrupomayor: Número de personas en el grupo más numeroso (GA o GB).

Entre más alto es el índice de discriminación, el reactivo diferenciará mejor a las personas con altas y bajas calificaciones. Si todas las personas del GA contestan correctamente un reactivo y todas las personas del GB contestan incorrectamente, entonces $D = 1$ (valor máximo de este indicador); si sucede lo contrario, $D = -1$ (valor máximo negativo); si ambos grupos contestan por igual, $D = 0$ (valor mínimo de discriminación).

Ebel y Frisbie (1986) plantearon la siguiente regla de “dedo” para determinar la calidad de los reactivos, en términos del índice de discriminación. A continuación se muestran los valores D y su correspondiente interpretación. Asimismo, se señalan las recomendaciones para cada uno de estos valores.⁴³

D	Calidad	Recomendaciones
> 0.39	Excelente	Conservar.
0.30 - 0.39	Buena	Posibilidades de mejorar.
0.20 - 0.29	Regular	Necesidad de revisar.
0.00 - 0.19	Pobre	Descartar o revisar a profundidad.
< -0.01	Pésima	Descartar definitivamente.

Para determinar la efectividad discriminativa de un reactivo se utilizó el coeficiente de discriminación específicamente la correlación del punto biserial (rpbis).⁴³

La correlación del punto biserial (rpbis) se emplea para saber si las personas adecuadas son las que obtienen las respuestas correctas, qué tanto poder predictivo tiene el reactivo, y cómo puede contribuir a las predicciones.⁴³

La ecuación para obtener este indicador, de acuerdo a Glass y Stanley (1986), es la siguiente:^{42, 43}

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

Donde:

X1 = Media de las puntuaciones totales de aquellos que respondieron correctamente el ítem.

X0 = Media de las puntuaciones totales de aquellos que respondieron incorrectamente el ítem.

SX = Desviación estándar de las puntuaciones totales.

n_1 = Número de casos que respondieron correctamente el ítem.

n_0 = Número de casos que respondieron incorrectamente el ítem.

$n = n_1 + n_0$

Según estándares internacionales la distribución de los resultados del coeficiente de correlación del punto biserial es:⁴³

$r_{pbis} < 0$ discriminan negativamente.

$r_{pbis} = 0 - 0.14$ discriminan pobremente.

$r_{pbis} = 0.15 - 0.25$ poder discriminativo regular.

$r_{pbis} = 0.26 - 0.35$ buen poder discriminativo.

$r_{pbis} > 0.35$ excelente poder de discriminación.

Otro de los indicadores de calidad aplicado fue el coeficiente alfa de Cronbach, que proporciona una medida de la consistencia interna de los exámenes. Para esto se utilizó el paquete estadístico profesional SyStat versión 12.0 para Windows.

En los requerimientos éticos de la investigación se consideró la voluntariedad de participación de los implicados y las informaciones obtenidas sólo fueron utilizadas con fines científicos, de igual forma se garantizaron las condiciones para desarrollar la entrevista a expertos, lo cual caracterizó los requerimientos éticos en la investigación que se presenta.

Para procesar los datos primarios utilizamos una computadora Pentium 4 con ambiente Windows Xp Profesional y el paquete estadístico profesional SyStat versión 12.0 para Windows. El informe final fue redactado en Microsoft Word y para los gráficos se utilizó el Microsoft Excel. Los resultados se muestran en tablas y gráficos

ANÁLISIS Y DISCUSIÓN DE LOS RESULTADOS

La asignatura Morfofisiología Humana II aborda dos temas, el primero sobre sistema osteomioarticular y el segundo sobre sistema tegumentario. Fueron impartidos en nueve semanas lectivas y al finalizar se aplicaron un examen práctico y uno teórico.

Descripción de las áreas de las ciencias básicas y tipo de preguntas utilizadas en el examen.

El examen de Morfofisiología Humana II en su momento teórico (anexo 2) se estructuró con un total de 7 preguntas con 58 ítems.

Explora contenidos básicos de Embriología, Anatomía Humana, Fisiología e Histología. En la tabla 1 se muestra la distribución de los contenidos de las ciencias básicas biomédicas a las que responde cada ítems.

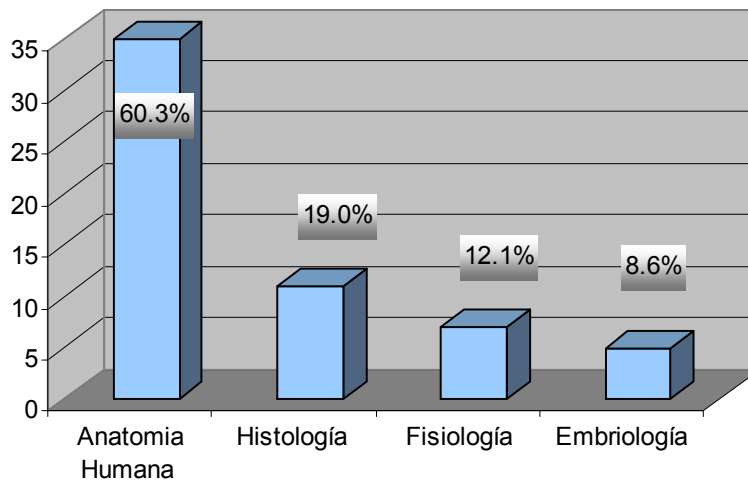
Tabla 1. Contenidos de las ciencias básicas biomédicas a las que responde cada ítem.

Contenido de las Ciencias Básicas Biomédicas	Ítems del examen
Anatomía Humana	I- 1.1, 1.2, 1.3, 1.5, 1.6 II- 2.1.1, 2.1.2, 2.1.3, 2.2, 2.3, 2.4, 2.5, 2.6.1, 2.6.2, 2.7 III- 3.1.1, 3.1.2, 3.2.1, 3.2.2 IV- 4.2.1, 4.2.2, 4.2.3, 4.3.1, 4.3.2, 4.3.3, 4.4.1, 4.4.2, 4.4.3 V- 5.1, 5.2, 5.3, 5.4, 5.5, 5.6, 5.7
Histología	I- 1.4, 1.7 VII- 7.1.1, 7.1.2, 7.2.1, 7.2.2, 7.2.3, 7.2.4, 7.2.5, 7.4.1, 7.4.2
Fisiología	VI- 6.1, 6.2, 6.3, 6.4, 6.5, 6.7
Embriología	III- 3.3 IV- 4.1.1, 4.1.2, 4.1.3 VII- 7.3

Fuente: Examen teórico.

En general el examen tiene 35 ítems de Anatomía, 11 de Histología, 7 de Fisiología y solo 5 ítems de Embriología, como se evidencia en el gráfico 1.

Gráfico 1. Contenidos de las ciencias básicas biomédicas exploradas en el examen.



Fuente: Examen teórico.

El mayor por ciento (60,3) tributa a la ciencia básica Anatomía Humana lo que se explica, según los expertos, que en esta asignatura estos contenidos tiene un gran peso por los temas que abarca. La autora coincide con lo planteado anteriormente ya que de los dos temas del programa, uno corresponde al sistema osteomioarticular que se imparte durante **nueve** semanas y comparte la última semana con el tema dos correspondiente al sistema tegumentario profundizándose fundamentalmente en Anatomía Humana.

Cuenta con 26 ítems de test objetivo, de ellos 14 (24,1 %) de verdadero y falso, 12 (20,7 %) de selección múltiple tipo complemento simple y 32 de ensayo; de estos 15 (25,9 %) fueron de ensayo o desarrollo, 7 (12,1 %) de respuestas cortas y 10 (17,2 %) de completamiento.

De los 58 ítems del examen 55,2 % (32) fueron de ensayo o desarrollo y solo 26 ítems que representa 44,8 %, fueron tipo test objetivo.

Las preguntas de ensayo son aquellas donde el sujeto aporta su propia respuesta ellas posibilitan la exposición de conocimientos, el ordenamiento de ideas y demuestran la originalidad, destreza en el uso del lenguaje y permiten determinar el grado de profundidad de conocimientos adquirido por el estudiante.

En las preguntas de test objetivo 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 la revisión realizada se pudo encontrar trabajos como los realizados por Zayas, en su tesina de terminación de diplomado en Pedagogía en el ISCM de Villa Clara, donde planteo que las preguntas más empleadas son las de test objetivo, específicamente de selección múltiple y resultaron las de mejor calificación, pero contradictoriamente las preguntas de mayor preferencia por los educandos fueron las de ensayo.⁴⁴

La autora considera que no debe abusarse de las preguntas de ensayo ni las de test objetivo, lo correcto sería utilizar ambos tipos, es decir preguntas con formatos mixtos y de esta forma aprovechar las ventajas de cada una y minimizar sus desventajas.

La primera pregunta explora los contenidos de Anatomía Humana en cinco ítems e Histología en dos, con un total de siete ítems de test objetivo variante verdadero y falso, que representan 12,1 % del examen. Esta pregunta abarcó los temas el cuerpo humano, artrología general, biomecánica y miología.

La segunda pregunta explora contenidos de Anatomía Humana, contiene 10 ítems de ensayo tipo completamiento que representan 17,2 % del examen, en la misma se exploran los contenidos artrología general y esqueleto de la cabeza.

La tercera pregunta contiene cinco ítems de tipo ensayo o desarrollo que representan 8,6 % del examen, evalúa contenidos de Anatomía Humana en cuatro ítems y de Embriología en uno, se tratan contenidos de esqueleto de cuello y tronco.

La cuarta pregunta tiene doce ítems que representan 20,7 % del contenido del examen, evalúa aspectos de Anatomía Humana en nueve ítems y de embriología en tres, trata los temas de artrología general y esqueleto apendicular y es una pregunta de selección múltiple.

La quinta pregunta tiene siete ítems de ensayo o desarrollo que representan 12,1 % del contenido del examen, en su totalidad explora temas de Anatomía Humana, relacionados con músculos de la cabeza, cuello, tronco y del esqueleto apendicular.

La sexta pregunta presenta siete ítems de test objetivo lo que muestra 12,1 % del examen, se abordaron contenidos relacionados con Fisiología específicamente miología y tejido muscular.

La séptima pregunta tiene diez ítems de ensayo o desarrollo lo que representa 17,2 %, aborda temas de Histología en nueve de ellos y uno de Embriología relacionado con el tema piel y faneras.

Relación entre el fondo de tiempo de cada una de los temas del plan calendario con el número de ítems presentes en el examen y grado de dificultad por tema.

La relación entre el fondo de tiempo de cada uno de los temas del plan calendario, con el número de ítem presentes en el examen y su grado de dificultad, se analizan en la tabla 2.

Tabla 2. Distribución del fondo de tiempo según contenidos del programa y presencia de los ítems en las temáticas.

Contenidos	Fondo de tiempo		Presencia en el examen		Grado de dificultad
	Horas	%	Ítems	%	
Tema 1.1-1.2	12	14,30	12	20,7	0,85
Tema 1.3-1.4	12	14,30	7	12,1	0,63
Tema 1.5-1.6	12	14,30	5	8,6	0,78
Tema 1.7-1.8	12	14,30	9	15,6	0,96
Tema 1.9-1.10	12	14,30	10	17,2	0,63
Tema 1.11-1.12	12	14,30	2	3,4	0,72
Tema 1.13	6	7,10	3	5,2	0,64
Tema 2.1	6	7,10	10	17,2	0,87
Total	84	100	58	100	

Fuente: Datos del autor.

Al analizar el fondo de tiempo por semanas según datos obtenidos del plan calendario del curso 2010-2011 de esta asignatura, a los temas de cada semana se le dedican 12 horas, distribuidos en cuatro horas de conferencia orientadora, cuatro horas al taller y cuatro horas de seminario, lo que representa 14,30 % del tiempo total; al tema 1.13 y 2.1 se le dedican seis horas a cada uno que representa 7,10 % del total.

Los temas que mayor cantidad de ítems presentó el examen fueron el 1.1- 1.2 con 12 ítems que representa 20,7 %, donde se abordaron los contenidos: El cuerpo humano, generalidades del aparato locomotor, osteología general, artrología y biomecánica; le siguieron en orden de frecuencia los temas 1.9-1.10 y 2.1, con 10 ítems cada uno que representan 17,2 % respectivamente, en los que se abordan contenidos de miología, músculos de la cabeza, cuello, piel y faneras. Llama la atención que el tema 1.11-1.12 solo estuvo presente en 2 ítems, este tema trata contenidos sobre músculos del tronco.

La autora considera que no existe uniformidad en la cantidad de ítems de los diferentes temas y el tiempo que se le dedicó en el plan calendario, es decir si a todos los temas se le dedica la misma cantidad de horas en su fondo de tiempo, debe existir igual cantidad de ítems aproximada en el examen.

Además se debe tener presente que las ciencias básicas son consideradas como bases cognoscitivas previas y hacen contribuciones importantes a la formación intelectual general de los educandos, por lo que todos los temas se consideran importantes.⁴⁸

Al analizar el cálculo del índice de dificultad por temas se puede observar que el de menor dificultad resultó ser el 1.7-1.8 correspondiente al esqueleto apendicular con un índice de 0,96, estos contenidos fueron evaluados en la pregunta cuatro la cual fue clasificada como fácil con una pequeña cantidad de suspensos. Los temas con mayor dificultad fueron: 1.3-1.4 y 1.9-1.10 con un índice de dificultad de 0,63.

El tema 1.3-1.4 aborda contenidos sobre esqueleto de la cabeza se exploró fundamentalmente en la pregunta dos que fue la única del examen que evidenció dificultad media, el tema 1.9-1.10 aborda contenidos de miología, tejido muscular, músculos de la cabeza y el cuello, la mayoría de estos aspectos se evaluaron en la pregunta seis que fue clasificada como pregunta medianamente difícil.

Calidad del instrumento de evaluación aplicado.

Índice de dificultad.

El Índice de dificultad se define como la proporción de estudiantes que responden correctamente el ítem de un examen. Cuanto mayor sea esta proporción, menor será

su dificultad, lo que quiere decir que se trata de una relación inversa, a mayor dificultad del ítem, menor será su índice.^{43, 45.}

En el anexo 3 aparecen los errores e índice de dificultad de cada ítem del examen, de los 58 ítems que presenta el examen, 18 (31,0 %) presentaron índice de dificultad de 91-100, 11(19,0 %) con índice de dificultad de 81-90 por lo se puede plantear que 50 % de los ítems del examen fueron fáciles y medianamente fáciles, sólo uno presento índice de dificultad menos de 0,40.

Estos resultados coinciden con los obtenidos por Abreu Reyes en su estudio sobre Caracterización del examen teórico final de Morfofisiología Humana del perfil Laboratorio Clínico, donde más del 50% de los ítems del examen analizado fueron fáciles.⁴

En la tabla 3 se observan los resultados obtenidos al agrupar los ítems de acuerdo a su Índice de dificultad según la escala establecida por los expertos.

Tabla 3. Ítems del examen según grado de dificultad esperado y real.

Ítem según su grado de dificultad	Ítems esperados		Ítems reales	
	Nº	%	Nº	%
Fáciles	3	5	21	36,3
Medianamente fáciles	12	20	10	17,2
Dificultad media	29	50	20	34,5
Medianamente difíciles	12	20	6	10,3
Difíciles	3	5	1	1,7
total	58	100	58	100

Fuente: Procesamiento estadístico.

Como se observa no se aprecia coincidencia entre los ítems fáciles esperados y los reales existentes en el examen, en este caso los superó en 18 ítems mientras que en las categorías de ítems medianamente fáciles la diferencia es solo de dos; en la categoría de ítems de dificultad media hubo 20 en el examen de 29 en el esperado, de forma similar se comportó la categoría medianamente difícil al encontrar en el examen 6 ítems de 12 esperado, existió una diferencia de dos ítems en la categoría difíciles pues de ellos hubo uno en el examen de tres esperados.

Para validar estos resultados se aplicó el Test de Correlación de Spearman y se comprobó la no correspondencia entre el número de ítems esperados y los encontrados en el examen, con un valor $r = 0,16$.

Estos resultados no coinciden con los descritos por Santiesteban E en su metodología para el análisis de los exámenes finales, donde existió una correspondencia entre los ítems esperados y los reales del instrumento analizado.⁴⁵

Al realizar un análisis del valor porcentual de los ítems reales en el examen, 53,5 % fueron clasificados en fáciles y medianamente fáciles, 34,5 % en dificultad media y sólo 12,0 % fueron medianamente difíciles y difíciles

Al revisar la literatura se encontró similitud con los resultados de Argudín Somonte Eglys en el trabajo sobre Índice de Dificultad del examen de Morfofisiología Humana I, donde predominó 75 % de los ítems entre las categorías medianamente fáciles y fáciles;⁴¹

A criterio de la autora no se tuvo en cuenta el índice de dificultad establecido en la elaboración de los ítems del examen.

Al analizar la dificultad por preguntas y el encontrado en el examen como se muestra en la tabla 4, se consideró fáciles las preguntas uno, tres, cuatro y siete; la pregunta cinco se clasificó como medianamente fácil, la dos de dificultad media y sólo una, la seis, clasificó como medianamente difícil.

Tabla 4. Índice de dificultad por preguntas y examen.

Preguntas	Índice de dificultad	Tipo de dificultad
Pregunta 1	0,98	Fácil
Pregunta 2	0,66	Dificultad Media
Pregunta 3	0,93	Fácil
Pregunta 4	0,98	Fácil
Pregunta 5	0,82	Medianamente fácil
Pregunta 6	0,60	Medianamente difícil
Pregunta 7	0,93	Fácil
Examen	0,96	Fácil

Fuente: Procesamiento estadístico.

En el instrumento analizado hubo un predominio de las preguntas fáciles lo que explica la gran cantidad de aprobados. Resultados similares obtuvo Moreno M, en su estudio realizado en la República Bolivariana de Venezuela sobre Caracterización del examen estatal escrito de la especialidad en Medicina General Integral.³¹

Backhoff E, Díaz Roig y otros^{20, 41, 43, 46} exponen que el nivel medio de dificultad de un examen debe oscilar entre 0,5 y 0,6. Para este examen el índice de dificultad fue de 0,96 considerándose un examen fácil.

Estos resultados coinciden con los obtenidos en el trabajo realizado por Argudín Somonte Eglis y colaboradores.⁴¹ Según Backhoff E.⁴⁵ en su estudio del examen EXHCOBA la dificultad fue de 0,56, resultados que difieren de lo encontrado en esta investigación.

Córica y colaboradores plantean que un examen que tenga dificultad media, es superior en confiabilidad de resultados a aquellos que tengan mucha dificultad, o que resulten muy fáciles. Un buen examen será aquel que sea superado por algo más de la mitad de los educandos.⁴⁷

Discriminación.

Si la prueba y un ítem miden la misma habilidad o competencia, se puede esperar que quien tuvo una puntuación alta en todo el test deberá tener altas probabilidades de contestar correctamente el ítem. También se debe esperar lo contrario, es decir, que quien tuvo bajas puntuaciones en el test, deberá tener pocas probabilidades de contestar correctamente el reactivo. Un buen ítem debe discriminar entre aquellos que obtuvieron buenas calificaciones en la prueba y aquellos que obtuvieron bajas calificaciones.⁴⁵

Se debe tener en cuenta que los 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 resultados muy bajos.⁴⁶

En el presente trabajo se utilizaron dos formas para determinar el poder discriminativo: El índice de discriminación y el coeficiente de discriminación.

En el anexo 4 se recogen los índices de discriminación para los diferentes ítems del examen.

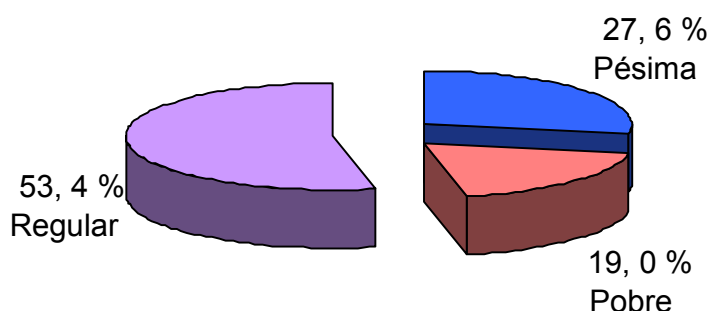
Según los valores D, se pueden agrupar los ítems de acuerdo a su poder discriminativo en tres grupos: los que discriminan en forma regular con valores entre 0,20-0,29, los ítems con discriminación pobre con valores entre 0,00-0,20 y con discriminación pésima con valores menores de -0,01 lo que se representa en la tabla 5 y en el gráfico 2.

Tabla 5. Índice de discriminación (D) por ítems.

Calidad	Cantidad de ítems.	% de ítems.
Regular	31	53,4
Pobre	11	19,0
Pésima	16	27,6
Total	58	100

Fuente: Anexo 4.

Gráfico 2. Por ciento de ítems según calidad del índice de discriminación (D).



Fuente: Tabla 5.

Como se puede apreciar 53,4 % (31) de los ítems discriminan de forma regular y la recomendación según Backhoff es que necesitan revisión, le siguen en orden de frecuencia 16 ítems que representa 27,6 % que tienen una calidad pésima los que se deben descartar definitivamente y 19,0 % (11) presentaron una pobre discriminación lo que se recomienda descartar o revisar a profundidad.

Estos resultados no coinciden con los encontrados con Backhoff donde la mayoría de las áreas temáticas analizadas presentaron un índice de discriminación D mayor de 0,30 siendo excelente y buena.⁴⁵

Al analizar el índice de discriminación D por preguntas (Tabla 6) todas clasificaron entre 0,00-0,20 lo que su índice de discriminación fue pobre.

El índice de discriminación medio del examen fue de 0,07. Carrazana en su tesis de maestría encontró un índice de discriminación medio de 0,38 que no coincide con lo encontrado en este estudio.⁷

Tabla 6. Índice de discriminación D por preguntas.

Preguntas	ID
Pregunta 1	0,06
Pregunta 2	0,17
Pregunta 3	0,08
Pregunta 4	0,06
Pregunta 5	0,13
Pregunta 6	0,17
Pregunta 7	0,08

Fuente: Procesamiento estadístico.

La autora de esta investigación considera que con este índice de discriminación tan bajo no permite definir al estudiante que tiene el conocimiento correcto de aquellos que no lo tienen ya que mientras más alto sea este índice, mejor será la posibilidad de diferenciar los estudiantes con altas y bajas puntuaciones.

El coeficiente de discriminación se utiliza para saber si los alumnos de más alto rendimiento obtienen las respuestas correctas, es decir qué tanto valor predictivo tiene la pregunta y como se puede contribuir a las predicciones.

En la tabla 7 se analiza el coeficiente de correlación del punto biserial (rpbis) de los diferentes ítems del examen, 50 (86,2 %) presentaron excelente poder discriminativo, cuatro discriminación regular e igual número buen poder discriminativo, que representó 6,9 % respectivamente.

Tabla 7. Correlación del punto biserial (rpbis) por ítems del examen.

rpbis	Cantidad de ítems.	% de ítems.
Menor 0	0	0
0-0,14	0	0
0,15-0,25	4	6,9
0,26-0,35	4	6,9
Mayor 0,35	50	86,20
Total	58	100

Fuente: Procesamiento estadístico.

La tabla 8 muestra la correlación del punto biserial (rpbis) por preguntas, clasificaron como excelente poder discriminativo la dos, tres, cinco y seis, buen poder discriminativo sólo la pregunta siete y regular poder discriminativo las preguntas uno y cuatro. El coeficiente de discriminación medio del examen fue de 0,37.

Tabla 8. Correlación del punto biserial (rpbis) por preguntas.

Preguntas	rpbis	Clasificación.
Pregunta 1	0,20	Regular
Pregunta 2	0,48	Excelente
Pregunta 3	0,37	Excelente
Pregunta 4	0,24	Regular
Pregunta 5	0,48	Excelente
Pregunta 6	0,47	Excelente
Pregunta 7	0,34	Buen poder discriminativo

Fuente: Procesamiento estadístico.

Al comparar estos resultados con los obtenidos en el índice de dificultad, las preguntas dos, cinco y seis fueron clasificadas de dificultad media, medianamente fácil y medianamente difícil respectivamente. Sólo la pregunta tres fue clasificada como fácil por lo que existe una correspondencia entre el índice de dificultad por preguntas y su coeficiente de discriminación, ya que las preguntas que más discriminan son aquellas que presentaron una dificultad mayor.

La literatura revisada plantea una ventaja del coeficiente de discriminación frente al índice de discriminación, con el primer método se toma en cuenta todas las personas

evaluadas, mientras que con el segundo, sólo 54 % de ellas (27 % más alto y 27 % más bajo). Se considera que el coeficiente de discriminación es más sensible para detectar el comportamiento de los reactivos de un examen.^{25, 49}

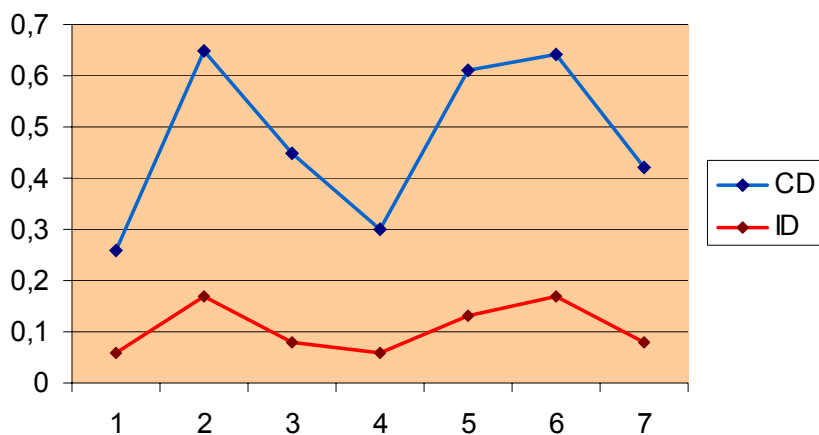
Resultados similares encontró Guadalupe Sánchez⁴⁹ en un estudio realizado en la Universidad Autónoma de México donde obtuvo un índice de discriminación pobre y un buen coeficiente de discriminación.

Una consideración importante a tener en cuenta en el examen objeto de estudio radica en que el mayor por ciento de preguntas es de tipo ensayo, en particular de ensayo o desarrollo (25,9 %), en este tipo de preguntas la objetividad es difícil de conseguir y sus resultados son poco confiables,¹⁹ por lo que su validez y confiabilidad es pobre ya que depende de la manera en que el estudiante responde la pregunta, del juicio del profesor y es muy difícil elaborar claves con precisión.

En el gráfico 3 se muestra el índice de discriminación y el coeficiente de discriminación por preguntas donde se puede observar que las preguntas dos, cinco y seis son las que más discriminan independientemente del indicador aplicado correspondiéndose con las de mayor dificultad.

La autora es del criterio que no obstante a lo descrito en la literatura y a los resultados encontrados en la correlación del punto biserial (rpbis) que benefician la discriminación del examen, no se debe despreciar lo obtenido en el índice de discriminación D que aporta resultados de los grupos extremos, por lo que al analizar la calidad de un instrumento evaluativo deben tenerse en cuenta ambos indicadores.

Gráfico 3. Índice de discriminación y coeficiente de discriminación por preguntas.



Fuente: Procesamiento estadístico.

Alfa de Cronbach.

En la tabla 9 se expresan los resultados del coeficiente α de Cronbach, para cada una de las preguntas y el examen como un todo. El valor total 0,71 demuestra que el instrumento tiene una confiabilidad aceptable. Este resultado concuerda con los estándares propuestos por el Colegio de Médicos de Canadá donde se consideran aceptables coeficientes entre 0,69 y 0,71³¹ y con los establecidos por la Comisión Nacional de Evaluación de la Competencia y el Desempeño del Ministerio de Salud Pública que propone valores superiores a 0,60.²⁰

Tabla 9. Coeficiente Alfa de Cronbach por preguntas y examen.

Preguntas	Alfa de cronbach
Pregunta 1	0,66
Pregunta 2	0,66
Pregunta 3	0,69
Pregunta 4	0,70
Pregunta 5	0,64
Pregunta 6	0,73
Pregunta 7	0,66
Examen	0,71

Fuente: Procesamiento estadístico.

Los métodos teóricos, empíricos y el procesamiento estadístico permitieron el cumplimiento de los objetivos y llegar a caracterizar la calidad del examen de la unidad curricular Morfofisiología Humana II, aplicado a estudiantes del primer año de la carrera de medicina de la Facultad de Ciencias Médicas de Holguín, curso escolar 2010 - 2011.

Caracterización del examen teórico de Morfofisiología Humana II curso 2010-2011.

I. Descripción de las áreas de las ciencias básicas y tipo de preguntas utilizadas en el examen.

1. El examen se estructuró en 7 preguntas y 58 ítems, con predominio de las preguntas de ensayo o desarrollo.
2. Explora contenidos básicos de Embriología, Anatomía Humana, Fisiología e Histología, donde predominan los contenidos de Anatomía Humana.

II. Relación entre el fondo de tiempo de cada una de los temas del plan calendario con el número de incisos presentes en el examen y su grado de dificultad.

1. No existió correspondencia entre el fondo de tiempo de cada uno de los temas y la cantidad de ítems presentes en el examen.
2. Todos los temas presentaron un índice de dificultad por encima de 0,61.

III. Calidad del instrumento de evaluación aplicado.

Índice de dificultad.

1. De los 58 ítems 50 % fueron fáciles y medianamente fáciles.
2. En el instrumento analizado predominaron las preguntas fáciles lo que explica el alto grado de aprobados.
3. El índice de dificultad del instrumento fue 0,96 considerándose un examen fácil.
4. No existió correspondencia entre los ítems esperados y los reales existentes en el examen, esto fue validado mediante el Test de Correlación de Spearman con una $r = 0,16$.

Índice de discriminación.

1. Discriminan de forma regular 53,4 % de los ítems y su recomendación según Backhoff es que necesitan revisión.

2. Todas las preguntas clasificaron entre 0,00-0,20 por lo que su índice de discriminación fue pobre.

3. El índice de discriminación medio del examen fue de 0,07.

Coeficiente de correlación del punto biserial.

1. El coeficiente de correlación del punto biserial demostró una discriminación excelente en 57,1 % de las preguntas.

2. El coeficiente de discriminación medio del examen fue 0,37.

Alfa de Cronbach

1. El análisis de la consistencia interna del examen aplicado fue satisfactorio con un valor de 0,71, este resultado concuerda con los establecidos por la Comisión Nacional de Evaluación de la Competencia y el Desempeño del Ministerio de Salud Pública.

CONCLUSIONES

1. La descripción de las áreas de las ciencias básicas manifestó que el instrumento explora contenidos de Embriología, Anatomía Humana, Fisiología e Histología, con predominio de los contenidos de Anatomía Humana. Según el tipo de pregunta utilizada las de mayor por ciento fue de ensayo o desarrollo.
2. En la identificación de la relación entre el fondo de tiempo de los contenidos de la unidad curricular y su exploración en el examen se demostró que no existió correspondencia.
3. La valoración de los indicadores de calidad mostraron que el instrumento según el índice de dificultad fue clasificado de fácil, donde existió correspondencia entre el índice de dificultad por preguntas con su discriminación. El coeficiente alfa de Cronbach mostró confiabilidad en su consistencia interna.

RECOMENDACIONES

Elaborar un sistema de acciones metodológicas para corregir las dificultades encontradas.

Realizar estudios similares en otras asignaturas de la Morfofisiología Humana con el objetivo de continuar el perfeccionamiento del sistema evaluativo.

REFERENCIAS BIBLIOGRÁFICAS

1. Gómez Álvarez AM, Cardellá Rosales LL, Hernández Fernández ML. Disciplina Morfofisiología Humana II: Problemas de la renovación educativa y sus requerimientos. 2008; 3(2).
2. Evaluación del Desempeño Profesional. Propuesta de un Modelo. Material de estudio Maestría en Educación Médica. En CD-ROM Maestría en Educación Médica. ENSAP: 2005.
3. Salas Perea RS. La calidad en el desarrollo profesional: Avances y desafíos. Material de estudio de la Maestría en Educación Médica. En CD-ROM Maestría en Educación Médica, ENSAP: 2005.
4. Abreu Reyes DR. Caracterización del examen teórico final de Morfofisiología Humana del perfil Laboratorio Clínico. Rev.Cubana Educ. Med. Sup. 2011; 25(1).
5. Rodríguez Arévalo L. Factores relacionados con los resultados del examen diagnóstico y el examen final de la asignatura Morfofisiología Humana I. Parroquia “Las Cocuizas”. Estado Monagas. Curso 2005- 2006. [Tesis para optar por el título de Master en Educación Médica Superior]. La Habana: ENSAP; 2007.
6. Águila Rodríguez Y. Caracterización del momento de evaluación del aprendizaje de la clase encuentro. Morfofisiología Humana II. Programa Nacional de Formación de Medicina Integral Comunitaria. [en línea] [Fecha de acceso 20 de junio de 2011] Disponible en: <http://www.portalesmedicos.com/publicaciones/articulos/2560/1>
7. Carrazana Lee A. Caracterización del segundo examen diagnóstico de Morfofisiología Humana I del nuevo programa de formación de médicos. Facultad “Félix Edén Aguada”, Cuba. [Tesis para optar por el título de Master en Educación Médica Superior]. La Habana: ENSAP; 2007.
8. Salas Perea RS. La evaluación en la educación superior contemporánea. Biblioteca de Medicina Volumen XXIV, 1998.Educación en salud. Competencia y desempeño profesionales. La Habana: Editorial Ciencias Médicas; 1999. En CD-ROM Maestría Educación Médica. ENSAP; 2005.

9. Gómez León K. Propuesta de guía para la evaluación del docente en la dirección de la autopreparación. Municipio Iribarren, Lara, Venezuela. [Tesis para optar por el título de Master en Educación Médica Superior]. La Habana: ENSAP; 2007.
10. Zubizarreta MM. Planeamiento y calidad del examen estatal escrito de Licenciatura en Enfermería. Año 2001. [Tesis para optar por el título de Master en Educación Médica Superior]. La Habana: ENSAP; 2001.
11. González González, M. La evaluación del aprendizaje. Rev. Cubana Educ. Med Sup. 2000; 3(1).
12. Santa Eulalia Mainegra JM. Caracterización de la Evaluación Frecuente durante el Encuentro en el Núcleo Docente, Modalidad Evaluación del Aprendizaje. Morfofisiología Humana III. Municipio Simón Bolívar, Estado Anzoátegui. Curso 2006- 2007. [Tesis para optar por el título de Master en Educación Médica Superior]. La Habana: ENSAP; 2007.
13. Vicedo Tomey A. La integración de conocimientos en la educación médica. Rev Cubana Educ. Med Sup. 2009; 23(4).
14. Arencibia Flores LG, Pernas Gómez M, Rivera Michelena N. Factores que influyen en la implementación de la asignatura Morfofisiología Humana I. Rev. Cubana Educ Med Sup. 2011; 25(3).
15. Ortiz Rodríguez F, Román Collazo CA. Disciplina morfofisiología como alternativa de integración curricular en la enseñanza de la medicina. Rev haban cienc méd. 2010; 9(2).
16. Lima Sarmientol L, Pérez Esquivell GJ, Díaz Molledall M, Rodriguez Mileral JD; Montano Luna JA. La aplicación del sistema de evaluación en la asignatura de Morfofisiología Humana III, desde la perspectiva del profesor. Rev Cubana Educ. Med Sup. 2010; 24(3).
17. Evaluación del aprendizaje. Módulo Evaluación y recursos del aprendizaje. Material de estudio Maestría en Educación Médica. En CD-ROM Maestría en Educación Médica. ENSAP: 2005.
18. Venturelli J. La evaluación en el proceso de formación de los profesionales médicos. Principios de evaluación formativa. En: Educación Médica: nuevos

- enfoques, metas y métodos. Capítulo 6. Serie PALTEX, Salud y Sociedad 2000 No. 5 Organización Panamericana de la Salud, Washington, 1997.
19. Salas Perea R. Planeamiento de un examen. Módulo Evaluación y recursos del aprendizaje Material de estudio de la Maestría en Educación Médica. En CD-ROM Maestría en Educación Médica, ENSAP: 2005.
 20. Marín Hernández D. Estructura y valoración de la calidad del examen final de Morfofisiología Humana I en el Estado Táchira. Curso 2005 – 2006. [Tesis para optar por el título de Master en Educación Médica Superior]. La Habana: ENSAP; 2007.
 21. Castro Pimienta OD. Evaluación y excelencia educativa personalizada Material de estudio de la Maestría en Educación Médica. En CD-ROM Maestría en Educación Médica, ENSAP: 2005.
 22. González Pérez M. La evaluación del aprendizaje.- tendencias y reflexión crítica. Rev. Cubana Educ. Med. Sup. 2000; 20(1): 47-62.
 23. Díaz Hernández RJ, Díaz Rodríguez L, Seiglie Rebolgar MA, García Enríquez, ME. Evaluación educativa en la formación laboral integral de estudiantes de Tecnología de la Salud, perfil Laboratorio Clínico. Edumecentro. 2011; 3(2). [en línea] [Fecha de acceso 16 de abril de 2012] Disponible en: dialnet.unirioja.es/servlet/fichero_articulo?codigo=3735171
 24. Reglamento trabajo docente y metodológico. Resolución ministerial No 210/2007.
 25. González OL, Portuondo Padrón R La evaluación del aprendizaje. Rev. Educ. Med Sup. 1998; 1: 49-63.
 26. Ruiz Gamio X. La evaluación del aprendizaje: un reto actual. [en línea] [Fecha de acceso 10 de abril de 2012] Disponible en: www.ucp.vc.rimed.cu/sitios/varela/articulos/rv1304.pdf
 27. Sierra S. La evaluación del aprendizaje en las ciencias básicas. En: Material de estudio de la Maestría de Educación Médica. [CD -ROM]. ENSAP: La Habana, 2005.
 28. Álvarez CM. La didáctica como ciencia. Su objeto. Los objetivos de la enseñanza. Enfoque sistémico de la didáctica en la educación superior. En:

- Material de estudio de la Maestría de Educación Médica. [CD-ROM]. ENSAP: La Habana, 2005.
29. Arencibia Flores LG, Valdés Moreno JI, Galvizu Díaz K, Corredera Guerra R. Evaluación y desarrollo de las habilidades Interpretar y Predecir en Fisiología I en los estudiantes de Medicina. *Rev Cubana Educ Med Sup.* 2006, 20 (1). [citado 10 Junio 2011], Disponible en la World Wide Web: <http://scielo.sld.cu/scielo.php?script=sci_arttext&pid=S086421412002000300002&lng=es&nrm=iso>. ISSN 0864-2141.).
 30. Hernández Nodarse M. Perfeccionando los exámenes escritos: reflexiones y sugerencias metodológicas. *Rev. Iberoamericana de Educación.* 2007; 41(4).
 31. Moreno Montañez M. Caracterización del examen estatal escrito de la especialidad en Medicina General Integral. 1ra convocatoria. República Bolivariana de Venezuela. Mayo 2006. [Tesis para optar por el título de Master en Educación Médica Superior]. La Habana: ENSAP; 2007.
 32. Salas Perea R. El examen escrito. Módulo Evaluación y recursos del aprendizaje. Material de estudio de la Maestría en Educación Médica. En CD-ROM Maestría en Educación Médica, ENSAP: 2005.
 33. Castillo Abreus DA, Carbonell Panequell SA, Barrios Herrero L, Vázquez Naranjo O. Bases teóricas para la integración de las ciencias básicas biomédicas en una disciplina. *Rev. Cubana Educ. Med Sup.* 2010; 24(3).
 34. Salas Perea R. Los instrumentos evaluativos. Módulo Evaluación y recursos del aprendizaje. Material de estudio de la Maestría en Educación Médica. En CD-ROM Maestría en Educación Médica, ENSAP: 2005.
 35. Moreno Montañez, M; Quintana Regalado, G. Caracterización del examen estatal escrito en la especialidad de Medicina General Integral. *Rev. Cubana Educ Med Sup.* 2009; 23 (3).
 36. Salas Perea R. Calidad de los instrumentos de evaluación. Módulo Evaluación y recursos del aprendizaje. Material de estudio Maestría en Educación Médica. En CD-ROM Maestría en Educación Médica. ENSAP: 2005.

37. Backhoff Escudero E, Tirado Segura F, Larrazolo Reyna N. Ponderación diferencial de reactivos para mejorar la validez de una prueba de ingreso a la universidad. *Revista Electrónica de Investigación Educativa*. 2001; 3(1).
38. Soler Cárdenas SF. Coeficientes de confiabilidad de instrumentos escritos en el marco de la teoría clásica de los tests. *Rev. Cubana Educ Med Sup*. 2008; 22 (2).
39. Ledesma R, Molina Ibáñez G, Valero Mora P. Análisis de consistencia interna mediante Alfa de Cronbach: un programa basado en gráficos dinámicos *Rev. Psico-USF*. 2002, 7(2): 143-152.
40. Soler Cárdenas, Rodríguez Fontes R, Fernández Oliva B, Moreno Lazo M. Análisis de confiabilidad del test de habilidades múltiples para el ingreso en las carreras de Ciencias Médicas. *Rev Cubana Educ Med Sup* 2002;16(4): 15-19.
41. Argudín Somonte E, Díaz Rojas P, Leyva Sánchez E. Índice de Dificultad del examen de Morfofisiología Humana I. *Rev. Cubana Educ. Med. Sup*. 2011; 25(1).
42. Pérez Tapia JH, Acuña Aguilar N, Arratia Cuela E. Nivel de dificultad y poder de discriminación del tercer y quinto examen parcial de la cátedra de citohistología de la carrera de medicina de la UMSA. *Rev. Cuad. Art. Originales*. 2008; 53(2).
43. Backhoff Escudero E, Larrazolo Reyna N, Rosas Morales M. Nivel de dificultad y poder de discriminación del Examen de Habilidades y Conocimientos Básicos (EXHCOBA) *Revista Electrónica de Investigación Educativa*. 2000; 2(1).
44. 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.
45. Santiesteban Leyva, E. Propuesta de metodología para el análisis de los resultados de los exámenes finales de las asignaturas de la disciplina Morfofisiología Humana del Programa Nacional de Formación de Medicina Integral Comunitaria en la República Bolivariana de Venezuela. [Tesis para

- optar por el título de Master en Educación Médica Superior]. La Habana: ENSAP; 2007.
46. Carrazana Lee A, Salas Perea RS; Ruiz Salvador AK. Nivel de dificultad y poder de discriminación del examen diagnóstico de la asignatura Morfofisiología Humana I. Rev. Cubana Educ Med Sup. 2011; 25 (1).
47. Córica J.L Hernández Aguilar M. L, Holloway Creed J, Dimos C. Propuesta de indicadores de calidad para la evaluación de la calidad de las preguntas en los exámenes de opción múltiple. IV Congreso internacional virtual de educación a distancia CIVE 2004. [en línea] [Fecha de acceso 10 de enero de 2012] Disponible en: http://www.ateneonline.net/datos/91_03%20PONENCIA%20CORICA-HERNANDEZ-HOLLOWAY-DIMOU.pdf.
48. Vicedo Tomey A. Papel de las ciencias básicas biomédicas en la formación de los profesionales de ciencias de la salud. Material de estudio de la Maestría en Educación Médica. En CD-ROM Maestría en Educación Médica, ENSAP: 2005.
49. Sánchez González G, Tapia Pérez GG. Descripción del nivel de facilidad y poder de discriminación del examen de inferencia estadística en métodos estadísticos en medicina veterinaria y zootecnia. Universidad Autónoma de México. Facultad de Medicina Veterinaria y Zootecnia. [en línea] [Fecha de acceso 2 de abril de 2012] Disponible en: http://www.facmed.unam.mx/sem/jem2010/Abstracts/trabajos_orales.pdf

Anexo 1

Entrevista a expertos

Estimado profesor:

La presente entrevista se realiza como parte del desarrollo de una investigación en el campo del Proceso de Enseñanza Aprendizaje.

Necesitamos de usted su sincera opinión en relación con las preguntas que les realizamos. Su opinión será importante para mejorar el trabajo docente metodológico de los profesores que imparten la asignatura Morfofisiología Humana II y contribuir de esta manera al perfeccionamiento del sistema evaluativo.

Guía de la entrevista

1. ¿Qué área de las ciencias están exploradas en cada uno de los incisos del examen?
2. ¿Qué tipo de pregunta fue el utilizado en la confección de cada inciso del examen?
3. ¿Qué contenidos temáticos de la asignatura fueron explorados en cada inciso del examen?
4. ¿Como clasificaría los inciso y preguntas del examen de Morfofisiología Humana II de acuerdo a: difíciles, medianamente difíciles, dificultad media, medianamente fáciles y fáciles?

Anexo 2

UNIVERSIDAD DE CIENCIAS MÉDICAS. HOLGUÍN
EXAMEN ORDINARIO DE MORFOFISIOLOGÍA HUMANA II
Curso 2010-2011

Nombre: _____ Grupo:

I. De acuerdo a los conocimientos adquiridos en el estudio del SOMA responda verdadero o falso según corresponda.

___ Los huesos largos presentan una diáfisis y dos epífisis.

___ Entre dientes y alvéolos dentales se presentan una variedad de articulaciones cartilaginosa.

___ Los músculos flexores y extensores son ejemplo de acción sinergista.

___ Las variedades de tejido cartilaginosa son elástico, hialino y fibroso.

___ En los huesos planos predomina el largo y el ancho sobre el grosor.

___ Las articulaciones complejas presentan discos o meniscos.

___ La osteogénesis membranosa presenta modelos de cartílago hialino.

II. Complete las afirmaciones siguientes:

1. En la vista lateral del cráneo encontramos las fosas _____,
_____ y _____.

2. La osificación de los huesos de la bóveda craneana es de tipo _____.

3. La silla turca es un detalle del cráneo que se observa en una vista
_____.

4. Las fontanelas son articulaciones de tipo _____.

5. El maxilar se clasifica como un hueso _____.

6. Las porciones de la mandíbula son _____ y
_____.

7. El canal óptico es un agujero e orbita que comunica a esta con la cavidad
_____.

III. Paciente que acude a su médico de familia después de haberse realizado un rayo X de columna cervical observándose la presencia de osteofito en la última vértebra de esta región. Mencione:

- a) Cantidad de vértebras que tiene esta región de la columna y curvatura que presenta.
- b) Detalles anatómicos que presenta esta vértebra que las diferencia de las demás.
- c) Origen embriológico.

IV. Paciente que presenta un defecto congénito denominado sindactilia, en el miembro superior derecho. Ahora es atendido por presentar una epicondilitis a nivel del cinturón de ese miembro. Seleccione la o las respuestas correctas.

1. La falta de desarrollo del defecto congénito presente se debe a: ___exceso de proliferación celular. ___ no ocurre la muerte celular programada. ___ falta de migración celular.
2. La articulación afectada es: ___ escapulohumeral. ___ radioulnar proximal. ___ radiocarpiana.
3. Esta articulación se clasifica en: ___ sinovial compuesta condilar y poliaxil. ___ sinovial simple esferoidal poliaxil. ___ sinovial simple combinada condilar poliaxil.
4. Huesos que forman la articulación: ___ radio. ___ húmero. ___ escápula.

V. Complete los espacios en blanco en el siguiente cuadro:

Nombre del músculo	Grupo muscular	Función (una)
Masetero		
	Superficiales del dorso	Son auxiliares de la respiración
Triceps braquial		
Recto del abdomen	Anterior del abdomen	
	Anterior de la región glútea	Flexor por excelencia del muslo, rotación lateral del muslo, flexión del tronco con el MI. fijo.

VI. Ordene cronológicamente los eventos de la contracción muscular.

- ___ Transmisión del potencial de acción a la fibra muscular.
- ___ Unión de las cabezas de la miosina con la actina.
- ___ Llegada del potencial de acción a la fibra muscular.
- ___ Separación de las cabezas de miosina de la actina.
- ___ Queda expuesto los sitios activos de la actina para la miosina.
- ___ Unión del calcio con la actina.
- ___ Liberación del calcio de las cisternas del retículo sarcoplasmático

VII. Paciente que acude a su Médico de Familia para realizar cura de herida profunda en la palma de la mano derecha. De acuerdo a sus conocimientos sobre sistema tegumentario, responda:

- a) Que capas de la piel se han dañado.
- b) Explique la organización microscópica de la capa más superficial.
- c) Mencione el origen embriológico para la capa superficial.
- d) Refiérase a dos funciones que pueden estar dañada en este órgano.

Anexo 3. Errores e índice de dificultad por ítems.

Ítems	Nº. de suspensos.	% Contra Presentados.	Índice de dificultad
1.1	2	1.09	0.99
1.2	17	9.34	0.91
1.3	50	27.47	0.72
1.4	28	15.38	0.85
1.5	9	4.94	0.95
1.6	18	9.89	0.90
1.7	28	15.38	0.85
2.1.1	46	25.27	0.75
2.1.2	73	40.10	0.60
2.1.3	79	43.40	0.56
2.2	64	35.16	0.65
2.3	90	49.45	0.50
2.4	79	43.40	0.56
2.5	43	23.62	0.76
2.6.1	51	28.02	0.72
2.6.2	54	29.67	0.70
2.7	72	39.56	0.60
3.1.1	4	2.19	0.98
3.1.2	30	16.48	0.83
3.2.1	60	32.96	0.67
3.2.2	68	37.36	0.63
3.3	34	18.68	0.81
4.1.1	12	6.59	0.93
4.1.2	15	8.24	0.92
4.1.3	12	6.59	0.93
4.2.1	8	4.38	0.96
4.2.2	5	2.74	0.97
4.2.3	3	1.64	0.98

Calidad del examen final de la asignatura Morfofisiología Humana II

4.3.1	10	5.49	0.94
4.3.2	18	9.89	0.90
4.3.3	11	6.04	0.94
4.4.1	7	3.84	0.96
4.4.2	4	2.19	0.98
4.4.3	6	3.29	0.97
5.1	27	14.83	0.85
5.2	56	30.76	0.69
5.3	54	29.67	0.70
5.4	55	30.21	0.69
5.5	51	28.02	0.72
5.6	47	25.82	0.74
5.7	92	50.54	0.49
6.1	111	60.98	0.39
6.2	71	39.01	0.61
6.3	105	57.69	0.42
6.4	64	35.16	0.65
6.5	67	36.81	0.63
6.6	61	33.51	0.66
6.7	53	29.12	0.71
7.1.1	2	1.09	0.99
7.1.2	6	3.29	0.97
7.2.1	29	15.93	0.84
7.2.2	33	18.13	0.82
7.2.3	35	19.23	0.81
7.2.4	37	20.32	0.80
7.2.5	29	15.93	0.84
7.3	36	19.78	0.80
7.4.1	11	6.04	0.94
7.4.2	13	7.14	0.93

Fuente: Procesamiento estadístico.

Anexo 4. Índice de discriminación (D) por ítems.

Ítems	ID
Ítem 1.1	0,19
Ítem 1.2	0,2
Ítem 1.3	0,22
Ítem 1.4	0,22
Ítem 1.5	0,2
Ítem 1.6	0,23
Ítem 1.7	0,21
Ítem 2.1.1	-0,085
Ítem 2.1.2	-0,03
Ítem 2.1.3	-0,01
Ítem 2.2	-0,07
Ítem 2.3	0,05
Ítem 2.4	-0,03
Ítem 2.5	-0,12
Ítem 2.6.1	-0,089
Ítem 2.6.2	-0,08
Ítem 2.7	-0,06
Ítem 3.1.1	0,13
Ítem 3.1.2	0,15
Ítem 3.2.1	0,25
Ítem 3.2.2	0,25
Ítem 3.3	0,16
Ítem 4.1.1	0,25
Ítem 4.1.2	0,26
Ítem 4.1.3	0,26
Ítem 4.2.1	0,26
Ítem 4.2.2	0,25
Ítem 4.2.3	0,25
Ítem 4.3.1	0,26

Ítem 4.3.2	0,26
Ítem 4.3.3	0,26
Ítem 4.4.1	0,25
Ítem 4.4.2	0,25
Ítem 4.4.3	0,25
Ítem 5.1	0,025
Ítem 5.2	0,08
Ítem 5.3	0,08
Ítem 5.4	0,08
Ítem 5.5	0,089
Ítem 5.6	0,27
Ítem 5.7	0,18
Ítem 6.1	0,071
Ítem 6.2	-0,09
Ítem 6.3	0,47
Ítem 6.4	-0,11
Ítem 6.5	-0,1
Ítem 6.6	-0,12
Ítem 6.7	-0,13
Ítem 7.1.1	0,21
Ítem 7.1.2	0,22
Ítem 7.2.1	0,26
Ítem 7.2.2	0,26
Ítem 7.2.3	0,25
Ítem 7.2.4	0,25
Ítem 7.2.5	0,26
Ítem 7.3	0,24
Ítem 7.4.1	0,23
Ítem 7.4.2	0,23

Fuente: Procesamiento estadístico.