

Calidad del examen final de la asignatura Psicología Médica I. Holguín. Curso 2015 – 2016.

Quality of the subject Medical Psychology I final test. Holguín. Academic year 2015 – 2016.

Sussel Expósito Rubio,¹ Gemma Margarita Ortiz Romero,² Yamela Aleida Pupo Mora..

1 Licenciada en Psicología, máster en educación médica. Profesor asistente. Policlínico Universitario Julio Grave de Peralta. Holguín.

2 Licenciada en educación, máster en educación médica. Profesor auxiliar. Universidad de Ciencias Médicas. Holguín.

3 Licenciada en Psicología, máster en educación médica. Profesor asistente. Universidad de Ciencias Médicas. Holguín.

Correspondencia a: suselerhg@infomed.sld.cu

RESUMEN

La evaluación del aprendizaje es una parte esencial del proceso de enseñanza-aprendizaje. La presente investigación se realizó con el objetivo de caracterizar la calidad del examen final de la asignatura Psicología Médica I aplicados a estudiantes de segundo año de la carrera de medicina del municipio Holguín, en el curso 2015 – 2016. Es una investigación de evaluación con enfoque mixto. Se emplearon métodos teóricos, empíricos y procedimientos estadísticos. Entre los empíricos se aplicó una entrevista a especialistas, se describieron las características de los dos temarios, se identificó la correspondencia entre los contenidos de la unidad curricular y su exploración en los temarios, se calculó el índice de dificultad, coeficiente de discriminación del punto de correlación biserial y el coeficiente alfa de Cronbach como indicadores de calidad de los instrumentos. Los resultados evidenciaron que ambos temarios fueron de formato mixto con predominio de las preguntas test objetivo. Se encontró correspondencia moderada entre los contenidos de la unidad curricular y su exploración en los instrumentos, corroborado por el coeficiente de correlación de Spearman. El temario II fue el de mayor grado de dificultad, con la temática de personalidad y salud como más difícil. Según el coeficiente de correlación del punto biserial (rpbis) todas las preguntas presentan un excelente poder predictivo. El temario II presenta mayor confiabilidad. La caracterización del examen demostró insuficiencias en la planeación de los instrumentos evaluativos.

Palabras clave: calidad, Psicología Médica, evaluación.

ABSTRACT

The evaluation of learning is an inherent process of the development of the educational-teaching process that allows its regulation and feedback. The investigation purpose was to value the quality of the subject Medical Psychology I final test. It was carried out an educational research work in Holguín University of Medical Sciences during the academic year 2015-2016. The study universe was constituted by the total of ordinary final exams of the subject Medical Psychology I at the municipality Holguín. There were used theoretical method, empiric method and stadistical procedures. Between empiric ones a survey in interview form was applied to key informants, they described the characteristic of the two test, set 1 and 2, the correspondence between the duration of the exam and the quantity of items explored in the both set was analyzed, it was calculated the index of difficulty and the coefficient of correlation of the difficulty level of the items on both set and the coefficient of Cronbach's alpha like indicators of quality. The results evidenced both set, set 1 and 2 came from mixed format with predominance of the questions objective test. The correspondence between the time dedicated to each topic and the quantity of explored ítems was moderated, corroborated for Spearman´s correlation coefficient. The set II was the one with bigger degree of difficulty with the subject matter of personality and health like more difficult. According to the correlation coefficient biserial point (rpbis) all the question presented an excellent predictive power. The set II presentes bigger reliability. The characterization of the exam demonstrated insufficiencies in the planning of the instruments evaluativos.

Key words: quality, Medical Psychology, evaluation.

INTRODUCCIÓN

El desarrollo de la educación médica como una necesidad que imponen los diferentes cambios sociales, entre otros factores, condiciona su perfeccionamiento y al mismo tiempo el de la evaluación, que es uno de los asuntos más polémicos en cualquier modelo pedagógico y uno de los medios para mejorar la calidad del proceso docente educativo.

En la evaluación del aprendizaje, los exámenes escritos, son una de las varias herramientas que forman parte de un amplio abanico de opciones evaluativas que tienen carácter de sistema, y tienen como función general la regulación del proceso docente educativo.¹

En el ámbito de las ciencias médicas los exámenes implican una enorme responsabilidad social, ya que a través de ellos se reconoce y autoriza una determinada práctica profesional vinculada directamente con la vida de las personas, por lo que su elaboración correcta es fundamental para la verificación de la calidad del proceso docente educativo y del futuro egresado.

A nivel internacional se le dedica gran importancia al tema. La Universidad Nacional Autónoma de

México y en Argentina, a través de la Dirección Nacional de Evaluación del Ministerio de Cultura y Educación han desarrollado importantes investigaciones en esta dirección²⁻³. Igualmente la Asociación Peruana de Facultades de Medicina (ASPEFAM) ha seguido esta línea investigativa.⁴

En Cuba se encuentran múltiples trabajos en Ciencias Médicas. Diversas facultades en el país han realizado investigaciones sobre el tema en especialidades de enfermería, estomatología y medicina.⁵⁻⁸ En la Universidad de Ciencias Médicas de Holguín se han desarrollado investigaciones sobre la calidad de exámenes en varias asignaturas sin embargo no existe referente de este tema dentro de la carrera de Medicina en la Disciplina Psicología Médica.

Esta disciplina esta compuesta por dos asignaturas, Psicología Médica I y II, y concibe una serie de habilidades de interés para la formación del médico general.

La disciplina es el producto de una nueva forma de pensamiento en salud, que considera la dimensión psicosocial en el proceso salud-enfermedad en los niveles conceptual, metodológico y en la organización de los servicios de atención a la población. Ha sido un resultado natural de las propias limitaciones del modelo biomédico, del fracaso de los sistemas sanitarios con una concepción restrictiva y descontextualizadora de la salud y de la posibilidad que se abre cada vez más de comprender cómo las variables sociales y personales mediatizan este proceso.⁹

El sistema de evaluación de la asignatura Psicología Médica I se desarrolla a través de evaluaciones frecuentes, orales y escritas, trabajo final por equipos y examen final escrito. No obstante en los colectivos de asignatura se han identificado una serie de problemas prácticos relacionados con la calidad de la evaluación del examen final de la asignatura.

Por todo lo expuesto, el **problema científico** que sustenta la presente investigación expresa: no existe una caracterización sobre bases científicas de la calidad de los instrumentos evaluativos en la asignatura Psicología Médica I y se plantea como **objetivo**: Caracterizar la calidad del examen final de la asignatura Psicología Médica I aplicado en la Facultad de Ciencias Médicas de Holguín en el curso 2015 -2016.

DISEÑO METODOLÓGICO

Se realizó una investigación de evaluación con un enfoque mixto en la Facultad de Ciencias Médicas de Holguín en el período de enero 2015 a marzo 2017, el campo de la ciencia fue el proceso docente educativo, el objeto de investigación la evaluación del aprendizaje en la asignatura Psicología Médica I y el campo de acción el examen final de la asignatura Psicología Médica I en la carrera de Medicina.

El Universo estuvo constituido por el total de exámenes finales ordinarios de Psicología Médica I aplicado a los estudiantes de segundo año de la carrera de Medicina del municipio Holguín en el curso 2015 – 2016 y los docentes que imparten la asignatura en el municipio Holguín.

Se seleccionó una muestra intencionada de todos (658) los exámenes de los estudiantes presentados en la primera convocatoria en el municipio Holguín: 328 en el temario I y 330 en el temario II, además, cinco especialistas de la asignatura con más de 10 años de experiencia docente, categoría docente superior, tres de los cuales son masters en Ciencias.

Se utilizaron métodos teóricos y empíricos.

Métodos Teóricos

Histórico lógico: para relacionar diferentes momentos y concepciones vinculadas con el desarrollo histórico de la disciplina, de la evaluación del aprendizaje e instrumentos de evaluación en los que se sustentó el marco teórico.

Análisis síntesis: para el procesamiento de la información teórica y empírica, al realizar el análisis de los exámenes, la elaboración de los resultados y conclusiones sobre la calidad del examen aplicado.

Inducción deducción: para arribar a generalizaciones acerca de las potencialidades y limitaciones que presentó el instrumento de evaluación aplicado.

Revisión documental: para la revisión del programa analítico y otros documentos de la asignatura y la carrera de Medicina, entre ellas bibliografía relacionada con el proceso docente educativo, trabajos investigativos, artículos sobre evaluación del aprendizaje, calidad de los instrumentos de evaluación y otras de interés para las autoras.

Métodos Empíricos

Se utilizó la investigación por encuesta a través de la entrevista. La entrevista fue programada con anterioridad y realizada personalmente. Se aplicó a los especialistas seleccionados compuesta por ocho preguntas que permitieron caracterizar los exámenes según tipos de preguntas, contenidos evaluados en cada uno de los incisos y preguntas, escala para clasificar las preguntas e incisos según su índice de dificultad, así como nivel de asimilación,

Para el análisis de la calidad de los exámenes se siguió la metodología propuesta por Díaz Rojas y Leyva Sánchez.¹⁰

Procesamiento de datos.

I. Análisis del examen aplicado.

Para la caracterización de los temarios se tuvo en cuenta:

- ✓ Tipo de examen.
- ✓ Cantidad de preguntas del examen.
- ✓ Tipos de preguntas que contiene el instrumento.
- ✓ Cantidad de inciso por preguntas.
- ✓ Contenido que aborda cada pregunta.
- ✓ Nivel de asimilación de las preguntas según la taxonomía cubana.

II. Correspondencia del fondo de tiempo asignado en el P1 a los contenidos y su exploración en los exámenes.

Se calculó el número de horas que se le dedica a cada contenido de la unidad curricular según el plan calendario de la asignatura y se determinó el por ciento que representa del total de horas dedicadas al tema.

Se clasificaron los incisos de los exámenes según el contenido temático que exploraron, se determinó el por ciento de incisos que exploraron cada temática en base al total de incisos o ítems de cada temario.

Se utilizó el coeficiente de correlación de Spearman para identificar la correspondencia entre el tiempo dedicado en el plan calendario a los contenidos de los temas y su representación en los incisos de los temarios. Los valores del coeficiente de correlación (r) próximos a 1 significan que la distribución de las preguntas en los exámenes se realizó en correspondencia con el tiempo que se le dedica a cada temática en el plan calendario de la asignatura.

Indicadores de Calidad.

III. Determinación del índice de dificultad para los incisos, temáticas, preguntas y exámenes.

A partir de los datos primarios se procedió al cálculo del índice de dificultad para cada inciso, pregunta, temática y examen.

Para identificar el índice de dificultad del inciso, se dividió el número de estudiantes que contestaron correctamente el inciso entre el número total que contestaron el inciso. Esta proporción se le denota como D_i y se utilizó la siguiente

Fórmula: ¹⁰

$$D_i = \frac{A}{N}$$

Donde:

D_i = Índice de dificultad del inciso.

A = Número de respuestas correctas del inciso.

N = Número total de estudiantes que contestaron el inciso.

Cuanto mayor sea esta proporción, menor será su dificultad. Lo que quiere decir que se trata de una relación inversa: a mayor dificultad del inciso, menor será su índice.

Para identificar el índice de dificultad de la pregunta, se dividió el número de respuestas correctas entre el número total de exámenes, esta relación se denomina con la letra D_p .⁵³

$$Dp = \frac{\sum(A_{p1} + A_{p2} + \dots + A_{pp})}{\frac{n_{tp}}{Nt}}$$

Donde:

Dp = Índice de dificultad de la pregunta.

Ap = Número de respuestas correctas de los incisos de la pregunta.

Ntp = Número total de incisos de la pregunta.

Nt = Número total de presentados en el examen.

Para calcular el índice de dificultad de cada temática explorada en el examen, se sumó el total de respuestas correctas de cada uno de los incisos que corresponden a la temática explorada y se dividió entre el número de incisos que exploró esa temática en el examen, obteniéndose la media de respuestas correctas para la temática, luego se calculó el índice de dificultad el cual se corresponde con el promedio de respuestas correctas entre el total de examinados, como se representa en la siguiente.¹⁰

$$Dt = \frac{\sum(A_{t1} + A_{t2} + \dots + A_{tt})}{\frac{n_{tt}}{Nt}}$$

Donde:

Dt = Índice de dificultad de la temática.

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

Ntt = Número total de incisos que exploran la temática.

Nt = Número total de presentados en el examen.

Determinación de los intervalos para la clasificación de los incisos según su Índice de dificultad.

Se realizó sobre la base del consenso del criterio de especialistas, y se definieron los intervalos según la proporción de estudiantes aprobados en ellos, se estableció que los incisos que tuvieron un Índice de dificultad:¹⁰

- Menor de 0.40, se consideran difíciles.
- Entre 0.41 y 0.50, se consideran medianamente difíciles.
- Entre 0.51 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.

Distribución de los incisos según el índice de dificultad pronóstico y el obtenido en cada examen.

Se distribuyeron los incisos de los exámenes según su índice de dificultad obtenido en cada uno de los intervalos de clasificación definidos previamente.

El Índice de dificultad pronóstico se determinó a partir del estándar propuesto por Backhoff¹⁴ para el análisis de dificultad en instrumentos de evaluación, el que establece que dichos instrumentos deben tener un:

- ✓ 5% de incisos fáciles.
- ✓ 20% de incisos medianamente fáciles.
- ✓ 50% de incisos de dificultad media.
- ✓ 20% de incisos medianamente difíciles.
- ✓ 5% de incisos difíciles.

Se elaboró una tabla de doble entrada en cuyas columnas se plasmaron las frecuencias absolutas y porcentaje de los índices de dificultad esperados y reales de dos exámenes aplicados. Se utilizó la prueba estadística de correlación de rangos de Spearman para determinar la correspondencia entre lo esperado y lo real.

VI. Determinación del Coeficiente de discriminación del punto biserial (rpbis):

Este indicador permitió calcular la efectividad discriminativa de cada pregunta en ambos exámenes.

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

Se calculó a través de la siguiente fórmula.¹⁰

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

Donde:

X_1 = Media de las puntuaciones totales de aquellos que respondieron correctamente el inciso.

X_0 = Media de las puntuaciones totales de aquellos que respondieron incorrectamente el inciso.

S_x = Desviación estándar de las puntuaciones totales.

n_1 = Número de estudiantes que respondieron correctamente el ítem.

n_0 = Número de estudiantes que respondieron incorrectamente el ítem.

$n = n_1 + n_0$

Según estándares internacionales la distribución de los resultados del coeficiente discriminación punto de correlación biserial (rpbis) es:

- < 0 discriminan negativamente.

- 0 0,14 discriminan pobremente.

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

VII . Deteminacion del coeficiente alfa de Cronbach.

El coeficiente alfa de Cronbach permite determinar la consistencia interna de un instrumento y brinda información relacionada con el aporte que hace cada pregunta a la consistencia general del examen. Se calculó para cada examen y al eliminar una pregunta del examen.

Su interpretación se realiza al tomar en cuenta cómo aumenta o disminuye el alfa sin la pregunta, si el valor es igual o inferior al alfa del examen la pregunta aporta consistencia interna al examen, pues al eliminarla el coeficiente desciende, en caso contrario, si el valor del alfa del examen se eleva al eliminar la pregunta, es un indicador de que la pregunta afecta la consistencia interna y disocia los resultados del examen.

Aspectos éticos de la investigación.

Se consideró la voluntariedad de participación de los implicados a los cuales se les comunicó que las informaciones obtenidas solo serían utilizadas con fines científicos.

ANÁLISIS Y DISCUSIÓN DE LOS RESULTADOS

I. Análisis del examen aplicado.

El instrumento de análisis fue confeccionado por especialistas de la Universidad de Ciencias Médicas "Mariana Grajales Coello" y se aplicó en el primer semestre del curso escolar 2015 - 2016. El temario I fue aplicado a 328 estudiantes y el temario II a un total de 330.

Ambos exámenes están compuestos por siete preguntas, el temario I tiene 45 incisos y el temario II, 41. Son de formato mixtos con preguntas de desarrollo y de test objetivo.

Los temarios contaron con los datos en el encabezamiento: Facultad, asignatura, fecha y año. Se omitió la carrera a quién sería destinado el examen y el tipo de examen.

En los dos instrumentos se exploraron contenidos de la totalidad (tres) de las temáticas de la asignatura. El tema 1 referido a las generalidades e introducción a la Psicología Médica. El tema 2 sobre personalidad y salud y el tema 3 sobre la psicología en la atención primaria de salud.

Las áreas de aprendizaje que estuvieron representadas en ambos temarios fueron los principios de la psicología como ciencia independiente, los procesos psíquicos, personalidad, tipos y modelos y su relación con la salud y la enfermedad, rasgos clínicos, las etapas del desarrollo psíquico y sus particularidades psicológicas, los factores de riesgo para la salud del individuo, la familia y la comunidad, y las vías de comunicación para la promoción y la prevención de salud.

La frecuencia de incisos con que se exploró cada área en los temarios aparece representada en el gráfico 1.

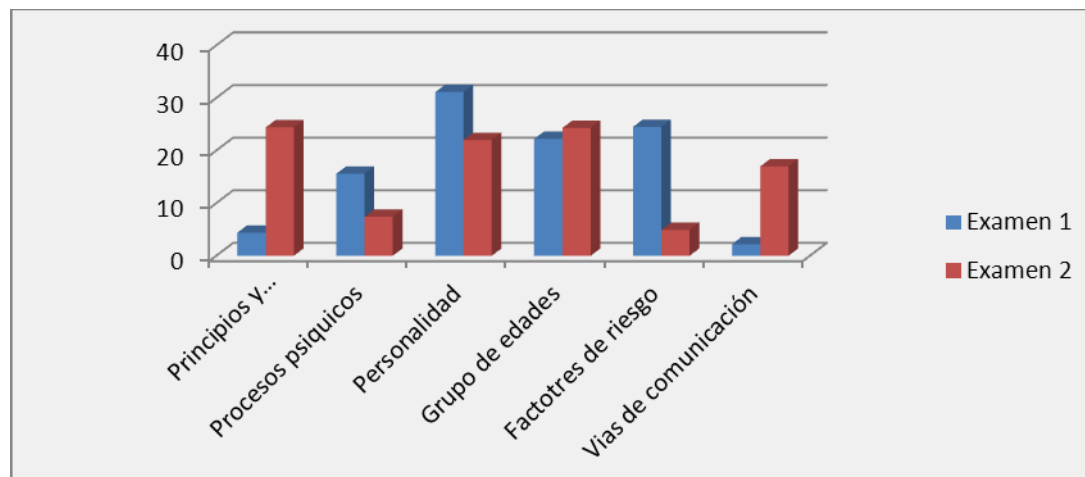
En el temario I aparece mayor representado el tema 2 específicamente contenidos de personalidad con un total de 14 incisos, lo que representa 31.1 %, seguido por los factores de riesgo para el individuo, la familia y la comunidad con 11 incisos para 24.5 % y luego los grupos de edades y sus particularidades psicológicas con 10 incisos para 22.2 % del total.

En el temario II aparecen, igualmente con 10 incisos, que representan 24.3 % los temas referidos a grupo de edades con sus particularidades y los principios de la psicología y generalidades. Luego el tema personalidad con 9 incisos para 21.9 %. En ambos temarios predominó un tema por pregunta y se aborda abundantemente el tema de personalidad y salud.

Esto es consistente con los objetivos instructivos de la asignatura que incluye la interpretación y predicción de conductas del individuo, si se conoce la interrelación de los subsistemas de la personalidad que las regulan, como organización sistémica más compleja y pluridimensional del psiquismo humano así como las particularidades de cada grupo etario lo que facilita el trabajo del médico.

GRÁFICO 1

PORCIENTO DE INCISOS POR ÁREA DE APRENDIZAJE QUE EXPLORA CADA TEMARIO



El tema personalidad se imparte por primera vez en la carrera, a diferencia de los factores de riesgo para el individuo, la familia y la comunidad que son retomados por esta asignatura, pero fueron tratados con anterioridad en asignaturas como Promoción y Prevención en Salud, y Medicina Comunitaria. Se pone de manifiesto la interdisciplinariedad con la asignatura rectora.

Otro de los objetivos del programa es realizar acciones de promoción y prevención encaminadas a lograr el incremento del estado de salud de la población. Las autoras coinciden con los especialistas entrevistados en que se debe incrementar las preguntas relacionadas con las acciones de promoción y prevención por relacionarse con los modos de actuación del futuro profesional, dándole especial importancia a las particularidades de cada etapa.

Sin embargo, en el temario II se le dedica tiempo a un contenido teórico con poca vinculación para el futuro profesional (Los principios de la Psicología) que no debieron ser explorados con 10 incisos pues aporta poco a la práctica profesional, no así a la cultura general de los futuros galenos.

Los dos temarios están formados por tres preguntas de desarrollo en su variante de respuesta corta y cuatro de test objetivo: dos de asociación simple y dos de verdadero /falso.

Hay similitud en cuanto al tipo de preguntas utilizadas en ambos exámenes. Las autoras se suma a la opinión de los entrevistados, quienes consideran hay monotonía en cuanto a su confección. No se explora otros tipos de preguntas que aportarían estética, novedad y complejidad a los exámenes tales como análisis progresivo, complemento agrupado, asociación compuesta. Las autoras consideran que se debe a que dichas preguntas conllevan mayor complejidad y meticuloso esfuerzo en su elaboración y desconocimiento en su confección.

Las preguntas utilizadas son, no obstante, muy ventajosas cuando se trata de un número elevado de estudiantes a evaluar pues sus calificaciones son más fáciles y permiten medir varios contenidos esenciales o básicos.

El temario II presenta menor cantidad de incisos en preguntas de desarrollo o ensayo aunque todas en su forma de respuesta corta. El menor por ciento de cada temario corresponde a las preguntas de desarrollo o ensayo con 18 y 17 % respectivamente.

Este tipo de preguntas permite que el estudiante exponga sus conocimientos, ordene sus ideas y demuestre su capacidad de expresión, así como una mayor evaluación de la instrucción 1/2009 acerca de los errores ortográficos y de redacción a tener en cuenta en todas las evaluaciones escritas que se realicen en los centros de Educación Superior . Las autoras consideran es una dificultad real a resolver por las diferentes asignaturas.

Aunque tiene como desventaja que limitan el número de áreas a explorar, su calificación consume tiempo y el resultado puede estar mediatizado por la subjetividad del calificador.

Dichos resultados coinciden con los encontrados por Roméu Escobar y Díaz Quiñones⁷ en la Facultad de Ciencias Médicas "Dr. Raúl Dorticós Torrado" en Cienfuegos, a los exámenes finales de Estomatología y Medicina. Los autores observaron que el formato de pregunta más utilizado en medicina fue respuesta alternativa (V/F), ensayo de respuesta corta y complemento simple.

En el nivel de asimilación de las preguntas se obtuvieron los siguientes resultados que se reflejan en el siguiente cuadro.

Se pudo observar que en ambos temarios hubo predominio de incisos a nivel de familiarización, en el examen I con 39 incisos (87 %) y en el examen II, 34 incisos (83 %). En ningún examen existió inciso a nivel creativo.

Las habilidades más utilizadas fueron identifique y mencione. Hay un solo inciso donde se solicita al estudiante explicar. Se repite el encabezado de las preguntas.

Las preguntas utilizadas fundamentalmente en este examen premian la memorización y no el pensamiento y análisis lógico, aspecto que fue considerado por el grupo de especialistas entrevistado. Los resultados anteriores difieren con el estudio realizado por Franco Pérez¹¹ donde predominan preguntas cuyas tareas se plantean a nivel reproductivo.

Estos resultados coinciden con los obtenidos por López Espinosa¹² y colaboradores en estudio realizado en la Universidad de Ciencias Médicas de Villa Clara, durante el curso 2011-2012 en las carreras de Medicina, Estomatología y Enfermería. Se observó ítems con dificultad en el diseño por el uso indiscriminado de términos como: menciona, diga, así como igual encabezamiento de preguntas en un mismo instrumento.

Las autoras consideran, que resulta imperioso abordar las fortalezas que ofrece el trabajo metodológico para la preparación pedagógica del claustro en la confección o diseño de los instrumentos, con especial atención a la materia de evaluación como mecanismo regulador del sistema y elemento de control de la calidad del producto resultante.

Al analizar los resultados del examen final de la asignatura se observa 89.5 % de estudiantes aprobaron el examen.

En el temario I predominan los estudiantes con calificación de excelente (5) y bien (4), 38 y 169 estudiantes para 11.5 % y 51.2 % respectivamente. Mientras en el temario II solo hay 123 estudiantes con calificación de bien (4) y 24 con calificación de excelente (5) para 37.2 y 7.2 % respectivamente.

En el temario II predominan los estudiantes con calificación de regular (3) y desaprobados (2) para 40.6 y 15 % respectivamente.

II. Correspondencia entre los contenidos por temas y su exploración en los temarios.

El cuadro 1 muestra el tiempo dedicado según plan calendario de la asignatura a los temas y la frecuencia con que se exploran en los temarios.

Se observa que el programa de la asignatura tiene un fondo de tiempo de 48 horas distribuidas en tres temas. El tema uno con 10 horas para 28.3 %. El tema dos referido a Personalidad y salud con 16 horas para 33.3 % y el tema tres referido a la psicología en la atención primaria con 22 horas que representa 45.9 %.

Cuadro 1. Distribución del fondo de tiempo según programa y presencia de los incisos en las temáticas de los temarios I y II.

Temas	Fondo de tiempo	Presencia en el temario I *	Presencia en el temario II **
--------------	------------------------	------------------------------------	--------------------------------------

	Horas	%	Inciso	%	Inciso	%
Generalidades	10	20.8	9	20	13	32
Personalidad y salud	16	33.3	24	53.3	19	46
En la atención primaria	22	45.9	12	26.7	9	22
Total	48	100	45	100	41	100

Fuente: Programa de la asignatura y examen.

* $r=0.67$

** $r=0.67$

Hay mayor correspondencia en el tema uno que en los dos restantes temas y su distribución en el examen.

Se calculó el test de correlación de Spearman. Se obtuvo que ambos temarios presentan un valor de 0.67. Demuestra que con la planeación de la evaluación en la asignatura se logró una correspondencia moderada entre cada temática y su exploración en el examen de acuerdo a la cantidad de ítems. Ambos con una $p < 0,05$ IC 95 %.

Estos resultados difieren con los obtenidos en la investigación de Quiñones Gutiérrez¹³ en la Carrera de Enfermería en la Universidad de Ciencias Médicas de Holguín donde no existe correspondencia entre el fondo de tiempo asignado para cada rotación y las preguntas elaboradas.

III. Determinación del índice de dificultad para los incisos, temáticas, preguntas y temarios.

El nivel de dificultad, llamado también como índice de facilidad, es la proporción de personas que responden correctamente un reactivo de una prueba. Este índice muestra qué tan fácil (valores cercanos a 1) o difícil (valores cercanos a 0) ha resultado la pregunta para el total de esos estudiantes.

El índice de dificultad del temario I fue de 0.76 y para del temario II fue de 0.74. Ambos datos ubican a los temarios con dificultad media. De acuerdo al manual del EXHCOBA, el nivel medio de dificultad del examen debe oscilar entre 0.5 y 0.6.¹⁴

Ambos temarios se acercan a lo planteado por Backhoff, aunque con una ligera disminución de la dificultad. Los resultados son considerados positivos por las autoras.

Lo referido anteriormente coincide con los valores obtenidos por Pupo Mora¹⁵ en su investigación de La calidad del examen estatal teórico de la Carrera de Psicología Mención Salud.

El cuadro 2 muestra el índice de dificultad por preguntas de los temarios I y II.

En el temario I, cinco preguntas presentaron dificultad media y dos preguntas medianamente fáciles. La pregunta tres y seis clasificaron como medianamente fácil mientras la uno, dos, cuatro, cinco y siete fueron de dificultad media.

En el temario II solo una pregunta fue medianamente fácil y el resto fue de dificultad media. La pregunta siete clasificó como medianamente fácil. En ninguno de los temarios hubo pregunta difícil o medianamente difícil.

Estos resultados coinciden con los obtenidos por Ortiz Romero¹⁶ y Ferrer Viant¹⁷ en investigaciones realizadas a los Exámenes finales de Sistemas reguladores del medio externo y de Metodología de la Investigación y Estadística en la Facultad de Estomatología "Raúl González Sánchez" y de Medicina "Mariana Grajales" respectivamente, donde no hubo pregunta difícil y hubo mayoría de preguntas con dificultad media.

Cuadro 2. Grado de dificultad por preguntas de los temarios.

Preguntas	Temario I		Temario II	
	Total de errores	Grado de dificultad	Total de errores	Grado de dificultad
Pregunta 1	98 (2)	0.70	79 (10)	0.76
Pregunta 2	92 (7)	0.72	120 (3)	0.64
Pregunta 3	63 (10)	0.81	104 (7)	0.68
Pregunta 4	71 (4)	0.78	92 (2)	0.72
Pregunta 5	101 (10)	0.69	86 (10)	0.74
Pregunta 6	36 (2)	0.89	105 (2)	0.68
Pregunta 7	74 (10)	0.77	59 (7)	0.82

Fuente: Temarios aplicados. n=328 (45 ítems) n=330 (41 ítems)

En el temario I la pregunta seis clasificó como medianamente fácil con el mayor índice de dificultad. Las autoras y los especialistas entrevistados consideran estuvo bien elaborada pues relaciona contenidos de la asignatura con los adquiridos en asignaturas precedentes, en especial con Prevención de salud lo que justifica ese resultado.

La interdisciplinariedad constituye una necesidad para la mejor preparación de los futuros egresados, aspecto en el que coinciden las autoras con la investigación de Caballero Cáceres³⁹.

La pregunta tres clasificada como medianamente fácil. Es una pregunta a nivel de familiarización.

En el temario II, la pregunta seis es la única clasificada como medianamente fácil. Es un contenido que ha sido abordado por otras asignaturas por lo que justifica esa clasificación. Pudiera vincularse más con la práctica del futuro profesional y en especial abordar la técnica dinámica de grupo, que es el contenido nuevo que aporta la asignatura. El resto de las preguntas clasifican como dificultad media.

La pregunta uno es de Verdadero o Falso fue la de mayor índice dentro de la dificultad media. Es un contenido teórico, que le permite al estudiante conocer e interpretar la concepción científica del

mundo así como la obtención de la información sobre la salud y la enfermedad, aporta conocimientos y cultura general al estudiante.

La pregunta cinco de test objetivo clasificó con 0.74 y explora contenidos que tiene gran vinculación con la práctica profesional. El conocimiento de los grupos de edades y sus particularidades biopsicosociales lo que permitirá caracterizar la salud individual y colectiva de su población para desarrollar actividades de promoción, prevención y orientación. Es de vital importancia y sin embargo, su clasificación informa que no se vencieron los objetivos. Los ítems que conforma la pregunta están bien elaborados.

La pregunta seis de ensayo vincula contenidos conocidos por otras asignaturas y acerca al estudiante a detectar factores de riesgo ambientales para la salud, pero se queda a un nivel de familiarización.

Por lo anterior referido, las autoras opinan que aún se tienen que perfeccionar cada vez más las preguntas y tener en cuenta la necesidad de elaborar mayor cantidad incisos difíciles para delimitar mejor entre los estudiantes.

En el cuadro 3 se muestra la distribución de incisos según grado de dificultad esperado y real de los temarios de acuerdo a la escala establecida por el criterio de los especialistas.

En los temarios no se cumplió con el pronóstico para los incisos fáciles. Se incrementaron en ocho más de lo esperado, para 22.2 % en el temario I y 24.4 % en el temario II. Mientras para los incisos difíciles en el temario II hubo correspondencia para 4.9 % y en el temario I hubo un ligero aumento de 8.9 % de 5% esperado. Valores considerados cercanos por las autoras.

En los incisos medianamente difíciles no hubo correspondencia en ningún temario. No hubo representatividad en el temario II y escasa en el temario I para 6.7 % del 20 % esperado.

En los incisos medianamente fáciles y dificultad media hubo correspondencia relativa el temario I hubo 11 incisos medianamente fáciles de nueve esperados para un 24.4 % y en el temario II hubo cinco incisos de ocho esperados para 12.2 %. Los valores están cercanos a los esperados.

Cuadro 3. Distribución de incisos de los temarios I y II según grado de dificultad esperado y real.

Incisos según grado de dificultad	Temario I				Temario II			
	Incisos esperados		Incisos reales		Incisos esperados		Incisos reales	
	No	%	No	%	No	%	No	%
Fáciles	2	5	10	22.2	2	5	10	24.4

Medianamente fáciles	9	20	11	24.4	8	20	5	12.2
Dificultad media	23	50	17	37.8	21	50	24	58.5
Medianamente difícil	9	20	3	6.7	8	20	0	0
Difícil	2	5	4	8.9	2	5	2	4.9
Totales	45	100	45	100	41	100	41	100

Fuente: Programa de la asignatura

En los incisos con dificultad media los valores reales se acercaron más a los esperados con un ligero aumento, de 21 a 24 incisos para un 58.5 % en el temario II mientras que en el temario I, hubo disminución para un 37.8 % no obstante las autoras consideran cercanos los valores.

En los incisos difíciles hubo correspondencia en ambos temarios. El temario II tuvo mejores valores que el temario I. Los resultados expuestos anteriormente corroboran los valores obtenidos por el coeficiente de correlación de Spearman ($r=0.67$)

Con una mejor planeación de la evaluación y capacitación de los profesores, en la elaboración correcta de preguntas e incisos se podrá lograr mejores instrumentos escritos en la asignatura. El análisis de la calidad de los instrumentos permite resultados cada vez mejores y proporciona un mayor perfeccionamiento de la evaluación como componente indispensable del proceso enseñanza aprendizaje.

VI. Determinación del Coeficiente de discriminación del punto biserial (rpbis):

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

Se sugiere que el rpbis es una medida que combina la relación entre el criterio del reactivo y el nivel de dificultad.

El Coeficiente de correlación punto biserial (rpbis) fue aplicado sólo a las preguntas.

El cuadro 4 muestra el coeficiente de correlación del punto biserial (rpbis) por preguntas en ambos exámenes.

Al analizar el coeficiente de correlación biserial (rpbis) por cada pregunta de los temarios en estudio se obtuvo que todas las preguntas presentan un excelente poder predictivo ($rpbis > 0,35$), resultados que coinciden con lo planteado por Backhof¹⁴ como las cifras promedio que se buscan en un examen para valorar la capacidad de predicción de un instrumento.

Los resultados obtenidos coinciden con los obtenidos por Rivera González¹⁸ en su investigación sobre la Calidad del instrumento evaluativo en el examen escrito de Medicina Interna en la Universidad de

Ciencia Médicas de Holguín donde cuatro de un total de cinco preguntas obtuvieron un excelente poder de discriminación.

En el temario I los valores oscilaron entre 0.37 y 1.00. El coeficiente de discriminación más alto corresponde a la pregunta cinco (1.00) seguido por la pregunta uno (0.93).

En el temario II el coeficiente de discriminación más alto corresponde a la pregunta dos con un valor de 0.85 seguido por la pregunta cuatro con un valor de 0.83. Los valores oscilaron en el temario entre 0.39 y 0.85.

Cuadro 4. Coeficiente de correlación del punto biserial (Rpbis) por preguntas.

Preguntas	Temario I		Temario II	
	Rpbis	Discriminación	Rpbis	Discriminación
Pregunta 1	0.93	Excelente	0.62	Excelente
Pregunta 2	0.37	Excelente	0.85	Excelente
Pregunta 3	0.43	Excelente	0.39	Excelente
Pregunta 4	0.76	Excelente	0.83	Excelente
Pregunta 5	1.00	Excelente	0.61	Excelente
Pregunta 6	0.41	Excelente	0.73	Excelente
Pregunta 7	0.46	Excelente	0.40	Excelente

Fuente: Examen.

n=328

n=330

En ambos temarios hay coincidencia entre el mayor valor obtenido por el coeficiente de correlación del punto biserial (rpbis) y la mayor dificultad en las preguntas, significa que las preguntas que discriminaron mejor fueron las más difíciles según su índice de dificultad.

En el temario I dicha relación se presenta además, para la pregunta uno con un valor de coeficiente de 0.93. El grado de dificultad y coeficiente de discriminación coinciden.

Backhoff¹⁴ considera que el coeficiente de discriminación es más sensible para detectar el comportamiento de los reactivos de un examen que el índice de discriminación. La ventaja de utilizar el coeficiente de discriminación en lugar del índice de discriminación (D) es que con el primer método se toman en cuenta todas y cada una de las calificaciones de las personas evaluadas, mientras que con el segundo, sólo se toma 54% de ellas (27% más alto y 27% más bajo).

Las autoras consideran necesario el papel que debe jugar el trabajo metodológico en el departamento de Psicología para el rediseño con calidad de los instrumentos evaluativos.

En el cuadro 5 se observan los valores del coeficiente alfa de Cronbach para cada una de las preguntas y de los temarios.

El coeficiente de confiabilidad alfa de Cronbach permite determinar la consistencia interna de un instrumento y brinda información relacionada con el aporte que hace cada pregunta a la consistencia general del examen.

Los valores totales de cada instrumento fueron de 0.58 y 0.71 para el temario I y II respectivamente. Según la Comisión Nacional de Evaluación de la Competencia y el Desempeño del Ministerio de Salud Pública proponen para este indicador valores superiores a 0.60. Se concluye que el temario II presenta mejor resultado que el temario I y se sugiere la revisión de este temario para su posterior aplicación en otro momento, pues no resulta confiable según los estándares nacionales e internacionales establecidos para este tipo de examen.

VII. Determinación del coeficiente alfa de Cronbach.

Cuadro 5. Coeficiente Alfa de Cronbach por preguntas y temarios.

Preguntas	Cronbach temario I	Cronbach temario II
Pregunta 1	0.55	0.68
Pregunta 2	0.53	0.68
Pregunta 3	0.52	0.65
Pregunta 4	0.50	0.69
Pregunta 5	0.51	0.70
Pregunta 6	0.60	0.64
Pregunta 7	0.57	0.69
Examen	0.58	0.71
n = 328		n = 330

En los resultados del coeficiente alfa de Cronbach por preguntas se muestra que en el temario I al eliminar la pregunta seis, el coeficiente aumenta ligeramente respecto al total obtenido. Este resultado evidencia que dicha pregunta no aporta lo suficiente a la coherencia interna del instrumento y debe ser revisada y perfeccionada para su incorporación a próximos instrumentos escritos.

En el temario II la totalidad de las preguntas muestran valores inferiores al valor total del coeficiente alfa de Cronbach del instrumento, lo que permite concluir que todas las preguntas contribuyen de igual manera a la confiabilidad del instrumento. El temario II presenta mayor confiabilidad.

CONCLUSIONES

- ✓ Los temarios analizados clasificaron como mixtos según el formato de las preguntas, cuentan con siete preguntas con predominio de la tipología test objetivo e incluyeron todas las temáticas de la asignatura.
- ✓ En ambos temarios existió correspondencia moderada entre cada temática y su exploración en el examen de acuerdo a la cantidad de ítems.
- ✓ Los dos temarios clasificaron de dificultad media, aunque los incisos fáciles estuvieron por encima de lo esperado.

- ✓ En ambos temarios las preguntas presentan un excelente poder predictivo. El temario I alcanzó mejores resultados que el temario II en cuanto al índice de discriminación. En el temario II predomina la clasificación de pobre discriminación.
- ✓ Los valores del coeficiente alfa de Cronbach muestran resultados aceptables. El temario II presenta mayor confiabilidad.
- ✓ El presente estudio permitió la caracterización del examen aplicado en la asignatura Psicología Médica I, a partir del cual se podrán trazar estrategia de perfeccionamiento del proceso de planeamiento y confección de los instrumentos evaluativos de la asignatura.

RECOMENDACIONES

- ✓ Desarrollar taller metodológico sobre Calidad de los instrumentos evaluativos, de forma sistemática en los colectivos de disciplina del departamento de Psicología con el objetivo de lograr una correcta planeacion y confeccion de instrumentos evaluativos.
- ✓ Divulgar los resultados obtenidos en publicaciones científicas.

REFERENCIAS BIBLIOGRÁFICAS.

1. Conde Fernández B. La evaluación de la calidad del proceso de desarrollo de las habilidades clínicas en los estudiantes de tercer año de la Carrera de Medicina. (Tesis) Sancti Spiritus: Universidad de Ciencias Pedagógicas "Capitán Silverio Blanco Núñez"; 2011.
2. Toranzos L. Evaluación y Calidad. Revista Iberoamericana de educación. 1996. [citado 02 Dic 2017];10: 63-78. Disponible en: <http://rieoei.org/oeivirt/rie10.htm>
3. Velásquez VM. La evaluación como recurso para elevar la calidad de la educación en México. Revista Iberoamericana de educación. 2013 [citado 09 Dic. 2017]; Disponible en: <http://rieoei.org/oeivirt/rie10a08.htm>
4. Torres-Noriega J. Los exámenes nacionales de medicina (ENAM) en el Perú. Rev Peru Med. Exp. Salud Publica. 2008 [citado 13 Ene 2017]; 25(3): 316-18. Disponible en: <http://www.redalyc.org/articulo.oa?id=36311611011>
5. López Cruz E; Legañoa Alonso J; Peraza Gutierrez L; Peraza Sanfeliz, E. Calidad de la evaluación estatal en estomatología. Rev Cubana Estomatol. 2014 [citado 18 Nov 2016]; 51(4): 435-443. Disponible en: <http://scielo.sld.cu/pdf/est/v51n4/est08414.pdf>
6. Romeu Escobar MR, Díaz Quiñones JA. Valoración metodológica sobre la confección de exámenes finales de la carrera de Medicina en Cienfuegos. EDUMECENTRO. 2015 [citado 18 Nov 2016]; 7(4):42-55. Disponible en: <http://scielo.sld.cu/pdf/edu/v7n4/edu04415.pdf>
7. Roméu Escobar MR, Díaz Quiñones JA. Valoración metodológica de la confección de temarios de exámenes finales de Medicina y Estomatología. Educ Med Super. 2015 Sep [citado 10

Ene 2017]; 29(3): 522- 531 Disponible en <http://scielo.sld.cu/pdf/ems/v29n3/ems11315.pdf>

8. López Espinosa GJ, Quintana Mugica R, Rodríguez Cruz O, Gómez López L, Pérez de Armas A, Aparicio Manresa G. El profesor principal y su preparación para diseñar instrumentos de evaluación escritos. Rev EDUMECENTRO. 2014 Ago [citado 11 Ene 2017]; 6(2): 94-109. Disponible en: <http://scielo.sld.cu/pdf/edu/v6n2/edu07214.pdf>
9. Martín Alfonso L. Aplicaciones de la psicología en el proceso salud enfermedad. Rev Cubana Salud Pública. 2003 Sep [citado 10 Nov 2016]; 29(3): 275-281. Disponible en: <http://scielo.sld.cu/pdf/rcsp/v29n3/spu12303.pdf>
10. Díaz Rojas P, Leyva Sánchez E. Metodología para determinar la calidad de los instrumentos de evaluación. Educ Med Super.2013 [citado 29 Nov 2016]; 27(2): 269-286. Disponible en: <http://scielo.sld.cu/pdf/ems/v27n2/ems14213.pdf>
11. Pérez Paz FM. Elaboración de la pregunta sobre Medicina Natural y Tradicional en exámenes estatales: reclamo de calidad. EDUMECENTRO 2016 [citado 30 Ene 2017]; 8(Suppl1): 18-30. Disponible en: <http://scielo.sld.cu/pdf/edu/v8s1/edu03116.pdf>
12. López Espinosa GJ, Quintana Mugica R, Rodríguez Cruz O, Gómez López L, Pérez de Armas A, Aparicio Manresa G. El profesor principal y su preparación para diseñar instrumentos de evaluación escritos. Rev EDUMECENTRO. 2014 Ago [citado 11 Ene 2017]; 6(2): 94-109. Disponible en: <http://scielo.sld.cu/pdf/edu/v6n2/edu07214.pdf>
13. Quiñones Gutiérrez D. Calidad de los exámenes estatales del curso 2011-2012 y 2012-2013 en la carrera de Licenciatura en Enfermería. (Tesis). Holguín: Facultad de enfermería "Arides Estévez Sánchez"; 2014.
14. Backhoff E., Larrazolo N., Rosas M. Nivel de dificultad y poder de discriminación del Examen de Habilidades y Conocimientos Básicos (EXHCOBA). Revista Electrónica de Investigación Educativa. 2000. [citado 18 Nov 2016]; 2 (1):12-28 Disponible en: <http://www.redalyc.org/articulo.oa?id=15502102>
15. Pupo Mora Y. Calidad del examen estatal teórico de la carrera de Psicología mención Salud. (Tesis) Holguín: Universidad de Ciencias Médicas; 2014.
16. Ortiz Romero G. Calidad del examen final de la asignatura Metodología de la Investigación y Estadística Carrera de medicina. (Tesis). Holguín: Universidad de Ciencias Médicas; 2012.
17. Ferrer Viant DV, González Sabín C, García Rodríguez I, Delfín Soto O, Solar Carballo O. Dificultad y discriminación del examen final de sistemas reguladores del medio interno. (Congreso Internacional 2015). 2-6 Noviembre. 2015; La Habana.
18. Rivera González M. Calidad del instrumento evaluativo en el examen escrito de Medicina Interna. (Tesis). Holguín: Universidad de Ciencias Médicas; 2014.

