

## **Comparación de la calidad de un instrumento evaluativo en la asignatura de Ortodoncia de la carrera de Estomatología**

Quality comparison of an assessment tool in Orthodontics subject of the Dentistry career

MSc. Yolanda Ayala Pérez <sup>1\*</sup> <https://orcid.org/0000-0003-0975-2981>

MSc. Kelia Batista Marrero <sup>2</sup> <https://orcid.org/0000-0001-6335-6774>

MSc. Maricelys Leonor Martínez Pérez <sup>1</sup> <https://orcid.org/0000-0002-9007-9271>

Esp. María del Carmen Torres Silva <sup>3</sup> <https://orcid.org/0000-0002-8840-5074>

Esp. José Antonio Estepa Taño <sup>1</sup> <https://orcid.org/0000-0002-9866-4651>

<sup>1</sup>Clínica Estomatológica Docente Mario Pozo. Holguín, Cuba.

<sup>2</sup>Universidad de Ciencias Médicas de Holguín, Cuba.

<sup>3</sup>Facultad de Ciencias Médicas Mariana Grajales. Universidad de Ciencias Médicas de Holguín, Cuba.

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

### **RESUMEN**

**Introducción:** la calidad de la educación superior se garantiza a través de la evaluación de los procesos docentes.

**Objetivo:** comparar la calidad de un instrumento evaluativo en dos cursos.

**Métodos:** investigación educacional en el campo de evaluación en la asignatura de Ortodoncia de 4to. año de la carrera de Estomatología, desde septiembre 2015 a junio 2017. Universo: 52 exámenes. Se aplicó la metodología para la calidad de instrumentos evaluativos y se compararon con los del curso anterior, en la Universidad de Ciencias Médicas de Holguín, Cuba.

**Resultados:** en 2015-16 se exploraron la mayoría de los objetivos instructivos, en el del 2016-17, todos. En el examen del 2015-16 el tema de más ítems fue el 2, con 23,3%, mientras que en el del 2016-17, el 3 (20%). Los ítems y el tiempo dedicado a cada temática mostraron escasa correlación en el 2015-16 con Spearman 0,39; en 2016-17 fue excelente, con Spearman 0,85. No existió correlación entre los ítems real y los esperados (Spearman 0,13) en el 2015-16 y sí en el curso 2016-17 con Spearman 0,95. Casi todas las preguntas tuvieron un excelente poder de discriminación en el 2015-16, en el 2016-17, todas. La confiabilidad del examen no fue aceptable en el 2015-16, con alfa de Cronbach 0,48 – 0,59; en el 2016-17 los valores de 0,60- 0,65 reflejaron buena confiabilidad.

**Conclusiones:** no se correspondió el fondo de tiempo con los ítems del 2015-16, mientras que el 2016-17, sí. En el 2015-16 no existió correlación entre el grado de dificultad de los ítems esperado y el real y en 2016-17 sí existió correlación. La mayoría de las temáticas estaban dentro de los rangos fáciles y medianamente fáciles en el 2015-16, en el 2016-17 predominó la dificultad media. Las preguntas tuvieron excelente poder de discriminación en ambos cursos. La confiabilidad fue aceptable en el 2016-17, pero no en el 2015-16.

**Palabras clave:** evaluación, educación superior, examen teórico, ortodoncia, estomatología.

## ABSTRACT

**Introduction:** the quality of higher education is guaranteed through the evaluation of teaching processes.

**Objective:** to compare the quality of an assessment tool in two school years.

**Methods:** an educational research in the evaluation field of Orthodontics subject, in the 4th year of the Dentistry career, from September 2015 to June 2017. Universe: 52 exams. The methodology for the quality of assessment tools was applied and they were compared with those of the previous year, at the University of Medical Sciences in Holguín, Cuba.

**Results:** most of the educational goals were explored in 2015-16; in 2016-17, all of them. In the 2015-16 exams the topic with more items was #2, (23.3%), while in those of 2016-17 was #3, (20%). Items and time spent on each theme showed little correlation in 2015-16, with Spearman 0.39, in 2016-17 it was excellent, with Spearman 0.85.

There was no correlation between actual and expected items (Spearman 0.13) in 2015-16; in 2016-17 there was correlation with Spearman 0.95. Almost all the questions had an excellent power of discrimination in 2015-16; in 2016-17, all of them. The reliability of the exam was not acceptable in 2015-16, with Cronbach's alpha 0.48 - 0.59; in 2016-17 the values 0.60- 0.65, showed good reliability.

**Conclusions:** the time fund did not correspond to the items of 2015-16, while in 2016-17, it did. In 2015-16 there was no correlation between the expected and actual degree of difficulty of the items, but in 2016-17 there was a correlation. Most of the themes were within easy and medium easy ranges in 2015-16; in 2016-17 the medium difficulty predominated. The questions had excellent discrimination power in both courses. Reliability was acceptable in 2016-17, but not in 2015-16.

**Keywords:** evaluation, higher education, theoretical examination, orthodontics, stomatology.

Recibido: 18/09/2019.

Aprobado: 01/12/2019.

## Introducción

La Cumbre Mundial de Educación Médica, celebrada en Edimburgo en 1993, desarrolló un movimiento internacional para un cambio entre la educación médica, la práctica médica y la organización de salud que posibilite la formación de un profesional que responda a las necesidades económico-sociales de sus países, pero a la vez sea capaz de enfrentar los retos científicos y tecnológicos que exige la actual centuria, sobre la base de una cooperación, tanto en los ámbitos nacionales como internacionales.<sup>(1)</sup>

Para lograr una adecuada preparación del estudiantado, el claustro de profesores deberá ser capaz de aplicar correctamente los instrumentos evaluativos.

La evaluación es un componente esencial de cualquier esfuerzo que se haga para producir algo de valor, así como para poder adoptar una buena decisión.<sup>(2)</sup>

Los procesos evaluativos se vienen desarrollando, cada vez más, en correspondencia con las necesidades sociales del mundo contemporáneo. La calidad en la educación superior está basada en una noción de cambio cualitativo, de transformación constante, que se utiliza como un término de referencia comparativo, dentro de un conjunto de elementos homologables, a partir de indicadores preestablecidos.

Para alcanzar este marco referencial necesario y enfrentar un cambio cualitativamente superior, estamos obligados a analizar de forma integral el desarrollo cualitativo de la educación médica y la investigación, junto con la práctica médica y la organización de salud, como un todo interdependiente.

Evaluar es emitir un juicio de valor que expresa la magnitud y calidad con que se han logrado los objetivos propuestos. Toda evaluación en salud, para que sea eficaz, tiene que estar comprometida con la competencia y el desempeño de los educandos, la eficiencia y efectividad del claustro, la calidad del curriculum y la calidad de los servicios donde se desarrolla el proceso educacional.

El principal factor determinante del desempeño académico es, quizá, la capacidad de evaluar y supervisar la calidad de los resultados – o impacto – de la enseñanza y la investigación en los servicios de salud. Los mecanismos de la autoevaluación promueven un sentido de verdadera responsabilidad institucional, por los resultados de los procesos educacionales en pregrado y posgrado.<sup>(3,4)</sup>

Los propósitos de la evaluación incluyen guiar el aprendizaje, la mejora de la enseñanza, el explicar por qué algunas cosas no se han aprendido, originar información válida que permita tomar decisiones racionales sobre cursos, carreras, entre otros, la predicción de resultados futuros y los sistemas para la certificación.<sup>(5)</sup>

La evaluación está determinada por los objetivos, pero su efectividad depende en gran medida del número, la frecuencia y la calidad de los controles que se apliquen, así como de la correcta y uniforme calificación que se realice de los resultados. Su carácter de continuidad permite la constante comprobación de los resultados y poder valorar su tendencia, que posibilita conformar un criterio dinámico sobre el aprovechamiento del estudiante. Es imprescindible realizar una adecuada evaluación, caracterizada por su objetividad y sus funciones educativas e instructivas.<sup>(6)</sup>

En la educación médica superior cubana, los objetivos educacionales constituyen la categoría rectora del proceso docente-educativo, pero la evaluación influye en lo que enseñamos y determina de manera directa lo que los educandos aprenden. Es esencial que las evaluaciones sean confiables. Estas exigen pruebas de validez de manera significativa, por lo que se requiere que los instrumentos y criterios utilizados en la medición de esta variable compleja tengan el mayor grado de validez.<sup>(7,8,9)</sup>

En los análisis que se efectúan al nivel de los colectivos de las asignaturas y disciplinas, generalmente no se incluyen qué preguntas deben responder los estudiantes; qué problemas deben resolver, cuántos y cómo; qué criterios e índices nos pueden mostrar y en qué grado; ni tampoco el logro de los objetivos propuestos para cada actividad docente, tema impartido y para la asignatura en general.

En la práctica todo esto se traduce en preguntas improvisadas que pueden estar acordes o no con los objetivos; ejercicios y problemas que no se corresponden con los niveles de asimilación ni se relacionan con el tipo de control empleado; en la aceptación de respuestas incompletas y en diferencias de criterios entre los profesores al calificar los instrumentos evaluativos, todo lo cual crea desconfianza entre los estudiantes.

Por tanto, para la construcción de un test de rendimiento, que constituya un verdadero instrumento de medición, se debe seguir el proceso de evaluación, el cual comienza con la definición de los objetivos educativos, luego utilizar los instrumentos de medida, interpretar los datos, formular juicios y adoptar las decisiones necesarias.

Toda esta situación demuestra que tanto el personal docente, como los estudiantes, no valoran ni comprenden en toda su magnitud la importancia, el papel y las funciones de la evaluación del aprendizaje dentro del proceso docente educativo en general y de enseñanza-aprendizaje en particular, lo que nos motivó a plantearnos: ¿Tendrán nuestros instrumentos evaluativos la calidad necesaria? Por ello se propuso como objetivo comparar la calidad de un instrumento de evaluación en dos cursos consecutivos en la asignatura de Ortodoncia de la carrera de Estomatología.

## **Métodos**

Se realizó una investigación educacional en el área del pregrado en el campo de la evaluación, en el período comprendido de septiembre 2015 a junio 2017, en la asignatura de Ortodoncia de 4to. año de la carrera de Estomatología de la Universidad de Ciencias Médicas de Holguín, Cuba.

La Ortodoncia está dirigida a brindar los fundamentos científicos a los estudiantes que les permitan realizar deducciones lógicas con acciones de salud referente a la prevención e intercepción de anomalías dentomaxilofaciales al menor de 19 años y dentro del grupo básico de salud del cual forma parte.

El universo es la totalidad de exámenes realizados en el período seleccionado: 52, a los que se aplicó la metodología para determinar la calidad de los instrumentos de evaluación diseñada por Díaz Rojas y Leyva Sánchez.<sup>(10)</sup> Posteriormente los resultados (curso 2016-17) se compararon con los obtenidos del curso anterior (2015-16).

Metodología para determinar la calidad de un instrumento de evaluación:

- I. Análisis del examen aplicado.
- II. Elaboración de la hoja de cálculo con los datos primarios de los resultados de cada estudiante en su examen.

III. Correspondencia del fondo de tiempo asignado en el P1 a los contenidos y la cantidad de ítems que lo exploran.

IV. Definición de la escala de clasificación de la dificultad, según expertos.

V. Obtención de indicadores de validez del instrumento.

- Índice de dificultad.

- Índice de discriminación de las preguntas.

VI. Análisis de la confiabilidad de las preguntas y el examen.

Se solicitaron los exámenes al Departamento Metodológico de la Facultad de Ciencias Médicas Mariana Grajales Coello, de la Universidad de Ciencias Médicas de Holguín, Cuba.

Se identificaron las características del examen, teniendo en cuenta:

- Temas de la asignatura obtenidos del programa analítico y P1.

- Tiempo en que se imparte la asignatura, obtenido P1.

- Estructura del examen, según cantidad de preguntas e ítems.

-Tipos de preguntas, contenidos temáticos y temas incluidos en el examen.

Se clasificaron las preguntas según su tipo. Se determinó el número de ítems evaluados por temas, de acuerdo con el tipo de pregunta. Se clasificaron los ítems y preguntas, según los contenidos temáticos y temas que se exploraron en el examen.

Correspondencia entre el fondo de tiempo de los contenidos temáticos y su exploración en el examen. Se determinó el número de horas dedicado a cada tema, según el plan analítico, P1 de la asignatura, y se definió el porcentaje que representó del total de tiempo dedicado a la asignatura.

Se determinó el índice de dificultad de cada uno de los ítems, dividiendo el número de estudiantes que contestó correctamente el ítem entre el número total que contestaron el ítem.

Esta proporción se le denota como  $D_i$  y se utilizó la siguiente fórmula:

$$D_i = A / N$$

Donde:

$D_i$  = Índice de dificultad del ítem.

A = Número de respuestas correctas del ítem.

N = Número total de estudiantes que contestaron el ítem.

- Índice de dificultad de las temáticas

En este caso se sumó el total de estudiantes que respondieron correctamente cada uno de los incisos de la temática y se dividió entre el número de ítems que exploraban el tema. Se calculó la media de respuestas correctas para el tema y se dividió entre el total de examinados, como se representa en la siguiente fórmula:

$$Dt = \frac{\sum (A_{i1} + A_{i2} + \dots + A_{in})}{\frac{n_{it}}{Nt}}$$

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.

Comparación de los resultados, según índice de dificultad obtenido (real) con la dificultad esperada.

- 5 % se consideran difíciles.
- 20 % se consideran medianamente difíciles.
- 50 % se consideran de dificultad media.
- 20 % se consideran medianamente fáciles.
- 5 % se consideran fáciles.

A partir de la obtención de los índices de dificultad, clasificamos cada ítem y temática, según la escala de clasificación definida en el colectivo.

Determinación de los intervalos para la clasificación de los ítems, según su índice de dificultad:

- 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.

Determinación del índice de dificultad del examen:

Para calcular el índice de dificultad del examen, se sumó el total de estudiantes que respondieron correctamente cada uno de los incisos del examen y se dividió entre el número de incisos del examen, se obtuvo la media de respuestas correctas para el examen; luego se calculó el índice de dificultad, el cual se corresponde con el promedio de respuestas correctas dividido por el total de examinados.

Análisis de la discriminación de las preguntas:

La discriminación expresa la medida en que la pregunta discrimina entre los buenos estudiantes y los peores. Este indicador permite predecir el comportamiento de la pregunta dentro del examen, para la elaboración de nuevos instrumentos con estructura similar.

Coefficiente de discriminación del punto biserial ( $r_{pbis}$ ):

Este indicador calcula la efectividad discriminativa de una pregunta. Esta correlación se emplea para saber si los estudiantes aplicados obtienen las respuestas correctas, qué tanto poder predictivo tiene el reactivo y cómo puede contribuir a las predicciones.

Se utilizó la siguiente fórmula:

$$r_{pbis} = \frac{\overline{X_1} - \overline{X_0}}{S_x} * \sqrt{\frac{n_1 n_0}{n(n-1)}}$$

Donde:

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

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

Sx = Desviación estándar de las puntuaciones totales.

n1= 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 rpbis es:

- $< 0$  discriminan negativamente.
- 0 a 0,14 discriminan pobremente.
- 0,15 a 0,25 discriminan regular.
- 0,26 a 0,35 buen poder discriminativo.
- $\geq 0,35$  excelente poder de discriminación.

El rpbis puede ser calculado para cada ítem y para cada pregunta.

Análisis de la confiabilidad de las preguntas y el examen:

Utilizamos el coeficiente alfa de Cronbach, que 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.

La fórmula que se emplea es la siguiente:

$$\alpha = \frac{nr}{1 + r(n - 1)}$$

Donde  $n$  es el número de elementos del examen y  $r$  su correlación media.

Tabulamos los resultados atendiendo al alfa del examen y de cada pregunta, que calculamos excluyendo los datos de esa pregunta.

Su interpretación se realizó tomando 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.

Los resultados se compararon con los obtenidos en el curso anterior.

## Resultados

Determinamos la calidad de un instrumento evaluativo en la carrera de Estomatología en la asignatura Ortodoncia, del 4to. año en el curso 2016-2017, y lo comparamos con igual análisis del examen del curso anterior.

### Análisis del examen aplicado

- Tipo de examen. El instrumento de evaluación analizado fue el examen teórico final de Ortodoncia, en el cual se exploraron todos los objetivos instructivos de la asignatura.
- Cantidad y tipos de preguntas del examen. El examen consta de 7 preguntas: de ellas 2 de test de ensayo o desarrollo, 1 de test de respuestas cortas, 1 de ensayo modificado, 1 de test objetivo de asociación simple y 1 de test objetivo de selección múltiple y 1 de asociación simple.
- Cantidad de ítems por preguntas y área temática.

La (tabla I) de comparación de exámenes de Ortodoncia, según ítems por área temática, muestra que los temas 4 y 8 no estuvieron presentes en el examen del curso 2015-16, mientras que en el examen del curso 2016-17 se incluyeron todos los temas del programa. El tema de más ítems en el examen del curso 2015-16 fue el tema 2. Crecimiento y desarrollo cráneo facial, con 7 ítems, que representan el 23,3 % del total de ítems en el examen, por la importancia que el mismo tiene dentro del sistema de conocimientos y de habilidades a desarrollar por el estudiante en su vida futura como profesional de la salud; sin embargo, en el examen del curso 2016-17 predominó el tema 3 Desarrollo de los dientes y la oclusión, con el 20 %. Estos resultados evidencian la evolución favorable en el planeamiento de los instrumentos evaluativos.

**Tabla I.** Comparación de exámenes de Ortodoncia, según ítems por área temática

Contenidos	Ítems confeccionados			
	Curso 2015-16		Curso 2016-17	
	n	%	n	%
1. Introducción a la Ortodoncia	1	3,3	3	6
2. Crecimiento y Desarrollo cráneo facial	7	23,3	7	14
3. Desarrollo de los dientes y la oclusión. Fisiología dentomaxilofacial	5	16,6	10	20
4. Diagnóstico de las anomalías dentomaxilofaciales	0	0	7	14
5. Biomecánica	6	20	6	12
6. Discrepancia hueso diente	5	16,6	4	8
7. Anomalías dentomaxilofaciales en el nivel primario de atención estomatológica	6	20	8	16
8. Aparatos ortodóncicos empleados en el nivel primario de Atención Estomatológica. Urgencias en Ortodoncia	0	0	5	10
Total	30	100	50	100

En la (tabla II) se comparan los exámenes de Ortodoncia, según ítems por tiempo en P1, se observa una escasa correlación entre los ítems del examen del curso 2015-16 con respecto al tiempo dedicado a cada temática dentro del programa de la asignatura, con un coeficiente de correlación de Spearman de 0,39.

El tema que más ítems tiene dentro del examen es el tema 2 (crecimiento y desarrollo craneofacial), con el 23,3%, lo que no coincide con el tema de más horas en P1. Al analizar el examen del curso 2016-17, existe una estrecha correlación con un índice de correlación de Spearman de 0,85; pues el tema que más horas tiene en P1 es el tema 3 (desarrollo de los dientes y la oclusión) y es también el de más cantidad de ítems dentro del examen, con 20%.

**Tabla II.** Comparación de exámenes de Ortodoncia, según ítems por tiempo en P1 del área temática

Contenidos	Ítems confeccionados					
	Tiempo en P1 (horas)		2015-16		2016-17	
	n	%	n	%	n	%
1. Introducción a la Ortodoncia	5	5,5	1	3,3	3	6

2. Crecimiento y Desarrollo cráneo facial	13	14,4	7	23,3	7	14
3. Desarrollo de los dientes y la oclusión. Fisiología dentomaxilofacial	17	18,8	5	16,6	10	20
4. Diagnóstico de las anomalías dentomaxilofaciales	10	11,1	0	0	7	14
5. Biomecánica	10	11,1	6	20	6	12
6. Discrepancia hueso diente	10	11,1	5	16	4	8
7. Anomalías dento maxilo faciales en el nivel primario de atención estomatológica	14	15,5	6	20	8	16
8. Aparatos ortodóncicos empleados en el nivel primario de Atención Estomatológica. Urgencias en Ortodoncia	11	12,2	0	0	5	10
Total	90	100	30	100	50	100

• **Obtención de indicadores de validez del instrumento**

- Índice de dificultad:

Al analizar el comportamiento de los Ítems de los exámenes, según grado de dificultad esperado y real, se puede apreciar una gran diferencia entre el grado de dificultad de los ítems esperado y el real, sin correlación entre ambas variables, con un coeficiente de correlación de Spearman de 0,13 en el curso 2015-16; sin embargo, en el examen del curso 2016-17 el coeficiente de correlación de Spearman fue de 0,95, lo que demuestra una excelente correlación (tabla III).

**Tabla III.** Comportamiento de los ítems de los exámenes, según grado de dificultad esperado y real

Ítems según su grado de dificultad	% Ítems esperados	% ítems reales 2015-16	% ítems reales 2016-17
Fáciles	5	33,3	10
Medianamente fáciles	20	26,6	22
De dificultad media	50	33,3	45,8
Medianamente difíciles	20	0	18
Difíciles	5	6,6	4,2
Totales	100	100	100

En la (tabla IV), comparación del coeficiente discriminación, según preguntas de los exámenes, observamos que en el examen del curso 2015-16 las preguntas 1, 3, 4 y 5 tienen un excelente poder de discriminación, pues sus valores son mayores que 0,35.

La pregunta 2 tiene buen poder discriminativo, ya que su coeficiente de discriminación está en el rango entre 0,26 a 0,35. Al realizar el análisis al examen del curso 2016-17, encontramos que todas las preguntas tienen un excelente poder de discriminación.

**Tabla IV.** Comparación del coeficiente discriminación según preguntas de los exámenes

Preguntas	Coeficiente discriminación	
	2015-16	2016-17
1	0,53	0,48
2	0,34	0,36
3	0,52	0,50
4	0,46	0,39
5	0,38	0,41
6	-	0,37
7	-	0,38

#### **Análisis de la confiabilidad de las preguntas y el examen**

La confiabilidad del examen y por preguntas del curso 2015-16 no es aceptable, pues los valores del alfa de Cronbach están entre 0,48 – 0,59, lo que demuestra poca consistencia interna en su estructura. No obstante estos resultados, los valores del alfa de Cronbach en las preguntas son menores o iguales al alfa de Cronbach del examen, lo que evidencia que le aportan consistencia interna. Cuando se analiza el examen del curso 2016-17, se le observa una adecuada consistencia interna, por tanto confiabilidad, pues los valores del alfa de Cronbach están entre 0,60 y 0,65, y las preguntas aportan consistencia interna al examen, por ser sus valores iguales o menores que el del examen (tabla V).

**Tabla V.** Comparación del coeficiente alfa de Cronbach de las preguntas y los exámenes

Preguntas	Coeficiente alfa de Cronbach	
	2015-16	2016-17
1	0,50	0,61
2	0,59	0,60
3	0,48	0,62
4	0,54	0,64
5	0,57	0,62

6	-	0,63
7	-	0,61
Examen	0,59	0,65

## Discusión

Al realizar el análisis de la calidad de los instrumentos evaluativos, podemos establecer las pautas para su perfeccionamiento. Dentro de las funciones de la evaluación se encuentra el establecer planes de mejora.<sup>(11)</sup> Estos planes incluyen todos los procesos docentes que se llevan a cabo en las universidades.

La asignatura es el proceso docente que se desarrolla en un cierto período de tiempo para garantizar la adquisición de determinados conocimientos y formación de habilidades y se encuentra estructurada mediante temas; este subsistema debe integrarse con los otros subsistemas de la disciplina en el sentido vertical del currículo.<sup>(12)</sup>

Cuando se realiza un correcto trabajo metodológico y una preparación adecuada del claustro, podemos obtener instrumentos evaluativos que constituyan verdaderos instrumentos para medir la calidad del proceso docente educativo.

El concepto de calidad se utiliza para definir un conjunto de cualidades del objeto de estudio.<sup>(13)</sup>

El trabajo de perfeccionamiento profesoral mediante cursos, diplomados y maestrías se pone de manifiesto en instrumentos evaluativos cada vez más cercanos a los parámetros establecidos.

El proceso de mejoramiento de la calidad de los sistemas educativos requiere la evaluación como un componente estratégico que atraviese todos sus ámbitos, brindando información útil para la toma de decisiones.<sup>(14)</sup>

Cuando se analiza el coeficiente de discriminación en un examen se entiende que está relacionado con el índice de dificultad, pues si las preguntas no se encuentran en los límites de dificultad, estas son capaces de discriminar entre estudiantes buenos y de malos rendimientos.

La confiabilidad es la expresión cuantitativa de la reproducibilidad con la que un instrumento mide un mismo atributo. A través de ella se logra la estabilidad de los resultados. Su utilización le brinda consistencia interna al instrumento evaluativo.

El Coeficiente alfa de Cronbach es el más utilizado mundialmente, pues proporciona una medida efectiva de la consistencia interna del examen y puede emplearse ante diferentes formatos de preguntas. Su valor se encuentra entre 0,60 –0,80. Es un concepto estrictamente estadístico que se expresa por un coeficiente de confiabilidad o a través del error típico de las mediciones efectuadas.

Al determinar la calidad de los instrumentos evaluativos, obtenemos informaciones sobre el rendimiento de los estudiantes que nos permitirán focalizar la atención sobre posibles problemas del aprendizaje e incidir en su seguimiento por medio de las comisiones verticales y horizontales. Este análisis arroja datos cuantitativos que permiten la identificación de ítems con deficiencias en su confección y por tanto exámenes con dificultades; de ahí se desprende la importancia para el desarrollo de las capacidades del docente para elaborar buenos ítems y exámenes.

## **Conclusiones**

En ambos exámenes teóricos de Ortodoncia existió variedad en el tipo de preguntas empleadas. No hubo correspondencia entre los objetivos instructivos de la asignatura, el fondo de tiempo y los contenidos, con el número de ítems explorados en el examen teórico final del curso 2015-16, pero sí existió correspondencia en el examen del curso 2016-17.

En el instrumento de evaluación analizado en el curso 2015-16 no existió correlación entre el grado de dificultad de los ítems esperado y el real; sin embargo, en el curso 2016-17 existió una correlación adecuada.



La mayoría de las temáticas se incluyen dentro de los rangos fáciles y medianamente fáciles en el curso 2015-16, mientras que en el examen del curso 2016-17 predominó el rango de dificultad media.

Las preguntas en su mayoría tienen un excelente poder de discriminación en ambos cursos y la confiabilidad del examen es aceptable en el curso 2016-17, pero no así en el 2015-16.

## Referencias Bibliográficas

1. Suárez Rosas L, Barrios Osuna I, González Espíndola ME. Metodología para la evaluación externa de la carrera de Estomatología. Educ Med Super. 2012[citado 15 oct 2018]; 26(2): 271-293. Disponible en: <http://www.ems.sld.cu/index.php/ems/article/view/31/27>
2. Borroto Cruz R, Salas Perea RS. El reto por la calidad y la pertinencia: la evaluación desde una visión cubana. Educ Med Super. 1999 [citado 15 oct 2018]; 13(1): 70-79. Disponible en: [http://scielo.sld.cu/scielo.php?script=sci\\_arttext&pid=S086421411999000100011&lng=es](http://scielo.sld.cu/scielo.php?script=sci_arttext&pid=S086421411999000100011&lng=es)
3. Salas Perea RS. La Evaluación en la Educación Superior Contemporánea. Biblioteca de Medicina. V 24. La Paz: Universidad Mayor de San Andrés; 1998. p. 30-35.
4. Ayala Pérez Y, Batista Marrero K, Martínez Pérez M, Díaz Morell JE. Calidad en la evaluación de la asignatura Ortodoncia en la carrera de Estomatología. CCM. 2018 [citado 18 sep 2019];23(1). Disponible en: <http://www.revcoemed.sld.cu/index.php/cocmed/article/view/2903>
5. Torres Guerra A, Ochoa Ardite Y, Iralda Oro JC, Machín Batista D, Robles García R. La evaluación del aprendizaje formativo contextualizado en la disciplina Filosofía y Sociedad. CCM. 2019 [citado 17 sep 2019];23(2). Disponible en: <http://revcoemed.sld.cu/index.php/cocmed/article/view/3131>

6. Quiñones Gutiérrez D, Díaz Rojas PA, Cabrera Diéguez L, Pérez Carralero L. Calidad de los exámenes estatales de los cursos 2011-2012 y 2012-2013 en la carrera de Licenciatura en Enfermería. CCM. 2017 [citado 17 sep 2019];21(2). Disponible en:

<http://www.revcoemed.sld.cu/index.php/cocmed/article/view/2261>

7. Gómez López VM, Rosales Gracia S, Vázquez Vázquez JA. Validez y consistencia de un instrumento para evaluar la práctica docente en clínica médica de pregrado. Rev Educ Desar. 2014 [citado 15 oct 2018];28: 15-20.Disponible en:

[www.cucs.udg.mx/revistas/edu\\_desarrollo/anteriores/28/RED\\_28\\_Completa\\_VF.pdf#page=16](http://www.cucs.udg.mx/revistas/edu_desarrollo/anteriores/28/RED_28_Completa_VF.pdf#page=16)

8. Sánchez Sanhuesa GA, Cisterna Cabrera F. La evaluación de los aprendizajes orientada al desarrollo de competencias en Odontología. Educ Med Super. 2014 [citado 15 oct 2018]; 28(1): 104-114. Disponible en: <http://www.ems.sld.cu/index.php/ems/article/view/72/147>

9. Valverde Berrocoso J, Ciudad Gómez A. El uso de e-rúbricas para la evaluación de competencias en estudiantes universitarios. Estudio sobre fiabilidad del instrumento. REDU.2014 [citado 15 oct 2018]; 12(1): 49-79. Disponible en:

<https://dialnet.unirioja.es/descarga/articulo/4691792.pdf>

10. Díaz Rojas PA, Leyva Sánchez E. Metodología para determinar la calidad de los instrumentos de evaluación. Educ Med Super. 2013 [citado 15 oct 2016]; 27(2): 269-286. Disponible en:

[http://scielo.sld.cu/scielo.php?script=sci\\_arttext&pid=S0864-21412013000200014&lng=es](http://scielo.sld.cu/scielo.php?script=sci_arttext&pid=S0864-21412013000200014&lng=es)

11. Glaría R, Carmona L, Pérez C, Parra P. Estructura factorial y consistencia interna de la escala de evaluación del currículum de programas universitarios en estudiantes de fonoaudiología de Chile. Rev Iberoam Diagn Eval Psicol. 2016 [citado 15 oct 2018]; 41 (1): 80-89. Disponible en:

<https://dialnet.unirioja.es/servlet/articulo?codigo=6474703>

12. Grau León IB, Barciela González Longoria MC, Rosales Reyes S, Peguero Morejón H. El trabajo docente metodológico en la Facultad de Estomatología de la Universidad de Ciencias Médicas de La Habana (UCMH). Educ Med Super. 2015 [citado 15 oct 2018]; 29(4): 763-776. Disponible en: <http://www.ems.sld.cu/index.php/ems/article/view/597/301>

13. Barciela González Longoria MC, Grau León IB, Fernández Pratt MJ, Rosales Reyes S. Calidad del claustro de profesores de la Facultad de Estomatología. Educ Med Super. 2015 [citado 15 oct 2018]; 29(1): 47-56. Disponible en: <http://www.ems.sld.cu/index.php/ems/article/view/367/227>

14. del Río Marichal S, Cuenca Arbella Y. Modelo pedagógico del proyecto de vida profesional para estudiantes de Medicina. CCM. 2019 [citado 17 sep 2019];23(2). Disponible en: <http://revcocmed.sld.cu/index.php/cocmed/article/view/2967>

### **Conflicto de intereses**

Los autores refieren no presentar conflicto de intereses.

### **Contribución de autoría**

YAP: diseñó el proyecto de investigación, recogida de datos y análisis y discusión de los resultados.

KBM: participó en el diseño del proyecto de investigación, recogida de datos y análisis y discusión de los resultados.

MLMP: recogida de datos y análisis y discusión de los resultados.

MCTS: recogida de datos y análisis y discusión de los resultados.

JAET: recogida de datos y análisis y discusión de los resultados.

Todos aprobaron el trabajo final.



Esta obra está bajo una [licencia de Creative Commons Reconocimiento-](https://creativecommons.org/licenses/by-nc/4.0/)

[No Comercial 4.0 Internacional.](https://creativecommons.org/licenses/by-nc/4.0/)