



## **Evolución de la calidad de un instrumento evaluativo en la Carrera de Estomatología.**

*Evolution of the quality of an instrument evaluative in the Dentistry Mayor.*

**Yolanda Ayala Pérez,<sup>1</sup> Kelia Batista Marrero,<sup>2</sup> Maricelys Leonor Martínez Pérez,<sup>3</sup> María del Carmen Torres Silva.<sup>4</sup>**

1 Dra. en Estomatología. Especialista de II grado en Ortodoncia. Máster en Atención de Urgencias en Estomatología. Profesor Auxiliar. Investigador agregado. Clínica Estomatológica Docente Mario Pozo. Holguín. Cuba.

2 Dra. en Estomatología. Especialista de II grado en EGI. Máster en Atención de Urgencias en Estomatología. Profesor Asistente. Investigador agregado. Facultad de Ciencias Médicas Mariana Grajales. Holguín. Cuba.

3 Dra. en Estomatología. Especialista de II grado en EGI. Máster en Atención de Urgencias en Estomatología. Profesor Auxiliar. Investigador agregado. Clínica Estomatológica Docente Mario Pozo. Holguín. Cuba.

4 Dra. en Estomatología. Especialista de II grado en EGI. Profesor Asistente. Facultad de Ciencias Médicas Mariana Grajales. Holguín. Cuba.

Correspondencia: [ayalahlg@infomed.sld.cu](mailto:ayalahlg@infomed.sld.cu)

### **RESUMEN**

La calidad de la educación superior se garantiza a través de la evaluación de los procesos docentes. Nuestro objetivo fue comparar la calidad de un instrumento evaluativo en dos cursos. Realizamos Investigación educacional en el campo de evaluación, marzo - septiembre 2017. Universo: 52 exámenes. Se aplicó la metodología para la calidad de instrumentos evaluativos y se compararon con los del curso anterior.

En el 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 7 (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 y en 2016-17 excelente con Spearman 0,85. No existió correlación entre los ítems real y los esperados, con Spearman 0,13 en el 2015-16 y si 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 y por preguntas no es aceptable en el 2015-16 con alfa de Cronbach 0,48 – 0,59, los valores en el 2016-17 de 0,60- 0,65 brindando buena confiabilidad.

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 si existió correlación. La mayoría de las temáticas dentro de los rangos fáciles y medianamente fáciles en el 2015-16, el 2016-17 predominó la dificultad media. Las preguntas tienen excelente poder de discriminación en ambos cursos. La confiabilidad es aceptable en el 2016-17, no en el 2015-16.

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

## **ABSTRACT**

The quality of higher education is guaranteed through the evaluation of the teaching processes. Objective: Comparing the quality of an instrument evaluative in two courses. Educational investigation at the field of evaluation, March - September 2017. Universe: 52 examens. The methodology for the quality of instruments applied over himself evaluative and they compared with give them the previous course.

They all explored the instructive objectives themselves the majority of, in the one belonging to the 2016-17 in the 2015-16. The theme of more items was the 2, with 7 in the 2015-16's exam (23.3 %) in the meantime than in the one belonging to the 2016-17 3 (20 %). Items and the time once each subject matter was dedicated, 0.39 showed scarce correlation in the 2015-16 with Spearman and in excellent 2016-17 with Sperman 0.85. Did not exist correlation between the real items and the hoped-for with Spearman 0.13 in the 2015-16 and if in the course 2016-17 with Spearman 0.95. Almost all the questions had an excellent power of discrimination in the 2015-16 in the 2016-17, everything. The reliability of the exam and for questions is not acceptable in the 2015-16 with alpha of Cronbach 0.48 0.65 offering good reliability – 0.59, the moral values in the 2016-17 of 0.60 -Time's bottom with the 2015-16's items was not reciprocated in the meantime than the 2016-17 yes. In the 2015-16 did not exist correlation between the degree of difficulty of the waited for items and the real and in 2016-17 if correlation existed . The majority of the thematic within the easy and fairly easy ranges in the 2015-16, the 2016-17 predominated the half a difficulty. The questions have excellent power of discrimination in both courses. Reliability is acceptable in the 2016-17, no in the 2015-16.

*Key word: evaluation, superior education, theoretical exam.*

## **INTRODUCCIÓN**

La Declaración Mundial sobre la Educación Superior de la UNESCO, en Paris 1998, plantea que “la calidad de la educación superior es un concepto pluridimensional que debe comprender todas las funciones de una institución de educación superior, prestando la

debida atención a las particularidades de los contextos regional, nacional e internacional. Es la adecuación del Ser y Quehacer de la Educación Superior a su Deber Ser".<sup>1</sup>

En la actualidad, en la inmensa mayoría de las universidades de prestigio en el mundo la evaluación y acreditación de carreras universitarias se encuentra impregnada de la concepción de la enseñanza superior como transformadora de la conciencia social.

Las acciones encaminadas a que el médico y el estomatólogo cubano adquirieran una proyección social en la práctica de su profesión, tiene en Cuba una larga tradición que rebasa el siglo y medio de existencia, pero aún es un problema la precisión de qué se entiende por un modelo educacional socio médico, lo cual es un tema a debate y ha transitado por la implementación de diversos planes de estudio.

La carrera de Estomatología no es ajena a estas realidades. El mejoramiento continuo de la calidad de los procesos y escenarios docentes que se encargan de la formación integral de los recursos humanos de Estomatología en las instituciones universitarias que asumen esta función es una necesidad que estas deben enfrentar para ser consecuentes con el compromiso adquirido con la sociedad que recibe en su seno a los egresados.<sup>2</sup>

La evaluación es un proceso integral que permite valorar los resultados obtenidos en términos de los objetivos propuestos, acorde con los recursos utilizados y las condiciones existentes.

Esto implica la obtención de informaciones que permitan la elaboración de juicios «válidos» acerca del alcance de determinado objetivo, de la eficiencia de un método, etc. Para el logro de esas informaciones la evaluación utiliza la medición, la cual garantiza datos más válidos y confiables en los cuales fundamentar los juicios.<sup>1, 3</sup>

Una de las funciones de la evaluación consiste en comprobar, de modo sistemático, en qué medida se han logrado los resultados previstos en los objetivos propuestos. Esto implica una congruencia entre las expectativas o conductas esperadas y sus logros. Las definiciones tradicionales de evaluación le han otorgado una interpretación temporal en el proceso enseñanza-aprendizaje que en la práctica se traduce a técnicas terminales que se limitan a una comprobación de productos o resultados.<sup>1, 2</sup>

Los criterios de evaluación tienen la función de establecer una definición explícita acerca de la forma de entender la calidad. Debido a la variedad de aproximaciones pueden ser operacionalizados en función de variables diversas. El papel central que desempeña la temática de la calidad y su evaluación, tanto a nivel nacional como regional, hace necesario profundizar el debate sobre las concepciones de calidad en la educación superior, definiendo con mayor precisión sus dimensiones, criterios e indicadores, sus enfoques metodológicos y las estrategias para su aseguramiento y mejoramiento permanente.<sup>1</sup>

Conceptualizar la evaluación, quiere decir estar ante una actividad que se desarrolla siguiendo pasos, que cumple múltiples funciones, que se apoya en una serie de ideas y formas de realizarla y que es la respuesta a unos determinados condicionamientos de la enseñanza institucionalizada.<sup>1</sup> *Ralph Tyler*, primer investigador de la evaluación educativa tradicional, en 1950 propuso una tesis conocida hoy como el modelo de "Evaluación por objetivos" donde plantea una constante comparación de los resultados del aprendizaje de los alumnos con los objetivos previamente determinados en la programación de la enseñanza. *Tyler* contemplaba que la evaluación se extendiera al proceso de aprendizaje y al currículo, no solo a los resultados, sin embargo en la práctica solo se utilizó para comparar los resultados con los objetivos, y de los resultados, solamente aquellos aprendizajes más fácilmente constatables.<sup>3,4</sup>

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. Para lograr estas propiedades debemos tener presente el análisis que se realice en el colectivo de profesores en el proceso de planeación y de elaboración de las preguntas e ítems, la forma de definir la clave de calificación y de determinar el éxito y el fracaso. La experiencia por parte del colectivo de profesores en la temática evaluación es esencial para el logro de un instrumento válido y confiable.<sup>5-9</sup>

Según el artículo 31 del Reglamento para el Trabajo Docente y Metodológico en los centros de Educación Superior, el trabajo metodológico que se realiza de forma colectiva, tendrá como rasgo esencial el enfoque en sistema y se llevará a cabo en cada uno de los niveles organizativos del proceso docente educativo, como vía para su perfeccionamiento en cada nivel.<sup>6</sup>

Estos niveles son los encargados de dirigir esta labor en los colectivos de carrera, de año, de disciplina y de asignatura, según corresponda.

En la Facultad de Ciencias Médicas de Holguín, en particular en la carrera de Estomatología, en este caso específico de la asignatura de Ortodoncia que se imparte en el primer semestre de cuarto año se realizó una investigación educacional para determinar la calidad de los instrumentos de evaluación, a partir de los resultados obtenidos en ese momento la dirección de la carrera puso énfasis en el desarrollo del trabajo metodológico a nivel de colectivos pedagógicos y de los distintos niveles de dirección, teniendo en cuenta las necesidades pedagógicas de los docentes y las dificultades identificadas en el proceso de formación de los estudiantes.

Uno de los aspectos que se priorizó tanto en su planificación, organización, dirección y control fueron los colectivos de años por considerarse que se concretan en el año los

objetivos previstos para ese nivel. Le corresponde, además, a este colectivo la integración horizontal del sistema de conocimientos y habilidades cuyos enfoques integrados se sustentan en la labor educativa, la cual tiene lugar a través del establecimiento de relaciones de carácter disciplinarios, multidisciplinarias, inter y transdisciplinarios. Teniendo en cuenta el análisis realizado en el curso anterior, luego de incrementar la preparación del claustro en la confección de los instrumentos evaluativos nos propusimos como objetivo comparar la calidad de un instrumento de evaluación en dos cursos consecutivos en la carrera de estomatología.

## **DISEÑO METODOLÓGICO**

Realizamos una investigación educacional en el área del pregrado en el campo de la evaluación, en el período comprendido de marzo a septiembre de 2017 con el objetivo de comparar la calidad de un instrumento de evaluación en dos cursos consecutivos en la carrera de estomatología en la asignatura Ortodoncia, que se imparte en el 4to año de la carrera. Esta asignatura 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 dentomáxilofaciales al menor de 19 años y dentro del grupo básico de salud del cual forma parte.

El universo estuvo constituido por la totalidad de exámenes que se realizaron en el período seleccionado: 52, a los que se aplicó la metodología para determinar la calidad de los instrumentos de evaluación diseñada por MSc. Pedro Augusto Díaz Rojas y MSc. Elizabeth Leyva Sánchez.<sup>10</sup> Posteriormente los resultados encontrados se compararon con los obtenidos del curso anterior.

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" de Holguín. Para el procesamiento de los datos utilizamos computadora con el programa Excel.

Como métodos teóricos se emplearon el análisis documental de la revisión bibliográfica de temas relacionados con el proceso enseñanza aprendizaje y evaluación, calidad de los instrumentos evaluativos, sus ventajas, desventajas de su empleo y otros materiales de interés para el desarrollo de la investigación.

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 al 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 horas dedicadas 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:

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

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.

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 que corresponden al 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 estudiantes. Este indicador nos permite además predecir el comportamiento de la pregunta dentro del examen, para la elaboración de nuevos instrumentos con estructura similar.

Coefficiente de discriminación del punto biserial (rpbis):

Este indicador permite calcular la efectividad discriminativa de una pregunta.

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 utilizó la siguiente fórmula:

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

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.

n1= Número de estudiantes que respondieron correctamente el ítem.

n0= Número de estudiantes que respondieron incorrectamente el ítem.

n = n1 + n0

Según estándares internacionales la distribución de los resultados del coeficiente discriminación punto de correlación biserial (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.
- > 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.

- Coeficiente alfa de Cronbach:

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 y se presentan en tablas para su mejor comprensión.

## **ANÁLISIS Y DISCUSIÓN DE LOS RESULTADOS**

Al realizar el análisis de la calidad de los instrumento 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>

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 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.

Cuadro 1. Comparación de exámenes de Ortodoncia según ítems por área temática.

Contenidos	Ítems confeccionados			
	Curso 2015-16		Curso 2016-17	
	Número	%	Número	%
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
<b>TOTAL</b>	<b>30</b>	<b>100</b>	<b>50</b>	<b>100</b>

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>

El Cuadro 1. 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 un 20 %. Estos resultados evidencian la evolución favorable en el planeamiento de los instrumentos evaluativos.

Cuadro 2. Comparación de exámenes de Ortodoncia según ítems por tiempo en P1 del área temática.

Contenidos	Ítems confeccionados					
	2015-16		2016-			
	17		17		17	
	Tiempo	%	No	%	No	%
	en P1					
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
<b>TOTAL</b>	<b>90</b>	<b>100</b>	<b>30</b>	<b>100</b>	<b>50</b>	<b>100</b>

### Scatter Plot Matrix

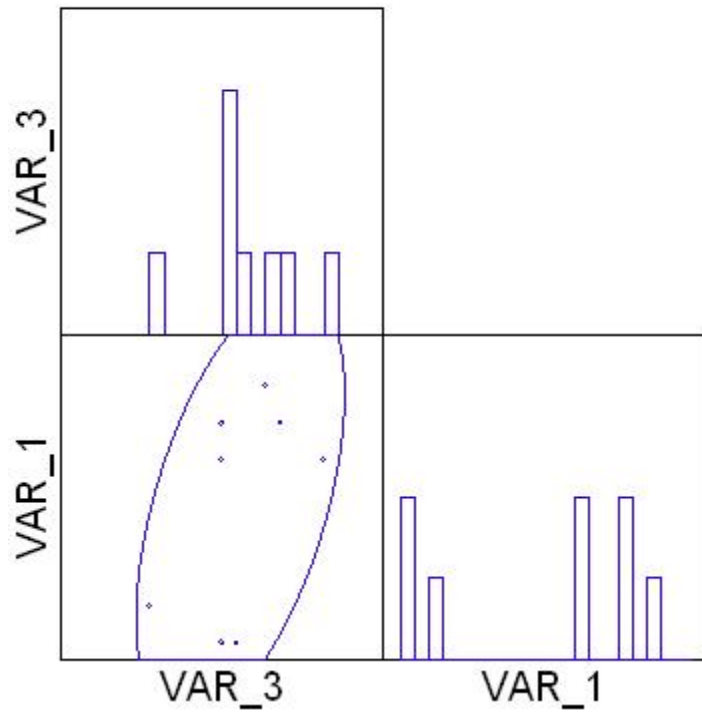


Gráfico 1. Coeficiente de correlación de Spearman del examen del curso 2015-16.

## Scatter Plot Matrix

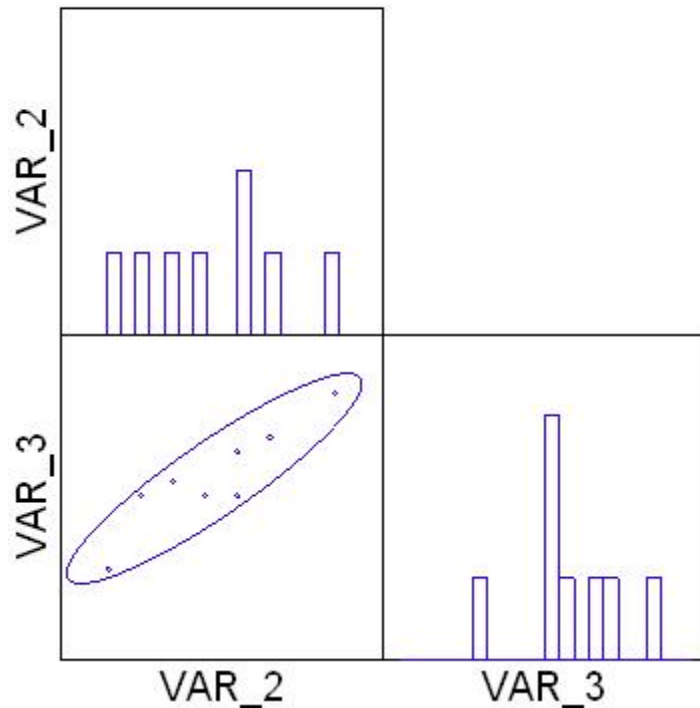


Gráfico 2. Coeficiente de correlación de Spearman del examen del curso 2016-17.

En el Cuadro 2 que evidencia la comparación de exámenes de Ortodoncia según ítems por tiempo en P1 del área temática, 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 % no coincidiendo con el tema de más horas en P1. Al analizar el examen del curso 2016-17 encontramos que 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, siendo también el de más cantidad de ítems dentro del examen con un 20 %.

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>

- Obtención de indicadores de validez del instrumento.

- Índice de dificultad:

Cuadro 3. Comportamiento de los Ítems de los exámenes según grado de dificultad esperado y real.

<b>Ítems según su grado de dificultad</b>	<b>% Ítems esperados</b>	<b>% ítems Reales 2015-16</b>	<b>% ítems Reales 2016-17</b>
<b>Fáciles</b>	<b>5</b>	<b>33,3</b>	<b>10</b>
<b>Medianamente fáciles</b>	<b>20</b>	<b>26,6</b>	<b>22</b>
<b>De dificultad media</b>	<b>50</b>	<b>33,3</b>	<b>45.8</b>
<b>Medianamente difíciles</b>	<b>20</b>	<b>0</b>	<b>18</b>
<b>Difíciles</b>	<b>5</b>	<b>6,6</b>	<b>4.2</b>
<b>Totales</b>	<b>100</b>	<b>100</b>	<b>100</b>

Al analizar el Cuadro 3. Comportamiento de los Ítems de los exámenes según grado de dificultad esperado y real, se puede apreciar que existe 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 al analizar el examen del curso 2016-17 encontramos un coeficiente de correlación de Spearman de 0,95 lo que demuestra una excelente correlación.

El trabajo de perfeccionamiento profesoral mediante cursos, diplomados, 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>

Cuadro 4. Comparación del coeficiente discriminación según preguntas de los exámenes.

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

En el Cuadro 4 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 pues 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. El coeficiente de discriminación tiene relación 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 malos.

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

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 correcta 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.

Cuadro 5. Comparación del coeficiente alfa de Cronbach de las preguntas y los exámenes.

<b>Preguntas</b>	<b>Coeficiente</b>	
	<b>Alfa de Cronbach</b>	
	<b>2015-16</b>	<b>2016-17</b>
<b>1</b>	<b>0,50</b>	<b>0, 61</b>
<b>2</b>	<b>0,59</b>	<b>0,60</b>
<b>3</b>	<b>0,48</b>	<b>0,62</b>
<b>4</b>	<b>0,54</b>	<b>0,64</b>
<b>5</b>	<b>0,57</b>	<b>0, 62</b>
<b>6</b>	<b>-</b>	<b>0,63</b>
<b>7</b>	<b>-</b>	<b>0,61</b>
<b>Examen</b>	<b>0,59</b>	<b>0,65</b>

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 la estructura del mismo. No obstante a 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 las mismas aportan consistencia interna al mismo. Cuando se analiza el examen del curso 2016-17 se observa una adecuada consistencia interna del mismo, 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.

Al determinar la calidad de los instrumentos evaluativos estamos obteniendo 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 mientras que si 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.
- La confiabilidad del examen es aceptable en el curso 2016-17, no siendo así en el curso 2015-16.

## REFERENCIAS BIBLIOGRÁFICAS

1. Alemañy Pérez E, Masjuán del Pino M, Gutiérrez Pérez M de los A. Estándares de calidad para la carrera de medicina. RevHabanCiencMéd [Internet]. 2009 Jun [citado 9 Oct 2017]; 8(2). Disponible en: [http://scielo.sld.cu/scielo.php?script=sci\\_arttext&pid=S1729-519X2009000200021&lng=es&nrm=iso&tlng=es](http://scielo.sld.cu/scielo.php?script=sci_arttext&pid=S1729-519X2009000200021&lng=es&nrm=iso&tlng=es).
2. Suárez Rosas L, Barrios Osuna I, González Espíndola M E. Metodología para la evaluación externa de la carrera de Estomatología. EducMedSuper [Internet]. 2012 Jun [citado 9 Oct 2017]; 26(2): 271-293. Disponible en: <http://www.ems.sld.cu/index.php/ems/article/view/31/27>.
3. Sánchez Sanhueza GA, Cisterna Cabrera F. La evaluación de los aprendizajes orientada al desarrollo de competencias en Odontología. EducMedSuper [Internet]. 2014 Mar [citado 8 Oct 2017]; 28(1): 104-114. Disponible en: <http://www.ems.sld.cu/index.php/ems/article/view/72/147>
4. Medina L. La evaluación en el aula: reflexiones sobre sus propósitos, validez y confiabilidad. 2013. *Revista Electrónica de Investigación Educativa*, [citado 8 Oct 2017]; 15(2): 34. Disponible en: <http://redie.uabc.mx/vol15no2/contenido-medina.html>
5. Delgado Ramírez M, Gómez Restrepo C. Evaluación en educación médica. Rev Colomb Psiquiat [Internet]. 2012 [citado 15 Oct 2016]; 41, (Supl1): 79-86. Disponible en: [www.scielo.org.co/pdf/rcp/v41s1/v41s1a09.pdf](http://www.scielo.org.co/pdf/rcp/v41s1/v41s1a09.pdf).
6. 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). EducMedSuper [Internet]. 2015 Dic [citado 9 Oct 2017]; 29(4): 763-776. Disponible en: <http://www.ems.sld.cu/index.php/ems/article/view/597/301>
7. Gómez López VM, Rosales Gracia S, Vázquez Vázquez JA. Validez y consistencia de un instrumento para evaluarla práctica docente en clínica médica de pregrado. Rev Educ Desar [Internet]. 2014 Ene-Mar [citado 15 Oct 2016]; 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 Sanhueza GA, Cisterna Cabrera F. La evaluación de los aprendizajes orientada al desarrollo de competencias en Odontología. Educ Med Super [Internet]. 2014 Mar [citado 15 Oct 2016]; 28(1): 104-114. Disponible en: <http://www.ems.sld.cu/index.php/ems/article/view/72/147>
  9. Valverde J, Ciudad A. El uso de e-rúbricas para la evaluación de competencias en estudiantes universitarios. Estudio sobre fiabilidad del instrumento. Rev Doc Univers [Internet]. 2014 abr [citado 15 Oct 2016]; 12 (1): 49-79. Disponible en: <https://dialnet.unirioja.es/descarga/articulo/4691792.pdf>
  10. Díaz Rojas Pedro Augusto, Leyva Sánchez Elizabeth. Metodología para determinar la calidad de los instrumentos de evaluación. Educ Med Super [Internet]. 2013 Jun [citado 2016 Oct 15] ; 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 [Internet]. 2016 [citado 15 Oct 2016]; 41 (1): 80-89. Disponible en: [www.aidep.org/03\\_ridep/R41/Art7.pdf](http://www.aidep.org/03_ridep/R41/Art7.pdf).
  12. Huapaya Yaya JM. Objetivos y evolución de la educación médica. Rev Horiz Med [Internet]. 2012 Oct – Dic [citado 15 Oct 2016]; 12(4): 43-47. Disponible en: [www.horizontemedicina.usmp.edu.pe/index.php/horizontemed/article/download/139/138](http://www.horizontemedicina.usmp.edu.pe/index.php/horizontemed/article/download/139/138).
  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 [Internet]. 2015 Mar [citado 15 Oct 2016]; 29(1): 47-56. Disponible en: <http://www.ems.sld.cu/index.php/ems/article/view/367/227>
  14. Alemañy Pérez E, Masjuán del Pino M, Gutiérrez Pérez M de los A. Estandares de calidad para la carrera de medicina. Rev Haban Cienc Méd [Internet]. 2009 Jun [citado 15 Oct 2016]; 8(2). Disponible en:

[http://scielo.sld.cu/scielo.php?script=sci\\_arttext&pid=S1729-519X2009000200021&lng=es&nrm=iso&tlng=es](http://scielo.sld.cu/scielo.php?script=sci_arttext&pid=S1729-519X2009000200021&lng=es&nrm=iso&tlng=es).