



**CUESTIONARIO DE OPINIÓN PARA LA EVALUACIÓN DEL
DESEMPEÑO DOCENTE EN LA UNAN-MANAGUA, NICARAGUA,
FAREM-CHONTALES**
QUESTIONNAIRE FOR EVALUATING TEACHER PERFORMANCE IN UNAN-MANAGUA,
FAREM-CHONTALES

Volumen 14, Número 2

Mayo - Agosto

pp. 1-28

Este número se publicó el 30 de mayo de 2014

Tonys Romero Díaz

Revista indizada en [REDALYC](#), [SCIELO](#)

Revista distribuida en las bases de datos:

[CATÁLOGO DE LATINDEX](#), [IRESIE](#), [CLASE](#), [DIALNET](#), [DOAJ](#), [E-REVIST@S](#),
[SHERPA/ROMEO](#), [QUALIS](#), [MIAR](#)

Revista registrada en los directorios:

[ULRICH'S](#), [REDIE](#), [RINACE](#), [OEI](#), [MAESTROTECA](#), [PREAL](#), [CLASCO](#)

Los contenidos de este artículo están bajo una licencia [Creative Commons](#)



CUESTIONARIO DE OPINIÓN PARA LA EVALUACIÓN DEL DESEMPEÑO DOCENTE EN LA UNAN-MANAGUA, NICARAGUA, FAREM-CHONTALES

QUESTIONNAIRE FOR EVALUATING TEACHER PERFORMANCE IN UNAN-MANAGUA, FAREM-CHONTALES

Tonys Romero Díaz¹

Resumen: El artículo expone la metodología empleada para desarrollar un cuestionario de evaluación docente para la Facultad Regional Multidisciplinaria de Chontales (FAREM), Nicaragua, al que nombramos CEDUCH, y los resultados obtenidos durante la investigación. El cuestionario consta de 30 preguntas distribuidas en ocho dimensiones o categorías las cuales se obtuvieron de un pilotaje mediante el Análisis Factorial exploratorio. Estas dimensiones mostraron estar altamente correlacionadas positivamente entre sí, confirmando que el cuestionario es multidimensional. Se utilizó una escala Likert de 5 adjetivos de respuesta, y para la valoración final se construyó una escala de 3 calificativos. Se aplicó una muestra probabilística de 453 estudiantes en la institución en donde los análisis realizados ítem-test e ítem-ítem mostraron correlaciones significativas que demuestran que las preguntas discriminan de acuerdo con el constructo medido. La validez de contenido del constructo arrojó un índice de 0.933; la consistencia interna mostró resultados altos e iguales a 0.967 el coeficiente theta y de 0.975 el coeficiente omega; la confiabilidad de la prueba se determinó mediante el Alfa de Cronbach donde se obtuvo un coeficiente de 0.966. Dado que el desempeño docente es multidimensional, se recomienda complementar el cuestionario con otro tipo de metodología que permita valorar holísticamente el actuar docente.

Palabras clave: ANÁLISIS FACTORIAL, DESEMPEÑO DOCENTE, EDUCACIÓN SUPERIOR, CUESTIONARIOS, EVALUACIÓN, NICARAGUA

Abstract: The article presents the methodology used to develop a teacher evaluation questionnaire for Multidisciplinary Regional School Chontales (FAREM), Nicaragua, which named CEDUCH, and the results obtained during the investigation. The questionnaire consists of 30 questions divided into eight dimensions or categories that were obtained from a pilot using exploratory Factor Analysis. These categories are shown to have a high positive correlation with one another, confirming that the questionnaire is in fact multidimensional. We use a Likert scale of five response adjectives, and for the final score we use a scale of three qualifiers. The questionnaire was given to a random sample of 453 students at CEDUCH; the item-test and item-item analyses showed significant correlations, which demonstrates that the questions discriminate in accordance with the construct being measured. The content validity of the construct yielded an index of 0.933; the internal consistency results were greater than or equal to 0.967 and 0.975 for the theta and omega coefficients, respectively. The test reliability was determined by Cronbach's alpha coefficient, which had a value of 0.966. As teacher performance is multidimensional, it is recommended to supplement the questionnaire with other holistic methodology to assess the teaching act.

KEY WORDS: FACTOR ANALYSIS, TEACHER PERFORMANCE, HIGHER EDUCATION, QUESTIONNAIRES, EVALUATION, NICARAGUA.

¹ Profesor de Matemática y Estadística en la Universidad Nacional Autónoma de Nicaragua, UNAN-MANAGUA. Estudiante del Doctorado en Educación e Intervención Social.

Dirección electrónica: tonyromerong@hotmail.com

Artículo recibido: 2 de setiembre, 2013

Devuelto para corrección: 4 de marzo, 2014

Aprobado: 15 de mayo, 2014

1. Introducción

Si bien es cierto que la evaluación de la docencia, por parte de los estudiantes y mediante el uso de cuestionario, es un tema muy estudiado en todas las universidades del mundo, en Nicaragua es un tema poco abordado desde el punto de vista de publicación. Está claro que la evaluación de la docencia en nuestras universidades se hace de algún modo, pues de otro modo carecería de sentido la premiación a profesores o el cambio de categoría docente, o la nueva contratación por el cese de actividades de algún compañero. Lo que no queda claro en algún documento oficial es el instrumento de evaluación y la metodología de aplicación, así como el uso que tiene la información.

En este contexto se propone la implementación del primer instrumento de evaluación docente llamado CEDUCH (Cuestionario de Evaluación Docente Unan Chontales), el cual servirá de apoyo al seguimiento y mejoramiento de la práctica docente en la Facultad. Este cuestionario presenta alta validez y confiabilidad, porque ha sido sometido a tres pruebas y dos evaluaciones de expertos (Anexo 1).

Será la primera vez que la Facultad cuente con un instrumento validado y probado en la misma Facultad, mediante el cual se obtendrán valoraciones confiables de los estudiantes y esto soportará decisiones del tipo académicas que se hagan con los docentes, en beneficio de un mejor ejercicio académico como es el de enseñar.

Este instrumento constituye el primer aporte teórico para nuestra universidad y además es novedoso en cuanto al estilo y la forma de recogida de la información, ya que el estudiante podrá evaluar a dos docentes simultáneamente y permitirá comparar sus aportes a la evaluación simultáneamente con lo cual el instrumento se vuelve más objetivo y eficaz.

Es así que ante las razones expuestas, se añade la necesidad de que los docentes podamos tener en cuenta los aportes de los estudiantes acerca de nuestro desempeño en el aula de clases y permita mejorar la misma práctica, ante lo cual surge la siguiente pregunta, ¿Por qué no utilizar la encuesta de medición del desempeño para estimar la satisfacción del estudiante sobre nuestro trabajo docente? Evidentemente, la negación de esta actividad haría que nuestra forma de trabajar siga siendo la misma de siempre y no encontraríamos puntos de inflexión en la docencia universitaria. Posturas como la negación de esta propuesta solo provocarían retrasos del orden competencia y actualidad universitaria que tanto necesita nuestra universidad en estos momentos de acreditación y competitividad internacional.

2. Breve referente teórico

La docencia es una de las actividades más complejas al momento de pretender evaluar el desempeño laboral de un docente, sea este administrativo o académico; la presente investigación propone un modelo de evaluación del tipo académica para docentes basado en un cuestionario de opinión por parte de los estudiantes. ¿Por qué solo por los estudiantes? En primer lugar, porque es la técnica de evaluación más usada por las universidades en el mundo y, al mismo tiempo, la más estudiada, existen más de dos mil trabajos publicados en relación a los cuestionarios de opinión docente (Luna y Torquemada, 2008). Aunque mucho se critique la centralidad que tienen los cuestionarios, hay muchas experiencias que muestran que la construcción de una adecuada escala para las preguntas y otra para la respuesta proporciona resultados confiables (González, Mendiola y Stack, 2010).

Los procesos de autoevaluación dentro de las instituciones universitarias son de los más importantes, ya que dan cuenta del funcionamiento interno de la organización, a través de ellos, se pueden establecer fortalezas y debilidades que retroalimentan la calidad de la educación ofrecida, y contribuyen significativamente a su mejora en todos los aspectos del quehacer universitario (Patiño, 2006); por tal motivo, este instrumento pretende ser un insumo para el proceso de autoevaluación institucional.

Hay muchas razones por las que se debiera tomar muy en cuenta la opinión de los estudiantes: primero, porque con ellos es con quienes el docente realiza todo el trabajo asignado; segundo, el tiempo que interactúan con el docente es el más largo posible que con cualquier otro involucrado; el alumno es un observador constante de la actividad docente; el alumno es afectado directamente por cualquier decisión que el docente realice en la asignatura correspondiente, se podría enumerar más razones y todas afirmarían que el estudiante es el actor más involucrado con la actividad docente.

La opinión de los alumnos sobre el desempeño de sus docentes representa un indicador muy importante a considerar en las tareas de evaluación. Por ello, el desarrollo sistemático de instrumentos como el que se presenta es una actividad no sólo deseable sino necesaria (Valenzuela, 2002).

Para la evaluación del desempeño docente existen muchas alternativas metodológicas, entre las más usadas están los cuestionarios de opinión, la entrevista cualitativa, los grupos de discusión con estudiantes, entre otras. Por ejemplo, se pueden usar pequeños grupos de

estudiantes para recoger información, por medio de entrevistas informales para analizar las percepciones de los estudiantes sobre el actuar docente.

Al respecto, Vásquez y Gabalán (2006) nos señalan que debe evitarse el medir la labor del docente en términos meramente cuantitativos, así como el establecer comparaciones con respecto a otros maestros, es decir, no abusar ni creer únicamente los resultados que arrojen los cuestionarios.

La evaluación debe orientarse a la obtención de un perfil del docente con respecto al modelo educativo, por ello, es conveniente hacer uso de varias evaluaciones al momento de considerar la implementación de un modelo de desempeño docente en la universidad: "La evaluación no sólo consiste en la recopilación de datos sobre su desempeño, sino que también conlleva la elaboración de juicios de valor sobre aspectos relevantes de la institución. Por eso, es una labor de carácter cuantitativo y cualitativo" (Aguilar, 2002, p. 10).

La inclinación al desarrollo de una metodología exclusivamente cuantitativa en el presente trabajo obedece a que constituye el primer acercamiento del problema en nuestra facultad, evidentemente como se señala en los párrafos anteriores, la metodología cuantitativa debe complementarse con la cualitativa para obtener un modelo integral al desempeño docente, sobre esta última se presentarán avances en los próximos meses, lo que le dará continuidad a la construcción del modelo de evaluación al desempeño docente en nuestra universidad.

Cuando se desea medir el desempeño es importante identificar que no se va a medir el desempeño en general, sino que se debe medir específicamente el desempeño de las funciones específicas, en este sentido como lo menciona Chiavenato (2000): "no estamos interesados en el desempeño general, sino en el desempeño del cargo, en el comportamiento de rol del ocupante del cargo", en esta medida es importante medir las actividades, tareas, procesos entre otros, lo que conlleva a un resultado propio de las funciones del cargo, pero este último no se mide en sí (p. 357).

Por su parte, Romero y González (2009) plantean que no se pretende que el docente sea un ser "perfecto" conocedor y solucionador de todos los problemas. Por el contrario, se requiere un ser humano, capaz de reflexionar en su propia práctica, corrigiendo errores, aceptando nuevas formas de concebir el mundo, dando apertura a todas las corrientes del pensamiento, utilizando el discurso argumentativo para erigirse en un líder de renovada

factura, convirtiéndose en un ejemplo de ciudadano mediante el cumplimiento de sus funciones y destacando con real humildad los elementales valores del hombre.

La evaluación del desempeño es un mecanismo que tiene como propósito la búsqueda de la calidad y del funcionamiento eficiente de las organizaciones orientadas a prestar servicios o producir bienes para el desarrollo de la sociedad. La dinámica social genera cambios profundos que logran alterar el orden socioeconómico, político, la tecnología y hasta los procesos de enseñanza y de aprendizaje, generando crisis, las cuales se deben solventar. He aquí la importancia de la evaluación para la autorregulación que incorporan las universidades, a fin de asegurar la calidad de los servicios que presta (Matos y Guadrón, 2008), más si esta evaluación está reglamentada permitiría que el proceso sea más efectivo.

En el diseño del cuestionario se ha considerado la experiencia de evaluación docente en universidades de países como México, España, los Estados Unidos y Venezuela. En estos países la aplicación de cuestionarios ha dado éxitos; sin embargo, es destacable que en estas universidades, aunque la evaluación se hace con fines de mejora de la práctica docente, también sirve como método para promoción y contratación de profesores. En este sentido, la propuesta en la FAREM Chontales es que sea un instrumento que permita evaluar el desempeño docente en el aspecto académico en beneficio de tener cambios que mejoren la práctica docente y el rendimiento en las asignaturas.

Por lo anterior, se ha concluido que la evaluación docente del departamento académico se haga por medio de un cuestionario. Dicho cuestionario ha pasado todas las pruebas de confiabilidad en sus tres etapas de validación.

3. Metodología

3.1. Tipo de estudio

El modelo adoptado en la investigación es de naturaleza exploratoria, descriptiva y correlacional. Esto es, se pretende explorar cuáles son las opiniones y criterios mediante el cual los estudiantes caracterizan el desempeño docente de la Facultad. Describir las dimensiones del trabajo académico en que se puede dividir el desempeño docente en la facultad por medio de este cuestionario. Correlacionar las diferentes dimensiones del instrumento, estableciendo una debida escala que justifique el uso del CEDUCH como instrumento de medición válido y confiable para evaluar la actuación docente.

La orientación predominante es correlacional, ya que tiene el propósito de evaluar la relación existente entre diversas variables y constructos acerca del desempeño docente (Hernández, Fernández y Baptista, 2004) y, sobre todo, se pretende dar a conocer cómo se comportan algunas de las variables involucradas, conociendo el comportamiento de las otras.

Por esta razón, optamos por una metodología de investigación centrada en el empleo de técnicas cuantitativas que se complementan y que nos permiten acceder al trabajo de campo con una visión más amplia, que se justifica con la aplicación del instrumento en cuestión y el respectivo análisis de datos, con el único objetivo de responder mejor al propósito de esta investigación.

3.2. Fase de diseño/adaptación

Primeramente, se aplicó el instrumento a cuatro grupos de clases que hacen un total de 34 estudiantes de distintos turnos. El objetivo en su primera versión, era entre cosas la de validar el instrumento, determinar si éste tenía un buen índice de discriminación y determinar mediante el análisis factorial exploratorio en cuántas dimensiones se podía agrupar el instrumento que medía el desempeño docente.

Inicialmente, el instrumento contaba con 20 preguntas, las que, de acuerdo a nuestro criterio, servían para medir el desempeño docente, así como otras preguntas de aspectos generales de las asignaturas, que en total hacían al instrumento tener 32 preguntas.

El procedimiento fue entregar en cada uno de los cuatros grupos el instrumento a los tres estudiantes con rendimiento más alto (más de 80 puntos) y otros tres cuestionarios a los estudiantes de rendimiento bajo (menos de 60 puntos) y dos cuestionarios a dos estudiantes de rendimiento medio, mediante una forma que ellos no podían saber, porque habían sido seleccionados (en dos de los grupos se pidió también la colaboración del presidente de sección). Se les dijo que era el resultado del muestreo y, una vez que entregaban el instrumento, el encuestador colocaba una clave al dorso del cuestionario e indicaba a qué grupo pertenecía. Este procedimiento se consideró así, porque se quería estar seguro que ninguna de las preguntas serviría de diferenciación para evaluar el desempeño docente². Una vez completado el instrumento, se obtuvieron los siguientes resultados:

² Queremos mencionar que este fue el procedimiento planificado antes de aplicar los cuestionarios, pero el día de la recolección de los datos hubo variación en la asistencia y en los tamaños de muestra, en algunos grupos no se

Tabla 1. Estadísticos de fiabilidad

Alfa de Cronbach	N de elementos
,916	32

Fuente: Elaboración propia del autor

Como se puede apreciar en la Tabla 1, la consistencia interna del instrumento resultó ser alta (Alfa de Cronbach = 0.916), lo que sirvió para proceder con el resto de análisis estadísticos.

Tabla 2. Preguntas del CEDUCH en su versión inicial (Algunas)

	PREGUNTAS DEL INSTRUMENTO (Versión No 1)	Correlación ítem-total	Alfa si se elimina
P3	Considera que el temario y contenido es Interesante	-,082	,934
P6	Considera que el temario y contenido son Aburridos	,000	,934
P7	Considera que el temario y contenido son Complicados pero merecen la pena	,098	,934
P8	Considera que el temario y contenido Me hacen ver otros puntos de vista	-,004	,934
P9	Considera que el temario y contenido es: N-D-L-A	,315	,934
P10	La asignatura ha aumentado mi interés en la materia	,761	,928
P11	Mi conocimiento sobre la materia ha aumentado gracias a esta asignatura.	,708	,929
P14	Demuestra sensibilidad ante las diferencias culturales	,351	,933
P15	El profesor es un experto en la asignatura	,400	,932
P16	Mi interés por esta asignatura ha aumentado a lo largo del curso cómo consecuencia de la docencia recibida por parte de este profesor.	,741	,928

Fuente: Elaboración propia del autor

La Tabla 2 muestra el alfa de Cronbach por cada ítem, así como la correlación de cada ítem respecto al instrumento en general; el criterio acá es que aquellos ítems que obtengan valores inferiores a 0.3 deben de ser eliminados. En nuestro caso, el criterio fue eliminarlos con valores inferiores a 0.4. ¿Por qué?, sencillamentem porque estos presentan dificultad en su llenado o la forma de responder no es la adecuada, por tal razón, se decidió eliminar del

pudo cumplir con lo establecido y se tuvo que completar con otros estudiantes de los mismos grupos seleccionados, Ver Anexo 2.

cuestionario las preguntas: P3, P4, P5, P6, P7, P8, P9, P14, P31 y P32; del mismo modo, se replanteó la pregunta número quince (P15) por estar en el valor del límite establecido como criterio.

En su criterio, Ebel y Frisbie (1991), a partir del análisis del comportamiento de ítems, sostienen que el coeficiente de dificultad y el índice de discriminación de un ítem están estrechamente relacionados y que aquellos ítems que tienen un coeficiente de dificultad medio (en el caso del ejemplo anterior entre 0.4 y 0.6) son los que mejor discriminan y, por tanto, proporcionan mayor información sobre las diferencias de nivel de competencia evaluado entre los participantes.

De esta forma, el cuestionario en su primera versión original y como propuesta contendría 21 preguntas. En el cuestionario el estudiante respondió en una escala numérica del 1 al 10, otorgando 1 al peor desempeño en cada pregunta y 10 al mejor de los desempeños, porque medir del 1 al 10 da mejores diferencias que de 1 al 5.

El nivel de confianza seleccionado fue del 0.05, por ser el más usado para las ciencias sociales. Después de esto se creó una variable que contenía la suma de las preguntas P10 a la P29, entonces, dividiendo por dos quedaba la calificación del estudiante dado a la asignatura impartida por un docente con una escala porcentual (del 1 al 100). Esta variable determinaría el tipo de distribución estadística que se usaría, con lo cual se probó la hipótesis de normalidad usando primero un gráfico P-P, Q-Q, y las pruebas de normalidad de Kolmogorov-Smirnov junto con la de Shapiro-Wilk para el caso en que el tamaño muestral es menor a 50 (Merino y Díaz, 2005).

Tabla 3 Pruebas de normalidad

Nombre de la Asignatura		Kolmogorov-Smirnov ^a			Shapiro-Wilk		
		Estadístico	Gl	Sig.	Estadístico	Gl	Sig.
Calificación	Estadística	,082	20	,200*	,975	20	,848
	Investigación de Operaciones	,332	5	,074	,797	5	,077
	Cálculo	,169	9	,200*	,903	9	,271

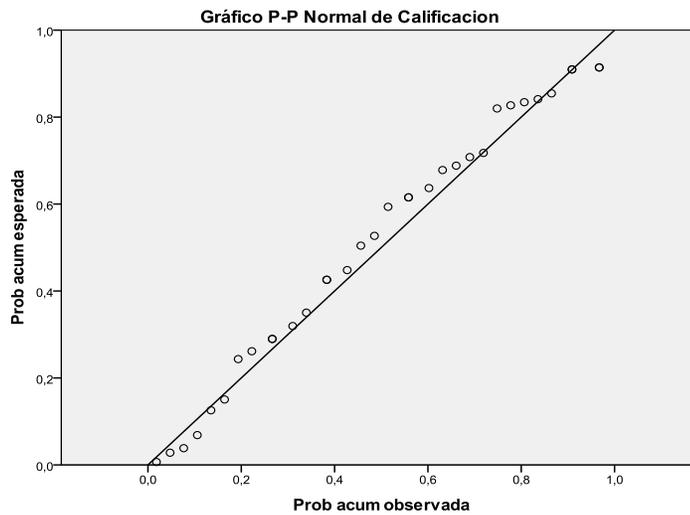
Fuente: Elaboración propia del autor

H₀: Los datos extraídos proceden de poblaciones normales.

H₁: Los datos extraídos no proceden de poblaciones normales.

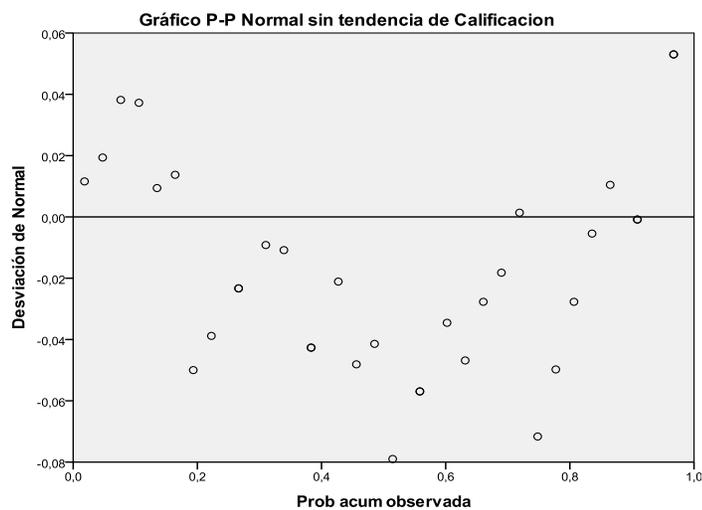
En la tabla 3 se aprecia que el estadístico de Shapiro-Wilk en los tres grupos fue superior a 0.05 con lo cual se acepta la hipótesis nula que las datos extraídos se comportan de acuerdo a la distribución normal (Se rechaza la hipótesis nula de normalidad en el caso que el valor de significancia sea inferior a 0.05). Mientras tanto, el gráfico 1 y el gráfico 2 confirman la hipótesis de normalidad en la variable calificación.

Gráfico 1. Normal de Calificación



Fuente: Elaboración propia del autor

Gráfico 2. Normal sin tendencia de Calificación



Fuente: Elaboración propia del autor

Entonces, una vez confirmado la normalidad de la variable calificación, se procedió a realizar el análisis de varianza de un factor. Para ello, se usó la variable calificación *versus* el rendimiento del estudiante. Esto para comprobar que el instrumento no diferencia lo suficiente en grupos diferentes y que en tales se pueden encontrar los mismos resultados para evaluar el desempeño, tanto para calificaciones bajas como altas.

Un supuesto necesario para el ANOVA es el homocedasticidad o la igualdad de varianza, así que se utilizó el estadístico de Levene, con el cual la el planteamiento de hipótesis quedaría de la siguiente manera:

H_0 : Las varianzas en los grupos son iguales.

H_1 : Las varianzas en los grupos son diferentes.

En suma, "La prueba de Levene consiste en llevar a cabo un análisis de varianza de un factor utilizando como variable dependiente la diferencia en valor absoluto entre cada puntuación individual y la media (o la mediana recortada) de su grupo" (Merino y Díaz, 2005, p. 275).

Tabla 4. Prueba de homogeneidad de varianzas

Estadístico de Levene	gl1	gl2	Sig.
,219	2	31	,805

Fuente: Elaboración propia del autor

Tal como lo muestra la Tabla 4, las variaciones en cuanto a las calificaciones obtenidas mediante el instrumento mostraron que son iguales en los tres grupos. Es decir, se aceptó la hipótesis nula.

Tabla 5. ANOVA

	Suma de cuadrados	Gl	Media cuadrática	F	Sig.
Inter-grupos	1344,493	2	672,247	2,317	,115
Intra-grupos	8995,514	31	290,178		
Total	10340,007	33			

Fuente: Elaboración propia del autor

El planteamiento para la hipótesis en el ANOVA es la siguiente:

H₀: Las medias en la calificación en los grupos son iguales.

H₁: Las medias en la calificación en los grupos son diferentes.

Del mismo modo, la Tabla 5, donde está contenido el resumen del ANOVA, muestra que el promedio obtenido en las calificaciones es igual en los diferentes grupos. De hecho, El nivel de significancia es mayor al valor crítico 0.05, por tanto se acepta la hipótesis nula.

Así mismo, se realizó un análisis *post hoc* para el ANOVA, específicamente la prueba de Tukey asumiendo homogeneidad en las varianzas para determinar si la prueba agrupaba en grupos diferentes los promedios de la calificación obtenidos. Como se puede apreciar, en la Tabla 6 solo un subconjunto puede ser extraído de todos los grupos, lo que confirma la prueba de hipótesis realizada con el ANOVA.

Tabla 6. Análisis Post-hoc
HSD de Tukey

Rendimiento en la asignatura	N	Subconjunto para	alfa =
		0.05	1
Bajo	10	68,5500	
Normal	15	74,9667	
Alto	9	85,2778	
Sig.		,073	

Fuente: Elaboración propia del autor

Con todo lo anterior, se superaba la primera etapa de determinar si el instrumento no diferenciaba claramente entre los estudiantes de distintos grupos y rendimiento, porque podría pensarse que si el instrumento cayese en manos de estudiantes de bajo rendimiento, estos calificarían mal el desempeño docente y viceversa, pero las pruebas han demostrado que tanto estudiantes con alto, medio o bajo rendimiento asignan puntajes en el instrumento independientes a su grupo de pertenencia, esto evitará que los resultados obtenidos al momento de aplicar el cuestionario estén sesgados anticipadamente.

La segunda parte de este análisis exploratorio era conocer en cuántos factores se pueden dividir el trabajo docente de acuerdo al cuestionario planteado. Aunque, debido a la muestra, los resultados deben de tomarse muy en consideración y aceptarlos solo si están bien alejados de los valores límites.

Uno de los requisitos que debe cumplirse para que el Análisis Factorial tenga sentido es que estas variables están altamente intercorrelacionadas. Por tanto, si las correlaciones entre todas las variables son bajas, el Análisis Factorial no es apropiado.

Por lo general, el análisis factorial exploratorio se utiliza para analizar las interrelaciones lineales dentro de un conjunto de indicadores observados, contados o medidos para cada individuo en un grupo, postulando que un conjunto reducido de factores comunes es suficiente para explicar las interrelaciones existentes entre indicadores. Debido a su función de exploración no asume modelo de medición. Su único propósito es reducir la dimensionalidad. Tampoco hay una cuota mínima de la varianza total explicada por el factor de estructura. Por lo tanto, diferentes criterios pueden ser usados como un criterio para un corte. Los métodos de rotación diferentes pueden dar lugar a diferentes interpretaciones (Barreto De Bem, Lanzer, Tambosi, Sanchez y Bernardi Junior, 2011).

Por tal razón, se aplicó el *test de esfericidad de Bartlett* que plantea la hipótesis nula de que la matriz de correlaciones es la identidad, y si eso es así, las intercorrelaciones entre las variables es cero. Para conformar el test, también se usó una medida de adecuación de la muestra, una de las formas más usadas es la Medida de adecuación de la Muestra KMO propuesta por Kaiser, Meyer y Olkin, (Vallejo, 2011) los cuales aconsejan que si $KMO \geq 0.75$ la idea de realizar un análisis factorial es buena, si $0.5 \leq KMO < 0.75$ es aceptable, y si es menor que 0.5 es inaceptable.

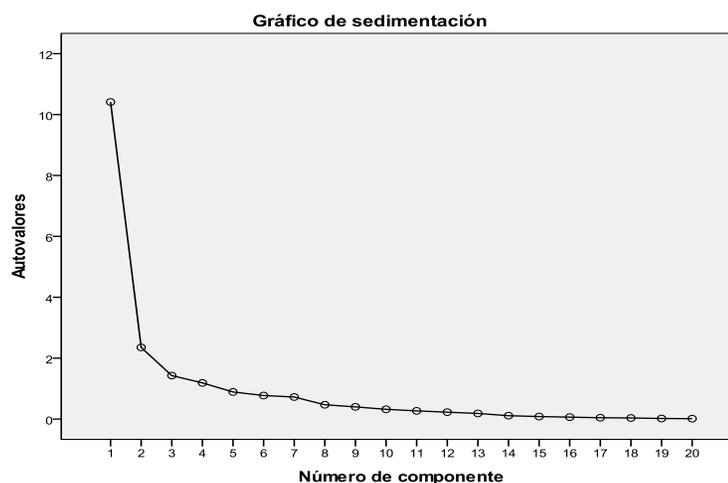
Tabla 7. KMO y prueba de Bartlett

Medida de adecuación muestral de Kaiser-Meyer-Olkin.		,712
Prueba de esfericidad de Bartlett	Chi-cuadrado aproximado	678,066
	GI	190
	Sig.	,000

Fuente: Elaboración propia del autor

La Tabla 7 muestra que tanto el test KMO como la prueba de Bartlett sugieren que se realice un análisis factorial, puesto KMO es mayor a 0.5 y el test de Bartlett rechaza la hipótesis de matriz de intercorrelaciones igual a la identidad, lo que significa que si hay intercorrelaciones entre los factores a extraer del cuestionario.

Gráfico 3. Sedimentación de los factores



Fuente: Elaboración propia del autor

El Gráfico 3 de sedimentación muestra que los seis primeros factores tienen autovalores superiores a 1, por lo tanto, de manera empírica podrían seleccionarse seis factores. Pero una vez aplicada la rotación de factores para su extracción mediante el análisis de componentes principales y el método de rotación Normalización Varimax con Kaiser, el resultado muestra claramente que son cuatro los factores que explican el desempeño docente mediante la aplicación de este instrumento.

3.3 Fase de validación del cuestionario

Para la segunda parte del diseño del cuestionario y superados los planteamientos de la primera etapa, se decidió revisar cuidadosamente la literatura respecto a la construcción de escalas para el diseño de cuestionarios de evaluación docente, encontrándose que casi todos usan la escala Likert con cinco puntuaciones y luego se tiene que crear una escala para la valoración del instrumento (Apodaca, 2002), por lo que el instrumento quedó determinado del mismo modo; además, al tomar la decisión sobre cuántos factores debería contener el instrumento, se encontró que el mínimo de factores era cuatro y lo más sobresaliente estaba entre seis y ocho factores, razón por la cual se agregaron más preguntas, para finalmente quedar en 30 preguntas divididas en ocho factores, el doble de lo planteado en el estudio exploratorio.

Esta vez, para validar el instrumento se sometió al juicio de cuatro profesores especialistas en temas de educación y que tienen vasta experiencia en la aplicación de instrumentos de evaluación tanto a estudiantes como a profesores.

Se puede hablar de dos tipos de validez del instrumento, uno relacionado a la consistencia interna (validez y confiabilidad) y otro relacionado a la validez del contenido, en tal caso, no se obtuvieron críticas negativas a la forma del instrumento y la cantidad de preguntas; se mejoró solo en cuestiones de redacción de cuatro preguntas, por lo que el CVC (índice referido a la validez del coeficiente de contenido) resultó ser muy alto, CVC = 0.9333, con lo cual, de acuerdo con Nieto (2000), el instrumento mide lo que debe medir.

La encuesta se aplicó en cinco grupos, con un tamaño de 52 encuestas de la modalidad sabatina del área de ciencias de la Educación, se seleccionó a estos grupos, porque la mayoría de ellos son docentes y están muy familiarizados con la evaluación mediante encuestas.

Tabla 8: Estadísticos de fiabilidad

Alfa de Cronbach	N de elementos
,933	30

Fuente: Elaboración propia del autor

Esta vez el instrumento resultó igual altamente confiable, de hecho, la Tabla 8 muestra que el alfa de Cronbach para este instrumento fue del 0.933; también se aplicó la prueba de dos mitades para medir el alfa de Cronbach:

Tabla 9. Estadísticos de fiabilidad

Alfa de Cronbach	Parte 1	Valor	,767
		N de elementos	17
	Parte 2	Valor	,910
		N de elementos	17
		N total de elementos	34
Correlación entre formas			,690
Coeficiente de Spearman-Brown	Longitud igual		,817
	Longitud desigual		,817
Dos mitades de Guttman			,768

Fuente: Elaboración propia del autor

Como se puede apreciar en la Tabla 9, los estadísticos del instrumento al dividirse este en dos partes (pares e impares) también resultó por encima de lo exigido por la estadística, (el alfa de Cronbach se considera óptima con resultados por encima de 0.7), junto con los otros estadísticos que acompañan la prueba.

Para la construcción del instrumento, como se mencionó, se rediseñó en ocho factores que miden, de la mejor forma, el desempeño docente; para constatar esto, se realizó otra vez el análisis factorial y se pudo comprobar que, efectivamente, estos ocho factores explicaban más del 80 % de la varianza compartida del instrumento. (Tabla 11)

También, para verificar la prueba de normalidad en la calificación obtenida en el cuestionario, se procedió nuevamente a aplicar la prueba Shapiro-Wilk, Tabla 10, donde se confirmó el test de normalidad para el instrumento.

Tabla 10. Pruebas de normalidad

Asignatura			Shapiro-Wilk		
			Estadístico	gl	Sig.
Calificación en el cuestionario	Asignatura 1	,864	4	,274	
	Asignatura 2	,961	6	,826	
	Asignatura 3	,919	4	,529	
	Asignatura 4	,947	7	,704	
	Asignatura 5	,922	5	,546	

Fuente: Elaboración propia del autor

Tabla 11. Determinación del número de dimensiones del instrumento

Componentes	Autovalores iniciales			Sumas de las saturaciones al cuadrado de la extracción			
	Total	% de la varianza	% acumulado	Total	% de la varianza	% acumulado	
Dimensiones	1	10,283	34,277	34,277	10,283	34,277	34,277
	2	4,135	13,785	48,061	4,135	13,785	48,061
	3	2,730	9,100	57,161	2,730	9,100	57,161
	4	2,018	6,726	63,887	2,018	6,726	63,887
	5	1,818	6,059	69,946	1,818	6,059	69,946
	6	1,566	5,221	75,167	1,566	5,221	75,167
	7	1,413	4,710	79,877	1,413	4,710	79,877
	8	1,052	3,505	83,382	1,052	3,505	83,382
	9	,934	3,115	86,497			
	10	,816	2,721	89,218			
	11	,716	2,387	91,605			
....				

Método de extracción: Análisis de Componentes principales.

Fuente: Elaboración propia del autor

Por todo lo mencionado hasta este momento, se puede argumentar que las bases para llevar a cabo la evaluación de los docentes, por medio del cuestionario (CEDUCH), mediante el número de factores propuestos y con la debida escala para cada ítem es adecuada, y se está en condiciones para aplicar el instrumento a un buen grupo de estudiantes de nuestra facultad de distintas carreras y departamentos académicos, confiadamente y con la seguridad que los resultados encontrados se corresponderán con los que el instrumento pretende medir.

3.4 Construcción de las escalas

Se construyó una escala Likert para el instrumento que constaba de cinco alternativas, con los puntajes, 1=Deficiente, 2=Regular, 3=Bueno, 4=Muy bueno, 5=Excelente o Sobresaliente para cada pregunta del cuestionario. Se escogió esta escala, porque es la más utilizada en los estudios de desempeño docente y porque nuestros estudiantes están muy familiarizados el medir con estos valores.

Así mismo, para los resultados finales se diseñó una escala para evaluar el buen o mal desempeño, la cual no es igual a la planteada, porque tenemos que recordar que la escala Likert presenta dificultades en cuanto a la tendencia al valor medio, por tanto autores como

González *et al.*, (2010), han construido y validado escalas a partir de una muestra de 20136 cuestionarios y 919 docentes/asignaturas para medir el desempeño final, donde una vez que se han procesado todos los instrumentos han desarrollado y validado el mismo de acuerdo al siguiente criterio: "Excelente o Muy Bien" si las calificaciones son 4.5 o mayores a este. Las calificaciones comprendidas entre 4 y 4.5 como "Bien o Aceptables" y todas aquellas inferiores a 4 puntos como "Insatisfactorias o Deficientes".

3.5 Fase de Aplicación y confirmación

El CEDUCH se aplicó en el séptimo encuentro del Primer Bimestre del Primer Semestre, esto porque los estudiantes ya estaban por finalizar el primer Bimestre (solo les quedaba pendiente un encuentro) y tenían conocimiento de todos los aspectos evaluados en el cuestionario. Se planificó encuestar a 22 grupos de clases del turno sabatino, pero al momento de aplicar el instrumento dos grupos de clases no se encontraban en la Facultad y se decidió completar con dos grupos de clases del turno dominical los cuales se seleccionaron aleatoriamente.

Se obtuvo respuesta de 453 cuestionarios, un número muy considerable para realizar todo tipo de estadística, sea ésta descriptiva o inferencial. En cada cuestionario el estudiante marcaba un número que se correspondía con la escala Likert construida de la siguiente manera:

- Deficiente o "1" cuando el docente exhibe este desempeño entre 0 a 19 %.
- Regular o "2" cuando el docente exhibe este desempeño entre 20 a 39 %.
- Bueno o "3" cuando el docente exhibe este desempeño entre 40 a 59 %.
- Muy bueno o "4" cuando el docente exhibe este desempeño entre 60 a 79 %.
- Excelente o "5" cuando el docente exhibe este desempeño entre 80 a 100 %.

El instrumento estaba compuesto por 30 reactivos, lo que requiere que el estudiante emita su opinión sobre el desempeño de sus profesores, distribuido en ocho dimensiones básicas del quehacer docente frente a un grupo de clases. Estas dimensiones del trabajo docente están sustentadas en el análisis factorial exploratorio realizado en el segundo pilotaje y en base al trabajo realizado por García (1997), donde hace una comparación de los principales instrumentos utilizados para medir el desempeño docente en 20 universidades distintas de España, con muestras de que van desde 3456 hasta 97249 estudiantes. Los

instrumentos analizados fueron: SEEQ, Endeavor, CEQ, E.E.P.A, CEDA-MOD; las dimensiones evaluadas en este trabajo concluye que varían entre cinco y nueve dimensiones, aunque las diferencias entre las distintas dimensiones, tienen que ver con el modo de ver la enseñanza universitaria, donde encontró que todos los instrumentos hacen referencia a:

Dominio de contenidos-dominio de la asignatura.

Claridad en la exposición y desarrollo del programa.

Comunicación con los alumnos tanto individual como grupal.

Programación- planificación- organización de la enseñanza. Adecuada carga de trabajo.

Evaluación-calificación-exámenes. (Rendimiento del alumno).

Entusiasmo del profesor-motivación al alumno-incremento del interés.

Es por eso que, de alguna manera, todos estos componentes están presentes en el diseño del cuestionario en esta investigación. En relación con el número de reactivos usados en los instrumentos de evaluación se sugiere que no sea menor de 20 ni mayor de 50 reactivos (Cantarero, Deus y Paz, 2002), donde proponen que una dimensión o componente del desempeño docente debe tener entre tres y nueve reactivos, por tal razón se justifica el número de reactivos usados en la construcción del CEDUCH.

La metodología para la aplicación fue tomar entre 10 y 12 estudiantes de cada grupo de clases, en cada aplicación se superó el 90% de la asistencia y tanto al estudiante como al docente que estaba no se les avisaba, sino hasta el momento de la aplicación, esto con el objetivo de evitar preparación anticipada a las respuestas del mismo.

El instrumento ha sido realizado también tomando en cuenta los trabajos de Marsh y Overall (1980) y Marsh (1982), donde presentan un modelo para evaluar el desempeño docente al que denominan "Evaluación de la Calidad de la Educación por el Estudiante" (SEEQ: Students' Evaluation of Education Quality), el cual ha sido usado por más de treinta años y ha dado excelentes resultados, a tal modo que se le considera el instrumento por excelencia para medir desempeños tanto de maestros y estudiantes, y se ha probado y validado en muestras hasta de hasta un millón de cuestionarios.

El modelo incluye un cuestionario conformado por reactivos diseñados para medir siete dimensiones de la enseñanza efectiva. Cada reactivo incluye un aspecto específico a evaluar. Estos se responden en una escala tipo Likert de cinco alternativas, diseñada para evaluar la calidad de algún aspecto expresado en cada reactivo. Los valores de la escala son "Muy mal",

"Mal", "Moderado", "Bien" y "Muy Bien", que como resultado obtienen siete factores que definen el desempeño docente: Entusiasmo/Interés, Amplitud en el enfoque, Organización, Interacción, Valor del aprendizaje, Exámenes y calificaciones, Carga de trabajo/Nivel de dificultad (Marsh, 1982).

Por esta razón, las ocho dimensiones en las que finalmente se estructuró el instrumento fueron: responsabilidad formal, metodología, comunicación, materiales, actitud ante la clase, evaluación, motivación y satisfacción del estudiante. La descripción de estas dimensiones se muestra a continuación:

Responsabilidad Formal: cumplimiento cabal de todas las actividades que involucran las funciones del profesor.

Metodología: habilidad para desarrollar los contenidos y buena organización para hacer cumplir lo programado en el curso.

Comunicación: destreza del profesor para hacer saber a sus alumnos de manera clara, eficiente y lógica las normas, reglas, políticas y actividades que se manejan en su clase y en la institución.

Materiales: uso de recursos didácticos-metodológicos que favorecen la claridad y desarrollo del curso.

Actitud ante la docencia: consideración de creencias y expectativas del profesor, su visión o posición sobre el contenido, su modo de enseñarlo, el clima del aula que genera su actuación, la opinión sobre sí mismo, sobre el proceso de enseñanza y sobre sus alumnos.

Evaluación: el uso de pruebas para acumular, la entrega a tiempo de los resultados, la correspondencia en los exámenes con lo desarrollado y la calificación justa de estos.

Motivación: contagia entusiasmo y crea un buen ambiente de trabajo, es personal y potencia todas las dimensiones restantes del instrumento.

Satisfacción del estudiante: consideración de creencias y expectativas planteadas sobre el profesor, su visión o posición sobre el contenido, su modo de enseñarlo, el clima del aula que genera su actuación, la opinión sobre sí mismo, sobre el proceso de enseñanza y sobre sus alumnos.

4. Resultados

La Tabla 12 muestra la correlación entre cada factor del instrumento, se observa que todos los factores están relacionados positivamente entre sí; todas las correlaciones son significativas bilateralmente al 0.05, lo cual es de esperarse, dado que para que exista una buena práctica docente todos los aspectos deben estar relacionados entre sí, ningún aspecto es más importante que otro, todos se complementan, lo que da lugar a un desempeño docente integral.

Tabla 12. Matriz de Correlaciones

	Responsabilidad	Metodología	Comunicación	Materiales	Actitud	Evaluación	Motivación	Satisfacción
Responsabilidad	*****							
Metodología	0.601	*****						
Comunicación	0.664	0.748	*****					
Materiales	0.579	0.661	0.619	*****				
Actitud	0.641	0.646	0.803	0.591	*****			
Evaluación	0.695	0.725	0.775	0.658	0.750	*****		
Motivación	0.652	0.798	0.779	0.690	0.748	0.758	*****	
Satisfacción	0.660	0.787	0.822	0.688	0.772	0.793	0.848	*****

Fuente: Elaboración propia del autor

Las correlaciones más altas se dan entre las dimensiones actitud-comunicación-satisfacción y la más alta correlación le corresponde a motivación-satisfacción. Estos resultados son consistentes con los trabajos desarrollados por Guzmán, Lara y García (2010) al referirse que el estudiante busca establecer una adecuada coordinación entre la actitud y la habilidad de comunicación del maestro como indicadores de una satisfacción general del curso.

Tabla 13. Promedio por categoría por carrera

Carrera que cursa el estudiante	Responsabilidad	Metodología	Comunicación	Materiales	Actitud	Evaluación	Motivación	Satisfacción
Carrera 1	4.4833	4.6500	4.6125	4.6000	4.7000	4.4400	4.2000	4.5750
Carrera 2	4.5261	4.4157	4.1867	4.3052	4.4137	4.3952	4.1536	4.1837
Carrera 3	4.2571	4.3429	<u>3.9429</u>	<u>3.9333</u>	4.1048	4.0400	4.1357	<u>3.7714</u>
Carrera 4	4.5194	4.4186	4.4651	4.2946	4.5891	4.3442	4.2035	4.1744
Carrera 10	4.6552	4.0862	<u>3.9181</u>	4.1092	4.1494	4.0276	<u>3.9267</u>	<u>3.6983</u>
Total	4.6336	4.4741	4.3996	4.3723	4.5467	4.4450	4.2958	4.2621

Fuente: Elaboración propia del autor

Para comprobar si el instrumento discriminaba adecuadamente, se extrajeron los rendimientos obtenidos por carrera por cada uno de los aspectos que medía el instrumento, donde, como puede observarse en la tabla 13, en muchas carreras los ocho aspectos son sobresalientes y aceptables. Sin embargo, las carreras 3 y 10 presentan desempeños deficientes en tres aspectos del instrumento, lo que indica que el instrumento y la escala constituyen buenos constructos para medir el desempeño docente en la facultad.

Tabla 14. Confiabilidad de los ocho aspectos que mide el instrumento.

	Media de la escala si se elimina el elemento	Varianza de la escala si se elimina el elemento	Correlación elemento-total corregida	Alfa de Cronbach si se elimina el elemento
Responsabilidad	30.7957	25.303	.727	.950
Metodología	30.9552	23.167	.822	.943
Comunicación	31.0297	22.143	.863	.940
Materiales	31.0569	23.431	.729	.949
Actitud	30.8825	23.165	.815	.943
Evaluación	30.9842	23.542	.850	.942
Motivación	31.1334	21.775	.878	.939
Satisfacción	31.1671	20.831	.896	.939

Fuente: Elaboración propia del autor

Se ha visto en los párrafos anteriores que el instrumento en lo general cumple todos los coeficientes de fiabilidad y validez, por tal motivo se analizaron factor por factor estos mismos índices, presentados en la tabla 14 y como resultado, una vez agrupadas las variables de cada aspecto en el instrumento junto con las pruebas del análisis factorial, se obtuvieron coeficientes altos y muy estables en relación a su medida, lo que nos confirma que la estructura propuesta es la más indicada. Asimismo, la Tabla 15 muestra la confiabilidad total de los treinta ítems que componen el instrumento, muy alta, razón por la cual hemos de esperar que los estadísticos descriptivos que arroje el instrumento una vez puesto en marcha como insumo para la autoevaluación en la facultad sean confiables y consistentes.

Tabla 15. Estadísticos de fiabilidad

alfa de Cronbach	N de elementos
.966	30

Fuente: Elaboración propia del autor

Una vez calculado el coeficiente alfa del instrumento, también, se calcularon los coeficientes theta (θ) y omega (Ω), que están basados en el análisis factorial de los ítems, son indicadores de la consistencia interna similares al coeficiente alfa de Cronbach. El coeficiente theta fue desarrollado por Carmines y Zeller en 1979, y el coeficiente omega fue desarrollado por Heise y Bohrnstedt en 1970, (Muñiz, 1998). Entre estos coeficientes se debe de cumplir la relación $\alpha < \theta < \Omega$.

El cálculo del alfa de Cronbach se realizó con el SPSS, mientras los otros dos coeficientes de fiabilidad se realizaron manualmente con el empleo de las fórmulas correspondientes.

Donde n es el número de ítems de la escala; λ_1 es el valor propio mayor, es decir la varianza explicada por el primer factor antes de la rotación correspondiente a la matriz de correlaciones; h^2 es la comunalidad de cada uno de los ítems; y r_{ij} la correlación entre cada par de ítems.

Por lo antes expuesto, se puede ver que este instrumento cumplió con los requisitos del análisis factorial, de consistencia interna y es altamente confiable al comprobarse la relación ($0.966 < 0.967 < 0.975$) ya citada.

5. Conclusiones

Con el proceso presentado para la construcción del cuestionario y las escalas de medidas, queremos hacer énfasis en el uso de este cuestionario como una necesidad para incluirlo como insumo en los procesos de autoevaluación.

Para la mejora de la práctica docente y el buen desempeño del profesorado es necesario contar con un buen modelo de evaluación. Ese ha sido el tema central de este trabajo, que ha girado en el desarrollo del cuestionario de evaluación docente por parte del alumno, dentro del paradigma positivista, donde ha sido necesario tomar en cuenta un sistema de fundamentos metodológicos, que han sido consultados en los trabajos más importantes que se desarrollaron en universidades de gran prestigio y que hasta el momento desarrollan este modelo de evaluación.

El instrumento y sus escalas fueron sometidas a un riguroso proceso de validación y fiabilidad, mediante tres etapas de implementación del mismo, tiempo durante el cual el CEDUCH mostró ser muy consistente y estable.

La escala propuesta para la medición de cada uno de los factores y del instrumento en particular debe servir para fomentar la cultura de evaluación del profesorado para mostrar desempeños por encima de la media esperada, con la finalidad de mejorar la práctica docente.

La estructura interna de la prueba construida presentó muchos elementos a favor para ser incluido como instrumento de evaluación interna de la facultad, entre los que destacamos:

Un coeficiente de Cronbach de 0.96 es muy alto y refleja la consistencia del cuestionario. En el análisis factorial se obtuvieron ocho factores que miden el desempeño docente de la facultad. En relación con el índice de discriminación medido por la correlación ítem-test fue superior a 0.4 en todas las preguntas del cuestionario, ($r > 0.4$) En este sentido, los reactivos que componen el instrumento tienen niveles adecuados de discriminación.

Comparando los ocho aspectos que mide el instrumento, los resultados mostraron claramente una alta confiabilidad que tiene la prueba, pues el menor valor obtenido al eliminarse una de estos aspectos es 0.939; así mismo, estas categorías discriminaron adecuadamente en el instrumento.

La estructura factorial indicada junto con la solución rotada de la misma aporta suficiente evidencia que son ocho factores los que explican más del 70 % de la varianza del instrumento; además, los coeficientes theta y omega que acompañan al análisis factorial se

corresponden con lo indicado en la literatura, los resultados de estos fueron 0.967 y 0.975, valores que junto al alfa de Cronbach mencionado afirman la confianza de utilizar este instrumento en la facultad para medir el desempeño docente.

El instrumento da números consistentes de fiabilidad: 0.91 para muestras entre 20 y 29 estudiantes; 0.933 para tamaños entre 30 y 50; finalmente 0.96 para muestras superiores a 50 estudiantes. Similares a los resultados que obtuvo Marsh (2011).

Por último, el instrumento comprueba que la práctica docente es multidimensional, por lo tanto, se destaca que se debe seguir profundizando al mismo tiempo en otras metodologías que incluyan técnicas cualitativas para la evaluación del desempeño docente, ya que el cuestionario no se justifica como única técnica que cubra todas las manifestaciones de una actividad tan compleja como la docencia.

6. Referencias

- Aguilar Hernández, Marielos. (2002). *La evaluación institucional de las universidades. Tendencias y desafíos*. Recuperado de http://www.6congreso.ucr.ac.cr/eval_inst.html.
- Apodaca, Pedro. (2002). Análisis dimensional de las opiniones de los alumnos universitarios sobre profesores: Comparación entre técnicas paramétricas y no paramétricas. *Redie*, 20(2), 385-409.
- Barreto De Bem, Amilton, Lanzer, Edgar Augusto, Tambosi Filho, Elmo, Sanchez, Octavio Próspero, y Bernardi Junior, Plinio. (2011). Validade e confiabilidade de instrumento de avaliacao de docencia sob a óptica dos modelos de equacao estructural. *Revista da Avaliacao da Educacao Superior*, (16), 375-401.
- Cantarero Muñoz, José Miguel; Deus Ríos, María Paula y Paz, Eduardo Abalde. (2002). Evaluación Docente vs. Evaluación dela Calidad. *Revista ELección de Investigación y EValuación Educativa*, 8(2), 103-134.
- Chiavenato, Ignacio. (2000). *Administración de recursos humanos* (5a. ed.). Bogotá, Colombia: McGraw-Hill.
- Ebel, Randolph Lee, y Frisbie, Dan August. (1991). *Essentials of Educational Measurement* (5a. ed.). New Jersey: Prentice Hall.
- García, José Manuel. (1997). Valoración de la competencia docente del profesor universitario. Una aproximación empírica. *Revista Complutense de Educación*, 8(2), 81-108.
- González Martínez, Adrián, Mendiola Sánchez, Melchor y Stack Martiínez, Jorge. (2010). Los cuestionarios de opinión del estudiantes sobre el desempeño docente. Una estrategia

institucional para la evaluación de la enseñanza en Medicina. *Revista Electrónica de Investigación*, 12(1), 7-8.

Guzmán Zambrano, Rogelio; Lara, Rosa Meda y García Lara, Baudelio. (2010). *Evaluación de profesores universitarios por parte de los alumnos mediante el Cuestionario de Evaluación de Desempeño Docente (CEDED)*. Guadalajara: Universidad.

Hernández Sampieri, Roberto, Fernández, Carlos y Baptista, Pilar. (2004). *Metodología de la Investigación* (3a. ed.). México D.F: Mc Graw-Hill.

Luna, Edna, y Torquemada, Alma. (2008). Los cuestionarios de evaluación de la docencia por los alumnos: balance y perspectivas de su agenda. *Revista Electrónica de Investigación Educativa* (Especial), 1-15.

Marsh, Herbert. (1982). Reliable, Valid and Useful Instrument for Collection Students' Evaluations Of University Teaching. *Brithish Journal of Education Psychology* (52), 77-95.

Marsh, Herbert. (2011). *Student's Evaluations of University Teaching: Ratings of Teacher, Universities and Departamentos*. Londres: Univesity of Oxford.

Marsh, Herbert, y Overall, John. (1980). Validaty of student's evaluations of teaching effectiveness: cognitive and effective criteria. *Journal of Educational Psychology*, 72(4), 468-475.

Matos Guerreo, Waldemar y Guadrón, Luis José. (2008). La evalaución en la praxis docente universitaria como mecanismo para asegurar la calidad. *ORBIS*(4), 110-142.

Merino Pardo, Antonio, y Díaz, Miguel Angel. (2005). *Análisis de datos con SPSS 13 Base*. Madrid: Mc Graw-Hill.

Muñiz, José. (1998). *Fiabilidad: Teoría Clásica de los test* (5a. ed.). Madrid: Pirámide.

Nieto Hernández, Rafael. (2000). *Instrumentos de Recolección de Datos en Ciencias Sociales y Ciencias Biomédicas*. Mérida: Universidad de los Andes.

Patiño, Luis Enrique. (2006). Análisis psicométrico de la escala de evaluación docente de la corporación universitaria iberoamericana. *Suma Pisológica* (194), 193-215.

Romero, Nerylena y González, Lorena. (2009). Competencias del docente de educación superior como mediador en los procesos de investigación y evaluación de los aprendizajes. *Omnia* (15), 39-55.

Valenzuela, Jesús Ernesto. (2002). *Evaluación del desempeño docente a partir de la opinión de los alumnos*. Sonora: Universidad de Sonora.

Vallejo Morales, Pedro. (2011). *El análisis factorial en la construcción e interpretación de test, escalas y cuestionarios*. Madrid: Universidad Pontificia Comillas.

Vásquez Rizo, Fredy Eduardo y Gabalán Coello, Jesús (2006). *Percepciones estudiantiles y su influencia en la evaluación del profesorado. Un caso en la Universidad Autónoma de Occidente, Cali – Colombia*. *RELIEVE*, 12(2). Recuperado en http://www.uv.es/RELIEVE/v12n2/RELIEVEv12n2_3.htm.

Anexo 1: Cuestionario CEDUCH versión final y confiabilidad por ítems.

LOS TREINTA REACTIVOS DEL CEDUCH	Correlación elemento-total corregida	Alfa de Cronbach si se elimina el elemento	Dimensión a la que pertenece
Asiste regularmente a clases y si falta justifica (Manda sustituto)	.484	.966	Responsabilidad
Cumple puntualmente con el horario de clases	.418	.966	Responsabilidad
En las horas de tiempo libre o tutoría es accesible con los estudiantes	.705	.965	Comunicación
Domina los contenidos de la materia que imparte	.617	.965	Metodología
Da a conocer el programa, metodología y evaluación de la asignatura al inicio del curso	.626	.965	Metodología
Explica de forma ordenada y clara sus clases	.795	.964	Metodología
Antes de empezar la clase empieza explorando y relacionando el nuevo contenido	.744	.964	Metodología
Motiva constantemente la participación y diálogo en clases	.775	.964	Motivación
Este profesor propicia que mi interés por la asignatura aumente	.788	.964	Motivación
La cantidad de trabajos y tareas asignadas las considero adecuadas a mi tiempo	.662	.965	Evaluación
Si le hago una consulta me brinda atención individual a la inquietud	.747	.964	Comunicación
Si los datos o información brindada son incorrectos, acepta correcciones del alumno	.733	.965	Comunicación
Es respetuoso con los estudiantes y sus ideas	.698	.965	Actitud
Utiliza esquemas, gráficos, medios audiovisuales para el desarrollo de la clase	.540	.966	Materiales
Relaciona el tema con la vida cotidiana o con el perfil de la carrera	.607	.965	Materiales
Los materiales utilizados y recomendados (folletos, guías,..) son de mucha utilidad	.674	.965	Materiales
El profesor es exigente con las tareas, trabajos y evaluaciones	.524	.966	Evaluación
El grado de exigencia en los exámenes se corresponde con los contenidos impartidos	.685	.965	Evaluación
Se muestra dispuesto a razonar la nota obtenida del estudiante que reclama revisión	.755	.964	Comunicación
Da a conocer las calificaciones en el plazo establecido	.706	.965	Evaluación
Estoy satisfecho con la labor de este profesor	.849	.964	Satisfacción
El nivel de exigencia académica de la asignatura ha sido:	.647	.965	Satisfacción
El docente realiza evaluaciones periódicamente de la asignatura. (trabajos, sistemáticos)	.606	.965	Evaluación
Entrega a tiempo los resultados de los trabajos independientes	.690	.965	Responsabilidad
Desarrolla la capacidad investigativa en los estudiantes	.708	.965	Motivación
Manifiesta un trato justo e imparcial con todos los estudiantes	.764	.964	Actitud

Es cuidadoso en su presentación personal	.597	.965	Actitud
Motiva a los estudiantes a asistir a clases	.717	.965	Motivación
¿Cuál es tu nivel de satisfacción general con el profesor?	.859	.964	Satisfacción
¿Te gustaría volver a cursar otra asignatura con este profesor?	.797	.964	Satisfacción

Fuente: Elaboración propia del autor

Anexo 2: Distribución inicial de la muestra, primera versión.

		Rendimiento en la asignatura			Total
		Bajo	Normal	Alto	
Nombre de la Asignatura	Estadística G1	3	5	2	10
	Investigación de Operaciones	2	1	2	5
	Cálculo	2	4	3	9
	Estadística G2	3	5	2	10
Total		10	15	9	34

Fuente: Elaboración propia del autor