

UNIVERSIDAD RAFAEL LANDIVAR  
FACULTAD DE HUMANIDADES  
DEPARTAMENTO DE PSICOLOGÍA

“CONFIABILIDAD DE UN INSTRUMENTO  
DE EVALUACIÓN DEL DESEMPEÑO  
PARA CATEDRÁTICOS UNIVERSITARIOS”

TESIS

Presentada al Consejo de la Facultad de Humanidades

Por:

JULIA MARÍA BALCONI VILLASEÑOR

Para optar al Título de

PSICÓLOGA INDUSTRIAL/ ORGANIZACIONAL

En el Grado Académico de

LICENCIADA

GUATEMALA DE LA ASUNCIÓN, JULIO DE 2002

## RESUMEN

El objetivo de la presente investigación fue determinar la confiabilidad de un instrumento de evaluación del desempeño para catedráticos universitarios, de una institución de educación superior, ubicado en la Ciudad Capital que se dedica fundamentalmente a formar hombres y mujeres civilizados, cultos y competentes en el campo de su profesión; enriquecer y divulgar la ciencia, la tecnología y la cultura, y colaborar con el desarrollo de la sociedad. La muestra fue conformada por un grupo de cuatrocientos estudiantes de los primeros como de los últimos años de carrera, tanto del sexo femenino como del sexo masculino, comprendidos entre los 18 años y 29 años de edad. El análisis se realizó por medio de un paquete estadístico (SPSS). El Alfa de Cronbach proporciona el índice de confiabilidad, el coeficiente alfa utilizado, se ubica dentro de las medidas de consistencia interna de los instrumentos. El instrumento que fue utilizado consta de 16 preguntas cerradas, donde se pretendió medir preparación y responsabilidad, metodología de enseñanza, evaluación, y relación interpersonal e investigación del catedrático. La confiabilidad de cada una de las áreas es: preparación y responsabilidad 0.88, metodología de enseñanza 0.89, evaluación 0.90 y relación interpersonal e investigación 0.92. el instrumento en conjunto mostró un coeficiente alpha de confiabilidad de 0.93.

## ACTO QUE DEDICO

A DIOS

A MIS PADRES:

- Julia Balconi Turcios
  - Lourdes Villaseñor de Balconi
- Por su apoyo y comprensión durante toda mi vida

A MIS HERMANAS

- Lourdes Claudette Balconi Villaseñor
- Nancy Lisette Balconi Villaseñor
- Diana Alessandra Balconi Villaseñor

A MI SOBRINO Y SOBRINAS

- Julio Alejandro Fernández Balconi
- Lourdes María Fernández Balconi
- Nancy María Turcios Balconi
- Fátima María Turcios Balconi

# ÍNDICE

Páginas

## CONTENIDO

## RESUMEN

I.	INTRODUCCIÓN.....	1
	1.1 Propósito general del estudio.....	1
	1.2 Antecedentes de estudio del problema.....	2
	1.3 Marco teórico.....	10
	1.3.1 Confiabilidad.....	10
	1.3.2 Análisis de ítem.....	15
	1.3.3 Evaluación de desempeño.....	17
	1.3.4 Evaluación del desempeño docente.....	25
	1.3.5 Información sobre la utilización de la evaluación del desempeño docente en tres universidades de la ciudad capital.....	30
II.	PLANTEAMIENTO DEL PROBLEMA.....	37
	2.1 Objetivo general de la investigación.....	37
	2.2 Objetivos específicos de la investigación.....	37
	2.3 Hipótesis de la investigación.....	38
	2.4 Definición de variables.....	40
	2.5 Alcances y límites de la investigación.....	40
	2.6 Aporte de la investigación.....	40
III.	MÉTODO.....	42
	3.1 Sujetos de investigación.....	42
	3.2 Instrumento de investigación.....	42
	3.3 Procedimiento de investigación.....	43
	3.4 Diseño y Metodología estadística.....	44
IV.	PRESENTACION Y ANALISIS DE RESULTADOS.....	45
V.	DISCUSIÓN DE RESULTADOS.....	60
	5.1 Conclusiones.....	64

5.2 Recomendaciones.....	65
VI. REFERENCIAS BIBLIOGRÁFICAS.....	66
VII. Anexos.....	70
7.1 Formato de entrevista.....	71

## I. Introducción

Actualmente en las instituciones educativas de nivel superior, se están utilizando una serie de instrumentos de opinión que muchas veces no están adecuadamente redactados o no están confiabilizados. Por tal razón no se conoce de forma válida si es constante en la medición de las variables deseadas, si los ítemes son claros y específicos, y si abarca y evalúa todos los puntos que se desean. Al mismo tiempo, se comete el error de aplicar las evaluaciones varias veces en un período corto de tiempo, lo que hace que la prueba pierda valor y consistencia. De ahí que sea necesario confiabilizar los instrumentos de evaluación educativa, para que los centros estudiantiles tengan una base real de conocimiento acerca del desempeño de sus catedráticos al impartir los cursos que le han sido asignados.

A la vez, al medir la confiabilidad al instrumento se observará si es necesario realizar algún tipo de cambio tanto en el contenido como en la redacción de los ítemes. Además se podrá sugerir concretamente cuáles son los ítemes que es necesario modificar y por qué razón, así como los temas que deben agregarse y cuáles deben eliminarse de la prueba ya preestablecida. Todo esto, con el fin que la prueba proporcione una adecuada retroalimentación del desempeño de los catedráticos a la institución, y asimismo sea de fácil comprensión para todos: alumnos, catedráticos y personal encargado de obtener los resultados de las mismas.

Aiken (1996) opina que cuando no ocurre ningún tipo de cambio en una persona; por el aprendizaje, enfermedad, accidente y en la etapa de crecimiento; y se da una variedad de resultados en las pruebas tanto en distancia como de tiempo en tiempo; es posible que la evaluación no sea confiable y no pueda usarse para explicar o hacer predicciones sobre el comportamiento de una persona.

Por otro lado, Anastasi y Urbina (1998) consideran que la confiabilidad de un test, señalan en qué porcentaje las diferencias personales en los resultados son consecuencia de

diferencias reales en las características estimadas y el nivel en que pueda ser producto de errores casuales o inesperados.

Asimismo es importante evaluar correctamente a los catedráticos universitarios, ya que se puede obtener una visión general de la percepción de los alumnos acerca de la forma en que planifican sus clases, su responsabilidad en todos los sentidos, los métodos que utilizan a la hora de enseñanza, formas que emplean para evaluar, el tipo de relación que tienen con sus alumnos y demás personas. También se puede obtener la opinión general de los alumnos con respecto al profesor, a la clase, y al curso. Esta evaluación se realiza por medio de una evaluación del desempeño, creada específicamente con el propósito de comprobar la manera en que los alumnos perciben como cumplen con sus actividades laborales.

También Werther y Davis (1997) exponen que la evaluación del desempeño establece imparcialmente la manera en que la persona ha cumplido con sus responsabilidades laborales. A la vez, se puede determinar que sujetos ejecutan o superan lo deseado y quienes no lo hacen. Por otro lado, Mondy y Noe (1997) discurren que la evaluación del desempeño es empleada con la intención de mejorar el trabajo de las personas. Al mismo tiempo conviene que sea profunda y permanente, más que un hecho que acontece en un período del año.

Por lo expuesto anteriormente la finalidad del presente trabajo es: confiabilizar por medio del análisis de ítems un instrumento de evaluación del desempeño de catedráticos universitarios.

En relación al tema del desempeño, Pinzón (1997) evaluó si la implementación de un programa de evaluación del desempeño en la pequeña y mediana empresa dentro de la industria del calzado, constituía una fortaleza en la administración del recurso humano. La muestra estuvo conformada por 23 pequeñas y medianas empresas ubicadas en el perímetro de la ciudad capital, de las cuales 8 eran pequeñas empresas y 15 medianas empresas con características similares, en cuanto al producto que fabricaban, el proceso de producción, y su localización. Recopiló la información por medio de un cuestionario, con 21 preguntas

cerradas y 2 preguntas abiertas, de estructura dicotómica y múltiple. Realizó una prueba piloto del cuestionario en dos empresas con características similares a la de la población, para observar si todas las preguntas eran explícitas y si cumplían con el objetivo deseado. Concertó citas previas con los gerentes o propietarios de las empresas, quienes después fueron las personas que contestaron el cuestionario. Procesó las respuestas agrupándolas por cada pregunta, graficó los resultados y calculó las desviaciones para cada pregunta. Desarrolló una investigación descriptiva. Concluyó que la poca importancia a la retroalimentación en las industrias del calzado, ha provocado que evaluaciones del desempeño anteriores no cumplan con el objetivo deseado. Recomendó efectuar evaluaciones del desempeño periódicas en las instituciones, a fin de poder tomar decisiones, de que exista retroalimentación, motivación, y que pueda ser una plataforma en la planeación de la administración de personal.

Del mismo modo, Bran (1999), realizó un plan de evaluación del desempeño para personal operativo, mandos medios y administrativos del Ingenio La Unión. Recopiló la información por medio de 3 cuestionarios estructurados en la modalidad de la escala de Likert, con preguntas dirigidas al rendimiento del personal, sus actitudes, satisfacción en el trabajo, estructura de la empresa, ubicación en el puesto de trabajo, sueldos y promociones. Uso una escala de calificación gráfica con la que evaluó a los empleados de las diferentes áreas. La muestra fue conformada por 456 trabajadores seleccionados de forma aleatoria; en donde 402 eran personal operativo, 25 personal del área administrativa, y 29 personal de mandos medios. Procedió a entrevistar a cada jefe de área, para determinar el organigrama, el número de empleados bajo su cargo, y sus principales funciones. Elaboró los formatos para la evaluación del desempeño, en base a las variables y sugerencias del Gerente Industrial del Ingenio. Luego dio forma al cuestionario que llenaron los trabajadores, a través de los indicadores de las variables enfocados en el área afectiva, cognoscitiva y psicomotora de los empleados. Determinó la población para poder aplicar los cuestionarios. Hizo la entrevista al azar, recopiló y ordenó la información de campo. Tabuló las respuestas de los cuestionarios agrupándolas según los indicadores. Generalizó las respuestas de los cuestionarios en excelente, bueno y regular; con el fin que tuvieran relación con las respuestas de las evaluaciones, las cuales eran: excelente, muy bueno,

bueno y regular. Comparó las entrevistas con las evaluaciones, agrupando los resultados de las evaluaciones en excelente y muy bueno. Fue usado el cálculo de porcentaje para visualizar y comparar los resultados de los cuestionarios y evaluaciones. Concluyó que era necesario que todo ingenio azucarero, aplique un proceso de evaluación del desempeño, ya que se puede establecer el desempeño del individuo en el cargo, potencial de desarrollo, calidad de trabajo, cualidades y actitudes del empleado. Recomendó que la evaluación del desempeño debe efectuarse por lo menos 2 veces al año, debido a que el personal ocupa distintos puestos en la empresa, en época de Zafra uno y en la Reparación otro.

En el trabajo descriptivo, García-Arroba (2000) elaboró un instrumento de evaluación del desempeño para medir cualitativa y cuantitativamente el trabajo de una secretaria contratada temporalmente. Con una población que estuvo formada por 28 sujetos, de los cuales 10 gerentes eran de recursos humanos y 17 jefes inmediatos, de las empresas a las cuales se les presta el servicio de personal temporal; como por el gerente de recursos humanos de la empresa que presta dicho servicio. Para recabar la información necesaria sobre los factores que califican el rendimiento laboral del empleado, elaboró un cuestionario que puede ser aplicado individual o colectivamente, conformado por 16 preguntas cerradas, el cual estaba dividido en dos partes que eran: la primera formada por 8 factores o características ideales que debe poseer una secretaria contratada temporalmente. Y en la segunda parte a los 8 factores anteriores se les debe colocar un porcentaje por orden de prioridad. Elaboró el cuestionario que luego aplicó a los gerentes de recursos humanos y jefes inmediatos, para recabar la información de los factores más importantes para calificar el rendimiento laboral de una secretaria. Tabuló los resultados del cuestionario, graficó los resultados, clasificó los factores más importantes según los resultados del cuestionario, ponderó los factores, las preguntas y las respuestas. Elaboró la propuesta del instrumento de evaluación del desempeño, realizó el manual del evaluador, diseñó un formato de entrevista de evaluación para dar a conocer los resultados, desarrollo un cuadro de resultados de la evaluación del desempeño y discutió y analizó los resultados. Tabuló los resultados del cuestionario por medio de la distribución de frecuencia relativa. Graficó las distribuciones de frecuencias a través de diagramas de barras y de pastel. Concluyó que la evaluación del desempeño esta orientada a detectar desviaciones en la planeación y

objetivos del trabajo, necesidades de capacitación y determinar el rendimiento del trabajador que labora en la empresa. Recomendó proporcionar la información necesaria a todos los empleados, sobre la evaluación del desempeño para evitar la resistencia a su implementación y lograr la cooperación de los mismos.

Estableció la validez y confiabilidad de la prueba de Conocimientos Generales para Oficinista I (CGO-1), Quintana (1982). Uso como muestra a alumnas que cursaron el cuarto año de la carrera de secretariado y oficinista en Institutos Nacionales, de las jornadas vespertinas, en la ciudad capital, comprendidas entre los 15 y 22 años de edad, con un nivel socioeconómico medio y medio bajo. El instrumento CGO-1, esta compuesto por 100 ítems de selección múltiple, dividido en 5 áreas, elaborado para medir conocimientos generales de oficina en personas con estudios de cuarto año de la carrera de secretariado y oficinista. En la USIPE se investigo el número y nombre de los institutos nacionales que impartían la carrera de secretariado y oficinista en la ciudad capital. Se le solicitó la autorización para aplicar la prueba CGO-1 a los directores de educación media, después se procedió a aplicar la prueba CGO-1 en base a 4 materias de la carrera, la secretaría de cada establecimiento proporcionó las calificaciones de las 4 materias que son: Redacción y correspondencia I, ortografía, matemáticas comerciales y organización, y práctica de oficina. Ordeno los datos, conforme a grado e institución; se calificaron las hojas de respuesta, por ultimo se elaboraron listados con los resultados obtenidos por las alumnas en Prueba CGO-1 y el promedio correspondiente de las cuatro asignaturas escogidas. El proceso estadístico aplicado fue la ecuación de regresión; la medida de tendencia central, el error standard de medición y la validez, empleando el Coeficiente  $r$  de Person. Confiabilidad, y el coeficiente de Equivalencia Racional fórmula Kuder Richardson 21. Uso el procedimiento de Thurstone para obtener rasgos percentiles. Concluyó que la prueba CGO-1 es confiable por la fórmula Kuder Richardson 21 ( $KR_{21}$ ), para el grupo de personas que conformaron la muestra. A la vez recomendó que se verificará la validez y confiabilidad de la prueba CGO-1 en estudiantes del interior de la república de Guatemala.

Además Carmona (1982) determinó la validez y confiabilidad del inventario Home para niños en Guatemala. La muestra fue constituida por: 80 niños de un año, seleccionados al

azar, que vivían en la colonia el Amparo, de estos 80 niños se eligieron 40 que llenaban el requerimiento de que el padre viviera en el hogar. Empleo dos instrumentos: El inventario Home para niños de 0 a 3 años edad, compuesto por 45 ítems, donde el punteo total esta formado por la suma de los punteos de cada sub-escala siguiente: 1. Respuesta verbal y emocional de la madre (11 ítems), 2. Ausencia de restricción y castigo (3 ítems), 3. Organización del ambiente físico y temporal del hogar (6 ítems), 4. Materiales de juego apropiados (9 ítems), 5. Relación madre-hijo (6 ítems), y 6. Oportunidades para variar la estimulación diaria (5 ítems). La segunda prueba que uso fue: el Cuestionario Socioeconómico desarrollado por las personas que hicieron el estudio “Correlaciones socioculturales en el desarrollo cognoscitivo y físico de niños pre-escolares del área urbana guatemalteca”. Este cuestionario acumula la información siguiente: edad, educación formal y ocupación del padre y la madre. Tiempo de vivir en la capital, estado civil, número de hijos y de personas que viven en el hogar, y la escolaridad o asistencia de los niños a la escuela o guardería. Procedió a pedirles autorización a las madres de los niños para la realización del estudio. Apliqué el cuestionario socioeconómico a las madres, el mismo día que se les pidió la autorización. Concerté citas para la posterior aplicación del Inventario Home, el cual necesitaba la presencia simultanea de la madre y el niño. Por ultimo realizo el análisis estadístico. La metodología estadística que empleo para la validez del Home fue,  $r$  en donde  $r$  es el coeficiente de correlación del momento de producto de Pearson. Para la confiabilidad del Home concluyó que el punteo global del Home esta conectado con el nivel de educación formal logrado por el padre (.54) y el número de personas que viven en el hogar. Recomendó que se ampliara el número y variedad de sujetos para comprobar la validez del Inventario Home para niños de cero a tres años.

Calvo (1984) determinó la confiabilidad y validez concurrente del test de uso de Lenguaje (T.A.D.). La población estuvo conformada por: 50 alumnos del Instituto Central de Varones y 50 alumnas del Instituto de Señoritas Centro América de la ciudad de Guatemala, comprendidos entre las edades de 15-5 y 13-6 años. Utilizo el test de Aptitud Diferencial (T.A.D.) uso de Lenguaje, forma A y B, ortografía y oraciones de Bennet y Seashore y Wesman. Estos test fueron traducidos y adaptados al idioma español, ya que formaban parte de un grupo de ocho baterías, de las cuales solo seis estaban traducidos y

adaptados por el Colegio Americano de Guatemala. Para lograr una adaptación satisfactoria de los test, realizo dos pruebas similares con el propósito de sustituir los ítemes que no cumplieran con el nivel de dificultad y de discriminación. Procedió a adaptar las formas A y B del test de Uso de Lenguaje, reviso las pruebas equivalentes de los tests originales que fue aplicada a especialistas en lingüística, selecciono los institutos donde aplico los test mediante una lista proporcionada por el programa de orientación y becas para estudios superiores (POES). Solicito la autorización para aplicar los test en los institutos seleccionados, aplico los test y analizo e interpreto los datos obtenidos. La metodología estadística que utilizo fue: la fórmula 20 de Kuder-Richardson. Concluyó que los test de ortografía son confiables según su coeficiente de confiabilidad, que es 0.5272. Recomendó realizar más estudios donde se pueda establecer la validez predictiva y de constructo de cada una de las pruebas.

De la misma forma, Pinzón (1999) desarrolló una investigación cuasi-experimental, donde evaluó el desempeño profesional de las docentes con especialidad y las docentes no especializadas en el área secretarial, de los Institutos Nacionales de la Ciudad Capital. Diseño 3 instrumentos conformados cada uno por 4 ítemes por cada subcategoría, lo que hace un total de 20 ítemes, que fueron aplicados a los directores, a las docentes, y a las alumnas. Las personas sujetas al estudio fueron: directores de los institutos nacionales del sector público de formación secretarial de la ciudad capital de las jornadas matutina y vespertina. Docentes sexo femenino, especializadas en ciencias comerciales, que imparten cursos en el área secretarial. Docentes sexo femenino, no especializadas en ciencias comerciales, que imparten cursos en el área secretarial, y poseen otras especialidades. Alumnas de las carreras de secretariado bilingüe y oficinistas. El cálculo estadístico se realizó separando los datos en dos grupos, que fueron: las docentes especializadas en ciencias comerciales y las docentes no especializadas; posteriormente construyó las gráficas de Perfiles de las docentes evaluadas, por medio del Método Estadístico del Coeficiente Serafín de Congruencia Simple. Primero validó la prueba piloto, aplicándola a una muestra de 10 alumnas, del Colegio el Rosario. Aplicó los cuestionarios a las personas de la muestra, se les asignó un promedio de punteo a cada prueba, realizó el cálculo estadístico en base a los punteos obtenidos, realizó el análisis de resultados, y desarrollo las gráficas de

perfiles de las docentes evaluadas. Concluyó que se logró comprobar el desempeño profesional de las docentes especializadas y las no especializadas en el área secretarial en un 99%. Recomendó la utilización del modelo de evaluación que uso para evaluar a las docentes especializadas y no especializadas en el área secretarial, de los institutos nacionales de formación secretarial de la ciudad capital.

Méndez (2000) propuso una evaluación del desempeño para el recurso humano de escuelas urbanas en el municipio de Escuintla. Realizó un estudio descriptivo, con una muestra de 21 directores de las escuelas oficiales primarias; donde 19 eran de escuelas mixtas, 1 era de una escuela para niñas y 1 era de una escuela para varones. El cuestionario fue validado en base al juicio de expertos, y constaba de 12 preguntas: 10 preguntas de selección múltiple y 2 preguntas cerradas. Se ordenaron e interpretaron los datos mediante la investigación de conceptos, opiniones y puntos de vista de directores escolares, y por medio de la observación de las prácticas existentes. Se solicitó por escrito la autorización a la Dirección Departamental para poder visitar las escuelas, luego fueron elaborados los cuestionarios, en base a los datos recabados se tabuló y se analizó los resultados obtenidos por medio de gráficas. Concluyó que la Supervisión de nivel primario del municipio de Escuintla, no cuenta con un programa de evaluación del desempeño que pueda evaluar objetivamente el desempeño de los directores. Recomendó que el programa, procedimiento y la periodicidad con que se realicen las evaluaciones del desempeño se comuniquen a los directores, así como los resultados de las mismas.

En el estudio descriptivo hecho por Mijangos (1998) evaluó el desempeño de los directores del nivel preprimario y primario que participaron en la escuela de Directores. En donde se uso para el trabajo estadístico, las tablas que presentan los resultados obtenidos, en forma de punteos absolutos y su conversión en punteos relativos, de acuerdo a los distintos sujetos de la muestra. Con una población que fue formada por: 90 directores participantes y 26 supervisores y capacitadores, para un total de 175 docentes de escuelas del departamento de Chimaltenango, escogidos de una forma no aleatoria, pero si en forma voluntaria. Para recopilar la información construyó una encuesta de opinión, basada en la escala del método de Likert, la cual fue validada antes en otro departamento con sujetos que tenían las mismas

características que las personas de la muestra. Procedió a construir los instrumentos que fueron aplicados para la investigación de campo, validó la encuesta en otro departamento, aplicó la encuesta a los sujetos que intervinieron en la investigación en el departamento de Chimaltenango, tabuló la información recolectada y construyó las tablas estadísticas necesarias. Concluyó que la evaluación del desempeño, debe realizarse en todos los programas o proyectos instituidos por todas las direcciones Departamentales de Educación o por el Ministerio de Educación. Recomendó evaluar el desempeño de los directores, con cierta frecuencia para establecer planes de retroalimentación en las áreas administrativas que presente debilidad.

Con el objeto de determinar la confiabilidad del instrumento utilizando el criterio de los alumnos al evaluar el desempeño de los docentes, Verastegui (1996) realizó un estudio con una muestra conformada por: 30 estudiantes de una sección del segundo ciclo de la carrera de Licenciatura en Ciencias Jurídicas y Sociales de la Universidad Rafael Landívar. El instrumento que se utilizó fue el Cuestionario de Evaluación de Catedráticos de la Universidad Rafael Landívar, dividido en dos partes; la primera compuesta por 20 preguntas, que agrupan 5 subescalas o bloques que son: Preparación y Responsabilidad, Metodología de Enseñanza, Evaluación, Relación Interpersonal e Investigación. La segunda consta de 5 ítems y se refiere a un sondeo del alumno con respecto a la clase, al profesor y al curso. Solicitó el permiso pertinente a las autoridades para trabajar con los resultados de la evaluación de catedráticos, identificó cada prueba numerándola previamente a la evaluación, posteriormente tomó las 150 hojas de respuesta dando a las respuestas una valoración de 1 a 4, y tabuló 3900 datos, con los cuales desarrolló cuadros de Análisis de Varianza para los primeros 20 ítems, los cinco bloques que los agrupan y el punteo global. El procedimiento estadístico que utilizó fue: el análisis de varianza para muestras relacionadas en planteamientos análogos. La investigación fue de tipo ex-post factum, ya que no se manipularon las variables, la asignación de los alumnos y los catedráticos. Concluyó que el punteo global fue más confiable que los ítems que lo componen y los bloques también fueron más confiables que los ítems que los conforman. Sugirió tomar en cuenta que tanto los ítems como los punteos globales aportan

información distinta; muy específica en el caso de los ítemes y muy genérica en el caso de los bloques y del punteo global.

Diferentes autores han escrito sobre confiabilidad, tema de este estudio, por tal razón se muestra a continuación el trabajo de varios autores de diversas procedencias.

Manifiesta Aiken (1996) que confiabilidad es el estado en el cual una herramienta de aproximación psicológica calcula algo en forma estable. Un documento confiable, está relativamente exento de fallas de comprobación, de modo que las notas que consiguen los individuos en la prueba son próximas en valoración numérica a sus calificaciones verídicas. La falta de confiabilidad es la consecuencia de las incorrecciones de medida ocasionados por situaciones internas transitorios, como disminución en la estimulación o malestar, o por cosas externas, como un medio perturbador o con distracciones para un test.

Al mismo tiempo Anastasi y Urbina (1998) afirman que confiabilidad atañe a la estabilidad en las calificaciones obtenidas por un grupo de personas evaluadas por la misma prueba varias veces, con grupos semejantes de elementos o en otro ambiente de evaluación.

Asimismo en otros términos la confiabilidad o fiabilidad de una prueba permite establecer que parte de la varianza de las puntuaciones se deben a la varianza de error. La varianza de error esta constituida por cualquier aspecto irrelevante para la finalidad de la prueba, es decir, son todas aquellas variabilidades casuales que se dan a la hora de realizar un test.

Los evaluadores tratan de disminuir la varianza de error, cuidando de mantener parecido las condiciones de utilización por medio del control del entorno, las instrucciones, los límites del tiempo, el rapport y otros elementos semejantes.

Por otro lado Nunnally (1991) sostiene que una escala es confiable en la medida que pueda repetirse y en la forma en que cualquier circunstancia casual pueda producir medidas diferentes.

Al mismo tiempo, Guilford (1936) estima que confiabilidad es la obtención de la misma calificación cuando se repite la prueba, en un ambiente similar o igual que en la primera vez que se paso la prueba. Hay tres formas en las que se puede lograr obtener la correlación de un test, que son: repetir el test en distintas ocasiones, utilizar dos formas de tests similares, dividir la prueba en dos partes iguales.

En la primera técnica, el intervalo de tiempo ayuda a evaluar si el paso del tiempo, el crecimiento de los individuos, el aprendizaje que han tenido en ese tiempo y los acontecimientos inesperados que suceden; pueden cambiar los resultados de la prueba. En la segunda, se puede detectar con más profundidad, si las características personales y el desarrollo interfieren o modifican los resultados. A la vez, se evita que los individuos utilicen su memoria para responder la prueba, ya que se utiliza una prueba parecida a la primera que se aplicó.

Hay diversos factores que influyen en la confiabilidad o fiabilidad de una prueba que son: la mayor cantidad de ítems en la prueba, mientras mayor sea el número de ítems de la prueba mayor será su confiabilidad, el test debe tener el mismo nivel de dificultad para todos los individuos que lo van a contestar, los ítems que tienen dependientes uno de los otros suelen reducir el nivel de confiabilidad, a más objetividad en las calificaciones mayor será la confiabilidad. A la vez, el cambio en el ambiente a la hora de pasar la prueba, puede ocasionar que se disminuya el nivel de confiabilidad de la prueba. La semejanza en el material, y que los individuos tengan experiencias similares o parecidas, ayuda a la fiabilidad del test.

Existen varios factores que ayudan a la mala interpretación de los ítems, y a la poca confiabilidad de una prueba, que son: usar palabras con tintes emocionales, la elección de las palabras, la mala construcción de las oraciones, la inapropiada redacción de las instrucciones, y darle un sentido erróneo a los ítems.

Indica Anastasi (1982) que todas las técnicas de confiabilización describen un nivel de proporción o estabilidad entre un grupo de calificaciones provenientes libremente, que se pueden expresar en aplicación de un coeficiente de correlación.

El coeficiente de correlación ( $r$ ) expone el nivel de afinidad, entre dos grupos de calificaciones. Se puede decir que, si un sujeto obtiene la mayor puntuación en la variable 1 y también obtiene la mejor calificación en la variable 2, asimismo el que obtiene la segunda mejor calificación en la variable 1 obtendrá el segundo puesto en la variable 2, y así alternativamente se va reduciendo la puntuación de los individuos hasta llegar al último lugar; entonces habrá una correlación perfecta entre la variable 1 y 2 y sería igual a +1.00.

La correlación positiva presenta que cada individuo ocupa la misma posición en ambas variables. La correlación negativa exhibe que cuando un sujeto obtiene la mejor calificación en la variable 1, obtiene la peor puntuación en la variable 2. Una correlación cero revela una separación de relación. En el caso de que se seleccionara al azar las posiciones de las personas en la variable 1 y de igual forma se hiciera en la variable 2, no se conocería que persona obtuvo la mayor puntuación en la variable 1 ni en la dos, y por ende no se sabría quien obtuvo la menor puntuación. También puede darse el caso que la persona que obtuvo mejor puntuación en una variable, pudo haber sacado la peor en la otra, o su puntuación puede estar en el promedio.

El modo más frecuente que se emplea para estimar la correlación es el coeficiente de correlación producto-momento de Pearson. Esto toma en consideración tanto el lugar del individuo en el grupo, como los resultados que tienen por encima por debajo de la media. Las puntuaciones serán altas y positivas cuando las notas sean de igual signo y casi iguales en ambas variables. En el caso que una persona tenga una puntuación alta en una variable y baja en otra los productos cruzados serán negativos.

Consideran Anastasi y Urbina (1998) que el nivel de significancia estadística de una prueba, es el nivel de error que se puede cometer a la hora de obtener las conclusiones de los datos. Existen distintas forma o técnicas para determinar la confiabilidad de una prueba

entre las cuales están: confiabilidad test-retest, de formas alternas, de división por mitades, de Kuder-Richardson y coeficiente alfa, entre calificadores. La confiabilidad test-retest aplica el mismo test dos veces. Se pretende confiabilizar los resultados de la misma persona en las dos veces que se empleo la prueba. La varianza de error que se mide son los cambios en puntajes a lo largo del tiempo que pasa entre la primera y la segunda aplicación de la prueba, en este período pueden presentarse acontecimientos casuales que no pueda controlar el examinador. Por tal motivo, el tiempo que pasa de la primera a la segunda aplicación debe ser corto, no debe superar a los seis meses.

Este tipo de confiabilidad presenta ciertas dificultades que son: los examinados pueden recordar las respuestas de algunas preguntas, esto puede darse si es muy corto el tiempo entre la primera y la segunda aplicación de la prueba.

En la confiabilidad de forma alterna o coeficiente de equivalencia, los sujetos de un grupo pueden ser evaluados por un tipo de prueba primero y luego con otra. Es decir que, el conjunto total de sujetos se divide en dos, en la primera aplicación del test, un grupo es examinado con la prueba A y el otro con la prueba B. En la segunda aplicación se invierten los test, es decir, el grupo que contesto la prueba A ahora responde la B, y el grupo realizo la test B ahora hace el test A. La correlación entre las calificaciones de ambas pruebas, es lo que da el coeficiente de confiabilidad, ya que no solo muestra la estabilidad transitoria, sino que el equilibrio de las respuestas ante otros tipo de pregunta.

En la confiabilidad de división por mitades, Downie y Heath (1973) exponen que es necesario aplicar únicamente una prueba una sola vez, donde se divide con el fin de que se obtenga dos puntuaciones por persona. La división se suele hacer de diferentes formas, que son: en ocasiones se cuentan las respuestas correctas con numeración par e impar. Otras veces se establece que la pregunta 1 y 2 entra en la primera puntuación, la pregunta 2 y 3 en la segunda, y así sucesivamente hasta que se termine con la división.

Este tipo de confiabilidad se conoce como coeficiente de consistencia interna. Muchas veces se considera que mientras más larga una prueba más confiable.

La confiabilidad de Kuder-Richardson, Anastasi y Urbina (1998) indican que sólo se aplica una vez de una única prueba, esta se basa en la consistencia de las calificaciones de la prueba. Esta consistencia entre preguntas esta influida por dos tipos de varianza de error que son: el contenido y la desigualdad del contenido. Mientras más uniforme sea el contenido mayor será la consistencia entre cuestiones.

En el coeficiente alfa, Aiken (1996) opina que la fórmula que se utiliza es:  $\alpha = (k/k-1)(1 - \sum (s_i^2) / s^2)$ . En donde  $k$  es la cantidad de reactivos,  $s_i$  la varianza de calificaciones en el  $i$  y  $s^2$  la varianza de las puntuaciones totales de la prueba. El coeficiente alfa es una fórmula que se utiliza para estimar la confiabilidad de un test que esta formado por reactivos donde  $k$  les dan dos o más valores de puntuación a las respuestas.

Por otro lado, la confiabilidad entre calificadores según Anastasi (1982) se refiere a las diferencias casuales entre los elementos equivalentes y envuelve cualquier tipo de inestabilidad entre los elementos. Existen dos puntos que son eliminados de la varianza de error que son: los factores en donde la varianza debe estar en la calificación, y los factores sin importancia que pueden ser controlados experimentalmente. La confiabilidad del calificador, se obtiene por medio de un grupo de pruebas examinadas por dos o mas evaluadores.

De la misma forma, Aiken (1996) reconoce que la puntuación lograda por un individuo en un test se forma por una nota real más alguna falla no constante de la medida. La puntuación real, es el promedio de las calificaciones que pudiera obtener una persona al hacer infinidad de veces una prueba, esta debe ser calculada a partir de la calificación obtenida por el sujeto en la prueba.

La confiabilidad del instrumento se enumera con cifras decimales positivas que pueden ir desde .00 hasta 1.00;  $r=1.00$  expresa la confiabilidad perfecta y  $r=.00$  indica la falta de confiabilidad. La confiabilidad se obtiene al analizar lo sucedido en la aplicación de la prueba y el contenido de la prueba en sus puntuaciones.

Así mismo Guilford, reconoce diversas técnicas para darle un nivel de confiabilidad a los test, entre los cuales están: El índice de confiabilidad, el error estándar de la medida, confiabilidad en distintos rangos de habilidad.

El índice de confiabilidad, estima el grado de correlación que existe entre la puntuación obtenida por los individuos, y la puntuación del test propiamente dicho. La fórmula que se utiliza es  $r_{\infty 1} = r$ . El coeficiente de confiabilidad es el coeficiente de determinación de  $X_1$  por  $X_{\infty}$ . Es decir que, que el punto de unión entre los resultados de un test, que se paso en un tiempo, y otro test similar o parecido, se le llama  $X_{\infty}$ .

El error estándar de la medida, aquí se utiliza el coeficiente de correlación para indicar que tan cerca están las predicciones. La fórmula utilizada para predecir  $X_1$  frente a  $X_{\infty}$  es:  $\sigma_{1\infty} = \sigma_1 \sqrt{1-r}$ .

Confiabilidad en distintos rangos de habilidad, demuestra en que punto o nivel puede ser más alto el nivel de correlación, entre los dos grupos de individuos a quienes se les aplique la prueba.

El análisis de ítems según Montero (2000), pretende estimar de una forma teórica el comportamiento de los datos que son resultado de la aplicación de un test o instrumento. Esta técnica permite juzgar la excelencia de cada uno de los ítems aisladamente. Por otro lado Hintze y Utah (1998) distinguen que el análisis de ítems investiga la confiabilidad interna de los instrumentos.

Este tipo de análisis, debe realizarse con la información obtenida de un grupo con 200 personas como mínimo. Cada respuesta a los reactivos o preguntas deben ser organizadas colocándoles una clave que puede ser de: otorgándole 0 a los ítems con menor interés y 1 a los ítems que sean de mayor interés. Al mismo tiempo, este tipo de análisis debe cumplir con dos supuestos básicos, que son: es el de unidimensionalidad, en donde se indica que el instrumento debe medir un solo reactivo o pregunta. Y el segundo que se denomina independencia local, donde se determina que la posibilidad de dar una respuesta

exacta a un ítem no se ve afectado por la forma o manera en que sean respondidos los demás.

Conceptos fundamentales relacionados con el análisis de ítemes son: la curva característica de ítem o CCI, el error típico de media o error estándar de medición y la función de información. En la CCI se acomodan una función matemática la conducta del ítem, partiendo de las respuestas dadas al instrumento por parte de los individuos que lo hacen. Cada punto de CCI, representa la posibilidad de que cada persona conteste adecuadamente cada una de las preguntas o reactivos.

El error estándar de medición, representa la inestabilidad en la apreciación del reactivo de interés por cada individuo. El valor se determina por medio del valor que se le otorga a la variable que se desea medir en cada individuo o al reactivo de interés. El error de medición al cuadrado se nombra función de información FI. La FI determina la exactitud que posee el ítem para valorar cada uno de los distintos valores de la variable que se desea medir en los individuos. La Curva Característica de Ítem (CCI) y la Función de Información (FI) admiten examinar las cualidades psicométricas de los ítemes y a la vez, poder escoger los ítemes que cumplan con el propósito del test o instrumento.

Montero (2000) diferencia varios modelos para el análisis de ítemes o la teoría de respuesta a los ítemes, que son: la función logística en uno, dos y tres parámetros. El modelo más usado es el de tres parámetros. La CCI se utiliza para especificar la posibilidad en darle una respuesta exacta a un ítem, por parte de los evaluados.

Confiabilizar un instrumento de evaluación del desempeño, es importante ya que se puede tener la completa seguridad, que cuando se va a utilizar el instrumento de nuevo los individuos van obtener una calificación similar a la primera vez que se utilizó. Al mismo tiempo, se podrá obtener una consistencia en los factores o aspectos que se desean medir. Todo esto es importante, ya que va hacer que las evaluaciones del desempeño sean más confiables y se obtengan resultados más objetivos.

En la actualidad se considera, que todo instrumento de evaluación del desempeño es útil para la obtención de información necesaria sobre el trabajo de un grupo de personas, para poder determinar que se necesita para alcanzar los objetivos deseados. Por tal razón, ya que lo que se hizo fue determinar el grado de confiabilidad del instrumento de evaluación docente universitario, se citaron algunos datos desarrollados por distintos autores.

La evaluación del desempeño, consideran Werther y Davis (1997) que determina imparcialmente la manera en que la persona ha realizado sus obligaciones laborales. Por otro lado, Chiavenato (1993) expone que la evaluación del desempeño sirve para valorar el eventual desarrollo del individuo en el puesto. Esta califica el valor, la excelencia y las capacidades de las personas.

Entre las ventajas de realizar una evaluación del desempeño están: la mejora del desempeño por medio de la retroalimentación, se establece quiénes son los que merecen un aumento de sueldos, se determina a quién se debe promover, transferir o despedir. Al mismo tiempo se puede identificar en qué aspectos se necesita desarrollo o capacitación. En base a la retroalimentación se puede planificar la carrera profesional, se tiene a la mano la información necesaria para establecer a quien se contrata, capacita o si hay necesidad de asesoría. También ayuda a identificar los errores en el diseño del puesto y se puede identificar que factores externos influyen en el desempeño.

Establece Arias (1979) que los requisitos para la evaluación del desempeño es: objetividad, validez y confiabilidad. Se pretende que a la hora de determinar los criterios a evaluar, estos sean seleccionados imparcialmente; deben expresarse concretamente la efectividad o ineffectividad de la persona evaluada, y los resultados deben ser parecidos siempre que se evalúe, siempre y cuando no cambie el trabajo de la persona.

Subraya Chiavenato (1993) que el objetivo general de la evaluación del desempeño, es mejorar los resultados de los individuos. Para lograr esto se debe dar cumplimiento a los objetivos fundamentales de la evaluación del desempeño que son: poder establecer la capacidad del individuo de desarrollarse; tratar a los individuos como parte de la institución

y que pueden desarrollarse dentro de la institución, y permitir que las personas evolucionen dentro de la corporación.

Al desarrollar una evaluación adecuadamente, se le debe dar a conocer a la persona si esta desarrollando adecuadamente sus labores, darle a entender lo que es un trabajo completo y que se desea de él, poder determinar cuales son las personas capacitadas para futuras promociones y poder establecer a quienes se les puede recompensar por su desempeño; (Sandel 1977).

Por otro lado, Chiavenato (1993) discurre que la evaluación del desempeño proporciona beneficios tanto a los jefes, subordinados y a la institución. Los beneficios para el jefe son los siguientes: valorar el trabajo y la conducta de los subalternos, plantear medidas para mejorar la actitud de los subalternos, y comunicar mejor el desempeño y la mecánica de la evaluación del desempeño, a sus empleados.

Los beneficios para los empleados son: saber qué es lo que valora más la institución, advertir las perspectivas que tienen sus jefes sobre su trabajo, saber qué van a realizar sus jefes para mejorar el desempeño de los empleados en la institución, y criticar ellos mismos cómo están haciendo su trabajo, para mejorar su autocontrol y auto desarrollo.

Los beneficios para la empresa son: examinar la capacidad humana a corto, mediano y largo plazo, establecer qué empleados necesitan capacitación en ciertas áreas, y a la vez determinar que tipo de estímulos necesitan los individuos.

Existen diversos métodos para la evaluación del desempeño, según Mondy y Noe (1997) entre los que están: la escala de calificación, incidentes críticos, ensayo, normas de trabajo, jerarquización, distribución forzada, informes de desempeño con lista de verificación ponderada y elección forzada, las escalas de calificación con ancla conductual, y la administración por objetivos.

La escala de calificación, se divide en cinco a siete categorías. Los aspectos que se suelen valorar son dos; lo relacionado con el cargo y las particularidades individuales. El

examinador llena la prueba especificando el factor que describe mejor al empleado y su trabajo. Entre sus desventajas es que se eliminan puntos cruciales de la prueba, para poderlo aplicar en otras áreas de la institución. A la vez, el empleado no recibe la adecuada retroalimentación, ya que es una prueba muy general.

En la técnica del incidente crítico, se debe mantener un registro de las acciones positivas o negativas en el trabajo. Estos datos se usan con otros para valorar el trabajo del individuo. Estima Chiavenato (1993), que en el método de incidentes críticos, el supervisor debe contemplar y apuntar las conductas extremas del individuo, tanto positivas como negativas, en el momento que realiza su trabajo.

El método de ensayo, indica que el examinador narre el desempeño del sujeto. Esto hace que se preste mayor atención en la actuación extrema en el trabajo del individuo, y no se le da la suficiente atención al desempeño diario de la persona.

Establecen Mondy y Noe (1997) que en el método de estándares de trabajo o normas de trabajo, se coteja el trabajo con un nivel de rendimiento deseado. Se utilizan distintas formas para establecer las normas de trabajo que son: estudios de tiempo y muestreo del trabajo. Para lograr que los empleados comprendan las normas de trabajo, se les debe explicar con claridad como y con que fin fueron instituidas.

En el método de jerarquización, el evaluador coloca los individuos en orden graduado de desempeño común. A la vez en la comparación por pares, el desempeño de cada persona se coteja con el desempeño de los demás individuos del área.

Por otro lado en el método de distribución forzada, se solicita al evaluador que le dé un número corto de rangos a los sujetos, de forma similar a una distribución normal de frecuencias. Aquí se considera que todos los empleados están en igualdad de circunstancias en la escala de excelente, promedio y pobres.

Igualmente Chiavenato (1993), considera que las ventajas del método de elección forzada son: que se elimina toda conclusión subjetiva por parte del evaluador, su aplicación es sencilla y no se necesita una capacitación extensa para el examinador. Al mismo tiempo este método tiene una serie de desventajas que son: la preparación e implementación son difíciles y necesita de un planeamiento muy detallado. Es una forma de evaluar muy comparativa y discriminatoria, y no brinda la suficiente información sobre el trabajo de los individuos.

El informe de desempeño con elección forzada, se necesita que el evaluador designe una enunciado que detalle el desempeño de un individuo. Al emplear este método se pueden encontrar una serie de dificultades, como: varios de los enunciados pueden tener un porcentaje de igualdad, y riesgo de que el evaluador emplee su opinión a la hora de examinar al individuo,(Mondy y Noe 1997), y (Werther y Davis 1997). Por otro lado, Chiavenato (1993) observa que en el método de frases descriptivas, el evaluador no se ve forzado a escoger un enunciado, sino que sólo se limita a marcar la oración que describe mejor el desempeño del individuo con una S o +, y marcar con una N o – cuando el trabajo de la persona sea diferente a lo que refiere el enunciado.

El informe de desempeño de lista de verificación ponderada, el examinador llena una hoja donde se encuentran una serie de enunciados que describen el desempeño de la persona, en donde cada frase tiene distinta puntuación para cada respuesta. El evaluador debe contestar de forma positiva o negativa cada expresión, (Mondy y Noe, 1997).

En la escala de calificación con ancla conductual, se trata de detallar con imparcialidad la conducta que debe tener el individuo en el puesto, como resultado de acontecimientos inesperados. En seguida se diferencian las actividades que debe tener un puesto en específico, después se detallan los elementos concretos que debe tener cada una de las actividades. Este método se utilizan frases que identifiquen de forma positiva o negativa, la conducta que se esta examinando.

Entre sus ventajas está: que se facilita la discusión de resultados, ya que se juzgan actuaciones establecidas del individuo. Una desventaja es que, las actitudes que se utilizan para determinar específicamente un comportamiento laboral, están guiadas a la acción en lugar de a los resultados.

También es preciso observar, que es necesario estipular claramente cuales son las actitudes laborales positivas o negativas de un individuo. Es decir, que lo que se determino que es positivo, sea positivo para todas las personas, y lo negativo sea de igual forma negativo para todos. Dichos términos deben ser claros, específicos y fáciles de entender o comprender para todos los evaluadores.

La administración por objetivos, examina y usa la colaboración de las personas de la organización. En este método la evaluación salta de las cualidades particulares al trabajo en el puesto. El evaluador pasa de ser juez a consejero, y el empleado pasa de ser simple espectador a participar en el proceso de evaluación.

La persona junto con su jefe establecen las metas a alcanzar. Al finalizar dicho proceso, tanto el supervisor como el empleado se reúnen para realizar una entrevista de evaluación. En esta entrevista se revisa cuanto se ha logrado alcanzar las metas, y cuales son las posibles vías a emplear para solucionar las dificultades sobrantes.

Una de las ventajas es que, durante el período de evaluación se mantiene una comunicación constante entre el examinador y el individuo. Se establecen mutuamente las metas a lograr, que vías deben o pueden tomarse para alcanzar las metas, y cuáles son las posibles actitudes que deben escogerse para solucionar los contratiempos que todavía quedan en el camino.

Por otro lado, Chiavenato (1993) identifica otros métodos de evaluación del desempeño, que son: Método de escala gráfica, investigación de campo, incidentes críticos, comparación por pares, de autoevaluación, y evaluación por resultados.

El método de escala gráfica, evalúa el desempeño de los individuos mediante elementos anticipadamente determinados y reconocidos. Cada elemento se limita con una explicación

breve, simple y justa, para impedir desviaciones. Se estipula el desempeño desde deficiente al más desfavorable y de lo excelente hasta el más beneficioso.

Esta técnica puede efectuarse por medio de diversos procedimientos de puntuación, entre los cuales están: escalas gráficas continuas, semicontinuas y discontinuas.

Esta técnica tiene diferentes ventajas, que son: es de fácil aplicación y entendimiento, permite observar de forma clara las actitudes laborales más importantes para la organización y como actúa cada individuo ante las mismas. También presenta una serie de desventajas, entre las cuales están: el examinador debe regirse exclusivamente al test y no a las peculiaridades del examinado, el evaluador puede interpretar el comportamiento de la persona de forma subjetiva, y se tiende a popularizar los resultados de los tests.

Observa Chiavenato (1993), que el método de investigación de campo, permite evaluar el desempeño de los empleados como las causas de la misma, a su vez, planear con los supervisores de que modo se puede desarrollar la evaluación. La persona que realiza la evaluación es el jefe, pero con ayuda de un experto en el área de evaluación del desempeño. Este experto sostiene una serie de entrevistas con los supervisores de cada área, preguntándoles cuál ha sido el desempeño de cada una de las personas bajo su mando.

Las ventajas al usar esta técnica son: permite al supervisor tener una visión amplia de las funciones de los puestos bajo su cargo, de las habilidades, capacidades y los conocimientos necesarios. Se evalúa a cada individuo de una forma extensa, equitativa y justa.

Pero a su vez, tiene una serie de desventajas que son: se retrasa el proceso debido a las entrevistas que se realiza con cada uno de los jefes de las áreas, y el costo es alto por la intervención de especialista en el proceso de evaluación del desempeño.

Sugieren Werther y Davis (1997), por otro lado, que el jefe debe llevar un libro, donde debe escribir las conductas más relevantes del empleado, tanto en el aspecto positivo como negativo.

Una de las ventajas más importantes, es que permite darle una retroalimentación clara y concreta al empleado de su desempeño, y evita las posibles desviaciones de la memoria de los últimos eventos.

En este método sólo se interesan por los hechos relevantes, tanto positivos como negativos, en el desempeño del individuo, pero no se interesan por otro tipo de acontecimientos, que aunque sean difíciles de detectar o sean considerados de poca importancia, afectan el trabajo que desarrollan las personas todos los días.

Subraya Chiavenato (1993), que en el método de comparación por pares, el examinador valoriza el trabajo de los individuos en pares. Es decir, se compara el trabajo hecho por el empleado A y el C, en las columnas siguientes se marca con una X el empleado que ha desarrollado mejor su labor.

En el método de autoevaluación, se le solicita a la persona que evalúe el mismo su trabajo. Esto permite que el individuo, se desarrolle individualmente, y evita en gran medida que se presenten actitudes defensivas ante la evaluación.

En el método de evaluación por resultados, se confrontan los resultados esperados, con los resultados logrados por cada individuo. Esta comparación que se realiza, permite determinar cuáles son las actividades que deben tomarse o seguirse, para lograr mejorar el desempeño de los empleados y alcanzar con más facilidad y éxito los objetivos planeados y deseados.

De igual forma, Mondy y Noe (1997), destacan que a la hora de realizar una evaluación del desempeño, se pueden suscitar una serie de circunstancias que pueden ocasionar problemas en el proceso de la evaluación del desempeño. Entre los cuales se pueden detectar: la carencia de objetivos, error de halo, indulgencia / rigor, tendencia central, prejuicio por comportamiento reciente, prejuicios personales y el papel de juez de evaluador.

En ocasiones las evaluaciones tienden a evaluar aspectos subjetivos del desempeño laboral, que acarrearán problemas tanto para el evaluador como para la compañía sobre el método usado para juzgar el trabajo de los empleados. Todo este problema de carencia de objetivos, puede solucionarse resaltando en la evaluación los aspectos relacionados con el puesto y no con el trabajador en sí.

Al mismo tiempo, se puede dar que el examinador realiza la evaluación del desempeño laboral del trabajador, en base a una característica que considera de mucha importancia, y basándose en esto califica mal o bien al empleado, sin juzgar como realiza su labor dentro de la institución donde trabaja, a esto se le llama error de halo.

Otro problema que se encuentra es, cuando la evaluación del desempeño juzga aspectos subjetivos del trabajo del área laboral de una corporación, y se le pide al examinador que discuta los resultados de la evaluación con los empleados; esto provoca que el evaluador califique la labor del empleado imparcialmente, debido a que desee evitar polémicas entre los trabajadores de la empresa, a esto se le denomina indulgencia. También se puede dar todo lo contrario, que se evalúe el desempeño de los empleados con mayor rigor de lo necesario.

De la misma manera, se puede exigir que los examinadores den una explicación sobre sus calificaciones muy altas o muy bajas dadas, en el proceso de evaluación laboral de los empleados. Esta situación se da cuando se juzga el trabajo de una persona cerca del promedio, a esto se le nombra tendencia central.

Por otra parte, se puede originar que los trabajadores de una corporación mejoren su desempeño laboral antes de la evaluación del desempeño, ya que en algunas ocasiones saben cuando será evaluado el trabajo dentro de la institución. En estos casos, el examinador suele recordar con mayor facilidad acontecimientos recientes que pasados. Pero para evitar problemas en la evaluación, el examinador debe recordar tanto los hechos recientes como los pasados, y debe evitar llevarse por sus prejuicios sobre sucesos recientes.

Igualmente, puede suscitarse que el evaluador tenga una serie de prejuicios personales sobre algunas personas que laboren en la institución y que vayan a ser evaluadas. Puede tender a discriminar a los empleados por el sexo, raza, religión, edad, etc.

La manipulación de los gerentes sobre el proceso, para poder argumentar sus decisiones sobre aumentar el sueldo o promover a algún empleado en particular, es un problema donde se obliga al evaluador a tener un papel de juez.

La evaluación del desempeño docente, es un instrumento útil para toda institución de estudios superiores, ya que ayuda a establecer cuales son las áreas donde el catedrático necesita mayor capacitación para mejorar su labor en el proceso de enseñanza. Por tal motivo, a continuación se describen una serie de datos sobre la evaluación del desempeño docente.

Valdés (2000), estipula que la evaluación del desempeño docente, es un proceso que habitualmente se realiza con el fin de determinar el efecto que produce en los alumnos las técnicas pedagógicas empleadas por el catedrático a la hora de impartir clases. Al mismo tiempo, se pretende detectar diferentes aspectos como: capacidades pedagógicas, emocionalidad, responsabilidad laboral, relaciones interpersonales, y resultados de su labor educativa.

Para determinar las capacidades pedagógicas se debe observar: el grado de dominio de los contenidos que imparte, capacidad de hacer su clase entretenida e interesante, capacidad de comunicación verbal y no verbal, contribuir a un adecuado clima de trabajo tanto dentro del aula como fuera de ella, una planificación adecuada del proceso docente-educativo, y la capacidad de poder identificar y comprender las distintas situaciones que surjan dentro del aula. Al mismo tiempo, la utilización de diversas técnicas educativas, el poder conocer y tratar las características psicológicas de cada uno de sus alumnos, saber como esta el nivel de aprendizaje de sus alumnos, como representa a la institución en la cual trabaja, poder contribuir a la formación de valores de sus alumnos, mantenerse actualizado y poder de autocrítica de su desempeño docente.

Así mismo, se puede establecer la emocionalidad de un docente, mediante su vocación pedagógica, autoestima, autoeficacia, el poder actuar con justicia y realismo, y su grado de satisfacción laboral y profesional. Otro aspecto que debe determinarse es la responsabilidad laboral, por medio de la observación de determinadas situaciones como: la puntualidad y asistencia a su clase, nivel de participación en sesiones de catedráticos, cumplimiento de normas y el grado profesional conseguido.

Las relaciones interpersonales del docente, se puede determinar a través de: el grado de preocupación y comprensión de los problemas de sus alumnos, expectativa sobre el desarrollo de sus estudiantes, y el modo de relacionarse con estudiantes, colegas, directivos y trabajadores de la institución. Al mismo tiempo, los resultados de su labor educativa, se pueden establecer mediante: el rendimiento académico obtenido por los estudiantes en la asignatura que imparte, y la orientación hacia las cualidades deseables de la personalidad de los alumnos.

Valdés (2000), destaca que se puede observar que toda evaluación del desempeño docente tiene una serie de funciones que deben cumplirse, que son: función de diagnóstico, instructiva, educativa y desarrolladora. En la función de diagnóstico, se pretende determinar el desempeño del catedrático en un período determinado de tiempo, donde se haga un compendio de los principales aciertos y desaciertos del mismo, con el fin de que ayude a determinar las posibles acciones a tomar para modificar dichos problemas en el trabajo desarrollado por el profesor.

La función instructiva, intenta realizar un resumen de los indicadores del desempeño del catedrático. La función educativa, lo que se pretende es lograr que el docente mejore sus errores en el proceso que emplea para enseñar, a partir de que conoce como es percibida su labor por su jefe inmediato, alumnos y los demás profesores.

El fin de la función desarrolladora, es lograr que el catedrático sea capaz de evaluar de una forma madura y crítica su propia labor como docente, que a la vez aprenda de sus errores,

todo esto para conseguir que trate de mejorar su trabajo futuro para obtener mejores resultados.

A la vez, se puede destacar que el objetivo de la evaluación docente es identificar las cualidades profesionales, la preparación y el rendimiento del catedrático. Así mismo, se puede destacar que tiene una serie de fines que son: conseguir un mejoramiento de la enseñanza dentro del aula, desarrollo y responsabilidad profesional, control administrativo y el pago por mérito.

Se pretende lograr una mejora en la enseñanza dentro del aula, por medio del aprendizaje de los distintos acontecimientos que se desarrollan dentro de clase, para poder determinar que es lo que debe modificarse, cambiar o mejorar.

También se puede observar, que en el control administrativo lo que se pretende es que por parte del área administrativa de la institución se supervise al docente a la hora de estar impartiendo el curso correspondiente. Por otro lado el pago por mérito, implica recompensar a los docentes de una forma monetaria por el buen rendimiento de los alumnos

En responsabilidad profesional, se reúnen una serie de datos con el propósito de observar si los catedráticos han alcanzado un grado aceptable de competencia y definido las metas que desean obtener en un futuro.

Por otro lado, se puede observar que el desarrollo profesional del docente, se puede ver estimulado por diferentes personas, como puede ser el mismo alumno, otros docentes y por el jefe inmediato. Así mismo, se puede establecer que en dicho proceso pueden influir tres factores que son: el factor contextual, que es el clima organizacional, recursos, liderazgo, entre otros. El factor relativo a procedimientos, las pruebas empleadas para recoger diversos datos, el empleo de otras fuentes de retroalimentación, etc. El factor relativo al profesor, como motivación, eficacia, etc.

Valdés (2000); recalca una serie de características que debe poseer un docente para estar relacionado con el desarrollo profesional, que son: fuertes expectativas profesionales, una orientación positiva hacia los riesgos, actitud abierta hacia los cambios, deseo de experimentar dentro de clase, actitud abierta ante las críticas, conceptos generales de los distintos métodos de enseñanza, conocimiento sobre su especialización, y experiencias positivas pasadas en la evaluación docente.

Así mismo, es importante destacar que en el proceso de evaluación es importante el proceso empleado en la evaluación y la retroalimentación para el catedrático. Los aspectos más importantes que deben tomarse en cuenta en el proceso de evaluación están: tener claros cuales son los aspectos de rendimiento establecidos por la institución, estar seguros si el catedrático esta de acuerdo y consiente de cuales son dichos aspectos de rendimiento, empleo de observaciones de clase, y examinar los datos del rendimiento de los estudiantes.

Dentro de la retroalimentación se destacan nueve factores que son: ideas claras sobre que se debe mejorar, una información clara y completa, ser específicos al brindar dicha información, resumir la información a dar, establecer cada cuanto se brindara la retroalimentación al docente, y la frecuencia de la retroalimentación.

Existen diversos modelos de evaluación del desempeño docente, entre los cuales se pueden destacar: Modelo centrado en el perfil del docente, modelo centrado en los resultados obtenidos, modelo de la práctica reflexiva y el modelo centrado en el comportamiento en el aula.

En el modelo centrado en el perfil del docente, se evalúa al catedrático en base a unas características previamente establecidas sobre lo que debe poseer un docente ideal. Este perfil, se puede desarrollar en base las percepciones que tienen los alumnos, catedráticos y jefes inmediatos de las características que deben tener todos los profesores.

Así mismo, en el modelo centrado en los resultados obtenidos, se realiza la evaluación a partir de los resultados obtenido por los estudiantes.

A la vez, en el modelo de la práctica reflexiva, se pretende juzgar la labor del docente con el fin que el catedrático mejore observando cuales son sus puntos débiles en el proceso de enseñanza. Y así puedan enfrentar, definir y resolver sus problemas prácticos. En este proceso, se observan tres etapas que son: observación y registro de actividades, conversación con la persona observada, y una segunda platica de seguimiento.

El modelo centrado en el comportamiento del docente en el aula, pretende evaluar al docente determinando aquellas características del profesor que estén relacionados con los logros de los estudiantes.

Valdés (2000); subrayan que la evaluación docente de distintas maneras, entre las que están: autoevaluación, observación de clases, encuestas, pruebas, portafolio y ejercicios de rendimiento profesional.

En la autoevaluación, se le solicita al catedrático que realice una crítica objetiva de su labor como educador. Esto se realiza con el objetivo, de estimular el proceso de que el mismo docente pueda realizar una crítica objetiva de su trabajo.

En la observación de clases, lo que se pretende es evaluar como el docente se desenvuelve cuando imparte sus cátedras y como reaccionan sus alumnos. Así mismo, en el portafolio se recopilan de manera escrita una serie de actividades que el catedrático realiza con el fin de poder testimoniar cual es el desempeño del mismo. En el portafolio se archivan tres tipos de documentos que son: los realizados por el profesor y otros colegas, los hechos por el docentes, y los elaborados por los directores.

En el ejercicio de rendimiento profesional, lo que se pretende evaluar es la capacidad que tiene el docente para preparar e impartir una clase. Existen diversos métodos empleados que son: que el docente imparta una clase siendo observado por algún colega, comentarios críticos sobre lo visto en una cinta de vídeo, responder preguntas de los alumnos por escrito ante la observación de otro docente, entre otros.

Por tal razón, es importante utilizar instrumentos con un grado de confiabilidad adecuado, ya que a la hora de utilizarlos, los puntajes mostrarán mayor consistencia en la medición de variables conductuales.

De las distintas universidades de la Ciudad Capital, se puede destacar que la universidad estatal, empezó a realizar evaluaciones docentes hace diez años atrás, utilizando una batería de pruebas creadas por OTEP (Oficina Técnica de Evaluación y promoción del personal docente), que actualmente es denominada DEPPA (Departamento de evaluación y promoción del personal académico) . Dicha batería, estaba constituida por 3 instrumentos, el que respondía el estudiante, el jefe inmediato y la autoevaluación. Aproximadamente hace doce años se crea un estatuto, donde se determina que cada área o facultad de la institución debe crear su propio instrumento de evaluación docente, (Cabrera, 2002).

El instrumento que se utilizó de 1991 al 2000, constaba de 30 preguntas. Tomando en cuenta los aspectos siguientes: el psicopedagógico, cultura general y el técnico académico. El aspecto de cultura general se evalúa en seis preguntas, el técnico académico en nueve preguntas y el psicopedagógico en quince preguntas. En la cultura general, se evalúa el conocimiento que tiene el catedrático en otras ciencias. En el técnico académico, si es claro y específico en su profesión y en el curso que imparte. Y en el psicopedagógico, el ambiente adecuado que debe mantener el docente para el buen proceso de enseñanza y aprendizaje de estudiantes.

En la actualidad, el DEPPA supervisa la evaluación docente y en casos extremos ellos mismos realizan las evaluaciones. Ya que en la actualidad, cada unidad académica selecciona una comisión que se encarga de realizar las pruebas respectivas, y después enviar al DEPPA en sobre cerrado e identificado con el nombre de catedrático los instrumentos, para su posteriormente obtener los resultados. Dicha comisión debe estar constituida por tres catedráticos y tres estudiantes, y esta es modificada cada dos años. Se puede decir que en el proceso de evaluación participan aproximadamente 180 personas, y se determina así: 30 unidades académicas por 6 personas por comisión, da un total de 180 personas.

Las evaluaciones docentes se suelen realizar cada semestre o una vez por año. Aproximadamente la institución cuenta con 3,000 docentes, de los cuales el año 2001 se evaluaron 2,500; pudiendo impartir de uno a seis cursos. Las funciones que se evalúan son: docencia, servicio o extensión, investigación y administración académica. En la docencia esta el 80% de sujetos, en servicio el 10%, en investigación el 6%, y en administración académica el 4%. En el campo de la docencia, se encuentra todos los docentes que imparten clases. En servicio están ubicados los supervisores de EPS (Ejercicio profesional supervisado). En investigación, las personas encargadas de realizar algún tipo de estudio para beneficio de la docencia en si. En la administración académica, los decanos de las facultades, el rector, etc.

A partir del año 2000, la mayoría de unidades académicas cuenta con su propio instrumento de evaluación docente, el cual puede estar formada por el número de ítems que desee la facultad, y pueden evaluar los aspectos y áreas que más se adecuen a su orientación o finalidad. La institución está formada por 10 facultades, que son: Agronomía, Arquitectura, Ciencias Económicas, Medicina, Farmacia, Humanidades, Ingeniería, Odontología y Veterinaria. Y 6 escuelas, que son: Ciencias de la Comunicación, Ciencia Política, Ciencias Psicológicas, Historia, Trabajo Social, y EFPEM (Escuela de Formación de Profesores de Enseñanza Media).

Para sacar los resultados finales de las evaluaciones, se tiene una ponderación específica que fue modificada de las puntuaciones anteriores. Estas son: de 1991 al 2000 de la evaluación que realizaba el estudiante, se obtenía el 35% del total de la puntuación del catedrático, de autoevaluación el 20% y de la evaluación del jefe inmediato el 20%, y de comisión el 25% por meritos académicos. Estos meritos académicos, eran los títulos obtenidos por el catedrático, publicaciones, diploma por haber participado en alguna conferencia, etc. En la puntuación actual, del estudiante se obtiene el 35% de la nota final del catedrático en la evaluación, del jefe inmediato el 15%, de la autoevaluación el 20%, y los meritos académicos constituyen el 30%.

Destaca Cabrera (2002), que entre las mayores ventajas de las evaluaciones docentes, esta que se detectan las deficiencias docentes, en base a la opinión de los estudiantes. Al mismo tiempo, participan en la evaluación las partes interesadas, los estudiantes, las mismas autoridades y el mismo docente expresando su opinión.

Como desventajas, se puede observar, que entre la población estudiantil no hay una conciencia sobre la evaluación docente, debido a que si se les indica un día antes que al día siguiente va tener lugar la evaluación del catedrático y él no se va presentar, lo más probable es que ningún estudiante llegue ese día a clases. También los estudiantes que participan de la comisión evaluadora, después de un tiempo ya no les interesa el proceso de evaluación por diferentes motivos. A la vez se puede notar, que si un catedrático es evaluado mal por el jefe inmediato, puede ser que el catedrático evalúe mal a su jefe inmediato. Otro caso es que el jefe inmediato evalúe bien a todos sus catedráticos, con el fin de que no lo evalúen mal, ya que puede ser que se esta en un año de elecciones.

La mayoría de las unidades académicas, realizaron estudios de validez y confiabilidad de sus instrumentos de evaluación. Pero hay facultades, que indicaron que no eran necesarios dichos estudios estadísticos y otras solo le dieron validez de constructo a las pruebas. Pero se puede destacar, que el instrumento de la Facultad de Farmacia, la prueba dio una confiabilidad de 0.98.

Mazariegos (2002), indica que en esta universidad privada, que esta ubicada en la zona 16, la evaluación docente empezó aproximadamente hace diez años atrás. Primero el Centro de Orientación Universitaria COU, se encargaba de pasar las evaluaciones de catedráticos. Después el COU realizaba dicho proceso en unión con las distintas facultades. A partir del año pasado, cada una de las distintas facultades se encargan de pasar las evaluaciones docentes.

Dichas evaluaciones se pasan tanto en los ciclos regulares como pre-ciclo e inter-ciclo. Aproximadamente se evalúan 1,000 profesores y 1,500 secciones. El instrumento anterior evaluaba los aspectos siguientes: responsabilidad, preparación, metodología y didáctica, y

relación con los estudiantes; y estaba compuesto por 26 preguntas. Esta evaluación era respondida únicamente por los estudiantes, y basándose en esto se realizaba la retroalimentación de los docentes.

Dicha retroalimentación la realizaba el Vicerrector Académico de la siguiente forma: al tener todos los resultados de las evaluaciones, el Vicerrector se encargaba de mandarle dicha información a cada uno de los Decanos de las distintas facultades por correo electrónico. Posteriormente cada uno de los Decanos se encargaba de informarles a sus respectivos catedráticos los resultados de las evaluaciones.

A partir del año pasado, la evaluación docente se realiza a través de tres instrumentos. El primero que responde el estudiante, el segundo es una autoevaluación hecha por el mismo catedrático, y el tercero que es hecho por los directores de las carreras. Los aspectos que evalúa el instrumento actual realizado por los estudiantes son: planificación y materiales, metodología, relaciones interpersonales y sobre el desempeño del mismo alumno. La autoevaluación fue hecha para que los docentes realicen una crítica de su desempeño dentro del aula. La evaluación que realizan los directores de las distintas carrera, evalúa los aspectos siguientes: relaciones interpersonales, enseñanza y evaluación, gestiones de clase y profesionalismo.

La evaluación hecha por los estudiantes esta constituido por 31 preguntas que están divididas de la siguiente manera: 6 ítemes en el área de planificación y materiales, 11 en el área de metodología, 8 en el área de relaciones interpersonales, y 6 en el área sobre el desempeño del mismo alumno. Esta es respondida vía Internet, (Mazariegos, 2002).

La autoevaluación esta formada por 16 preguntas abiertas, donde el catedrático expresa su opinión sobre su labor como catedrático. A la vez, debe rodear con un círculo las opciones que reflejen mejor su opinión que son: totalmente de acuerdo (TA) y totalmente en desacuerdo (TD).

La evaluación hecha por cada uno de los directores de la distintas carrera, esta conformada por 30 preguntas que están divididas de la forma siguiente: 8 ítemes en el aspecto de relaciones interpersonales, 8 en el aspecto de enseñanza y evaluación, 7 en el aspecto de gestiones de clases, y 7 en el aspecto de profesionalismo. Después de cada uno de los bloques de preguntas correspondientes a cada una de las áreas, está un espacio donde el director escribe los esfuerzos, preocupaciones y recomendaciones de cada docente que evalúa.

Anteriormente los resultados se leían a través de un lector óptico, debido a que era una hoja electrónica. Actualmente como la evaluación docente es vía Internet, el mismo sistema realiza la calificación y saca los resultados del test.

Indica Mazariegos (2002), que el grado de aceptación que los alumnos le otorgan a la prueba es poca, debido a que sienten que el objetivo que ellos consideran que tiene la prueba no se cumple.

Por otro lado, los catedráticos no están muy conformes con el proceso de evaluación, debido a que no le ven ningún beneficio, por que consideran, que muchas veces no reflejan realmente como ellos imparten sus cursos.

A la vez, se puede decir que la primordial ventaja, es que todos tienen un conocimiento acerca de lo que sucede en el aula, y brinda una orientación hacia donde deben ir enfocada la planificación de los cursos a impartir por el catedrático. A la vez, se esta revisando el proceso completo para cambiar la dinámica, con el fin de que todo mundo participe; alumnos, catedráticos y las facultades; para que todos estén compenetrados.

Entre sus desventajas se puede encontrar, que muchas veces los alumnos pueden calificar mal a un catedrático por represalia, debido a que les cae mal, y por tal razón muchas veces la prueba no refleja exactamente lo que esta pasando en clase.

También si no se explican los objetivos de realizar una evaluación docente, puede crear falsas expectativas dentro de los estudiantes. Los estudiantes pueden creer que calificando mal a un catedrático, este ya no será contratado por la universidad, pero cuando lo ven dando clases se pueden frustrar, debido a que consideran que su opinión no fue tomada en cuenta. Pero no se explica, que para lo que sirve una evaluación docente, es para ver cuales son las fortalezas y debilidades del catedrático a la hora de impartir un curso, y poder estipular medios para mejorar dichos aspectos deficientes que tiene en su metodología.

El único estudio que hay sobre confiabilidad del instrumento de evaluación docente, fue realizado por Verastegui en 1996, donde utilizó el instrumento antiguo compuesto por 26 preguntas.

Por otro lado, en otra universidad privada, ubicada en la zona 10, Bonilla (2002) destaca que se empezó a realizar las evaluaciones docentes aproximadamente hace dieciocho años atrás.

Pero en el transcurso del tiempo la evaluación de docentes se ha ido modificando, destacando que actualmente el instrumento que se emplea esta compuesto por nueve preguntas, donde las respuestas pueden ser: excelente, bueno, regular y deficiente. Estas evaluaciones se realizan antes de cada examen parcial y antes del examen final.

El departamento que se encarga de realizar el proceso de evaluación docente es: la unidad de atención al estudiante. Son seis personas que se encargan de realizar la evaluación de los catedráticos en toda la institución.

El instrumento esta orientado a evaluar los aspectos siguientes: conocimiento del catedrático, organización del curso y la forma de evaluar el curso. Aproximadamente se han evaluado a 173 catedráticos, y 240 cursos; donde los docentes imparten de uno a seis cursos en distintos semestres o trimestres, y en distintas carreras y facultades.

Los resultados se obtienen por medio del programa DELPHI. Bonilla (2002) indica que entre las mayores ventajas de evaluar al docente están: que se evalúa la calidad del docente

y de la cátedra. Y entre sus mayores desventajas está: que la opinión que externen los estudiantes no sea objetiva.

También se puede destacar, que por el momento no se ha realizado ningún tipo de estudio de confiabilidad del instrumento de evaluación docente.

El confiabilizar un instrumento de evaluación del desempeño, ayuda a establecer adecuadamente si el mismo es suficientemente congruente en sus distintos elementos. Al mismo tiempo, se puede determinar si los ítemes están adecuadamente identificados y establecidos para obtener los resultados deseados.

Se puede observar, que el proceso de dar confiabilidad a un instrumento es muy amplio e importante para la evaluación del desempeño laboral, tanto a nivel empresarial como docente. Ya que permite determinar, si el trabajo desarrollado por los docentes es el más apto para capacitar a los futuros profesionales en sus áreas específicas.

## **II. Planteamiento del Problema**

Para mejorar la evaluación de los docentes, se han desarrollado una serie de estudios en donde se crearon los tests específicamente para evaluar el desempeño en ciertos cursos o clases que impartían en la institución. Estas pruebas fueron constituidas con ítems donde se tocaban los puntos y contenidos que se deseaban evaluar. Simultáneamente los instrumentos fueron validados o confiabilizados con un grupo de personas con características similares a la muestra de la investigación.

No obstante, es importante no sólo darle confiabilidad y validez a pruebas donde se evalúe el desempeño de los docentes de nivel diversificado. También es necesario realizar este tipo de estudios con instrumentos de evaluación del desempeño de los catedráticos universitarios, ya que en este tipo de test se pretende juzgar diferentes áreas o aspectos que comprenden un desempeño docente adecuado.

Por ello, se pretenderá determinar si la evaluación del desempeño de catedráticos que emplea un centro de estudios universitarios es confiable o no. Para lograr establecer lo anterior se limitará la investigación a responder la pregunta siguiente ¿Es confiable el instrumento de evaluación para catedráticos universitarios?

### **2.1 Objetivo General**

Establecer la confiabilidad por medio del análisis de ítems de un instrumento de evaluación del desempeño para catedráticos universitarios.

### **2.2 Objetivos específicos**

- Establecer el grado de confiabilidad de las áreas de preparación y responsabilidad, metodología de enseñanza, evaluación, relación interpersonal e investigación.
- Correlacionar las áreas preparación y responsabilidad, metodología de la enseñanza, evaluación y relación interpersonal e investigación del instrumento de evaluación del desempeño para catedráticos universitarios.

- Determinar el porcentaje obtenido en cada una de las opciones de los ítemes del instrumento de evaluación del desempeño para catedráticos universitarios.

### **2.3 Hipótesis**

Hipótesis de Investigación:

Existe confiabilidad en la prueba de evaluación del desempeño docente utilizada por una universidad privada ubicada en la zona 10.

Ho.1 No existe confiabilidad en la prueba de evaluación del desempeño docente utilizada por una universidad privada ubicada en la zona 10.

Hipótesis Estadísticas:

1.Existe correlación estadísticamente significativa a nivel de 0.05 entre el área de metodología de enseñanza y el área de Preparación y responsabilidad del instrumento de evaluación del desempeño para catedráticos universitarios.

Ho.1 No existe correlación estadísticamente significativa a nivel de 0.05 entre el área de metodología de enseñanza y el área de Preparación y responsabilidad del instrumento de evaluación del desempeño para catedráticos universitarios.

2. Existe correlación estadísticamente significativa a nivel de 0.05 entre el área de evaluación y preparación y el área responsabilidad del instrumento de evaluación del desempeño para catedráticos universitarios.

Ho.2 No existe correlación estadísticamente significativa a nivel de 0.05 entre el área de evaluación y preparación y el área responsabilidad del instrumento de evaluación del desempeño para catedráticos universitarios.

3. Existe correlación estadísticamente significativa a nivel de 0.05 entre el área de relación interpersonal e investigación y el área de preparación y responsabilidad del instrumento de evaluación del desempeño para catedráticos universitarios.

Ho.3 No existe correlación estadísticamente significativa a nivel de 0.05 entre el área de relación interpersonal e investigación y el área preparación y responsabilidad del instrumento de evaluación del desempeño para catedráticos universitarios.

4. Existe correlación estadísticamente significativa a nivel de 0.05 entre el área de metodología de enseñanza y el área de evaluación del instrumento de evaluación del desempeño para catedráticos universitarios.

Ho.4 No existe correlación estadísticamente significativa a nivel de 0.05 entre el área de metodología de enseñanza y el área de evaluación del instrumento de evaluación del desempeño para catedráticos universitarios.

5. Existe correlación estadísticamente significativa a nivel de 0.05 entre el área de metodología de enseñanza y el área de relación interpersonal e investigación del instrumento de evaluación del desempeño para catedráticos universitarios.

Ho.5 No existe correlación estadísticamente significativa a nivel de 0.05 entre el área de metodología de enseñanza y el área de relación interpersonal e investigación del instrumento de evaluación del desempeño para catedráticos universitarios.

6. Existe correlación estadísticamente significativa a nivel de 0.05 entre el área de evaluación y el área de relación interpersonal e investigación del instrumento de evaluación del desempeño para catedráticos universitarios.

Ho.6 No existe correlación estadísticamente significativa a nivel de 0.05 entre el área de evaluación y el área de relación interpersonal e investigación del instrumento de evaluación del desempeño para catedráticos universitarios.

## **2.4 Variables**

Variable: Evaluación del Desempeño

Definición conceptual: Evaluación del desempeño es un procedimiento que se desarrolla habitualmente con el fin de determinar el rendimiento de una persona o un grupo de personas en su ambiente laboral, (Mondy y Noe, 1997).

Definición operativa: Evaluación del desempeño es el punteo obtenido en la sumatoria de las áreas de preparación y responsabilidad, metodología de enseñanza, evaluación, y relación interpersonal e investigación de la boleta de evaluación de la universidad ubicada en la zona 10.

## **2.5 Alcances y Limites**

El estudio estableció que tanto es confiable la evaluación de catedráticos a la hora de evaluar el desempeño de los mismos. A la vez, se identificarán cuáles son los ítemes que correlacionan mejor al juzgar el desempeño de los profesores.

La carencia de información sobre si el instrumento estaba validado o si poseía algún grado o nivel de confiabilidad, es una gran limitante.

## **2.6 Aporte**

Se estableció la confiabilidad del instrumento, lo que permitirá que en el futuro la empresa pueda recopilar la información que desea, con plena seguridad de que la prueba es constante en la medición de la o las variables deseadas y para lo cual fue creada.

Además se podrá observar que ítemes discriminan mejor la labor de los catedráticos. Lo que promoverá identificar en que puntos o aspectos se pueden realizar cambio, y así mejorar el instrumento y lograr una adecuada retroalimentación sobre el desempeño de los catedráticos.

Por otro lado, se intenta que se tome con más importancia evaluar el trabajo de los docentes, con el propósito de que cada día las personas que preparan profesionalmente a los individuos sean más preparados, eficaces y eficiente.

### **III. Método**

#### **3.1 Sujetos**

La muestra estuvo conformada por un grupo de cuatrocientos estudiantes de la jornada matutina, seleccionados al azar, de los primeros como de los últimos años de carrera, de ambos sexos, comprendidos entre los 18 años y 29 años de edad.

Es un centro de educación superior cuya misión es formar hombres y mujeres civilizados, cultos y competentes en el campo de su profesión; enriquecer y divulgar la ciencia, la tecnología y la cultura, y colaborar con el desarrollo de la sociedad.

De la totalidad de estudiantes que están inscritos en el centro de estudios superiores, se sacó la muestra al azar. La totalidad de estudiantes según su sexo son los siguientes: de primer ingreso; sexo masculino 232 y femenino 101. Reingreso: sexo masculino 1349, y femenino 447, que da un conjunto de 2,129 estudiantes. En el proceso de evaluación docente, se juzgó el desempeño de 173 catedráticos, donde cada uno puede impartir de uno a seis cursos respectivamente, en las diferentes carreras tanto semestral como trimestralmente, de las siguientes facultades: Facultad de Ingeniería de Sistemas, Informática y Ciencias de la Computación (FISICC), Facultad de Ciencia, Tecnología e Industria (FACTI), y la Facultad de Educación (FACED).

Por motivos de guardar la confidencialidad de la institución, tanto a nivel de cátedras como de catedráticos, no se detalla más la población. Debido a que la evaluación, se realiza por cátedra que imparte cada uno de los docentes que labora en dicha institución.

#### **3.2 Instrumento**

Se analizó un instrumento, conformado por preguntas cerradas, que esta dividido en 2 partes, la primera consta de 13 preguntas y la segunda 3 preguntas, donde se pretende medir preparación y responsabilidad, metodología de enseñanza, evaluación, investigación y relaciones interpersonales del catedrático. En la parte inferior se encuentra una nota, donde se indica que los comentarios se escriben en el reverso de la evaluación.

Los estudiantes puntúan el trabajo y el desempeño de sus catedráticos, en la primera parte guiándose por los criterios siguientes: siempre, con frecuencia, muy pocas veces y nunca. Y en la segunda parte se guían por los criterios siguientes: excelente, bueno, regular y deficiente.

El medio que se utilizó para calificar cada instrumento es un escáner programado con la clave cuantitativa del instrumento. El escáner envía los resultados de la prueba en una hoja de cálculo de Excel, estos resultados son posteriormente empleados para determinar los puntajes finales, que son hechos manualmente por la secretaria de recursos humanos. Dicha secretaria transcribe finalmente las preguntas con sus respectivas puntuaciones, el puntaje final y cada uno de los comentarios hechos por los estudiantes.

El instrumento que se utilizó para las entrevistas, fue hecho con el fin de obtener información sobre cómo era manejado el proceso de evaluación docente en las distintas universidades de la ciudad capital. Dicho instrumento está conformado por 12 preguntas abiertas, colocando al final el nombre y el cargo de la persona entrevistada.

Dicho instrumento se empleó para realizar las entrevistas en tres universidades diferentes que son: las universidades privadas ubicadas en la zona 16 y en la zona 10, y la universidad estatal.

### **3.3 Procedimiento**

- Se solicitó la información a la directora del departamento de Recursos Humanos de la institución ubicada en la zona 10 para establecer los criterios de la investigación.
- Después se empezó a recabar información teórica en las bibliotecas de las distintas universidades.
- Se entrevistó a representantes involucrados en el proceso de evaluación docente de tres universidades; una pública y dos privadas.

- Se pasó el instrumento a los alumnos de las distintas carreras, para que evalúen el desempeño de sus catedráticos.
- Se calificó las pruebas por medio de un escáner.
- Posteriormente se realizó un análisis de ítemes, del instrumento que actualmente se utiliza en la institución.
- El análisis se realizó por medio del paquete estadístico para las ciencias sociales versión 7.5 (SPSS), en donde los datos arrojados en Excel por el escáner pueden ser utilizados para interpretar mejor y claramente la información.
- Por medio del Alfa de Cronbach se obtuvo el índice de confiabilidad, con la fórmula siguiente:  $\alpha = \frac{K}{K-1} \left[ 1 - \frac{\sum \sigma_{ii}}{\sum \sigma_{ij}} \right]$   
 $K$  es el número del ítem.  
 $\sigma_{ij}$  es la estimación de covarianza entre el ítem  $i$  y el  $j$ .  
 $\sigma_{ii}$  es la varianza del ítem  $i$ .
- Posteriormente se realizó una evaluación, respecto a la pertinencia de los ítemes a partir de los datos cuantitativos correspondientes a los factores que se desean evaluar.
- Elaboración de conclusiones y recomendaciones.

### 3.4 Diseño y Metodología estadística

La investigación fue de tipo ex-post factum, ya que no se manipularon ningún tipo de variables. Tampoco se interfirió en la selección de los estudiantes que respondieron la prueba, ni en la selección del día y la hora en que se aplicó la prueba.

Según Achaerandio (1995), la investigación ex-post factum estudia la probable influencia y relación entre dos variables, en situaciones en donde no se puede operar la variable independiente, o ya acontecieron los sucesos y las permisibles influencias entre las mismas.

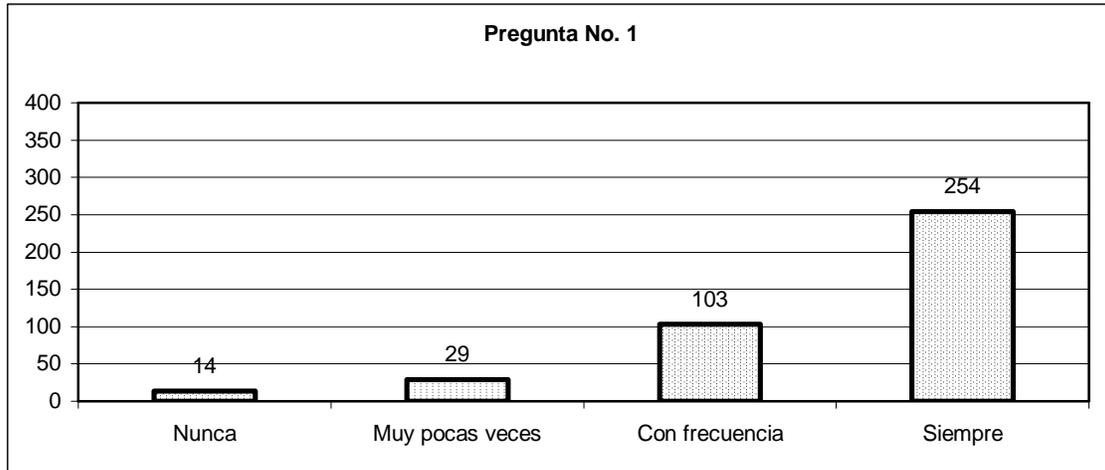
Para medir el nivel de discriminación de los ítemes se utilizó el análisis de ítemes. A través del Alfa de Cronbach se obtuvo el índice de confiabilidad, usando el paquete estadístico SPSS versión 7.5

#### IV. Presentación y Análisis de Resultados

##### Área No. 1: Preparación y responsabilidad

1. Presenta al inicio del curso un programa claro de evaluación y bibliografía.

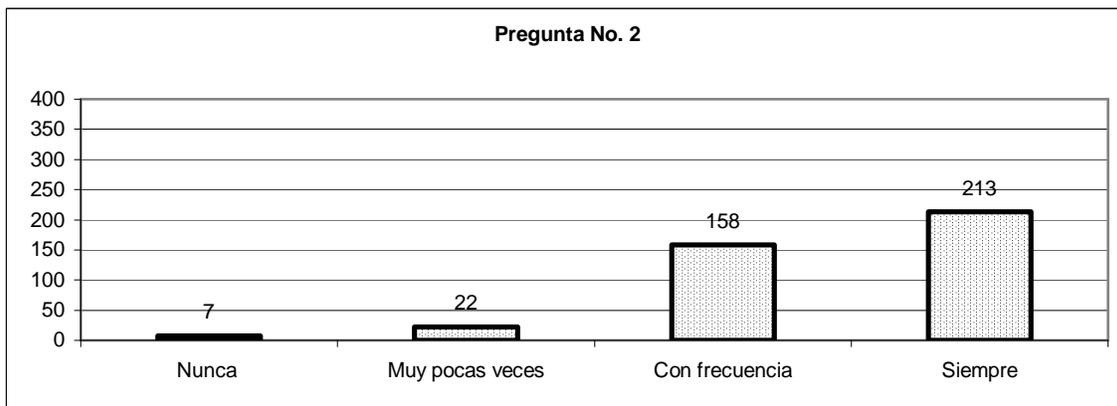
Gráfica No.1



En la pregunta No. 1, en donde se pretende determinar si los catedráticos presentan al inicio de sus cursos el programa correspondiente, del 100 por ciento de la muestra un 3.5 % contestaron que nunca, un 7.3% respondieron que muy pocas veces, un 25.8 % replicaron que con frecuencia y un 63.5 % contestaron que siempre.

2. Cumple con el programa

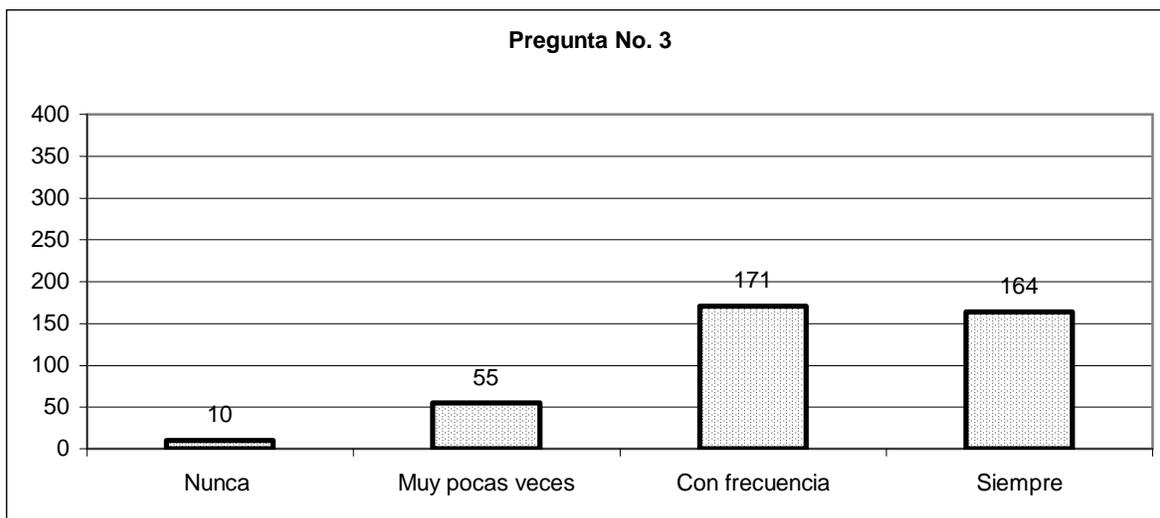
Gráfica No.2



De la muestra tomada al azar, un 1.8% respondieron que nunca, un 5.5% contestaron que muy pocas veces, el 39.5% indico que con frecuencia, y un 53.3% puntualizo que siempre.

### 3. Logra mantener la atención en clase

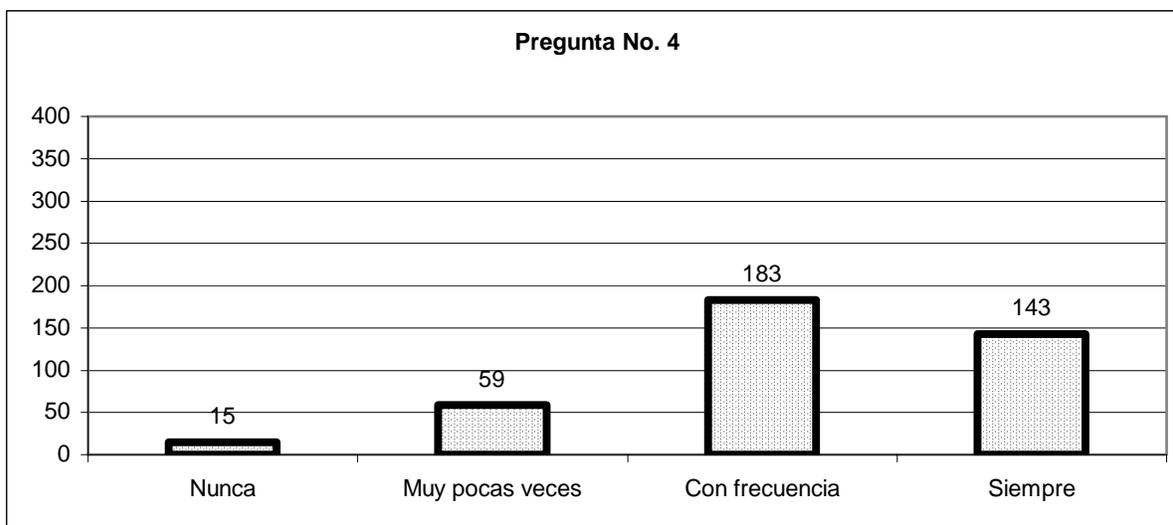
Gráfica No.3



De un total de 400 estudiantes que conformo la muestra, un 2.5% contesto que nunca, el 13.8% indico que muy pocas veces, el 42.8% replico que con frecuencia y el 41.0% respondió que el docente siempre logra mantener la atención en clase.

### 4. Expone con claridad los temas en el curso

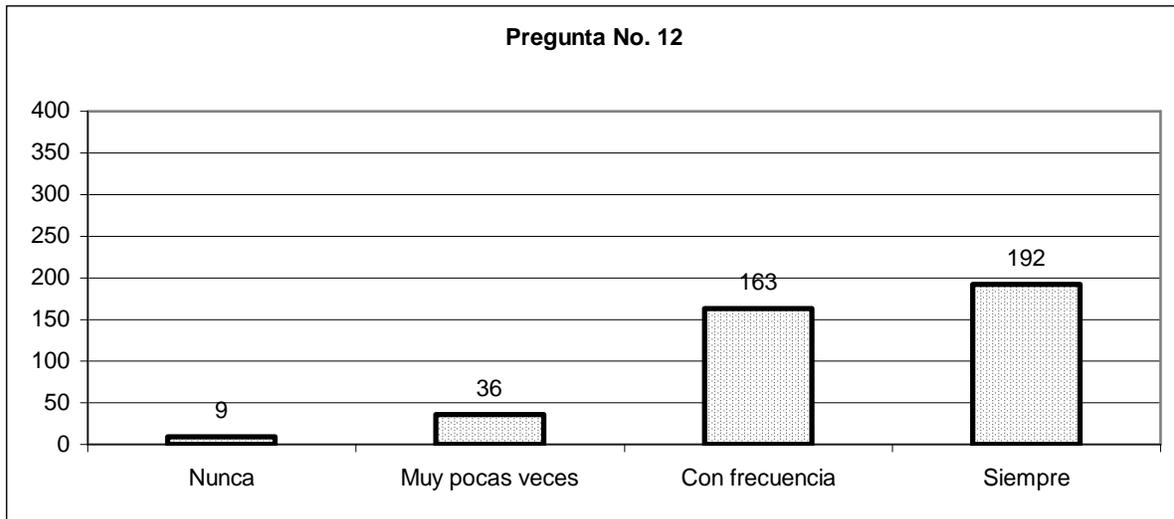
Gráfica No. 4



Del 100% de la muestra, un 3.8% respondió que nunca, el 14.8% muy pocas veces, 45.8% con frecuencia y 35.8% que siempre, expone con claridad los temas en el curso.

## 12. Cumple con el horario establecido

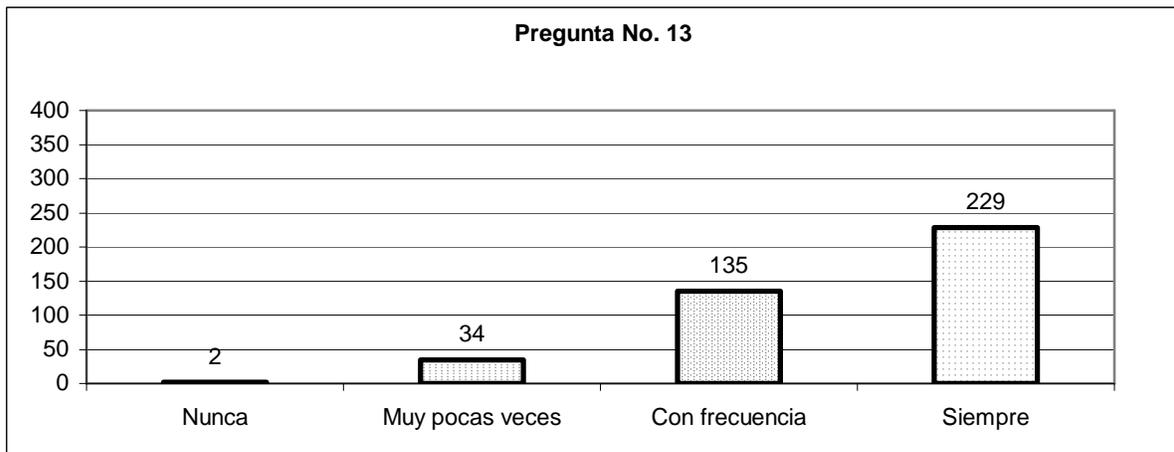
Gráfica No. 5



De la muestra tomada al azar, el 2.3% afirmó que nunca, el 9.0% muy pocas veces, el 40.8% con frecuencia y el 48.0% siempre.

## 13. Asiste con regularidad

Gráfica No.6

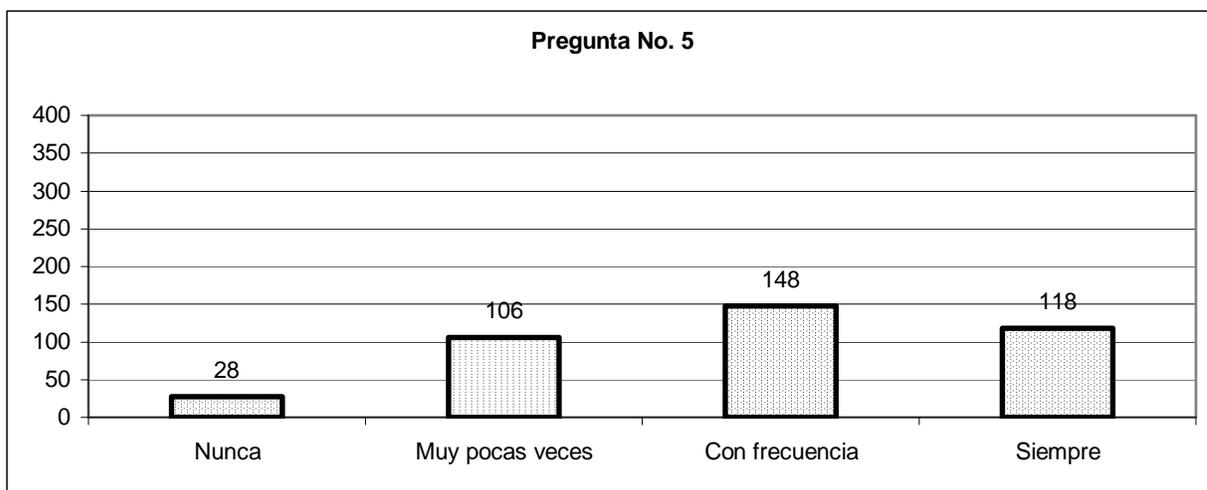


Del 100% de la muestra, un 0.5% indicó que el docente nunca asiste con regularidad, el 8.5% muy pocas veces, 33.8% con frecuencia y el 57.3% siempre.

Área No. 2: Metodología de la enseñanza

5. Relaciona los temas del curso con los de otras asignaturas

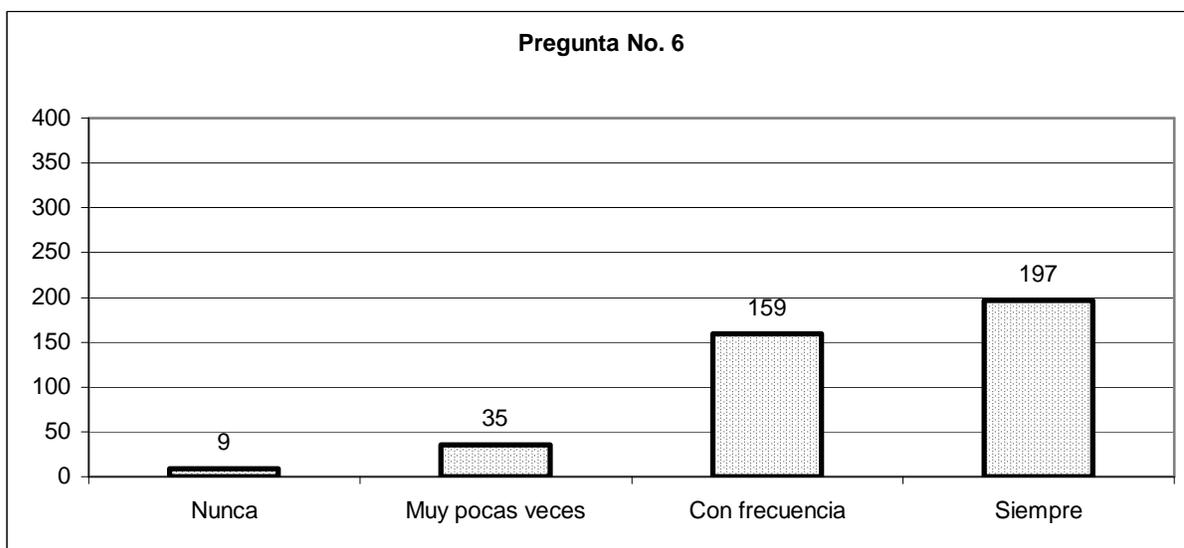
Gráfica No. 7



Del total de la muestra, un 7.0% indico que nunca, el 26.5% que muy pocas veces, el 37.0% con frecuencia y el 29.5% siempre.

6. Permite la intervención de los estudiantes

Gráfica No.8

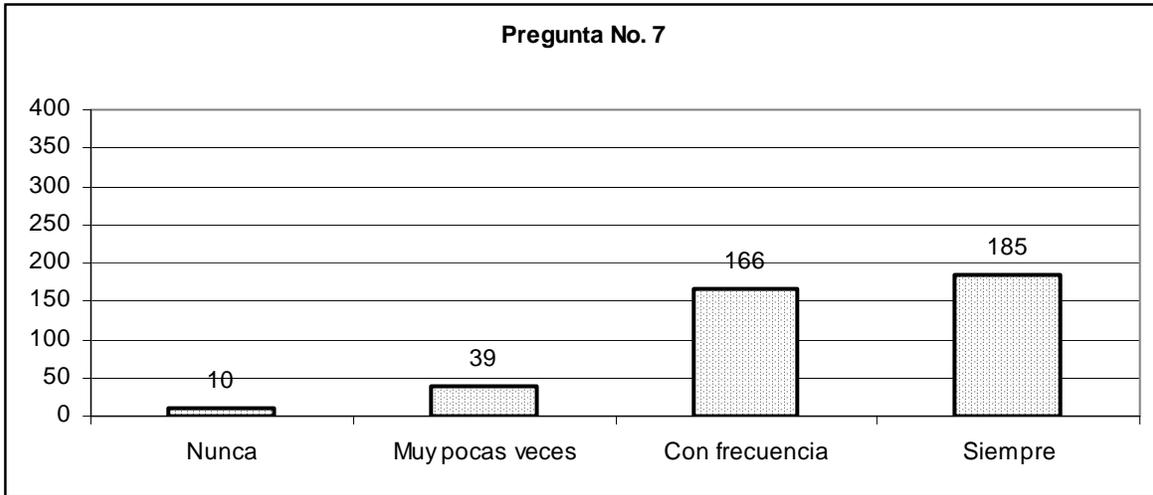


Del 100% de la muestra, el 2.3% contesto que nunca, el 8.8% muy pocas veces, el 39.8% con frecuencia, y el 49.3% siempre.

Área No. 3: Evaluación

7. Se ajustan a su evaluación, casos, problemas o proyectos al contenido del curso

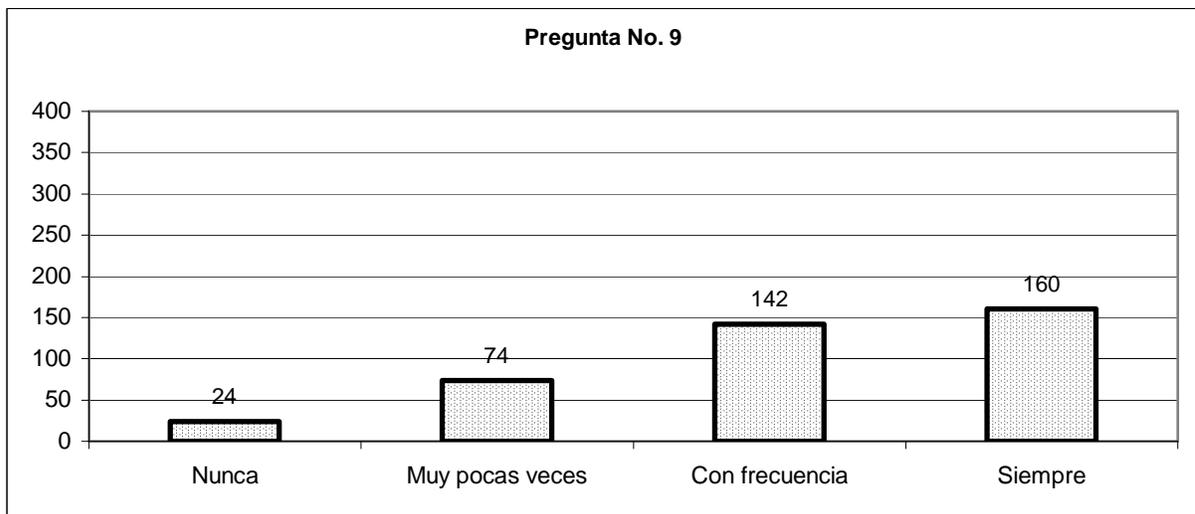
Gráfica No.9



Del total de la muestra, un 2.5% respondió que nunca, el 9.8% muy pocas veces, el 41.5% con frecuencia, y el 46.3 % siempre.

9. Da retroalimentación sobre evaluaciones y tareas en clase

Gráfica No. 10

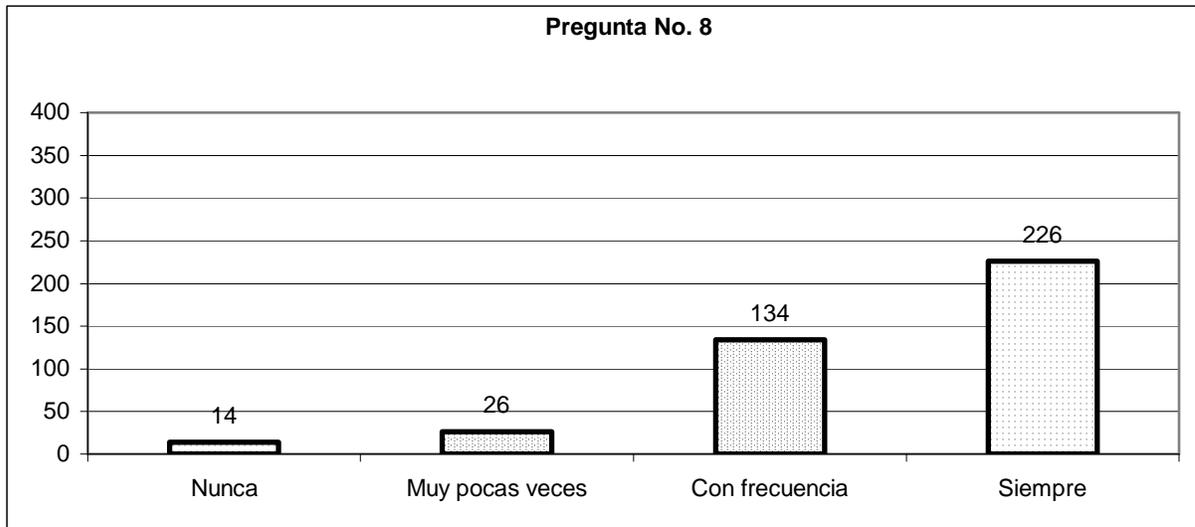


Del 100% de la muestra, el 6.0% manifestó que nunca, 18.5% muy pocas veces, 35.5% con frecuencia, y 40.0% siempre.

Área No. 4: Relaciones interpersonales

8. Su relación con los estudiantes es respetuosa o cordial

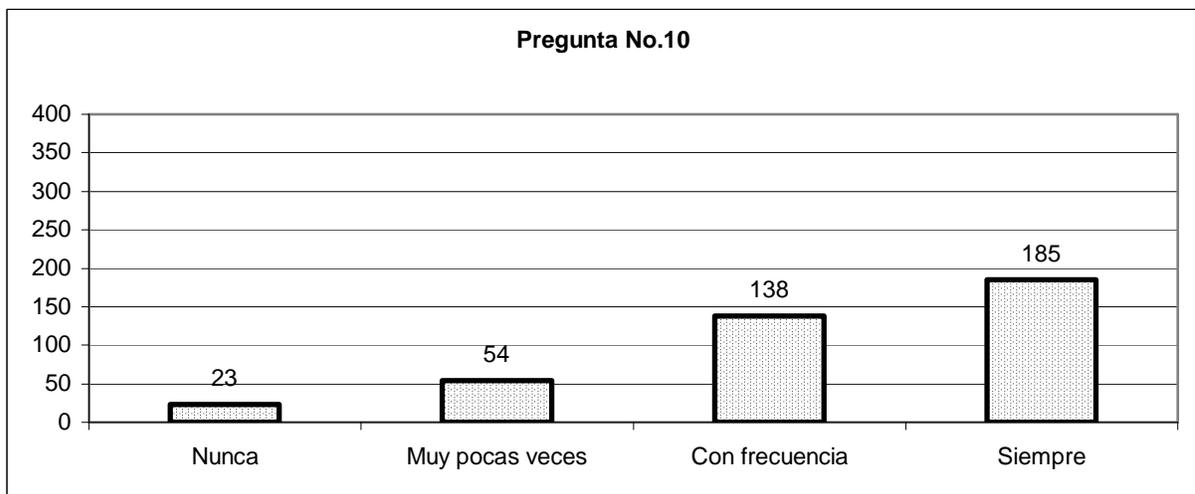
Gráfica No. 11



Del total de la muestra, el 3.5% afirmó que nunca, 6.5% muy pocas veces, 33.5% con frecuencia, y 56.5% siempre.

10. Es accesible para aclarar dudas dentro y fuera del aula

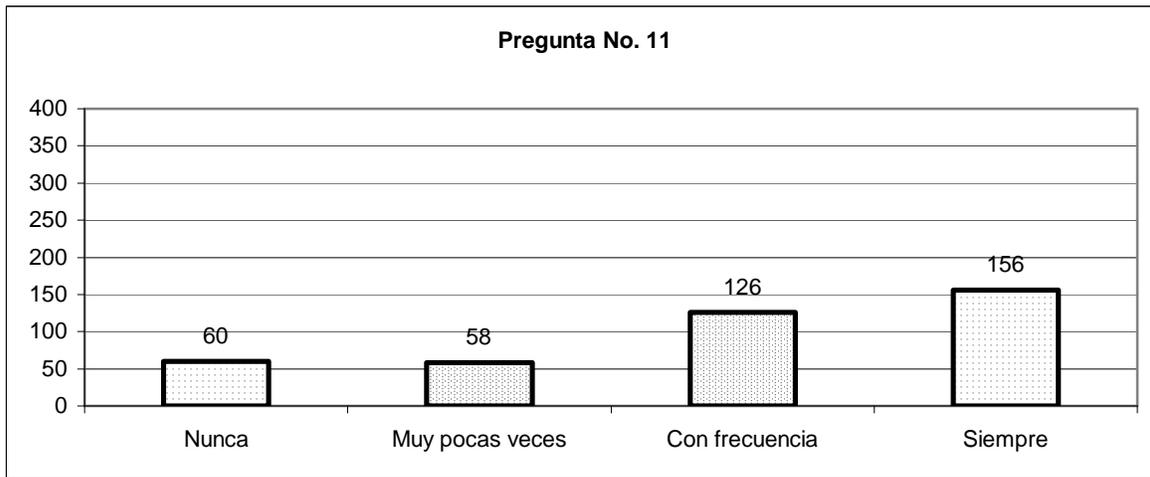
Gráfica No. 12



Del 100% de la muestra, el 5.8% contestó que nunca, 13.5% muy pocas veces, 34.5% con frecuencia, y el 46.3% siempre.

11. Le gustaría recibir otro curso con el mismo catedrático.

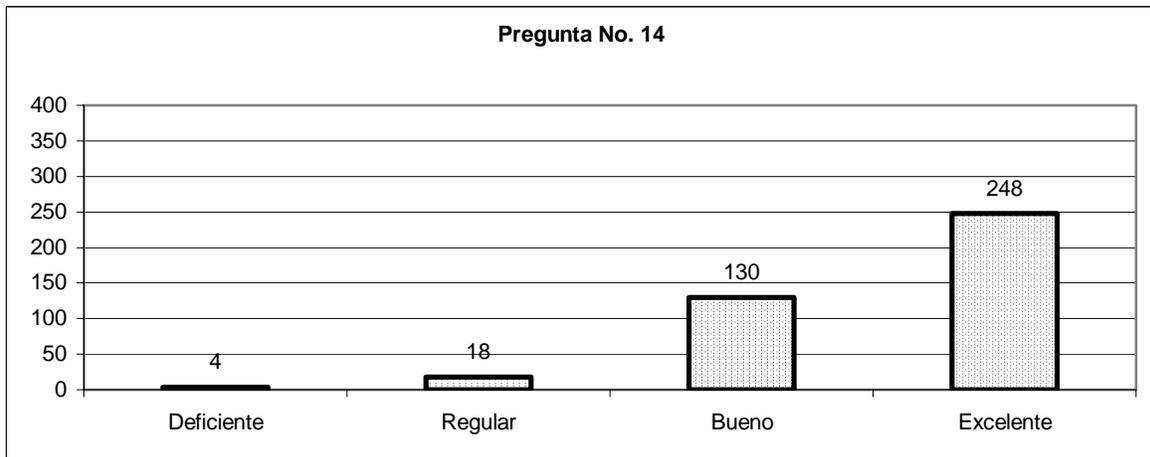
Gráfica No. 13



Del total de la muestra, el 15.0% afirmó que nunca, el 14.5% muy pocas veces, el 31.5% con frecuencia, y el 39.0% siempre.

14. El grado de conocimiento que usted considera que tiene el catedrático sobre el curso es

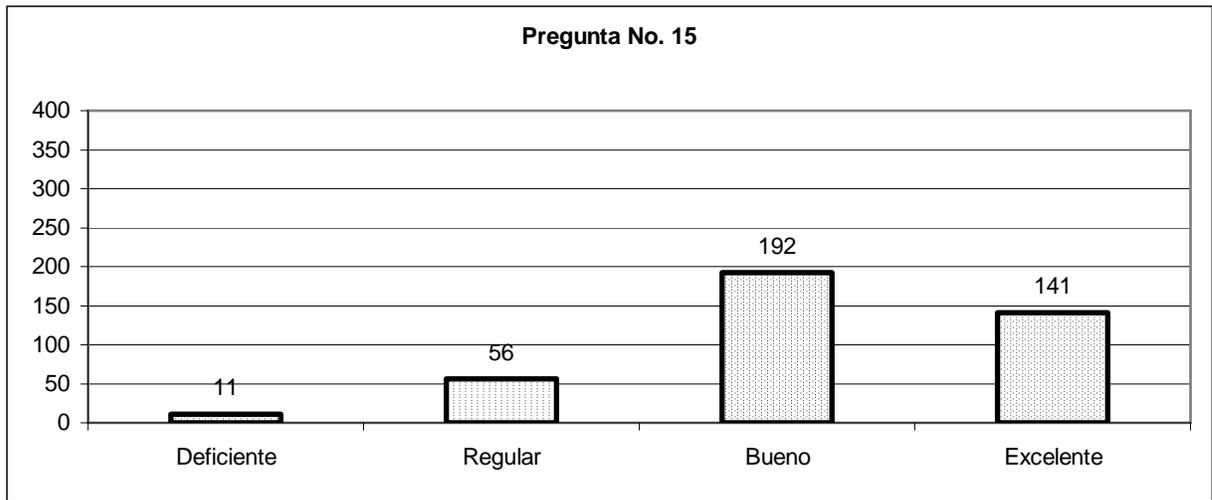
Gráfica No. 14



Del 100 % de la muestra, un 1.0% contestó deficiente, 4.5% regular, 32.5% bueno, y el 62.0% excelente.

15. Mi calificación sobre el curso es

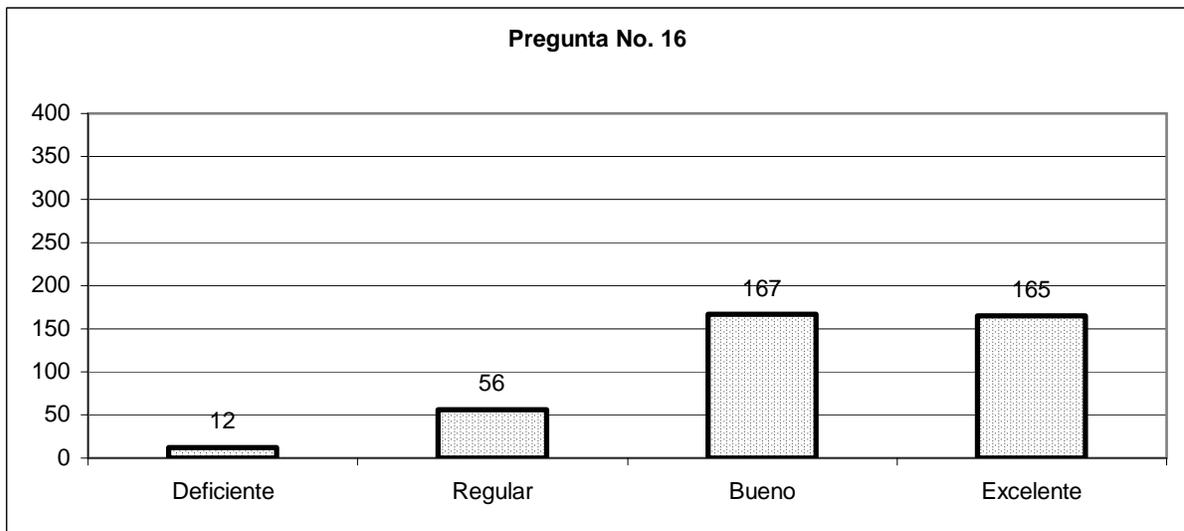
Gráfica No. 15



Del total de la muestra, el 2.8% respondió deficiente, el 14.0% regular, el 48.0% bueno y el 35.3% excelente.

16. Mi calificación sobre el catedrático es

Gráfica No. 16



Del 100% de la muestra, un 3.0% afirmó que deficiente, 14.0% regular, 41.8% bueno, y el 41.3% excelente.

Tabla No. 1

	0	2	5	8	10
Pregunta 1	<i>14 (3.5)</i>	<i>0 (0.0)</i>	<i>29 (7.3)</i>	<i>103 (25.8)</i>	<i>254 (63.5)</i>
Pregunta 2	<i>7 (1.8)</i>	<i>0 (0.0)</i>	<i>22 (5.5)</i>	<i>158 (39.5)</i>	<i>213 (53.3)</i>
Pregunta 3	<i>10 (2.5)</i>	<i>0 (0.0)</i>	<i>55 (13.8)</i>	<i>171 (42.8)</i>	<i>164 (41.0)</i>
Pregunta 4	<i>15 (3.8)</i>	<i>1(0.3)</i>	<i>58 (14.5)</i>	<i>183 (45.8)</i>	<i>143 (35.8)</i>
Pregunta 5	<i>28 (7.0)</i>	<i>0 (0.0)</i>	<i>106 (26.5)</i>	<i>148 (37.0)</i>	<i>118 (29.5)</i>
Pregunta 6	<i>9 (2.3)</i>	<i>0 (0.0)</i>	<i>35 (8.8)</i>	<i>159 (39.8)</i>	<i>197 (49.3)</i>
Pregunta 7	<i>10 (2.5)</i>	<i>0 (0.0)</i>	<i>39 (9.8)</i>	<i>166 (41.5)</i>	<i>185 (46.3)</i>
Pregunta 8	<i>14 (3.5)</i>	<i>0 (0.0)</i>	<i>26 (6.5)</i>	<i>134 (33.5)</i>	<i>226 (56.5)</i>
Pregunta 9	<i>24 (6.0)</i>	<i>0 (0.0)</i>	<i>74 (18.5)</i>	<i>142 (35.5)</i>	<i>160 (40.0)</i>
Pregunta 10	<i>23 (5.8)</i>	<i>0 (0.0)</i>	<i>54 (13.5)</i>	<i>138 (34.5)</i>	<i>185 (46.3)</i>
Pregunta 11	<i>60 (15.0)</i>	<i>0 (0.0)</i>	<i>58 (14.5)</i>	<i>126 (31.5)</i>	<i>156 (39.0)</i>
Pregunta 12	<i>9 (2.3)</i>	<i>0 (0.0)</i>	<i>36 (9.0)</i>	<i>163 (40.8)</i>	<i>192 (48.0)</i>
Pregunta 13	<i>2 (0.5)</i>	<i>0 (0.0)</i>	<i>34 (8.5)</i>	<i>135 (33.8)</i>	<i>229 (57.3)</i>
Pregunta 14	<i>4 (1.0)</i>	<i>0 (0.0)</i>	<i>18 (4.5)</i>	<i>130 (32.5)</i>	<i>248 (62.0)</i>
Pregunta 15	<i>11 (2.8)</i>	<i>0 (0.0)</i>	<i>56 (14.0)</i>	<i>192 (48.0)</i>	<i>141 (35.3)</i>
Pregunta 16	<i>12 (3.0)</i>	<i>0 (0.0)</i>	<i>56 (14.0)</i>	<i>167 (41.8)</i>	<i>165 (41.3)</i>
Total	<i>252 (3.9)</i>	<i>1 (0.0)</i>	<i>756 (11.8)</i>	<i>2415 (37.7)</i>	<i>2976 (46.5)</i>

La tabla No. 1 presenta las frecuencias y los porcentajes obtenidos en cada una de las opciones de los ítems del instrumento de evaluación del desempeño para catedráticos universitarios. En donde los números en cursiva son las frecuencias obtenidas, y la cifra entre paréntesis representan el porcentaje de respuestas.

Tabla No. 2

		Media	Desviación Estándar	No. de Casos
1	Pregunta 1	8.77	2.20	400.0
2	Pregunta 2	8.76	1.78	400.0
3	Pregunta 3	8.21	2.11	400.0
4	Pregunta 4	7.97	2.27	400.0
5	Pregunta 5	7.23	2.73	400.0
6	Pregunta 6	8.54	1.98	400.0
7	Pregunta 7	8.43	2.04	400.0
8	Pregunta 8	8.65	2.16	400.0
9	Pregunta 9	7.76	2.66	400.0
10	Pregunta 10	8.06	2.60	400.0
11	Pregunta 11	7.14	3.43	400.0
12	Pregunta 12	8.51	1.98	400.0
13	Pregunta 13	8.85	1.63	400.0
14	Pregunta 14	9.02	1.57	400.0
15	Pregunta 15	8.06	2.10	400.0
16	Pregunta 16	8.16	2.19	400.0
17	Total	132.15	24.44	400.0

En la tabla No. 2 se encuentra las medias y desviaciones estándar por pregunta del instrumento de evaluación del desempeño para catedráticos universitarios, en donde se observa que el promedio más alto obtenido (9.02) corresponde a la pregunta número 14; y el promedio menor corresponde a la número 11 con (7.14) de una puntuación máxima de 10 puntos.

### Gráfica de Promedios por ítem

Gráfica No. 17

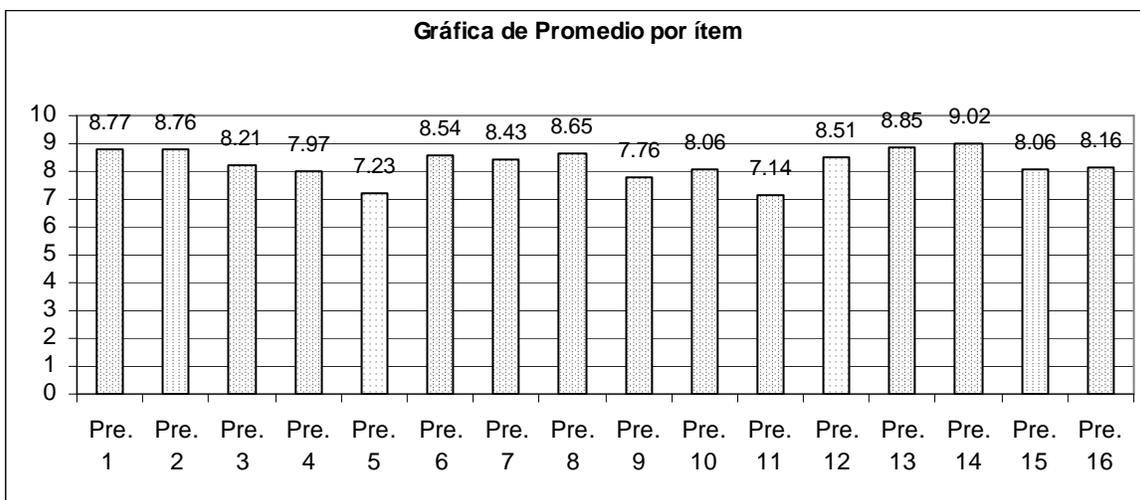


Tabla No3

Área		Alfa de Cronbach $\alpha$	No. de ítems
1	Preparación y Responsabilidad	0.88	6
2	Metodología de Enseñanza	0.89	2
3	Evaluación	0.90	2
4	Relación Interpersonal e Investigación	0.92	6
Total		0.93	16

La tabla No.3 muestra la confiabilidad obtenida utilizando el alfa de Cronbach para cada una de las áreas y la prueba completa.

La confiabilidad de la prueba fue estimada mediante el coeficiente Alfa de Cronbach. Este coeficiente estima la consistencia interna de una prueba, es decir, la relación que existe entre cada ítem y el total de la prueba.

Tabla No.4

No. de Ítem	Media Total	Desviación Estándar Total	Coefficiente Alpha	Correlación Total	R2
Pregunta 1	123.38	23.29	0.92	0.49	0.50
Pregunta 2	123.40	23.41	0.92	0.55	0.55
Pregunta 3	123.95	23.18	0.92	0.57	0.40
Pregunta 4	124.19	22.81	0.91	0.69	0.55
Pregunta 5	124.92	22.66	0.92	0.62	0.45
Pregunta 6	123.61	23.11	0.92	0.65	0.52
Pregunta 7	123.72	23.08	0.92	0.64	0.45
Pregunta 8	123.5	22.94	0.91	0.67	0.51
Pregunta 9	124.39	22.52	0.91	0.69	0.57
Pregunta 10	124.10	22.57	0.91	0.69	0.57
Pregunta 11	125.01	21.91	0.92	0.70	0.60
Pregunta 12	123.65	23.32	0.92	0.53	0.62
Pregunta 13	123.31	23.54	0.92	0.52	0.59
Pregunta 14	123.13	23.48	0.92	0.59	0.47
Pregunta 15	124.09	23.03	0.92	0.65	0.54
Pregunta 16	123.99	22.65	0.91	0.80	0.72
Total	132.16	24.44	0.92		

En la cuarta columna de la tabla No.4 se muestra cuan fiable es el instrumento aplicado, ya que el resultado está basado en el alpha de Cronbach en donde los coeficientes pueden oscilar entre 0 y 1; es decir un coeficiente 0 significa nula confiabilidad y 1 representa un máximo de confiabilidad. Por lo tanto mientras más se acerque el coeficiente a 0 hay mayor error en la confiabilidad; así mismo en la presente tabla la puntuación obtenida en todas las preguntas esta arriba de 0.90.

Tabla No. 5

Preguntas	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Pregunta 1	1.00															
Pregunta 2	0.67	1.00														
Pregunta 3	0.37	0.40	1.00													
Pregunta 4	0.26	0.31	0.48	1.00												
Pregunta 5	0.26	0.30	0.47	0.47	1.00											
Pregunta 6	0.27	0.32	0.38	0.61	0.46	1.00										
Pregunta 7	0.34	0.38	0.38	0.51	0.43	0.52	1.00									
Pregunta 8	0.32	0.34	0.33	0.54	0.40	0.56	0.52	1.00								
Pregunta 9	0.39	0.36	0.39	0.53	0.52	0.49	0.53	0.54	1.00							
Pregunta 10	0.29	0.30	0.36	0.52	0.50	0.54	0.46	0.55	0.66	1.00						
Pregunta 11	0.25	0.31	0.45	0.58	0.54	0.51	0.45	0.52	0.54	0.58	1.00					
Pregunta 12	0.53	0.58	0.28	0.31	0.28	0.27	0.39	0.33	0.35	0.33	0.28	1.00				
Pregunta 13	0.51	0.54	0.36	0.31	0.23	0.29	0.36	0.29	0.34	0.32	0.27	0.73	1.00			
Pregunta 14	0.23	0.26	0.37	0.38	0.43	0.41	0.39	0.44	0.37	0.44	0.49	0.29	0.31	1.00		
Pregunta 15	0.28	0.37	0.43	0.49	0.40	0.37	0.42	0.48	0.38	0.42	0.58	0.38	0.39	0.56	1.00	
Pregunta 16	0.32	0.40	0.43	0.59	0.53	0.57	0.52	0.61	0.58	0.63	0.71	0.45	0.39	0.63	0.66	1.00

La tabla No. 5 muestra la correlación entre cada una de las preguntas en donde se observa que la mayor correlación (0.73) se ubica entre las preguntas número 13 y 12; y entre las preguntas número 16 y 11 (0.71). En consecuencia las correlaciones mencionadas de acuerdo al criterio de (Hernández, Fernández y Baptista, 1998) poseen una correlación positiva considerable.

Tabla No.6

<b>Variable</b>	<b><math>\xi</math></b>	<b>S</b>	<b>n</b>	<b>r</b>	<b>Significancia</b>	<b>Decisión</b>
Metodología de Enseñanza	15.78	4.05	400	0.56	0.0001	Rechazar
Preparación y Responsabilidad	51.07	8.69	400			Ho.

Tabla No.7

<b>Variable</b>	<b><math>\xi</math></b>	<b>S</b>	<b>n</b>	<b>r</b>	<b>Significancia</b>	<b>Decisión</b>
Evaluación	16.2	4.12	400	0.62	0.0001	Rechazar
Preparación y Responsabilidad	51.07	8.69	400			Ho.

Tabla No.8

<b>Variable</b>	<b><math>\xi</math></b>	<b>S</b>	<b>n</b>	<b>r</b>	<b>Significancia</b>	<b>Decisión</b>
Relación Interpersonal e Investigación	49.12	11.24	400	0.64	0.0001	Rechazar
Preparación y Responsabilidad	51.07	8.69	400			Ho.

Tabla No.9

<b>Variable</b>	<b><math>\xi</math></b>	<b>s</b>	<b>n</b>	<b>r</b>	<b>Significancia</b>	<b>Decisión</b>
Metodología de Enseñanza	15.78	4.05	400	0.65	0.0001	Rechazar
Evaluación	16.2	4.12	400			Ho.

Tabla No.10

<b>Variable</b>	<b><math>\xi</math></b>	<b>s</b>	<b>n</b>	<b>r</b>	<b>Significancia</b>	<b>Decisión</b>
Metodología de Enseñanza	15.78	4.05	400	0.71	0.0001	Rechazar
Relación Interpersonal e Investigación	49.12	11.24	400			Ho.

Tabla No.11

<b>Variable</b>	<b><math>\xi</math></b>	<b>s</b>	<b>n</b>	<b>r</b>	<b>Significancia</b>	<b>Decisión</b>
Evaluación	16.2	4.12	400	0.71	0.0001	Rechazar Ho.
Relación Interpersonal e Investigación	49.12	11.24	400			

## V. **Discusión de Resultados**

Verastegui (1996), determinó la confiabilidad del instrumento utilizando el criterio de los alumnos al evaluar el desempeño de los docentes; pero en el presente estudio lo que se intenta es darle un nivel de confiabilidad al test de evaluación docente en base a los punteos dados por los estudiantes a la hora de evaluar a sus catedráticos. Se pudo determinar que la prueba tiene un nivel de confiabilidad mayor en su punteo total, que en cada una de las áreas que evalúa el instrumento.

Pero al mismo tiempo, se puede establecer que el nivel de confiabilidad del área de Relaciones Interpersonales e Investigación, que fue de 0.92; esta cerca del grado de confiabilidad obtenido en el punteo total de la prueba, que fue 0.93, que determina la confiabilidad del instrumento.

La confiabilidad total de la prueba analizada es igual a la confiabilidad encontrada por Verastegui (1996).

Al mismo tiempo; Aiken (1996), Anastasi y Urbina (1998), y Nunnally (1991), destacan que confiabilidad es determinar la estabilidad de instrumento a nivel de puntuación. Es decir, si el test va obtener la misma o similar puntuación a la anterior, pasándola en distintos períodos de tiempo, con distinta población en un ambiente similar al anterior donde fue pasada la prueba en un principio.

Por otro lado se observa, que es necesario realizar las evaluaciones del desempeño de una forma periódica, debido a que las personas evaluadas necesitan tener una retroalimentación sobre su desempeño laboral dentro de la institución, (Pinzón 1997, Bran 1999 y Mijangos 1998).

Esta retroalimentación puede ayudar a que se implemente algún tipo de programa de retroalimentación a los individuos, con el fin de mejorar su trabajo dentro de la organización, y así poder obtener un mejor resultado de sus actividades internamente.

Se puede notar, que tanto García-Arroba (2000) y Chiavenato (1993), coinciden en que la evaluación del desempeño fundamentalmente detecta como el empleado realiza su trabajo dentro de la empresa, y en base a esto en que áreas deben ser capacitado para mejorar su nivel de productividad en todo sentido.

Es importante poder determinar en que aspectos de sus labores los individuos necesitan capacitación, con el fin de poder desarrollar actividades que logren un alto nivel de actualización en áreas donde la persona realiza su trabajo.

En coincidencia con Chiavenato (1993), los resultados manifiestan los beneficios que la evaluación del desempeño brinda a los jefes, subordinados y la institución. Debido a que se puede determinar y desarrollar programas donde se les brinde a los individuos la retroalimentación necesaria de su labor dentro de la institución, (Mijangos 1998).

García-Arroba (2000) y Chiavenato (1993), comentan que habiéndose detectado la necesidad de capacitación por este instrumento se hace necesario darle al docente sus resultados, para implementar cursos de capacitación con el fin de mejorar su desempeño futuro.

Es importante que los resultados del procedimiento de evaluación sean comunicados a los evaluados, con el fin de que ellos sepan como fueron calificados en las distintas áreas que toma en cuenta el instrumento, y así puedan determinar y detectar cuales son los aspectos en los que tienen cierto grado de debilidad a la hora de realizar sus actividades dentro de la empresa, ( Méndez 2000 y Cabrera 2002).

Lo que se pretende al informar al evaluado de todo el proceso de la evaluación, es que logre comprender el motivo por el cual debe realizarse dichas evaluaciones, y así poder obtener su cooperación en el momento que se juzgue y se valore su trabajo dentro de la institución. Pero al mismo tiempo, no se debe olvidar, que las personas que fueron evaluadas deben saber los resultados de la misma, con el fin de que comprendan cuales son sus debilidades y

fortalezas en el momento de desarrollar su trabajo dentro de la organización, (García-Arroba 2000 y Cabrera 2002).

Tanto la universidad estatal como una de las privadas, que esta ubicada en la zona 16, coinciden en que se realizan tres evaluaciones, la evaluación que realiza el estudiante, la autoevaluación, y la que realiza el director de cada carrera. Al mismo tiempo, cada facultad se encarga de pasar sus propias evaluaciones docentes, (Cabrera 2002 y Mazariegos 2002).

En lo que difieren una de la otra, es que en la estatal cada facultad desarrolló sus propios instrumentos de evaluación docente; en cambio la privada todas las facultades utilizan los mismos 3 instrumentos, lo único que es que cada facultad estipula quienes son las personas que llevan a cabo dicho proceso.

Otra diferencia es que, en la estatal todos los instrumentos llegan al DEEPA y ahí se encargan de sacar los resultados en base a lector óptico. Pero en la privada, como el instrumento fue puesto en Internet, el mismo sistema califica y saca los resultados de la prueba.

Curiosamente coinciden ambas universidades, que entre sus mayores ventajas es la participación de las partes interesadas; alumnos, catedráticos y los directores de cada área o carrera.

También se puede notar, que el instrumento que se utilizó para este estudio de confiabilidad, tiene preguntas similares a la prueba utilizada por la universidad privada, ubicada en la zona 16, hasta el año anterior.

Una diferencia entre las tres universidades es que: dos de ellas utilizan tres instrumentos el que responde el alumno, la autoevaluación, y la evaluación que responde el director de cada carrera. Al mismo tiempo, se puede notar que una de las universidades privadas, ubicada en la zona 10, solo utiliza el instrumento que responden los estudiantes, (Cabrera 2002, Mazariegos 2002 y Bonilla 2002).

De acuerdo a la entrevista realizada las tres universidades, tanto la estatal como las universidades privadas ubicadas una en la zona 16 y la otra en la zona 10, coinciden en que una de las mayores desventajas, es que los resultados pueden ser parcializados. Ya que, los estudiantes pueden evaluar pésimamente a un catedrático simplemente porque les cae mal, debido a que no los trata bien dentro de clases.

## 5.1 Conclusiones

- La información recabada por los estudiantes es confiable, ya que el índice de alfa de confiabilidad del instrumento de evaluación es:
  - a. Preparación y Responsabilidad 0.88
  - b. Metodología de Enseñanza 0.89
  - c. Evaluación 0.90
  - d. Relación Interpersonal e Investigación 0.92
  - e. El instrumento de evaluación del desempeño para catedráticos universitarios aporto una confiabilidad de 0.93
  
- Las dos áreas con mayor punteo de confiabilidad fueron relación interpersonal e investigación y evaluación.
  
- Se aceptan las hipótesis alternas de las áreas preparación y responsabilidad, metodología de enseñanza, evaluación y relación interpersonal e investigación. Por lo tanto se rechazan las hipótesis nulas de Ho1 a la Ho6.
  
- Uno de los mejores ítemes que tiene la prueba es el No. 16, ya que tiene relación con la mayor parte de los ítemes que conforman el instrumento del estudio.

## 5.2 Recomendaciones

- Revisar los ítems 4, 5, 6, 7, 8, 9 y 10 ya que el valor que obtuvieron en la correlación total es mayor de 0.5.
- El ítem 11 debe ser borrado, ya que el valor que obtuvo en la correlación total es mayor de 0.6.
- Las opciones “con frecuencia” y “siempre” que son puntuadas con 8 y 10 respectivamente podrían ser unidas, ya que la mayoría de respuestas se encuentran en estas dos opciones.
- Tener varias formas paralelas del test para evitar que la manipulación y conocimiento de este test influya en el nivel de confiabilidad.
- Continuar aplicando el instrumento, ya que tiene un grado de confiabilidad de 0.93 considerado con alto.
- Seguir utilizando la prueba, debido a que hay correlación entre las áreas de: preparación y responsabilidad, metodología de enseñanza, evaluación, relación interpersonal e investigación.
- Que las diferentes universidades trabajen en conjunto, compartiendo sus conocimientos y experiencias, referentes al proceso de evaluación docente.

## VI. Referencias Bibliográficas

- Achaerandio L. (1995), **Iniciación a la práctica de la Investigación**. Guatemala: Universidad Rafael Landivar
- Aiken, L. (1996), **Test Psicológicos y Evaluación**. Octava edición. Traducción Parres V. México: Prentice Hall Hispanoamérica, S.A.
- Anastasi A. (1982), **Test Psicológicos**. Sexta edición. Madrid: Aguilar ediciones
- Anastasi A. y Urbina S. (1998), **Test Psicológicos**. (7ma. ed.) México: Prentice Hall.
- Arias F. (1979), **Administración de recursos humanos**. México: Editorial Trillas.
- Bran de López, M. (1999), **Beneficios que proporciona la evaluación del desempeño en una empresa azucarera. Caso Ingenio La Unión, S.A.** Tesis, Universidad Rafael Landivar. Guatemala.
- Bonilla, J. (2002, marzo, 7), **Entrevista personal**. Directora del Departamento de Recursos Humanos, Guatemala: Universidad Galileo.
- Caballeros de Mazariegos, H. (2002, marzo, 5), **Entrevista personal**. Directora de Administración Académica, Guatemala: Universidad Rafael Landivar.
- Cabrera, A. (2002, febrero, 13), **Entrevista personal**. Auxiliar de Investigación Científica, Guatemala: Universidad de San Carlos.
- Calvo, B. (1984), **Confiabilidad y validez concurrente del test uso de Lenguaje (T.A.D.)**. Tesis, Universidad del Valle de Guatemala. Guatemala.

- Carmona, M. (1982), **Validez y confiabilidad del inventario home para niños de un año de edad**. Tesis, Universidad del Valle de Guatemala. Guatemala.
- Chiavenato I. (1993), **Administración de recursos humanos**. (2da. ed.) Traducción Villa mizar G. México: McGraw-Hill Interamericana, S.A.
- Downie N. y Heath R. (1973), **Métodos estadísticos aplicados**. México: Harla, S.A.
- García-Arroba, M. (2000), **Propuesta de un instrumento de evaluación del desempeño para secretarías contratadas temporalmente**. Tesis, Universidad Rafael Landívar. Guatemala.
- Guilford, J. P. (1936), **Psychometric methods**. New York: McGraw-Hill
- Hernández R., Fernández C., y Baptista P. (1998), **Metodología de la Investigación**. México: McGraw-Hill.
- Hintze J., y Utah K. (1998), **User's Guide-II**. Estados Unidos: Number Cruncher Statistical Systems (NCSS).
- Méndez, M. (2000), **Propuesta de Evaluación del desempeño al recurso humano en escuelas primarias urbanas del municipio de Escuintla**. Tesis, Universidad Rafael Landívar. Guatemala.
- Mijangos, J. (1998), **Evaluación del desempeño de la escuela de directores, del nivel preprimario y primario en el departamento de Chimaltenango**. Tesis, Universidad Rafael Landívar. Guatemala.
- Mondy R., y Noe R. (1997), **Administración de Recursos Humanos**. (6ta. ed.) México: Prentice-Hall Hispanoamericana S.A.

- Montero E. (2000), **La teoría de respuesta a los ítemes: Una moderna alternativa para el análisis psicométrico de instrumentos de medición.** Costa Rica: Artículo para SIMMAC, Simposio Internacional de Métodos Matemáticos Aplicados a las Ciencias.
- Nunnally J. (1991), **Teoría Psicométrica.** Traducción González E. México: Editorial Trillas.
- Pinzón, M. (1997), **La evaluación del desempeño como una fortaleza en la administración de personal para la pequeña y mediana empresa dentro de la industria de calzado en la ciudad de Guatemala.** Tesis, Universidad Rafael Landivar. Guatemala.
- Pinzón, N. (1999), **Evaluación del desempeño profesional de las docentes especializadas y no especializadas de los institutos nacionales de formación secretarial de la ciudad capitales.** Tesis, Universidad Rafael Landivar. Guatemala.
- Quintana, O. (1982), **Validez y confiabilidad de la prueba de conocimientos generales para oficinista I (CGO-1) (en una muestra de estudiantes que cursaron el cuarto año de la carrera de Secretariado y Oficinista, en Institutos Públicos, jornada vespertina, durante el ciclo escolar 1981, en la ciudad capital).** Tesis, Universidad Rafael Landivar. Guatemala.
- Sandell R. (1977), **Administración de recursos humanos, su efectividad en la organización moderna.** Universidad Confederada de los Estados Unidos. Guatemala-Centroamérica.
- Valdés, H. (2000), **Encuentro Iberoamericano sobre Evaluación del Desempeño Docente.** México. [www. Campus-oei.org/de/rifad01.htm](http://www.Campus-oei.org/de/rifad01.htm).

- Verastegui, L. (1996), **Confiabilidad de la opinión de los alumnos al evaluar el desempeño de sus catedráticos**. Tesis, Universidad Rafael Landivar. Guatemala.
- Werher W. Y Davis (1997), **Administración de personal y recursos humanos**. Cuarta edición. México: McGraw-Hill.

## **Anexos**

## Entrevista

1. ¿Cuándo inició la Universidad ..... su evaluación del desempeño para catedráticos?
2. ¿Qué departamento aplica actualmente la evaluación del desempeño y cuantas personas participan en el proceso?
3. ¿Cada cuánto tiempo realizan las evaluaciones de catedráticos?
4. ¿Qué aspectos toman en cuenta en la evaluación docente?
5. ¿Cuántos catedráticos y cursos evalúan normalmente en la universidad?
6. ¿Cuáles son las técnicas que se utilizan para evaluar el desempeño docente?
7. ¿Qué áreas evalúa el instrumento de evaluación del desempeño?
8. ¿Cuántos ítems tiene la prueba?
9. ¿Cómo es el proceso de calificación de la prueba?

10. ¿Cuál es el grado de aceptación que le otorgan alumnos y docentes a la prueba?

11. ¿Cuáles son las ventajas y desventajas de aplicar la evaluación de docentes?

12. Si han realizado algún estudio de confiabilidad y validez de la prueba de evaluación de docentes.

Nombre de la persona entrevistada: \_\_\_\_\_

Cargo: \_\_\_\_\_