

**“CONFIABILIDAD Y VALIDEZ DE LA ENCUESTA DE OPINIÓN
REALIZADA A LOS ALUMNOS PARA EVALUAR Y RETROALIMENTAR
EL DESEMPEÑO DE LOS PROFESORES DEL ITESM”**

TESIS

**MAESTRÍA EN CIENCIAS CON ESPECIALIDAD EN ESTADÍSTICA
APLICADA**

**INSTITUTO TECNOLÓGICO Y DE ESTUDIOS SUPERIORES
DE MONTERREY**

POR

JOSÉ ENRIQUE MONTEMAYOR GALLEGOS

MAYO 2002

INSTITUTO TECNOLÓGICO Y DE ESTUDIOS SUPERIORES DE MONTERREY

**DIVISIÓN DE ELECTRÓNICA, COMPUTACIÓN, INFORMACIÓN Y
COMUNICACIONES**

**PROGRAMAS DE POSGRADO EN ELECTRÓNICA, COMPUTACIÓN,
INFORMACIÓN Y COMUNICACIONES**

Los miembros del comité de tesis recomendamos que la presente tesis del Lic. José Enrique Montemayor Gallegos sea aceptada como requisito parcial para obtener el grado académico de Maestro en:

CIENCIAS CON ESPECIALIDAD EN ESTADÍSTICA APLICADA

Comité de tesis:

Dr. Christian Garrigoux Michel
Asesor

Dra. Olivia Carrillo Gamboa
Sinodal

Dra. Rebeca Romero Álvarez
Sinodal

David Alejandro Garza Salazar, PhD.
Director de los Programas de Posgrado en Electrónica
Computación, Información y Comunicaciones

MAYO 2002

**“CONFIABILIDAD Y VALIDEZ DE LA ENCUESTA DE OPINIÓN
REALIZADA A LOS ALUMNOS PARA EVALUAR Y RETROALIMENTAR
EL DESEMPEÑO DE LOS PROFESORES DEL ITESM”**

POR

JOSÉ ENRIQUE MONTEMAYOR GALLEGOS

TESIS

Presentada a la División de Electrónica, Computación,
Información y Comunicaciones
Este trabajo es requisito parcial para obtener el Título de
Maestro en Ciencias con Especialidad en Estadística Aplicada

INSTITUTO TECNOLÓGICO Y DE ESTUDIOS SUPERIORES DE MONTERREY

MAYO 2002

DEDICATORIA

Quisiera dedicar esta tesis:

A Dios

Por estar siempre conmigo y por colmarme a mi y a mis seres queridos con salud y amor.

A mis padres

Les agradezco todo su apoyo; su cariño y confianza siempre han sido claves en todos mis logros, con amor los llevo siempre en mi corazón.

A mis hermanos

Gracias por tolerar cuando llegaba cansado o atareado y no me era posible platicar con ustedes, su preocupación por mí y sus ánimos fueron siempre alentadores para seguir adelante.

A mi novia Amy

Que tu angelito de la guarda te acompañe a donde quiera vayas y que siempre que te vayas a dormir sueñes con los angelitos barrigones. Gracias por todo tu apoyo bebe => sin ti hubiera sido más difícil.

A mis profesores

En especial a Christian Garrigoux, Jorge Sierra, Olivia Carrillo y Rebeca Romero por brindarme sus conocimientos, pero sobre todo su amistad.

RESUMEN

La actitud negativa hacia las evaluaciones de los alumnos por parte de los profesores, basada en experiencias propias o de compañeros, puede ser mitigada con la evidencia encontrada aquí, porque sugiere que las opiniones de los alumnos son confiables y válidas para evaluar y retroalimentar su desempeño.

El coeficiente de consistencia interna que cuantificó la confiabilidad de las opiniones de los alumnos, indicó que en promedio el 89.1% de la variabilidad de las evaluaciones de los alumnos es atribuible a la varianza verdadera y el 10.9% a la varianza de los errores no sistemáticos (ocasionales), provocados por variables como el estado de ánimo del alumno al contestar la encuesta, el tiempo disponible, el entorno donde la contesta, etc.

Los resultados informaron también que la confiabilidad se incrementa conforme aumenta el número de alumnos que contesta la encuesta en un grupo, lo que incentiva a realizar más promoción de ésta en el periodo de aplicación.

A su vez, se encontró evidencia que señala a las preguntas específicas de la encuesta que son objetivas como las más consistentes para evaluar al profesor. Reportaron menor variabilidad en las evaluaciones de los alumnos gracias a que poseen un parámetro base que utilizan los alumnos para externar sus opiniones.

Además, se observó que la información que aportaron las preguntas de la encuesta a pesar de ser confiable fue muy similar, debido a que fue muy común encontrar que los alumnos evaluaron a sus profesores con el mismo valor de la escala en todas las preguntas, aspecto de la encuesta que debe de ser mejorado, con el fin de brindar una buena retroalimentación al profesor.

Las evaluaciones de los alumnos estuvieron más relacionadas con variables internas al proceso de enseñanza-aprendizaje que con las variables externas. Por este motivo, se concluyó que el indicador utilizado para evaluar al profesor y que es elaborado a partir de las opiniones de los alumnos, la Opinión Global Promedio del Profesor (OGP) es válido.

Las variables internas, que midieron aspectos específicos del desempeño del profesor, fueron modeladas con las preguntas específicas de la encuesta y explicaron en un 92.9% la evaluación global del profesor, OGP. Al igual, dos factores estimados que resumieron la información de estas preguntas explicaron en un 89.4% dicha evaluación.

Por otro lado, las variables externas, que nada tienen que ver con el desempeño del profesor, fueron creadas de acuerdo a ciertas características del curso (tipo de disciplina), grupo (rediseñado o tradicional, tamaño y calificación promedio) y del profesor (género, años de experiencia en el Instituto, edad, tipo de contrato y grado máximo de estudios), y explicaron entre un 4.3% y un 7.2% la evaluación global del profesor, OGP. Si bien, el sesgo de estas variables sobre las evaluaciones de los alumnos fue significativo, no fue representativo dada la magnitud de los porcentajes de explicación.

CONTENIDO

	Página
ÍNDICE DE TABLAS.....	I
ÍNDICE DE GRÁFICOS.....	II
INTRODUCCIÓN.....	1
CAPÍTULO 1. ANTEPROYECTO.....	3
1.1 DEFINICIÓN DEL PROBLEMA	3
1.2 HIPÓTESIS A PROBAR	3
1.3 OBJETIVO DE LA INVESTIGACIÓN	4
1.4 ALCANCES Y LIMITACIONES.....	5
1.4.1 ALCANCES	5
1.4.2 LIMITACIONES.....	5
1.5 ANTECEDENTES SOBRE LA EVALUACIÓN DEL PROFESORADO EN EL ITESM.....	5
CAPÍTULO 2. MARCO TEÓRICO REFERENCIAL.....	7
2.1 EVIDENCIA SOBRE LA CONFIABILIDAD Y VALIDEZ DE LAS OPINIONES DE LOS ALUMNOS PARA EVALUAR Y RETROALIMENTAR EL DESEMPEÑO DEL PROFESOR.....	7
2.1.1 VARIABLES QUE SE DESEA SEAN FUENTES DE EXPLICACIÓN DE LA EVALUACIÓN GLOBAL DEL PROFESOR.....	8
2.1.2 VARIABLES QUE SE DESEA “NO” SEAN FUENTES DE EXPLICACIÓN DE LA EVALUACIÓN GLOBAL DEL PROFESOR.....	9
2.1.2.1 RENDIMIENTO ACADÉMICO DE LOS ESTUDIANTES.....	10
2.1.2.2 CARACTERÍSTICAS DEL CURSO Y DEL GRUPO.....	11
2.1.2.3 CARACTERÍSTICAS DEL PROFESOR (GÉNERO, EDAD, AÑOS DE EXPERIENCIA EN EL INSTITUTO Y TIPO DE CONTRATO).....	12
2.2 METODOLOGÍAS UTILIZADAS PARA COMPROBAR LAS HIPÓTESIS DEL ESTUDIO.....	14
2.2.1 METODOLOGÍA EMPLEADA PARA COMPROBAR LA CONFIABILIDAD DE LAS OPINIONES DE LOS ALUMNOS.....	14
2.2.1.1 COEFICIENTE DE CONSISTENCIA INTERNA.....	15
2.2.2 METODOLOGÍA EMPLEADA PARA COMPROBAR LA VALIDEZ DE LAS OPINIONES DE LOS ALUMNOS.....	15
2.2.2.1 EL MODELO DE REGRESIÓN LINEAL MÚLTIPLE.....	16
2.2.2.2 ANÁLISIS DE FACTORES.....	18
CAPÍTULO 3. CONFIABILIDAD DE LAS OPINIONES DE LOS ALUMNOS PARA EVALUAR Y RETROALIMENTAR EL DESEMPEÑO DE LOS PROFESORES.....	21
3.1 ANÁLISIS DE LAS PREGUNTAS DE LA ENCUESTA	21
3.1.1 PATRONES DE RESPUESTA DE LAS EVALUACIONES DE LOS ALUMNOS POR PREGUNTA.....	23
3.1.2 EVALUACIÓN REALIZADA POR LOS ALUMNOS POR PREGUNTA	26

3.1.3 VARIABILIDAD EN LAS EVALUACIONES DE LOS ALUMNOS POR PREGUNTA.	29
3.2 DETERMINACIÓN DE LA CONFIABILIDAD DE LAS OPINIONES DE LOS ALUMNOS.....	30
3.3 RESUMEN DE LOS RESULTADOS ENCONTRADOS ACERCA DE LA CONFIABILIDAD.....	32
CAPÍTULO 4. VALIDEZ DE LAS OPINIONES DE LOS ALUMNOS PARA EVALUAR EL DESEMPEÑO DE LOS PROFESORES	35
4.1 RELACIÓN DE LA EVALUACIÓN GLOBAL DEL PROFESOR CON LAS VARIABLES QUE SE DESEA SEAN FUENTES DE EXPLICACIÓN.....	35
4.1.1 EVALUACIÓN GLOBAL DEL PROFESOR MODELADA CON LAS PREGUNTAS ESPECÍFICAS DE LA ENCUESTA.....	36
4.1.2 EVALUACIÓN GLOBAL DEL PROFESOR MODELADA CON LOS FACTORES QUE RESUMEN LA INFORMACIÓN DE LAS PREGUNTAS ESPECÍFICAS.....	40
4.2 RELACIÓN DE LA EVALUACIÓN GLOBAL DEL PROFESOR CON LAS VARIABLES QUE SE DESEA “NO” SEAN FUENTES DE EXPLICACIÓN.....	44
4.2.1 RELACIÓN GRÁFICA DE LA EVALUACIÓN GLOBAL DEL PROFESOR Y LAS VARIABLES QUE SE DESEA “NO” SEAN FUENTES DE EXPLICACIÓN.....	45
4.2.1.1 EFECTO DE LA DISCIPLINA QUE IMPARTE EL PROFESOR.....	45
4.2.1.2 EFECTO DEL TIPO DE GRUPO (REDISEÑADO O TRADICIONAL).....	53
4.2.1.3 EFECTO DEL TAMAÑO DEL GRUPO.....	54
4.2.1.4 EFECTO DEL REDIMIENTO ACADÉMICO DE LOS ESTUDIANTES.....	56
4.2.1.5 EFECTO DEL GÉNERO DEL PROFESOR.....	57
4.2.1.6 EFECTO DE LOS AÑOS DE EXPERIENCIA DEL PROFESOR EN EL INSTITUTO.....	58
4.2.1.7 EFECTO DE LA EDAD DEL PROFESOR.....	60
4.2.1.8 EFECTO DEL TIPO DE CONTRATO DEL PROFESOR.....	62
4.2.1.9 EFECTO DEL GRADO MÁXIMO DE ESTUDIOS DEL PROFESOR.....	63
4.2.2 EVALUACIÓN GLOBAL DEL PROFESOR MODELADA CON LAS VARIABLES QUE SE DESEA “NO” SEAN FUENTES DE EXPLICACIÓN.....	64
4.3 RESUMEN DE LOS RESULTADOS ENCONTRADOS ACERCA DE LA VALIDEZ.....	69
CAPÍTULO 5. CONCLUSIONES Y RECOMENDACIONES	73
5.1 CONCLUSIONES.....	73
5.1.1 CONFIABILIDAD.....	73
5.1.2 VALIDEZ.....	74
5.2 RECOMENDACIONES.....	76
TRABAJOS FUTUROS.....	78
BIBLIOGRAFÍA.....	79
ANEXO 1.....	82
ANEXO 2.....	84
ANEXO 3.....	85
VITA.....	87

ÍNDICE DE TABLAS

CAPÍTULO 3

Tabla	Título	Página
	Cuestionario aplicado a los alumnos de profesional para cuantificar el desempeño de los profesores en el ejercicio académico de Agosto-	
3.1	Diciembre 2001.....	22
3.2	Parámetros de comparación de las preguntas específicas objetivas.....	23
3.3	Cálculos para obtener la desviación estándar promedio de las evaluaciones de los alumnos en todas las preguntas de la encuesta.....	28
3.4	Cálculo de la variabilidad global de las preguntas de la encuesta.....	30
3.5	Cálculo del coeficiente de consistencia interna por grupo.....	31
3.6	Coefficientes de consistencia interna de acuerdo al número de alumnos que contestaron la encuesta.....	32

CAPÍTULO 4

Tabla	Título	Página
4.1	Estimación del modelo de regresión, OGP versus preguntas específicas.....	37
4.2	Matriz de correlaciones entre las preguntas específicas y la pregunta 14, OGP.....	38
4.3	Estimación del modelo de regresión (OGP versus preguntas específicas) corrigiendo por multicolinealidad y por heterocedasticidad.....	39
4.4	Estimación de los factores que resumen la información de las preguntas específicas.....	41
4.5	Estimación del modelo de regresión, OGP versus los factores estimados.....	43
4.6	Disciplinas consideradas en el nivel de profesional.....	46
4.7	Composición del total de grupos de profesional por disciplina.....	47
4.8	Composición del total de grupos por tipo de grupo (rediseñado o tradicional).....	54
4.9	Composición del total de grupos por el tamaño del grupo.....	55
4.10	Composición del total de grupos de acuerdo a la calificación promedio del grupo.....	56
4.11	Composición del total de grupos de acuerdo al género del profesor.....	57
4.12	Composición del total de grupos de acuerdo a los años de experiencia del profesor en el Instituto.....	59
4.13	Composición del total de grupos de acuerdo a la edad del profesor.....	60
4.14	Composición del total de grupos de acuerdo al tipo de contrato del profesor.....	62
4.15	Composición del total de grupos de acuerdo al grado máximo de estudios del profesor.....	63
4.16	Estimación del modelo de regresión, OGP versus las variables externas al proceso de enseñanza-aprendizaje.....	66

ÍNDICE DE GRÁFICOS

CAPÍTULO 3

Gráfico	Título	Página
3.1	Patrones de respuesta de las preguntas de la encuesta.....	24
	Porcentaje de respuestas por pregunta en el valor de 7 de la escala con	
3.2	respecto a su total de respuestas.....	25
3.3	Evaluación promedio de las preguntas de la encuesta.....	26
3.4	Variabilidad global en las opiniones de los alumnos por pregunta.....	30

CAPÍTULO 4

Gráfico	Título	Página
4.1	Patrones de respuesta de la pregunta global por disciplina.....	48
	Patrones de respuesta de la pregunta global por tipo de grupo	
4.2	(tradicional o rediseñado).....	54
4.3	Evaluación en la pregunta global por tamaño de grupo.....	55
	Patrones de respuesta de la pregunta global de acuerdo al tamaño del	
4.4	grupo.....	55
	Evaluación en la pregunta global de acuerdo a la calificación promedio	
4.5	obtenida por el grupo.....	56
	Patrones de respuesta de la pregunta global de acuerdo a la calificación	
4.6	promedio obtenida por el grupo.....	57
	Patrones de respuesta de la pregunta global de acuerdo al género del	
4.7	profesor.....	58
	Evaluación en la pregunta global de acuerdo a los años de experiencia	
4.8	del profesor dentro del Instituto.....	59
	Patrones de respuesta de la pregunta global de acuerdo a los años de	
4.9	experiencia del profesor en el Instituto.....	59
4.10	Evaluación en la pregunta global de acuerdo a la edad del profesor.....	61
	Patrones de respuesta de la pregunta global de acuerdo a la edad del	
4.11	profesor.....	61
	Patrones de respuesta de la pregunta global de acuerdo al tipo de	
4.12	contrato del profesor.....	62
	Patrones de respuesta de la pregunta global de acuerdo al grado	
4.13	máximo de estudios del profesor.....	64

INTRODUCCIÓN

En toda institución educativa la preocupación por brindar servicios de calidad es de interés primordial. Por este motivo, se han creado dentro de las mismas, departamentos en donde la actividad principal es la de elaborar indicadores que permitan evaluar el desempeño de la Institución, en todos los aspectos relacionados con la educación que brinda a sus alumnos.

Uno de estos indicadores, esta relacionado con el desempeño de los profesores. Si bien existen muchas metodologías para evaluar el desempeño de éstos, en este trabajo solamente se analizará una, la evaluación que hacen los alumnos.

El sistema Tecnológico de Monterrey ha mostrado interés por conocer la opinión de los alumnos de cada Campus sobre el desempeño de sus profesores. Debido a ello, se comenzaron a aplicar encuestas a los alumnos desde 1974 en Campus Monterrey y desde 1994 a nivel sistema y se han realizado semestre a semestre hasta la fecha. La finalidad de estas encuestas es la de evaluar a los profesores desde la perspectiva de los alumnos.

El presente estudio tiene como objetivo iniciar el análisis de este instrumento de evaluación, así como de aquellos aspectos cuya optimización proporcionará un impulso efectivo a la docencia.

La evaluación, es una etapa muy importante dentro del proceso de enseñanza-aprendizaje, proporciona información útil sobre el desempeño de los profesores. Esta puede ser formativa, cuyo propósito es mejorar aspectos en la forma de enseñar del profesor, e igualmente sumativa, utilizada para medir el rendimiento o desempeño de éste con el fin de tomar decisiones administrativas.

El proceso de enseñanza-aprendizaje puede ser visualizado como un sistema en el que interactúan dos subsistemas fundamentales, Díaz (1997), los profesores y los alumnos. Sin embargo, también participa otro más, las personas administrativas, quienes pueden propiciar el mejor escenario en donde interactúen eficientemente dichos subsistemas.

Actualmente, la evaluación del profesor con base en las opiniones de los alumnos, es una práctica muy común en las instituciones de educación superior. Así, es de gran importancia conocer los aspectos que influyen en ella, no sólo dentro del proceso de enseñanza-aprendizaje sino también fuera de él.

Pueden existir variables externas a este proceso que contaminan tanto a la evaluación formativa como a la sumativa del profesor, es decir, las evaluaciones de los alumnos sobre sus profesores pueden estar sesgadas por variables que nada tienen que ver con el desempeño del profesor.

Por tanto, para que la evaluación del profesor pueda contribuir al mejoramiento del aprendizaje, debe de estar aislada de cualquier factor contaminante, ya que el aspecto que

apoya de manera más eficaz a la docencia es el propio conocimiento de los profesores sobre su propio desenvolvimiento en clase.

De esta forma, surge la inquietud de analizar la confiabilidad y validez de las opiniones de los alumnos utilizadas para evaluar y retroalimentar el desempeño de los profesores.

La confiabilidad de las opiniones de los alumnos indica la medida en que las diferencias individuales en las evaluaciones de éstos sobre el desempeño de los profesores, pueden atribuirse a “verdaderas diferencias” en las características consideradas a evaluar y el grado en que pueden deberse a errores fortuitos. En términos más técnicos, la medición de la confiabilidad permite estimar qué proporción de la varianza total de las evaluaciones se debe a la varianza verdadera y que tanta a la del error, Anastasi y Urbina (1998). De este modo, a mayor consistencia en las opiniones de los alumnos sobre el desempeño del profesor, mayor será su confiabilidad.

Por otro parte, la validez de las opiniones se define como el grado en el cual éstas miden el desempeño del profesor y que tan bien lo hacen, para después hacer inferencias sobre los resultados, Anastasi *et al* (1998). Existen muchas líneas de investigación en las cuales se ha aportado evidencia para demostrar la validez de las opiniones de los alumnos sobre el desempeño de los profesores. Dentro de éstas, existe una que enfoca su área en determinar si las opiniones están sesgadas por variables que nada tienen que ver con el desempeño del profesor, (Aubrecht, 1981 y Cashin 1988 y 1995).

Entonces, la medición de la validez permite identificar que porcentaje de la variabilidad total de las opiniones de los alumnos, es explicado por variables internas al proceso de enseñanza-aprendizaje (variables relacionadas con el desempeño del profesor) y que tanto por variables externas (aquellas que no están bajo el control del profesor) interpretándose este último porcentaje como el sesgo en las opiniones.

En resumen, lo que se pretende en este estudio es investigar la confiabilidad y validez de las opiniones de los alumnos obtenidas de la encuesta actualmente utilizada en el ITESM, para determinar si se está retroalimentando eficazmente a los profesores y cuantificando correctamente su desempeño. También, investigar cuáles variables tanto internas como externas al proceso de enseñanza-aprendizaje determinan la evaluación que hacen los alumnos, con el fin controlarlas y asegurar que los tres participantes en este proceso, alumnos, profesor y personas administrativas interactúen óptimamente.

CAPÍTULO 1

ANTEPROYECTO

1.1 DEFINICIÓN DEL PROBLEMA

El Instituto Tecnológico y de Estudios Superiores de Monterrey (ITESM), es actualmente una de las universidades con mayor prestigio en México y Latinoamérica. Entre los factores que posicionan al ITESM en un lugar privilegiado, se encuentra la calidad académica que oferta a los estudiantes, preocupándose siempre por tener un nivel de excelencia, contando con una plantilla de profesores que cubran dicho objetivo.

Ya que el Tecnológico, así como cualquier institución de educación superior privada, se enfoca en la satisfacción del estudiante tanto en términos académicos como de servicio, surge la necesidad de evaluar la labor de los profesores y para ello se aplica en cada periodo escolar una encuesta a los alumnos desde 1974 en el Campus Monterrey y desde 1994 en todo el Sistema ITESM.

En la actualidad, la encuesta se compone de 15 preguntas enfocadas a los alumnos de preparatoria y profesional, y 16 para los de posgrado. Las preguntas permiten evaluar el desempeño del profesor dentro del curso en diferentes ámbitos específicos, sin embargo, existe una que engloba la opinión que tiene el alumno acerca del desempeño del catedrático, es la pregunta número 14 con la cual se elabora el indicador utilizado para evaluar su labor, denominado la Opinión Global Promedio del profesor (OGP).

La Opinión Global Promedio, provee la información más importante de la encuesta, ya que es utilizada para evaluar al profesor con la finalidad de tomar decisiones administrativas. La trascendencia de estas decisiones pueden ir desde asignarle la titularidad de un grupo, hasta despedirlo si su desempeño no es el esperado.

No obstante la naturaleza de la OGP (sumamente importante por lo que cuantifica y las acciones que se toman) no se conocen los variables que la afectan (ya sea que estén relacionadas o no al proceso de enseñanza-aprendizaje), y a su vez, no se ha realizado un análisis estadístico para determinarlas y poder con ello propiciar una atmósfera adecuada para que el profesor se desempeñe mejor.

1.2 HIPÓTESIS A PROBAR

La evaluación del docente es un tema complejo que representa un área de atención cada vez más importante para las instituciones educativas. Lo importante de la evaluación es ayudar al sistema; darle fuerza y orientación mediante la retroalimentación entre alumnos y profesores.

El intercambio de experiencias entre profesores y alumnos constituye una fuente importante de aprendizaje para los primeros; ello implica un concepto de retroalimentación que pertenece fundamentalmente al docente cuya orientación profesional se ubica en la enseñanza. El estudiante se alimenta en el sistema durante una generación, el docente enriquece su ejercicio durante una continuidad de generaciones (Díaz, 1997).

Debido a esto, surge la necesidad de plantear la primera hipótesis de este trabajo, que implica probar que tan confiable es la información aportada por la encuesta actual para retroalimentar al profesor.

En consecuencia, se probará si la encuesta ofrece o no un diagnóstico que, mediante las preguntas de la encuesta, permita al profesor obtener una retroalimentación de sus alumnos para posteriormente hacer los ajustes necesarios en su desempeño, con el fin de lograr los objetivos primordiales de la institución, brindar una educación de calidad.

Esta modalidad de evaluación (la basada en la opinión de alumnos) ha suscitado una preocupación constante en el personal docente referida a la validez que pueda otorgarse a los resultados de dichas encuestas de opinión para evaluar su desempeño.

Esto da paso a la segunda y última hipótesis de la investigación, la que consistirá en probar que tan válido es el indicador utilizado para evaluar el desempeño global del profesor, la Opinión Global Promedio, OGP, elaborado a partir de las opiniones de los alumnos. Esto es, qué tanto este indicador está cuantificando realmente su objetivo y si está afectado o no por variables externas al proceso de enseñanza-aprendizaje.

Ambas hipótesis tienen la misma importancia ya que en conjunto ayudan a mejorar el sistema en donde profesores y alumnos interactúan.

1.3 OBJETIVO DE LA INVESTIGACIÓN

Los objetivos de esta investigación se resumen en los siguientes puntos:

1. Realizar una investigación bibliográfica de los estudios existentes sobre la evaluación de profesores con base a las opiniones de los alumnos, para analizar las aportaciones que dichos estudios contengan.
2. Realizar un análisis para determinar la confiabilidad de la retroalimentación brindada al profesor con la información recabada en las preguntas de la encuesta.
3. Determinar la validez del indicador utilizado para evaluar al profesor, lo que permitirá identificar si la OGP está siendo sesgada por variables que no están relacionadas con el desempeño del profesor, las que son llamadas variables externas al proceso de enseñanza-aprendizaje.

1.4 ALCANCES Y LIMITACIONES

1.4.1 ALCANCES

Los alcances de esta investigación consistirán en verificar la confiabilidad y validez de las opiniones de los alumnos obtenidas de la encuesta aplicada en el ejercicio académico de Agosto-Diciembre de 2001, en el nivel de profesional y sólo para los cursos teóricos, utilizando como unidad de análisis la evaluación promedio del grupo en las preguntas de la encuesta y no la evaluación individual de los alumnos.

1.4.2 LIMITACIONES

En un estudio de esta naturaleza, es necesario establecer limitaciones ya que no es posible abarcar todos los tópicos que comprende el área de interés, por tanto, el presente trabajo no abarcará:

- El mismo análisis pero para otros niveles académicos,
- Para otros Campus del Sistema,
- Para otros indicadores que sean alternos al que actualmente se esta utilizando para medir el desempeño del profesor.

1.5 ANTECEDENTES SOBRE LA EVALUACIÓN DEL PROFESORADO EN EL ITESM

Actualmente, la finalidad primordial de los resultados obtenidos en la encuesta del ITESM, es la de evaluar al profesor a mayor grado que retroalimentarlo, ya que la encuesta es aplicada casi al final del curso, entre el tercer parcial y el examen final, por lo que la retroalimentación que recibe el profesor, sólo puede aplicarla en el siguiente curso.

Adicionalmente, el indicador utilizado para evaluar al profesor, se centra en la variabilidad generada por las evaluaciones de los alumnos, esto es, se plasman las opiniones de los alumnos en una medida de tendencia central, esto implica que a mayor dispersión en las opiniones, el indicador tendrá una confiabilidad menor, ya que no habrá consenso (consistencia) sobre el desempeño del profesor y a su vez, carecerá de validez debido a condiciones ajenas a la enseñanza que sesgan el valor del indicador.

Lo anterior indica la gran necesidad de estudiar éste instrumento de evaluación y retroalimentación en cuanto a la confiabilidad y validez de sus resultados, porque pueden tomarse decisiones erróneas si las opiniones de los alumnos no cuantifican realmente el desempeño del profesor.

Es por esto que existe preocupación por parte de los profesores sobre la confiabilidad y validez de los resultados obtenidos de dichas encuestas, pero hay pocos estudios en el Sistema ITESM relacionados con el tema.

Los estudios cualitativos existentes se han enfocado principalmente en proponer preguntas nuevas en la encuesta (Campus Monterrey, 1998), pero no han investigado si las existentes son confiables, es decir, si están aportando información relevante para retroalimentar al profesor.

A su vez, se han realizado estudios basados en metodologías etnográficas para percibir el acto de enseñanza a plenitud en el salón de clases, con el fin de conocer el perfil del docente en cuanto a principios didácticos, habilidades y estrategias de enseñanza que identifican al profesor excelente, pero no se ha determinado estadísticamente si existen diferencias significativas entre el profesor excelente y el mejor evaluado, lo cual determinaría la validez de la encuesta, González (1996).

Asimismo, se han analizado métodos alternos diseñando un programa integral para evaluar el desempeño del docente, en donde la retroalimentación de éste viene no sólo de los alumnos sino también de colegas y de personas administrativas (Campus Ciudad de México, 2001). Pero antes de proponer algún método alternativo, es importante verificar la confiabilidad y validez de la metodología actual para en dado caso que sea deficiente mejorarla o cambiarla.

En cuanto a los estudios cuantitativos, éstos han investigado mediante coeficientes de correlación o diagramas de dispersión, cómo esta relacionada la evaluación global del profesor con la evaluación en ciertos indicadores de exigencia, metodologías docentes y relaciones maestro-alumno (Rectoría Zona Sur, 2000). No obstante, no se ha aprovechado el valor agregado que brindan análisis más completos que permiten ver la relación de la evaluación global del profesor considerando el efecto simultáneo de varias variables.

Estos puntos mencionados, serán considerados a continuación en el presente estudio, para investigar la confiabilidad y validez de las opiniones de los alumnos utilizadas para evaluar y retroalimentar el desempeño de los profesores.

CAPÍTULO 2

MARCO TEÓRICO REFERENCIAL

2.1 EVIDENCIA SOBRE LA CONFIABILIDAD Y VALIDEZ DE LAS OPINIONES DE LOS ALUMNOS PARA EVALUAR Y RETROALIMENTAR EL DESEMPEÑO DEL PROFESOR

El uso de la evaluación, lejos de retroalimentar al profesor, alumno e institución en busca de una mejor enseñanza, con frecuencia sólo se utiliza para efectos de filtro y control, en beneficio únicamente de las decisiones administrativas.

El propósito de obtener las opiniones de los alumnos a través de una encuesta, no debe ser solamente para fines sumativos, las evaluaciones que los alumnos hacen sobre sus profesores deben ser utilizadas para dos cosas, evaluar al profesor y proveer información de diagnóstico para mejorar su desempeño. En efecto, no debe de pasarse por alto la evaluación formativa, con la que se retroalimenta al profesor para mejorar su desempeño.

Dentro de éste punto surge la pregunta con relación a la confiabilidad de las opiniones de los alumnos para retroalimentar y evaluar a sus profesores; la confiabilidad radica en qué tan consistente es la información que se obtiene de dicha encuesta.

Aubrecht (1979 y 1981) y Cashin (1988, 1990 y 1995) en un resumen de la literatura de Estados Unidos encontraron que la confiabilidad, medida con coeficientes de consistencia interna (los cuales miden el grado de acuerdo entre las opiniones de los alumnos para evaluar a los profesores), variaba conforme se incrementaba el número de evaluadores (alumnos), en promedio la confiabilidad que reportaron fue de 0.69 cuando participaban 10 alumnos, 0.83 con 15, 0.83 con 20, 0.88 con 30 y 0.91 con 40 alumnos.

Ruedas y Díaz (2000) en su trabajo de compilación incluyen un artículo de García Garduño, el cual hace una revisión detallada de la bibliografía de Estados Unidos sobre la confiabilidad y validez de los cuestionarios de opinión estudiantil que se emplean en la evaluación de la docencia, encontrando que las encuestas a los alumnos son instrumentos confiables y validos que permiten medir la efectividad del docente, resultados similares a los encontrados por Aubrecht (1979 y 1981) y Cashin (1988, 1990 y 1995).

En el contexto del ITESM, la imprecisión de la metodología con la que se evalúa al profesor, tiende a la búsqueda del profesor mejor evaluado, la preocupación es tan grande por tener profesores bien evaluados que se pasa por alto que este concepto no es sinónimo de profesor excelente, González (1996).

En búsqueda del profesor mejor evaluado, entra la pregunta sobre si el indicador utilizado está cuantificando realmente el desempeño del profesor, por lo que en este estudio el término de validez se refiere a si la pregunta global de la encuesta aplicada a los

alumnos, realmente mide el desempeño de los profesores, si esto es así, entonces será posible predecir la eficacia del profesor para enseñar, así como el desempeño del alumno.

Al considerar la medición del desempeño del profesor en el contexto de la validez, es necesario especificar dentro de la teoría y la evidencia cuáles variables deberían de relacionarse o no con el indicador utilizado. De esta manera, las variables internas al proceso de enseñanza-aprendizaje se desea sean fuentes de explicación de la evaluación global del profesor; en cambio, las variables externas al proceso son orígenes no deseados de explicación, (Díaz, 1997).

Tomando estos resultados como base, es importante investigar si en las instituciones mexicanas, en este caso el ITESM, se confirman los resultados observados en la evidencia de Estados Unidos, esto porque el efecto de la cultura puede estar presente y no necesariamente se encuentre que las opiniones de los alumnos obtenidas a través de encuestas sean un instrumento confiable y valido para evaluar y retroalimentar al profesor.

Por esto, en el presente estudio para probar la confiabilidad de dicho instrumento se calcularan coeficientes de consistencia interna para indagar que tanta proporción de la variabilidad de las opiniones de los alumnos es explicada por el error, lo cual permitirá también conocer que tan consistentes son las opiniones de los alumnos.

Por otro lado, para probar la validez se investigará mediante un modelo de regresión, no sólo qué variables internas al proceso de enseñanza-aprendizaje están determinando la evaluación de los alumnos, sino también qué variables externas a este proceso, éstas últimas esperando arrojen una influencia pequeña sobre las evaluaciones.

2.1.1 VARIABLES QUE SE DESEA SEAN FUENTES DE EXPLICACIÓN DE LA EVALUACIÓN GLOBAL DEL PROFESOR

Las variables internas al proceso de enseñanza-aprendizaje miden aspectos específicos del desempeño del profesor y se desea sean fuentes de explicación de la evaluación global del profesor. Estas variables, comúnmente están cuantificadas con las preguntas específicas de la encuesta, o con factores que se estiman mediante un análisis de factores y que son una combinación lineal de las preguntas específicas, que si bien se conocen o no a priori, pueden ser interpretados para verificar el perfil del profesor establecido por los alumnos.

Para el caso de México, Díaz (1997) realiza un estudio en la Universidad Iberoamericana mediante un análisis de factores para investigar la confiabilidad y validez de la encuesta utilizada por ésta institución para evaluar y medir el desempeño de sus profesores, encontrando un bajo grado de confiabilidad en la encuesta así como de validez.

En cuanto a la interpretación de los factores estimados en las sub-poblaciones consideradas (Divisiones), encontró que éstos sólo diferían en la organización jerárquica con las preguntas de la encuesta, interpretándose de la siguiente manera: el factor exposición, cumplimiento e interacción social.

En el análisis de regresión que utilizó para identificar las preguntas que predeterminan la evaluación global del profesor, observó que las únicas significativas fueron las relacionadas con la imagen asertiva del profesor, con la exigencia, esfuerzo y cumplimiento, y con la evaluación objetiva basada en el razonamiento.

Para Estados Unidos Cranton y Smith (1990), examinaron la estructura fundamental de las evaluaciones de los alumnos para investigar las dimensiones del profesor en diferentes unidades de análisis: la media de clase, las evaluaciones individuales de los alumnos y la desviación de las evaluaciones de los alumnos con respecto a la media de clase. Estos autores, encontraron que la estructura aportada por la media de clase fue diferente de la encontrada en las otras dos unidades de análisis. Específicamente, encontraron para cada unidad de análisis, dos factores que explicaron aproximadamente entre el 86% y el 91% de la variabilidad total de las preguntas específicas de la encuesta, los cuales tuvieron las siguientes interpretaciones: el interés y atmósfera que transmitió el profesor, así como su organización para impartir el curso.

2.1.2 VARIABLES QUE SE DESEA “NO” SEAN FUENTES DE EXPLICACIÓN DE LA EVALUACIÓN GLOBAL DEL PROFESOR

Primeramente, es importante mencionar que la evaluación del profesor es afectada por el sentimiento de poder que tienen algunos alumnos al evaluar, esto es, es posible que la evaluación obtenida por el profesor este afectada por evaluaciones extremas que resultan de resentimientos de la relación profesor-alumno; en consecuencia, si la medida de tendencia central utilizada para resumir las opiniones de los alumnos, es afectada por este tipo de evaluaciones, entonces se estará cuantificado erróneamente el desempeño del profesor, formulándose en consecuencia decisiones erróneas acerca de su futuro en la Institución.

Sí las evaluaciones de los estudiantes reflejaran sólo el contenido enseñado en el curso por el profesor, entonces no habría de qué preocuparse acerca de variables externas o del sesgo, como comúnmente se hace mención. Sin embargo, se sabe que existen ciertas variables que se correlacionan con las evaluaciones, como lo son la calificación obtenida por el alumno, el tamaño del grupo, etc., lo que hace importante y necesario cuantificar el efecto de estas variables, esto es, la magnitud del sesgo que ocasionan en las evaluaciones que hacen los alumnos.

Cashin (1995) compilando estudios para verificar la validez de las opiniones de los alumnos obtenidas a través de encuestas, encontró que las evaluaciones de los alumnos estuvieron sesgadas (siendo mínimo el sesgo) por variables relacionadas con el profesor como: la edad, la experiencia como docente, el género, nacionalidad, personalidad y productividad como investigador. Aunado a ello, por variables relacionadas con el alumno como: la edad, el género, el semestre, la calificación del examen de admisión, la calificación esperada en el curso y la personalidad. Y por último, con características del grupo como: el tamaño y la hora en que se imparte.

Considerando los resultados del estudio de Cashin (1995), a continuación se repasan las variables identificadas en la literatura como que sesgan las evaluaciones de los alumnos, las cuales serán utilizadas en este estudio con el fin de validar el indicador utilizado en el ITESM, la Opinión Global Promedio, OGP.

2.1.2.1 RENDIMIENTO ACADÉMICO DE LOS ESTUDIANTES

En la literatura de Estados Unidos, ha sido bastante estudiado el sesgo que ocasionan las calificaciones de los alumnos sobre las evaluaciones que realizan éstos a sus profesores y se han encontrado resultados contradictorios. Una obvia relación, pero no necesariamente correcta, se ha investigado “los profesores compran las evaluaciones de los alumnos al asignarles buenas calificaciones”.

Feldman, (1976) demostró que la calificación obtenida por los alumnos no puede señalarse como un factor de sesgo en las evaluaciones que hacen ellos sobre sus profesores. Lo anterior porque encontró que las calificaciones estuvieron fuertemente correlacionadas con la evaluación del profesor en diferentes sentidos, es decir, en algunos cursos la relación fue positiva, en otros negativa y también sin relación alguna en otros más. Además, porque el sesgo incurrido si bien fue significativo, no explicó más del 9% de la varianza total de las evaluaciones de los alumnos.

Stumpf, Freedman y Aguanno (1979), encontraron que la evaluación del profesor, reportó una correlación menor a 0.15 con la calificación promedio de los estudiantes, concluyendo que su sesgo era mínimo. Marsh (1980 y 1984), señala que tal sesgo es mínimo si las evaluaciones de los alumnos son obtenidas con un instrumento apropiado.

Giglotti y Buchtel (1990), utilizando como unidad de análisis las evaluaciones individuales de los alumnos y la media del grupo, encontraron una relación positiva entre las calificaciones de los alumnos y la evaluación del profesor siendo más aparente en el nivel individual que en el grupo. También, encontraron que la discrepancia entre la calificación esperada del alumno y la obtenida fue un factor de sesgo para la evaluación del profesor; si bien la relación que encontraron fue significativa, explicó solamente entre el 7% y el 13% de la varianza de las dimensiones utilizadas para evaluar al profesor, por lo cual consideraron que el sesgo fue mínimo.

Cashin (1995) encontró en su compilación evidencia de que existe una correlación positiva entre las calificaciones esperadas por los alumnos en el curso y las evaluaciones que asignan a sus profesores, sin embargo ésta no fue relevante dado que se encontraba en el intervalo de 0.1 a 0.3.

En conclusión, en la literatura presentada se encontró que la relación entre la evaluación del profesor y las calificaciones de los alumnos fue positiva y significativa; no obstante, no fue muy representativa ya que esta variable explicó una proporción pequeña de la variabilidad del indicador utilizado para evaluar al profesor.

2.1.2.2 CARACTERÍSTICAS DEL CURSO Y DEL GRUPO

Feldman (1978) en cuanto a la disciplina del curso, encontró que los profesores impartiendo grupos de humanidades, artes finas e idiomas, obtenían mejores evaluaciones que profesores en disciplinas del área de ciencias sociales, física, matemáticas e ingeniería. En esto coincide Cashin (1995), su evidencia indicó que los profesores impartiendo grupos de humanidades recibieron mejores evaluaciones que los de ciencias sociales, adicional a ello, descubrió que estos profesores recibieron mejores evaluaciones que los de matemáticas.

Cranton y Smith (1986) demostraron que los efectos del tamaño de grupo y de la disciplina tuvieron gran influencia sobre la evaluación del profesor. En este punto Stumpf, *et al* (1979), encontraron en un análisis de causalidad que el tamaño de grupo afectaba negativamente la evaluación del profesor, sin embargo, esta variable junto con otras contempladas explicaron el 9% de la varianza total. Al igual, Cashin (1995), descubrió que los profesores que impartieron clases en grupos pequeños recibieron mejores evaluaciones que los que impartieron grupos grandes.

Aubrecht (1979), en su estudio de compilación, encontró que las correlaciones entre las evaluaciones de los alumnos sobre sus profesores y el tamaño de clase, estuvieron en un rango de 0.1 a 0.3 con un efecto negativo. Asimismo, descubrió que algunos estudios reportaban una relación no lineal donde las evaluaciones de los profesores impartiendo grupos de tamaño mediano eran menores a las obtenidas por profesores impartiendo grupos de tamaño pequeño o grande, al igual que Feldman (1978).

Wigington, Tollefson, Nona y Rodríguez, (1989), orientados en la evaluaciones individuales de los estudiantes, encontraron que los profesores impartiendo clases pequeñas recibieron mejores evaluaciones que los profesores de clases de tamaño mediano o grande.

Cranton y Smith (1990), encontraron en dos de sus unidades de análisis contempladas (las evaluaciones individuales de los alumnos y la media de clase), que el tamaño de grupo afectó significativamente y en forma negativa la evaluación del profesor. No obstante, estas variables explicaron no más del 3% y 4% de la varianza total.

Feldman (1984) en un análisis extenso de la literatura, reveló que en promedio la relación entre el tamaño del grupo y la evaluación del profesor fue negativa pero débil. Adicionalmente, encontró resultados interesantes cuando se correlacionaban con esta variable, dos factores estimados con un análisis de factores e interpretados como: la habilidad del profesor para presentar el material de clase, factor 1, así como para comunicar información, factor 2, encontrando también una relación negativa pero mucho más grande.

En resumen, la literatura reportó resultados consistentes en cuanto a la relación de esta variable con las evaluaciones de los alumnos, esto es, se encontró una relación negativa si bien fue significativa no se consideró representativa ni tampoco un factor de sesgo considerable.

2.1.2.3 CARACTERÍSTICAS DEL PROFESOR (GÉNERO, EDAD, AÑOS DE EXPERIENCIA EN EL INSTITUTO Y TIPO DE CONTRATO)

GÉNERO

Stumpf *et al* (1979), en un análisis de causalidad observaron que el efecto del género del profesor no estuvo presente, es decir, no afectaba las evaluaciones de los profesores. En contraparte, Wigington *et al* (1989) encontraron que las profesoras recibían mejores evaluaciones que los profesores y que a su vez éstas obtenían mejores evaluaciones en grupos pequeños, reportando lo contrario en grupos grandes.

Freeman (1994), mediante un análisis de varianza encontró que los alumnos preferían profesores andróginos, los cuales son profesores que sin importar su género poseen atributos tanto femeninos como masculinos en su forma de enseñar, sobre todo en el área de ciencias (naturales y sociales); este resultado indicó que el género del profesor fue menos importante que los atributos del profesor al enseñar.

Basow (1995) encontró una relación entre el género del profesor y el del alumno; específicamente el género del alumno presentó influencia sobre la evaluación de las profesoras pero no en la evaluación de los profesores, reportando que las alumnas evaluaban mejor a sus profesoras, mientras que los alumnos las evaluaban peor.

De la misma forma, Cashin (1995) encontró en su análisis de compilación una relación negativa entre el género del profesor y el del alumno, esto es, los alumnos evaluaban mejor a los profesores y las alumnas evaluaban mejor a las profesoras.

La evidencia presentada por la bibliografía revisada mostró resultados diversos sobre el efecto del género del profesor en las evaluaciones de los alumnos, el cual estuvo relacionado con el género del alumno.

EDAD DEL PROFESOR

Feldman (1983) observó en los estudios relacionados con esta variable, que la edad del profesor afectó negativamente la evaluación de los alumnos, esto es, los alumnos evaluaban peor a los profesores de edad avanzada. Estos resultados también fueron evidenciados por Cashin (1995).

Sí bien existe poca literatura sobre este tema, la relación encontrada es negativa, es decir, la edad del profesor afecta negativamente la evaluación del profesor, lo cual indica que esta variable no cuantifica los años de experiencia del profesor a lo largo de su carrera profesional ya que si fuera así, su relación sería positiva.

AÑOS DE EXPERIENCIA EN EL INSTITUTO

Stumpf, *et al* (1979) encontraron que la correlación de esta variable con la evaluación del profesor fue negativa, pero no significativa, por lo que en su análisis de causalidad realizado no concluyeron que sesgaba a la evaluación del profesor.

Feldman (1983) observó que la variable de experiencia del profesor estaba relacionada positivamente con las evaluaciones de los alumnos. Aún y cuando esta relación fue débil, fue suficientemente robusta para mantener su efecto cuando se incluían otras variables como el tamaño del grupo y el género del profesor. También, indicó que la relación puede ser subestimada por considerar sólo una relación lineal ya que esta puede ser curvilínea.

Cashin (1995) al analizar las evaluaciones de una muestra de profesores durante 13 años, encontró que la experiencia del profesor no es una variable determinante de las evaluaciones de los alumnos, ya que no hubo cambios sistemáticos a través de los años.

Estos estudios señalan que la relación entre la evaluación del profesor y sus años de experiencia, no fue consistente. No obstante, se espera que esta variable afecte positivamente a la evaluación del profesor, esto porque la experiencia le da bases sólidas para que ofrezca una mejor enseñanza.

TIPO DE CONTRATO

Se encontró sólo un estudio donde se analizó el efecto del tipo de contrato sobre la evaluación del profesor, Cranton y Smith (1990) incluyeron variables binarias para cuantificar el efecto de los profesores de tiempo completo y parcial, observando que los de tiempo completo obtuvieron peores evaluaciones que los de tiempo parcial, a este respecto indicaron que el sesgo incurrido fue mínimo, porque éstas variables explicaron muy poco la varianza total.

Los estudios presentados hasta aquí, han permitido identificar las variables que comúnmente han sido utilizadas para cuantificar el sesgo en las evaluaciones de los alumnos, las cuales serán utilizadas en este estudio para determinar la validez del indicador utilizado en el ITESM para evaluar a los profesores.

Es por este motivo, que en este trabajo se estudiará el efecto de las variables internas sobre la OGP y aunado a ello el de las siguientes variables externas, las que permitirán identificar mediante un análisis de regresión si la OGP es un indicador válido para evaluar el desempeño de los profesores:

- Variables externas relacionadas al curso y grupo: disciplina a la que pertenece el curso, el tamaño de grupo, el efecto del tipo de grupo (rediseñado o tradicional) y la calificación promedio del grupo.
- Variables externas relacionadas al profesor: el género, años de experiencia en el Instituto, edad, tipo de contrato y grado máximo de estudios.

2.2 METODOLOGÍAS UTILIZADAS PARA COMPROBAR LAS HIPÓTESIS DEL ESTUDIO

La metodología empleada en el presente estudio es cuantitativa y se basa en tres tipos de análisis: descriptivo, de regresión y de factores. Con los que se probarán las dos hipótesis establecidas en este estudio:

Hipótesis 1: *Las opiniones de los alumnos son confiables para evaluar y retroalimentar el desempeño del profesor*

Hipótesis 2: *El indicador utilizado para evaluar al profesor, OGP, es válido, esto es, las opiniones de los alumnos son validas para evaluar el desempeño del profesor.*

Los tres tipos de análisis se realizaran para el nivel de profesional con datos obtenidos del Campus Monterrey, utilizando como unidad de estudio la media aritmética de las evaluaciones de los alumnos en un grupo, pero sólo se consideraran los cursos teóricos.

2.2.1 METODOLOGÍA EMPLEADA PARA COMPROBAR LA CONFIABILIDAD DE LAS OPINIONES DE LOS ALUMNOS

La teoría clásica sobre la confiabilidad supone que la evaluación obtenida por un profesor por parte de sus alumnos, se compone de una evaluación real más un error no sistemático de medida. Si bien la evaluación real nunca puede medirse con exactitud, puede estimarse a partir de la evaluación observada. De igual forma, se supone que la varianza de las evaluaciones del profesor observadas en un grupo (S_{obs}^2) es igual a la varianza de las evaluaciones reales (S_{real}^2) más la varianza de errores no sistemáticos de medición (S_{err}^2), Aiken (1996):

$$S_{obs}^2 = S_{real}^2 + S_{err}^2 \quad (1)$$

De este modo, la confiabilidad de las opiniones de los alumnos en un grupo (r) se define como la proporción de la varianza observada que es explicada por la varianza real, Aiken (1996):

$$r = \frac{S_{real}^2}{S_{obs}^2} \quad (2)$$

Esta proporción, es estimada de diversas formas y a continuación se presenta la que será utilizada en este estudio.

2.2.1.1 COEFICIENTE DE CONSISTENCIA INTERNA

Para encontrar la confiabilidad de las opiniones de los alumnos en un grupo, este estudio se basa en la consistencia de las evaluaciones de los alumnos en todas las preguntas de la encuesta (consistencia interna), que puede ser medida mediante el coeficiente alfa de Cronbach (Anastasi et al, 1998 y Aiken, 1996). El procedimiento consiste en calcular el coeficiente con base a la siguiente formula:

$$r_j = \left(\frac{n}{n-1} \right) \frac{DE_t^2 - \sum(DE_i^2)}{DE_t^2} \quad (3)$$

donde,

- r_i es el coeficiente de confiabilidad en el grupo “j”
 n es el número de preguntas en la encuesta
 DE_t^2 es la varianza de las evaluaciones de los alumnos en toda las preguntas en el grupo “j”
 $\sum(DE_i^2)$ es la suma de las varianzas de las evaluaciones de los alumnos en cada pregunta de la encuesta.

Así, por ejemplo, un coeficiente de confiabilidad de 0.85 significa que el 85% de la varianza de las evaluaciones de los alumnos sobre un profesor en un grupo, depende de la varianza verdadera en el rasgo medido y que el 15% depende de la varianza del error.

Una vez obtenido el coeficiente de confiabilidad en cada uno de los grupos considerados, se podrá realizar un promedio para obtener la confiabilidad de las opiniones de los alumnos del Campus Monterrey. Entonces, si el promedio es bajo, se podrá rechazar la primera hipótesis planteada en este estudio, concluyendo que las opiniones de los alumnos no son confiables para evaluar y retroalimentar el desempeño del profesor, en cambio si es alto la conclusión será que si son confiables.

Aunado a lo anterior, con los cálculos realizados se podrá determinar cuales preguntas de la encuesta fueron más precisas para retroalimentar al profesor, que serán las preguntas en donde las opiniones de los alumnos sean muy similares, es decir, donde la variabilidad de las evaluaciones sea cercana a cero. A su vez, se podrá indagar cuales fueron las preguntas mejor o peor evaluadas en el ejercicio académico considerado.

2.2.2 METODOLOGÍA EMPLEADA PARA COMPROBAR LA VALIDEZ DE LAS OPINIONES DE LOS ALUMNOS

Para probar la validez de las opiniones de los alumnos se utilizaran dos análisis: el modelo de regresión múltiple y el análisis de factores.

Con el modelo de regresión, se podrá investigar qué tanto las variables internas y externas al proceso de enseñanza-aprendizaje están relacionadas con la evaluación global del profesor, la Opinión Global Promedio, OGP. Además, se podrá determinar qué variables internas predeterminan las evaluaciones de los alumnos, así como aportar evidencia para rechazar o no la segunda hipótesis de este estudio, la relacionada con la validez del indicador utilizado para evaluar al profesor, esto es, la evidencia consistirá en demostrar si este indicador está siendo sesgado o no por variables externas al proceso de enseñanza-aprendizaje.

Por otro lado, el análisis de factores es comúnmente utilizado para resumir la información de “p” variables, en este caso “p” preguntas específicas en un número pequeño de factores. De esta forma, mediante una combinación lineal puede resumirse la información de todas las preguntas específicas de la encuesta dentro de unos cuantos factores, que pueden ser interpretados con base a las preguntas que están más representadas en ellos. Por ejemplo, si la información de un factor, está representada fuertemente por preguntas que se relacionan con la comunicación del profesor hacia sus alumnos, entonces el factor puede interpretarse como el factor comunicación.

Lo anterior permite conocer el perfil del profesor sugerido por los alumnos y hace posible, una vez identificado, retroalimentar al profesor.

Además, como los factores estarán representados por las preguntas específicas de la encuesta y como éstas cuantifican aspectos internos del proceso de enseñanza-aprendizaje, se espera que sean fuentes de explicación de la evaluación global del profesor, OGP. Así, al utilizarlos en un modelo de regresión como variables explicativas, se desea modelen de mejor forma a la OGP que las preguntas específicas, siempre y cuando hayan resumido bien la información de éstas.

2.2.2.1 EL MODELO DE REGRESIÓN LINEAL MÚLTIPLE

El modelo de regresión lineal para la observación “i” puede ser escrito de la siguiente manera:

$$y_i = \beta_1 + \beta_2 x_{i2} + \dots + \beta_k x_{ik} + \dots + \beta_K x_{iK} + \varepsilon_i \quad (4)$$

donde “y” es la variable dependiente (a explicar), en este estudio, la Opinión Global Promedio, OGP; las “x’s” son las variables independientes (explicativas) en este caso, las variables tanto internas como externas al proceso de enseñanza-aprendizaje y “e” es el error de la ecuación. β_2 hasta β_K son los parámetros que indican los efectos de “x” sobre “y”; β_1 es el intercepto. En notación matricial:

$$y = X\beta + \varepsilon \quad (5)$$

donde “ y ” y “ e ” son vectores de tamaño $N \times 1$, X una matriz de $N \times K$, y “ β ” un vector de $K \times 1$. Para la fila “ i ” de X , el modelo puede ser escrito como:

$$y_i = \mathbf{x}_i \beta + \varepsilon_i \quad (6)$$

A su vez, los siguientes supuestos concluyen el modelo:

1. Linealidad: y esta linealmente relacionada a las x 's en todos los parámetros β .
2. Colinealidad: las x 's no son linealmente dependientes.
3. Valor esperado de e : $E(e_i \mid \mathbf{x}_i) = 0$ para todo “ i ”.
4. Homocedasticidad: Para una colección dada de \mathbf{x}_i , los errores tienen varianzas constante, es decir, $\text{Var}(e_i \mid \mathbf{x}_i) = \sigma^2$ para todo “ i ”.
5. Errores no correlacionados: Para dos observaciones “ i ” y “ j ”, la covarianza entre e_i y e_j es cero.

Con estos supuestos, el estimador de Mínimos Cuadrados Ordinarios, MCO, es

$$\hat{\beta} = (X'X)^{-1} X'y, \quad (7)$$

el cual tiene la siguiente matriz de varianzas y covarianzas.¹

$$\text{Var}(\hat{\beta}) = (X'X)^{-1} X'\Phi X(X'X)^{-1} \quad (8)$$

donde Φ es una matriz con $\phi_{ii} = \text{Var}(\varepsilon_i)$ y debido a que es satisfecho el supuesto de que los errores son homocedásticos, Φ puede ser escrita como $\Phi = \sigma^2 I$. Con este supuesto, la matriz de varianzas-covarianzas de los estimadores de MCO puede ser reducida, como:

$$\text{Var}(\hat{\beta}) = \sigma^2 (X'X)^{-1} \quad (9)$$

Con la cual pueden hacerse inferencias sobre los parámetros estimados mediante MCO.

Una vez establecido el modelo, es preciso explicar como se aplicará al análisis de este estudio.

¹ Para una explicación detallada del modelo básico de regresión, sus supuestos y su estimación mediante Mínimos Cuadrados Ordinarios, MCO, ver Gujarati (1995) o Maddala(1992).

Primero, para conocer el efecto de las variables internas del proceso de enseñanza-aprendizaje, se estimará un modelo de regresión donde la variable dependiente, a explicar, será la Opinión Global Promedio, OGP, y las variables explicativas las preguntas específicas de la encuesta. Esto permitirá determinar cuáles son las preguntas de la encuesta que intervienen en el proceso de decisión cuando los alumnos expresan su opinión global sobre sus profesores.

Segundo, para verificar la validez de la encuesta, es decir, investigar si la OGP realmente está cuantificando el desempeño del profesor y no está siendo sesgada por variables externas al proceso de enseñanza-aprendizaje, se especificará un modelo de regresión en el que se buscará explicar a la OGP con variables externas, las que fueron identificadas en la evidencia.

2.2.2.2 ANÁLISIS DE FACTORES

El modelo de factores utilizado en este estudio puede ser escrito de la siguiente forma para la pregunta “i” donde ésta va desde la pregunta 1 hasta la “p” (sólo se utilizan las preguntas específicas de la encuesta en este análisis).²

$$x_i - \mu_i = \lambda_{i1}F_1 + \lambda_{i2}F_2 + \dots + \lambda_{im}F_m + \varepsilon_i \quad (10)$$

El vector aleatorio observable \mathbf{x} , con “p” componentes, tiene media μ y matriz de varianzas y covarianzas Σ . El modelo postula que \mathbf{x} es linealmente dependiente de unas cuantas variables aleatorias no observables, F_1, F_2, \dots, F_m llamadas factores comunes, y “p” recursos adicionales de variación $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_p$, llamados errores o factores específicos.

En particular, el modelo de análisis de factores matricialmente es:

$$\underset{(px1)}{\mathbf{x} - \mu} = \underset{(pxm)}{\mathbf{L}} \underset{(mx1)}{\mathbf{F}} + \underset{(px1)}{\boldsymbol{\varepsilon}} \quad (11)$$

donde ,

- μ_i es la media de la pregunta “i”
- ε_i es el factor específico “i”
- F_j es el factor común j
- λ_{ij} es la carga de la pregunta i sobre el factor j

Los supuestos que completan el modelo son:

$$E(\mathbf{F}) = \mathbf{0}, \text{Cov}(\mathbf{F}) = \mathbf{I}$$

² Para una explicación detallada sobre el análisis de factores ver Johnson y Wichern (1992) o Jobson (1992).

$E(\varepsilon) = 0$, $Cov(\varepsilon) = \Psi$, donde Ψ es una matriz diagonal.

La estructura de varianzas y covarianzas del modelo es:

$$Cov(x) = LL' + \Psi \quad (12)$$

lo que es igual, a:

$$\begin{aligned} \text{Var}(x_i) &= \lambda_{i1}^2 + \dots + \lambda_{im}^2 + \Psi_i \\ \text{Cov}(x_i, x_j) &= \lambda_{i1}\lambda_{j1} + \dots + \lambda_{im}\lambda_{jm} \end{aligned}$$

$$Cov(x, F) = L \quad (13)$$

o simplemente:

$$Cov(x_i, F_j) = \lambda_{ij}$$

Con esta información se puede explicar que la proporción de la varianza de la pregunta “i” aportada por los “m” factores comunes es llamada comunalidad. Por otra parte, la proporción de la varianza, debido al factor específico, es llamada la varianza específica.

Esto es,

$$\text{Var}(x_i) = \underbrace{\lambda_{i1}^2 + \dots + \lambda_{im}^2}_{\text{Comunalidad}} + \underbrace{\Psi_i}_{\text{Varianza específica}} \quad (14)$$

El modelo de factores asume que la matriz de varianzas y covarianzas de las “p” preguntas, puede ser explicada por las p x m cargas λ_{ij} y por las “p” varianzas específicas, Ψ_i .

El análisis de factores obtiene estimaciones únicas de L y Ψ y una vez que se han estimado las cargas y las varianzas específicas, se elabora con ellas la matriz de cargas la cual es rotada (multiplicada por una matriz ortogonal) con el fin de identificar los factores, Además, se utilizan para estimar los valores de los factores, los que resumen la información de las preguntas específicas de la encuesta.

Es posible distinguir dos tipos de análisis de factores: el análisis factorial exploratorio y el confirmatorio. En el primero, no se conocen los factores “a priori”, sino que se determinan mediante el análisis factorial y, por el otro lado, en el segundo, se propone “a priori” un modelo, según el cual hay unos factores que representan a las variables originales y se somete a comprobación el modelo.

En este estudio, se utilizará el análisis factorial exploratorio, debido a que en el ITESM no existen estudios que hayan identificado las dimensiones del profesor que son importantes para los alumnos, es decir, los factores no observables que se conforman como combinaciones lineales de las preguntas de la encuesta.

El análisis de factores permitirá concentrar la información aportada por las preguntas específicas en unos cuantos factores, los que al ser identificados ayudarán a determinar el perfil del profesor sugerido por los alumnos, después al utilizarlos en un modelo de regresión donde se pretenda explicar la evaluación global del profesor, OGP, se podrá corroborar si ésta es modelada mejor con las preguntas específicas o con los factores estimados.

CAPÍTULO 3

CONFIABILIDAD DE LAS OPINIONES DE LOS ALUMNOS PARA EVALUAR Y RETROALIMENTAR EL DESEMPEÑO DE LOS PROFESORES

3.1 ANÁLISIS DE LAS PREGUNTAS DE LA ENCUESTA

Para investigar la confiabilidad de la encuesta utilizada para retroalimentar y evaluar al profesor con base en las opiniones de los alumnos, se utilizó la información reportada por la que fue aplicada en el Campus Monterrey en el nivel de profesional y en el ejercicio académico de Agosto-Diciembre de 2001.

Se consideró un total de 12,135 alumnos que contestaron la encuesta y que estuvieron inscritos en 2882 grupos de cursos teóricos impartidos por 1136 profesores.³

La encuesta analizada constó de 12 preguntas específicas y 1 global.⁴ Las preguntas específicas están relacionadas con el desempeño del profesor en ciertos puntos que permiten al alumno tener una visión para prepararlo a contestar la pregunta global, donde se pretende cuantificar el desempeño global del profesor.

Los alumnos para emitir sus opiniones utilizaron una escala de Likert de 1 a 7 en cada pregunta (donde el valor de 1 significa total acuerdo sobre el desempeño y el valor de 7 total desacuerdo). Así, si las respuestas de los alumnos en un grupo en la pregunta global de la encuesta se promedian con una simple media aritmética, se calcula el indicador con el cual se evalúa el desempeño del profesor, la Opinión Global Promedio (OGP).⁵

Dentro de las preguntas específicas el tesista identificó a dos tipos de preguntas, las objetivas y las subjetivas. En las primeras, el alumno tiene una percepción tangible del desempeño del profesor y es difícil que la visualización de los alumnos sea diferente, ya que hay parámetros específicos que se toman como base para evaluar al profesor. En las segundas, las opiniones de los alumnos pueden ser muy diferentes ya que se pretende cuantificar el desempeño del profesor pero sin brindar al alumno ningún parámetro, lo que hace que su percepción varíe mucho. Sin embargo, en ambos tipos de preguntas las

³ Para el análisis de esta investigación se tomaron en cuenta sólo los cursos teóricos, se excluyen los grupos de cursos de la División de Ciencias de la Salud, de la Dirección de Programas de Apoyo a la Misión y de la Dirección de Programas Internacionales.

⁴ La definición de las preguntas fue obtenida del “Manual de ECOAS” de la Dirección de Información Institucional, Sistema ITESM (2001).

⁵ Al decir “respuestas” se hace referencia a las evaluaciones que los alumnos hicieron sobre sus profesores en cualquier pregunta de la encuesta.

opiniones de los alumnos pueden ser distorsionadas por variables externas al proceso de enseñanza-aprendizaje que nada tienen que ver con el desempeño del profesor.

En la tabla 3.1, pueden observarse las preguntas específicas de la encuesta, ésta contiene 4 objetivas (la 2, 3, 9 y 10) y 8 subjetivas (1, 4, 5, 6, 7, 8, 11 y 12), además de la pregunta 14 la global, que se consideró subjetiva, con la que es elaborada la OGP.⁶

Tabla 3.1 Cuestionario aplicado a los alumnos de profesional para cuantificar el desempeño de los profesores en el ejercicio académico de Agosto-Diciembre 2001

Pregunta	Descripción	Tipo de pregunta	
1	Manifiesta dominio de la disciplina académica correspondiente al curso teórico que imparte o al curso que coordina.	Específica	Subjetiva
2	Asiste a las sesiones establecidas, es puntual y cumple con el servicio de asesoría acordado.	Específica	Objetiva
3	Ha cumplido hasta ahora con el programa del curso, así como con las políticas generales que se establecieron al inicio.	Específica	Objetiva
4	Fomenta una comunicación efectiva, se mantiene abierto(a) al diálogo y crea un ambiente de confianza y respeto.	Específica	Subjetiva
5	Evalúa de manera justa y objetiva mi trabajo y cumple con los criterios de evaluación establecidos al inicio del curso.	Específica	Subjetiva
6	Promueve que los alumnos y las alumnas participen activamente con sus ideas y opiniones en el desarrollo del curso.	Específica	Subjetiva
7	Fomenta el aprendizaje individual y colaborativo, de conocimientos relevantes, conectados con situaciones reales y sustentado en información actualizada y confiable.	Específica	Subjetiva
8	Promueve la reflexión, el cuestionamiento, el análisis crítico y la creatividad a través de las actividades del curso	Específica	Subjetiva
9	Practica y fomenta el ejercicio de aquellos valores, habilidades y actitudes de la Misión ITESM-2005 que estableció como relevantes para el curso.	Específica	Objetiva
10	Utiliza y fomenta el empleo de recursos tecnológicos adecuados para enriquecer y facilitar el proceso de aprendizaje del curso.	Específica	Objetiva
11	Muestra una actitud comprometida y motivadora en la impartición y coordinación del curso.	Específica	Subjetiva
12	Organiza e imparte el curso de manera que favorece el aprendizaje.	Específica	Subjetiva
14	Tomando en cuenta todos los aspectos anteriores, el profesor o la profesora se desempeña de manera excelente en este curso.	Global	Subjetiva

Dentro de las preguntas específicas objetivas los parámetros de comparación que se consideraron son presentados en la tabla 3.2.

⁶ Se excluye la pregunta 13 que es para laboratorios y la 15 que trata acerca del valor agregado del curso.

Tabla 3.2 Parámetros de comparación de las preguntas específicas objetivas

Pregunta	Parámetros de comparación
2	Asistencia a clases a la hora establecida y cumplimiento de las horas de asesoría.
3	Cumplimiento con el contenido del curso establecido en los planes analíticos y sintéticos entregados al alumno al inicio del curso.
9	Objetivos establecidos en la Misión de ITESM-2005
10	Uso de programas computacionales, biblioteca digital, etc.

Una vez presentado el cuestionario, es importante conocer los resultados que aportó esta encuesta, esto es, la distribución de las evaluaciones de los alumnos en cada una de las preguntas así como la evaluación promedio, con el fin de determinar los valores de la escala más utilizados.

A su vez, es relevante investigar la variabilidad de las evaluaciones de los alumnos en las preguntas, ya que aquellas con menor variabilidad serán las que brinden una mejor retroalimentación al profesor.

3.1.1 PATRONES DE RESPUESTA DE LAS EVALUACIONES DE LOS ALUMNOS POR PREGUNTA

Los patrones de respuesta de las evaluaciones de los alumnos en cada pregunta son presentados en el gráfico 3.1, estos patrones se obtuvieron al sumar todas las respuestas (evaluaciones de alumnos) reportadas en los valores de la escala utilizada para evaluar al profesor, en todos los grupos de profesional considerados.

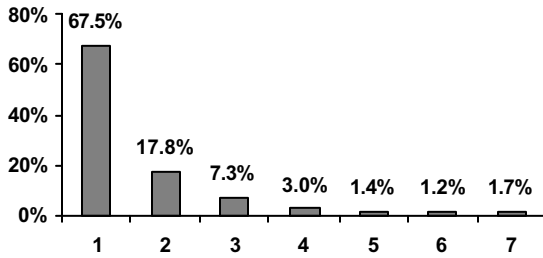
El patrón de respuestas se grafica en porcentajes, por ejemplo, en la pregunta 1 el porcentaje en el valor de 1 de la escala corresponde al 67.5% del total de respuestas sumadas en profesional (67378) en esa pregunta, lo que equivale a 45494 respuestas.

Como puede observarse en los gráficos, predominaron las respuestas en el valor de 1 de la escala, disminuyéndose éstas a medida que este valor se acerca a 6, reportando un ligero incremento en el de 7. Este ligero incremento encontrado al final de la escala, puede considerarse como anormal debido a la tendencia decreciente que se observa y es posible que sea atribuible a que muchos alumnos están expresando su inconformidad al evaluar con 7 a sus profesores, ya sea por su mal desempeño, por resentimientos o por confundir el significado de la escala.

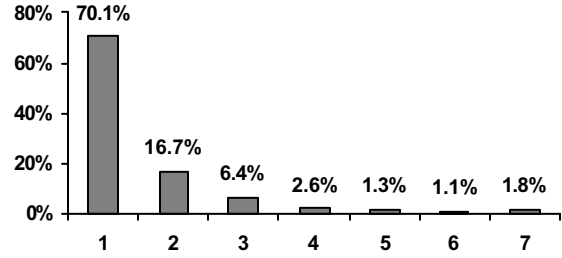
Las tres preguntas con patrones de respuestas más afectados por evaluaciones en el valor de 7 de la escala fueron la 12, 4 y la 11 en orden descendente y las menos afectas la 9, 7 y la 3, también en orden descendente, como se muestra en el gráfico 3.2 donde se presentan los porcentajes reportados en el gráfico 3.1 en el valor de 7 de la escala de los patrones de respuesta de cada una de las preguntas.

Grafico 3.1 Patrones de respuesta de las preguntas de la encuesta

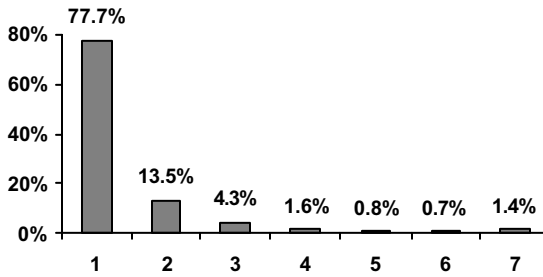
Pregunta 1, Total de Respuestas 67378



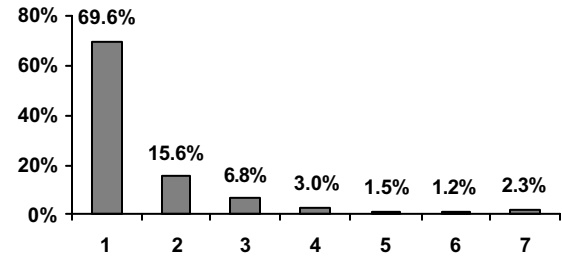
Pregunta 2, Total de Respuestas 67206



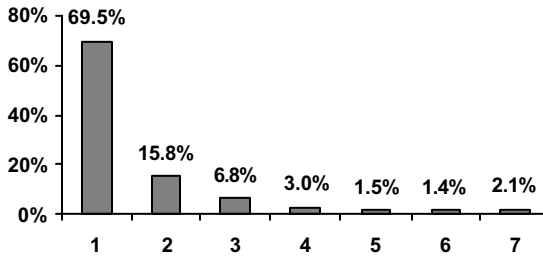
Pregunta 3, Total de Respuestas 67246



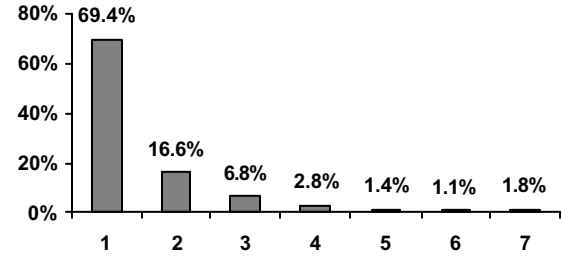
Pregunta 4, Total de Respuestas 67196



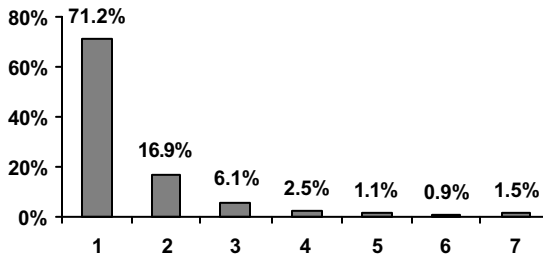
Pregunta 5, Total de Respuestas 66974



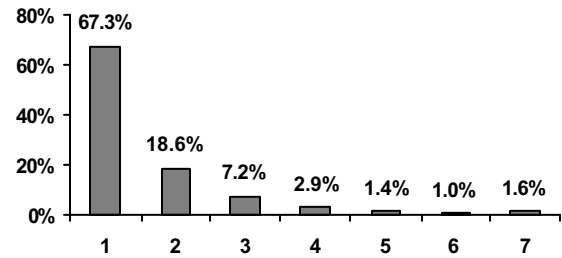
Pregunta 6, Total de Respuestas 66814



Pregunta 7, Total de Respuestas 66831



Pregunta 8, , Total de Respuestas 66579



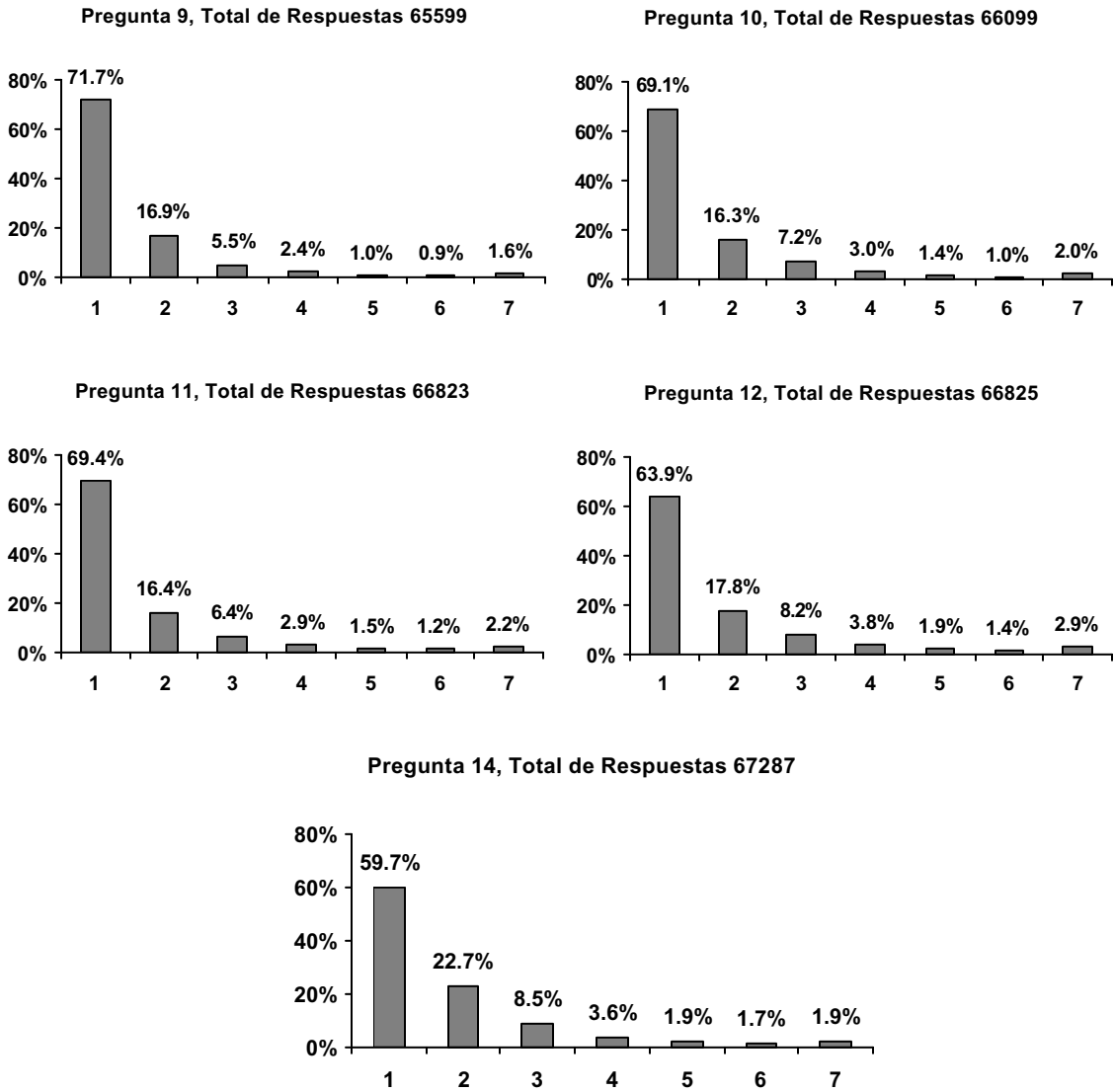
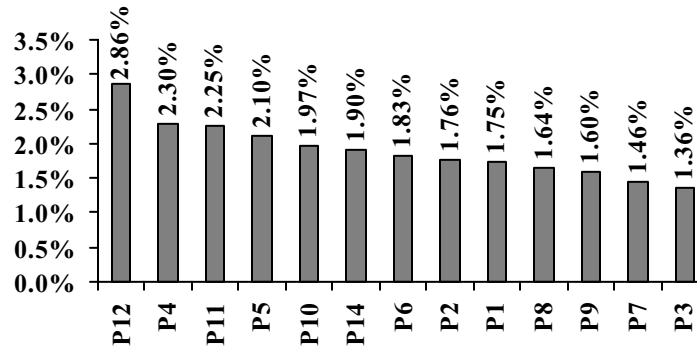


Gráfico 3.2 Porcentaje de respuestas por pregunta en el valor de 7 de la escala con respecto a su total de respuestas



Si se ubican las preguntas objetivas en el gráfico 3.2 puede observarse que la mayor parte de ellas la 2, 9 y la 3 se encuentran del lado derecho del gráfico, esto es, dentro de las preguntas con patrones de respuesta menos afectados por evaluaciones en el valor de 7 de la escala.

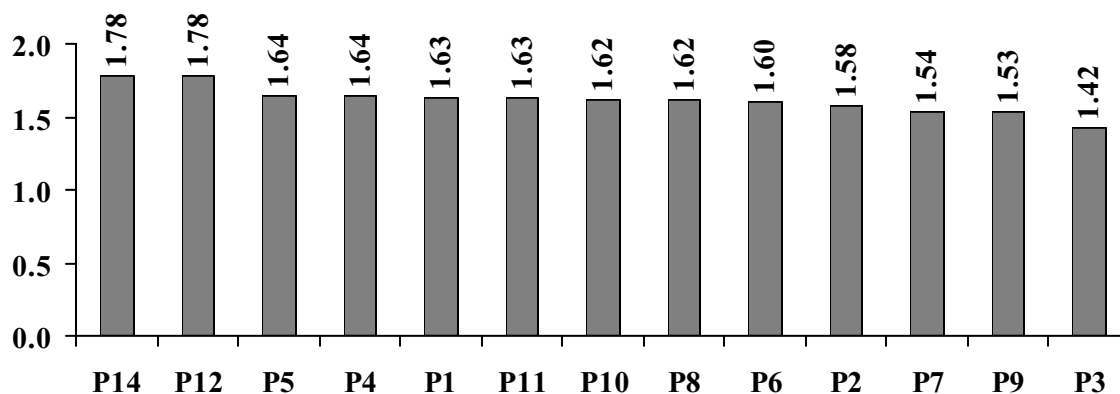
Este incremento al final de la escala tiene efectos negativos en la evaluación global del profesor, esto porque la OGP es la simple media aritmética de los respuestas reportadas en los valores de la escala en un grupo o en cualquier división o agrupación de grupos, la cual es afectada cuando existen observaciones extremas.

Por este motivo, es conveniente analizar si el incremento en el valor de 7 es significativo y sobre todo si es ocasionado por variables que nada tienen que ver con el desempeño del profesor, lo cual se investigará en el siguiente capítulo al probar la validez de las opiniones para evaluar el desempeño del profesor.

3.1.2 EVALUACIÓN REALIZADA POR LOS ALUMNOS POR PREGUNTA

Para comenzar este análisis se presenta el gráfico 3.3 donde se reportan los promedios de las preguntas elaborados con base en las opiniones de todos los alumnos que contestaron la encuesta en los grupos de profesional. Así, en el caso particular de la pregunta 14, el número reportado es la OGP del Campus.

Gráfico 3.3 Evaluación promedio de las preguntas de la encuesta



Es posible observar, que el promedio de las evaluaciones de los alumnos en las preguntas de la encuesta fue muy similar, la peor evaluación fue de 1.78 y la mejor de 1.42. Considerando esta información, se distingue que las preguntas objetivas se encuentran en su mayoría concentradas en el lado derecho del gráfico indicando que fueron las que resultaron mejor evaluadas.

Más aún, con esta información y considerando las tres preguntas mejor y peor evaluadas se puede decir que en el nivel de profesional:

- Los profesores se desempeñaron mejor cumpliendo con el programa y las políticas del curso (pregunta 3); fomentando la misión del ITESM (pregunta 9) y promoviendo el aprendizaje individual y colaborativo (pregunta 7).
- En cambio, necesitan mejorar la manera en que organizan e imparten el curso (pregunta 12) y su forma de evaluar (pregunta 5), aspectos que afectan la evaluación global del profesor.

Observando las evaluaciones de las preguntas y calculando el promedio de éstas, que es de 1.62, surge la pregunta del porqué los alumnos utilizan por lo general los valores de la escala cercanos a 1 y 2 para evaluar a sus profesores.

Esto puede deberse a que los alumnos conocen la evaluación que las personas administrativas tienen como límite para considerar un buen desempeño, 2.5, la cual da poco margen para que los alumnos exploten todos los valores de la escala.

También, puede ser el resultado de que no están definidos todos los valores de la escala, en el periodo analizado sólo estaba definido el valor de 1, totalmente de acuerdo, y el de 7, en total desacuerdo, lo que puede evitar aprovechar el valor agregado de los restantes valores.

O bien, simplemente puede estar sucediendo que los alumnos no son muy críticos para evaluar a sus profesores ya sea por falta de interés o porque en su mayoría existen preguntas subjetivas, las que son más difíciles de contestar.

Adicionalmente, surge la pregunta del porqué los alumnos evalúan a sus profesores de una manera muy similar en todas las preguntas de la encuesta. Para verificar esto, se calcula por alumno la desviación estándar de las evaluaciones realizadas en las preguntas de la encuesta, para después obtener un promedio de las desviaciones de todos los alumnos, el cual fue de 0.52 (ver tabla 3.3 para una ejemplificación de los cálculos) valor que indica que la variabilidad en las evaluaciones de los alumnos asignadas en las preguntas es baja.

Los motivos que se consideran en este estudio que ocasionan que los alumnos no se esfuercen al evaluar críticamente a sus profesores, son: la falta de interés ya sea porque las preguntas son muy subjetivas o porque los alumnos no encuentran las preguntas que realmente desean contestar, o simplemente debido a que invierten demasiado tiempo al contestar la encuesta.

Scriven (1994) señala ciertas circunstancias que pueden afectar la confiabilidad de la encuesta, dentro de éstas indica que la encuesta puede no contener las preguntas que los alumnos desean responder, lo cual genera descontento y falta de voluntad de contestar correctamente las preguntas.⁷

⁷ Este autor es citado por Díaz (1997).

Tabla 3.3 Cálculos para obtener la desviación estándar promedio de las evaluaciones de los alumnos en todas las preguntas de la encuesta

Grupos	Alumnos	P1	P14	Desviación estándar
Grupo 1	Alumno 1,1	P1,1,1	P14,1,1	Desviación estándar Alumno 1,1

	Alumno 1, n ₁	P1,1,n ₁	P14,1,n ₁	Desviación estándar Alumno 1,n ₁
....
Grupo m	Alumno m,1	P1,m,1	P14,m,1	Desviación estándar Alumno m,1

	Alumno m,n _m	P1,m,n _m	P14,m,n _m	Desviación estándar Alumno m,n _m
					Promedio de desviaciones

Se cuenta con un total de 2882 grupos ($m = 2882$) y dentro de cada grupo con n_i alumnos, donde el subíndice “i” indica que el número de alumnos por grupo difiere. También, se cuenta con las evaluaciones realizadas por los alumnos en cada pregunta, esto es, por ejemplo el término P1,1,1 representa la evaluación realizada en la pregunta 1, en el grupo 1, por el alumno 1 de un total de n_1 alumnos que contestaron la encuesta.

De esta manera, sería muy importante un estudio como el realizado por el Campus Monterrey (1998) para investigar cuales son las preguntas que los alumnos desean que se incluyan en la encuesta, con el fin de mejorar la retroalimentación del profesor.

A su vez, es recomendable que la encuesta se enfoque únicamente en obtener evaluación y retroalimentación de los alumnos sobre el desempeño de sus profesores, donde las preguntas estén formuladas considerando obtener la mayor (y mejor) retroalimentación (y evaluación) posible sobre su labor, sin que el tiempo que el alumno tenga que invertir en su respuesta sea excesivo.

En la actualidad, además de las 15 preguntas de la encuesta se incluye lo siguiente en el cuestionario aplicado a los alumnos:

- Un apartado para que los alumnos emitan comentarios a cada uno de los profesores de sus grupos.
- Si el alumno lleva algún grupo de Laboratorio, 10 preguntas más.
- Un apartado para enviar comentarios al almacenista e instructor del Laboratorio.
- 12 preguntas acerca de servicios que el ITESM ofrece (deportes, difusión cultural, prevención de adicciones) y del desempeño de los principales directivos (director de carrera, división o Campus).
- Un apartado para emitir comentarios a los principales directivos (director de carrera, división o Campus, Rector de Zona y Rector del Sistema).
- Una sección de preguntas para evaluar los servicios internos del Campus (mantenimiento a áreas verdes, seguridad, escolar, tesorería, etc) con un total de 7 reactivos.

- Además, si el alumno lleva alguna materia satelital, 15 preguntas más sobre los servicios que en específico ofrece la Universidad Virtual.

Como se puede observar, debido a las partes anexas a la evaluación de los profesores de cursos teóricos, el tiempo que el alumno invierte en responder se incrementa, tomando aproximadamente desde 40 minutos hasta 1 hora 15 minutos, lo que puede estar afectando negativamente a la evaluación que los alumnos otorgan a sus catedráticos al hacérselos la encuesta tediosa o cansada.

3.1.3 VARIABILIDAD DE LAS EVALUACIONES DE LOS ALUMNOS POR PREGUNTA

Las opiniones de los alumnos sobre el desempeño de un profesor se pretende sean muy similares, esto es, sean consistentes con lo que se está intentando medir. Si éstas difieren puede deberse a que las preguntas son muy subjetivas o a que hay factores contaminantes que afectan las opiniones de los alumnos.

Una pregunta que reporte mucha variabilidad no brindará una buena retroalimentación al profesor, debido a que no abra consenso en las opiniones de los alumnos.

En este sentido, para determinar las preguntas que brindan una mejor retroalimentación al profesor, es posible utilizar una medida de variabilidad de las opiniones de los alumnos en un grupo para después obtener una medida global en cada pregunta.

En la tabla 3.4 se presenta como se calcula la medida global de variabilidad y en el gráfico 3.4 se muestra esta medida para cada pregunta, reportando de manera decreciente las preguntas con mayor variabilidad.

Esta información indica que las tres preguntas con mayor variabilidad fueron la 12, 14 y la 5 todas subjetivas; en cambio, las tres con menor variabilidad fueron la 3, 2 y la 7, dos objetivas y una subjetiva.

Esto parece indicar que las preguntas objetivas, gracias al parámetro en el cual se basan los alumnos para emitir sus opiniones, brindaron al profesor una retroalimentación más exacta que las subjetivas, debido a que las opiniones fueron menos dispersas.

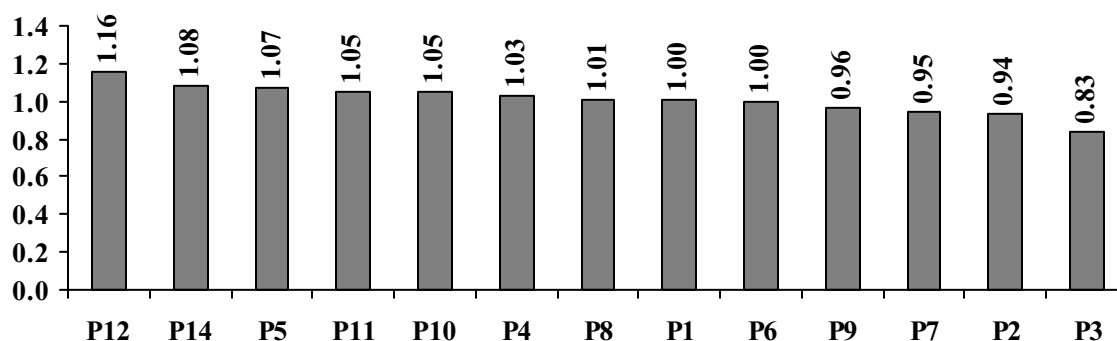
Por este motivo, sería conveniente introducir más preguntas objetivas en la encuesta o pedir a los alumnos que evalúen el desempeño de sus profesores tomando como base al mejor profesor que han tenido (o al mejor profesor que tuvieron en el semestre) o simplemente establecerles en la encuesta el perfil del profesor ideal, para que el alumno pueda tener un parámetro de comparación al evaluar a sus profesores, ver L'Hommedieu, Menges y Brinko (1990) y Rotem (1978).

Tabla 3.4 Cálculo de la variabilidad global de las preguntas de la encuesta

Grupos	Pregunta 1	Pregunta 14
	Desviación estándar	Desviación estándar
Grupo 1	Desviación estándar del Grupo 1 en P1	Desviación estándar del Grupo 1 en P14
....
Grupo m	Desviación estándar del Grupo m en P1	Desviación estándar del Grupo m en P14
	Desviación estándar promedio en P1	Desviación estándar promedio en P14

Se cuenta con un total de 2882 grupos ($m = 2882$) y con las evaluaciones realizadas por los alumnos en un grupo se obtiene una desviación estándar por pregunta para después promediar por pregunta todas las desviaciones calculadas en los grupos.

Gráfico 3.4 Variabilidad global en las opiniones de los alumnos por pregunta



Hasta aquí, se han analizado las preguntas de la encuesta en forma individual, en la siguiente sección se analizará si la información obtenida de la encuesta, es globalmente confiable para retroalimentar y evaluar el desempeño del profesor

3.2 DETERMINACIÓN DE LA CONFIABILIDAD DE LAS OPINIONES DE LOS ALUMNOS

Para probar la confiabilidad de las opiniones de los alumnos obtenidas en la encuesta utilizada para retroalimentar y evaluar el desempeño del profesor, se calculará un coeficiente de consistencia interna llamado alfa de Cronbach, Aiken (1996) y Anastasi et al (1998).

Este coeficiente de consistencia interna se calculará por grupo, en total serán 2882, después se dividirán los grupos de acuerdo al número de alumnos que contestaron la encuesta para obtener un promedio de estos coeficientes calculados, esto con el fin de averiguar como cambia la confiabilidad a medida que se incrementa el número de opiniones acerca del desempeño del profesor.

Para calcular el coeficiente de consistencia interna se procede a realizar lo siguiente en cada uno de los grupos considerados, ver tabla 3.5. Primero, para cada alumno que contestó

la encuesta en el grupo, se obtiene una suma de sus evaluaciones en las preguntas de la encuesta. Segundo, se calcula por grupo una desviación estándar total que mide la variabilidad de la suma de las evaluaciones de los alumnos en las preguntas de la encuesta.

Tabla 3.5 Cálculo del coeficiente de consistencia interna por grupo

Grupos	Alumnos	P1	...	P14	Suma de evaluaciones
Grupo 1	Alumno 1,1	P1,1,1	P14,1,1	P1,1,1 +....+ P14,1,1

	Alumno 1,n1	P1,1,n1	P14,1,n1	P1,1, n1 +....+ P14, 1,n1
		Desv. Est. del Grupo 1 en P1	Desv. Est. ... del Grupo 1 en P14		Desviación Estándar total del Grupo 1
....
Grupo m	Alumno m,1	P1,m,1	P14,m,1	P1,1,1 +....+ P14,1,1

	Alumno m,nm	P1,m,nm	P14,m,nm	P1,1, n1 +....+ P14, 1,n1
		Desv. Est. del Grupo m en P1	Desv. Est. ... del Grupo m en P14		Desviación estándar total del Grupo m

Después, para calcular el coeficiente de consistencia interna en el grupo se suman las desviaciones estándar calculadas para cada pregunta y se utiliza la siguiente formula.

Coeficiente de consistencia interna en el Grupo 1

$$r_1 = \left(\frac{n}{n-1} \right) \frac{DE_t^2 - \sum (DE_i^2)}{DE_t^2}$$

donde

r_1 es el coeficiente de consistencia interna del grupo 1
 n es igual a 13, ya que existen 13 preguntas consideradas de la encuesta (12 específicas, de la 1 a la 12, y una global, la 14)

DE_t^2 es el cuadrado de la desviación estándar total del grupo 1 (la varianza)

$\sum (DE_i^2)$ es la sumatoria del cuadrado de las desviaciones estándar de las preguntas de la encuesta en el grupo 1.

Por último, una vez calculados los coeficientes de consistencia interna para cada uno de los grupos, se procedió a agruparlos de acuerdo al número de alumnos que contestaron la encuesta para después obtener un promedio en cada agrupación, los que son presentados en la tabla 3.6.

Tabla 3.6 Coeficientes de consistencia interna de acuerdo al número de alumnos que contestaron la encuesta

Alumnos que contestaron la encuesta en el grupo	Coeficiente de consistencia interna (*)	Número de Grupos
Menos de 10	0.772	354
10 o más pero menos de 20	0.887	732
20 o más pero menos de 30	0.916	956
30 o más pero menos de 40	0.924	704
40 o más pero menos de 50	0.923	86
50 o más pero menos de 189	0.927	50

(*) Es un promedio de los coeficientes calculados para cada grupo.

Cashin (1988 y 1995) indica que deben de interpretarse con cautela los coeficientes de consistencia interna calculados para los grupos en donde contestaron la encuesta menos de 10 alumnos, esto se debe a que en esos grupos la variabilidad es muy grande.

Los resultados encontrados en este estudio son similares a los reportados por Cashin (1988 y 1995), la confiabilidad de las opiniones de los alumnos se incrementa a medida que en el grupo contestan más alumnos la encuesta.

En promedio, se observa un coeficiente de consistencia interna de 0.891 lo que significa que el 89.1% de la varianza de las evaluaciones de los alumnos sobre un profesor en un grupo depende de la varianza verdadera en el rasgo medido y el 10.9% depende de la varianza del error.

En conclusión, no es posible rechazar la primer hipótesis establecida en este estudio “Las opiniones de los alumnos son confiables para evaluar y retroalimentar el desempeño del profesor”, esto porque la varianza de las evaluaciones de los alumnos es explicada principalmente por la varianza verdadera y en mucho menor magnitud por la varianza de los errores no sistemáticos, relacionados con las condiciones en las cuales los alumnos contestan la encuesta, esto es, con el tiempo que tienen para contestarla, con el lugar en donde la contestan, con su estado de animo etc.

3.3 RESUMEN DE LOS RESULTADOS ENCONTRADOS ACERCA DE LA CONFIABILIDAD

En este capítulo fue posible estudiar la encuesta aplicada a los alumnos para evaluar el desempeño de sus profesores en profesional y de los resultados encontrados los más importantes son los siguientes.

Los patrones de respuesta de las evaluaciones de los alumnos en las preguntas de la encuesta fueron decrecientes en los valores de 1 a 6 de la escala con un ligero incremento en el de 7. Este incremento al final de la escala se consideró anormal de acuerdo a la forma del patrón y puede deberse a errores sistemáticos en las evaluaciones de los alumnos,

ocasionados por variables que nada tienen que ver con el desempeño del profesor, lo cual se estudiará en el siguiente capítulo.

La información aportada por los patrones de respuesta indicó que los valores más utilizados para evaluar a los profesores fueron el 1 y el 2, pero también se encontró que las evaluaciones de los alumnos fueron muy similares en todas las preguntas de la encuesta.

Lo anterior se consideró fue atribuible a los siguientes puntos:

- a) Posiblemente porque los alumnos conocen el valor que las personas administrativas establecen como una buena evaluación para los profesores, este es de 2.5, lo que podría resultar en una poca utilización de los restantes valores de la escala.
- b) También, es posible que sea debido a que no están definidos todos los valores de la escala, en el periodo analizado sólo estaba definido el valor de 1, total de acuerdo, y el de 7, en total desacuerdo, lo que puede evitar aprovechar el valor agregado de los restantes valores.
- c) Por falta de interés de los alumnos por evaluar a sus profesores, ya sea porque existen muchas preguntas subjetivas que hacen difícil la evaluación; o tal vez debido a que invierten mucho tiempo para contestar la encuesta, como resultado de incluir preguntas adicionales que nada tienen que ver con el profesor sino más bien con el desempeño de directivos o sobre los servicios que ofrece el Campus; o simplemente porque la encuesta no incluye los aspectos que el estudiante considera importantes medir.

Las preguntas de la encuesta que aportaron la información más consistente para retroalimentar al profesor fueron las objetivas, es decir, aquellas donde los alumnos tuvieron una percepción similar ya que contaron con un parámetro base para evaluar al profesor.

Dentro de las preguntas con mayor variabilidad se encontraron la 12, 14 y la 5 todas subjetivas; en cambio, las preguntas con menor variabilidad fueron la 3, 2 y 7.

Esto sugiere modificar el cuestionario actual para incluir preguntas más objetivas o si se continúa con las mismas, es necesario proporcionarle al alumno un parámetro que tenga como base al evaluar a sus profesores, ya sea indicándole que asigne sus evaluaciones considerando al mejor profesor que tuvo durante el semestre o bien brindándole un perfil ideal que sea establecido por estudiosos de técnicas didácticas para que lo compare con el de su profesor.

Aún y cuando hay cosas que mejorar de la encuesta, se concluye que las opiniones de los alumnos son confiables para retroalimentar y evaluar el desempeño de los profesores, ya que la variabilidad en sus opiniones se explica en promedio en un 89.1% por la varianza verdadera y en un 10.9% por la varianza del error, esto es, las diferencias en las opiniones de los alumnos se atribuyen principalmente a percepciones diferentes de los alumnos y no a

errores no sistemáticos que suceden casualmente ocasionados por el ambiente en que los alumnos contestan la encuesta o por su estado de animo.

Por último, los resultados indican que la consistencia en las opiniones de los alumnos así como su confiabilidad, mejora al incrementarse el número de alumnos que opinan en un grupo, por lo que hay que incentivar más a los alumnos en el periodo de encuestas para mejorar el porcentaje de respuesta en el Campus.

CAPÍTULO 4

VALIDEZ DE LAS OPINIONES DE LOS ALUMNOS PARA EVALUAR EL DESEMPEÑO DE LOS PROFESORES

En este capítulo se pretende indagar sobre la validez de las opiniones de los alumnos para evaluar el desempeño de los profesores del Campus Monterrey.

Nuevamente se contará con un total de 2882 grupos de cursos teóricos en el nivel de profesional, en el ejercicio académico de Agosto-Diciembre 2001, y la unidad de análisis que se utilizará será el promedio de las evaluaciones individuales de los alumnos realizadas en un grupo en cada una de las preguntas de la encuesta. De esta manera, el valor reportado en la pregunta 14 para cualquier grupo será la “Opinión Global Promedio”, OGP, del profesor, indicador utilizado para evaluar su desempeño.

4.1 RELACIÓN DE LA EVALUACIÓN GLOBAL DEL PROFESOR CON LAS VARIABLES QUE SE DESEA SEAN FUENTES DE EXPLICACIÓN

Se desea que el indicador utilizado para evaluar el desempeño del profesor, esté relacionado con variables internas al proceso de enseñanza-aprendizaje, aquellas que miden aspectos específicos del desempeño del profesor y que pueden ser modeladas con la información aportada por las preguntas específicas de la encuesta.

A este respecto si el indicador está más relacionado con las variables internas que con las externas, las que nada tienen que ver con el desempeño del profesor pero que distorsionan las opiniones de los alumnos, entonces se concluirá que el indicador es válido para evaluar el desempeño del profesor. Por esto, este análisis comenzará estimando un modelo de regresión donde se modele a la OGP mediante las preguntas específicas de la encuesta, así se podrá identificar la relación que existe con estas variables.

Complementando lo anterior, con la ayuda del análisis de factores, se resumirá la información de las preguntas específicas en unos cuantos factores estimados que se utilizarán para estimar otro modelo de regresión en el que se modelará también a la OGP.

Los resultados que se obtengan con este análisis aportarán información para determinar no sólo la relación del indicador con las variables internas al proceso de enseñanza-aprendizaje, sino también como se modela mejor a la OGP, y a su vez, si existe un perfil para el profesor establecido con las opiniones de los alumnos obtenidas mediante las preguntas específicas de la encuesta, el que será determinado por los factores que se estimen mediante el análisis de factores.

4.1.1 EVALUACIÓN GLOBAL DEL PROFESOR MODELADA CON LAS PREGUNTAS ESPECÍFICAS DE LA ENCUESTA

El modelo de regresión estimado en esta sección, incluye a todas las preguntas específicas de la encuesta.⁸ La estimación del modelo fue realizada mediante Mínimos Cuadrados Ordinarios, MCO, y es presentada en la tabla 4.1, sin embargo para entender mejor los resultados reportados, es conveniente revisar el anexo 1 donde se brinda una explicación detallada de los estadísticos utilizados.⁹

Se observa que la bondad de ajuste del modelo, es decir, la magnitud de la variabilidad de la OGP explicada por las preguntas específicas, fue del 92.9%, como lo indica la R-cuadrada ajustada, esto significa que el modelo fue bueno, esto es, las preguntas explicaron bien a la OGP.

La F conjunta y su probabilidad indican que el modelo fue significativo, lo que contrasta con los resultados de la R-cuadrada ajustada, ya que fue elevada.

Puede observarse que sólo dos preguntas resultaron no significativas con un nivel de confianza del 95% y con respecto al signo de los coeficientes de cada una de las preguntas específicas, se encontró, con excepción de la constante, que todos fueron positivos, esto significó que las preguntas específicas tuvieron una relación positiva con la OGP.

Podría pensarse con estos resultados que las preguntas 7 y 8 no aportaron información para explicar a la OGP debido a que no fueron significativas, sin embargo al observar la matriz de correlaciones presentada en la tabla 4.2, se puede pensar lo contrario ya que los coeficientes de correlación de estas preguntas con la OGP fueron de 0.83, esto señala que posiblemente existieron problemas de multicolinealidad en el modelo estimado.

Esto se confirma con el diagnóstico de multicolinealidad, como demuestra el índice de condición donde se reportan valores mayores a 30. Bajo estas condiciones, los estimadores de MCO serán ineficientes, no tendrán la varianza mínima, esto significa que las pruebas de significancia estadística no serán válidas ya que puede suceder que no se rechace la hipótesis nula de que un coeficiente sea igual a cero, cuando realmente no lo es.

No obstante, este no es el único problema que presentó el modelo estimado, debido a que la prueba de White para heterocedasticidad rechazó la hipótesis nula de que los errores tuvieron varianza constante, con este problema presente los estimadores de MCO también serán ineficientes. Pero además, lo más importante será que las estimaciones de los errores estándar utilizados para hacer las pruebas de significancia estadística, serán inconsistentes,

⁸ Las estimaciones presentadas aquí, fueron elaboradas con el programa estadístico SPSS versión 10. Por otra parte, para elaborar las pruebas que determinan la existencia o no de heterocedasticidad y para corregir este problema, se utilizó el programa econométrico E-views, versión 2.

⁹ Para una explicación detallada del método de Mínimos Cuadrados Ordinarios, MCO, ver Gujarati (1995) y Maddala (1992).

es decir, resultarían o muy grandes o muy pequeñas aportando en cualquier caso inferencias erróneas.

Tabla 4.1 Estimación del modelo de regresión, OGP versus preguntas específicas

Variable dependiente: OGP				
Desviación				
Variable	Coefficiente	estándar	Estadístico t	Probabilidad
Constante	-0.156	0.013	-11.939	0.000
P1	0.147	0.010	14.348	0.000
P2	0.077	0.007	10.950	0.000
P3	0.032	0.013	2.425	0.015
P4	0.080	0.011	7.009	0.000
P5	0.071	0.009	7.640	0.000
P6	0.033	0.014	2.312	0.021
P7	0.023	0.018	1.312	0.190
P8	0.015	0.016	0.927	0.354
P9	0.075	0.014	5.436	0.000
P10	0.021	0.009	2.276	0.023
P11	0.165	0.014	11.877	0.000
P12	0.421	0.011	37.814	0.000
R-cuadrada		0.929	Estadístico F	3137.163
R-cuadrada ajustada		0.929	Probabilidad	0.000

Prueba de White para Heterocedasticidad			
Estadístico F	11.371	Probabilidad	0.000
2882*R-cuadrada	251.262	Probabilidad	0.000

Diagnóstico de Multicolinealidad (o Colinealidad)					
Variable	Tolerancia	VIF	Dimensión	Valor propios	Índice de Condición
P1	0.334	2.991	1	12.641	1
P2	0.496	2.016	2	0.085	12.197
P3	0.291	3.436	3	0.068	13.621
P4	0.201	4.965	4	0.041	17.612
P5	0.32	3.126	5	0.038	18.284
P6	0.178	5.62	6	0.026	22.045
P7	0.166	6.03	7	0.021	24.684
P8	0.169	5.922	8	0.018	26.192
P9	0.249	4.011	9	0.018	26.856
P10	0.464	2.156	10	0.015	28.96
P11	0.154	6.506	11	0.011	33.495
P12	0.18	5.554	12	0.010	35.089
			13	0.008	40.909

Tabla 4.2 Matriz de correlaciones entre las preguntas específicas y la pregunta 14, OGP

	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11	P12	OGP	
P1	1	0.517	0.703	0.664	0.617	0.682	0.749	0.747	0.661	0.528	0.735	0.754	0.800	P1
P2	0.517	1	0.685	0.477	0.463	0.483	0.520	0.519	0.569	0.467	0.610	0.544	0.623	P2
P3	0.703	0.685	1	0.644	0.647	0.640	0.708	0.694	0.714	0.590	0.745	0.728	0.775	P3
P4	0.664	0.477	0.644	1	0.792	0.839	0.744	0.732	0.772	0.603	0.810	0.781	0.826	P4
P5	0.617	0.463	0.647	0.792	1	0.726	0.680	0.656	0.705	0.551	0.745	0.737	0.779	P5
P6	0.682	0.483	0.640	0.839	0.726	1	0.831	0.824	0.768	0.651	0.816	0.805	0.831	P6
P7	0.749	0.520	0.708	0.744	0.680	0.831	1	0.882	0.776	0.676	0.802	0.803	0.830	P7
P8	0.747	0.519	0.694	0.732	0.656	0.824	0.882	1	0.768	0.670	0.810	0.812	0.830	P8
P9	0.661	0.569	0.714	0.772	0.705	0.768	0.776	0.768	1	0.672	0.813	0.771	0.820	P9
P10	0.528	0.467	0.590	0.603	0.551	0.651	0.676	0.670	0.672	1	0.659	0.622	0.661	P10
P11	0.735	0.610	0.745	0.810	0.745	0.816	0.802	0.810	0.813	0.659	1	0.872	0.903	P11
P12	0.754	0.544	0.728	0.781	0.737	0.805	0.803	0.812	0.771	0.622	0.872	1	0.931	P12
OGP	0.800	0.623	0.775	0.826	0.779	0.831	0.830	0.830	0.820	0.661	0.903	0.931	1	OGP
	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11	P12	OGP	

Debido a esto, es necesario atacar estos problemas antes de definir cuales fueron las preguntas específicas que explicaron en mayor magnitud a la OGP. El problema de multicolinealidad puede atacarse de dos formas, la primera consiste en eliminar del modelo a tantas preguntas específicas como sea necesario hasta obtener un modelo con un índice de condición menor a 30, eligiendo sólo a aquellas que aporten un incremento significativo en la R-cuadrada ajustada y que no estén correlacionadas lo suficiente para que los estadísticos de diagnóstico detecten que existen problemas de multicolinealidad.

La eliminación de preguntas se hace porque no es posible separar la información que aporta una pregunta individualmente de la que aportan las demás, sin embargo existe una manera de no eliminar la información de las preguntas identificadas como las más correlacionadas, pero sí erradicando el problema de multicolinealidad. Esto es, la segunda manera de eliminar este problema, es utilizar como variables explicativas a los factores estimados mediante un análisis de factores, los que son una combinación lineal de las preguntas específicas de la encuesta y que pueden concentrar la mayor parte de la información reportada por éstas.

En cambio, el problema de heterocedasticidad es posible resolverlo mediante el estimador consistente de la matriz de varianzas y covarianzas propuesta por White (1980), la cual provee estimadores consistentes de las desviaciones estándar de los coeficientes de la regresión, sin hacer ningún supuesto acerca de la forma de la heterocedasticidad, ver anexo 2.¹⁰ Esto permitirá que los resultados de las pruebas de significancia estadística sean válidos, siempre y cuando también sea corregido el problema de multicolinealidad.

¹⁰ Ver también Greene (1991) y Judge et al. (1985).

De esta manera, al erradicar el problema de multicolinealidad, será posible comparar las estimaciones realizadas con las preguntas específicas y con los factores, para determinar con que modelo se explica mejor a la OGP.

A continuación se presenta en la tabla 4.3 un modelo que no presentó los problemas mencionados, esto se logró eliminando las preguntas específicas que estuvieron muy correlacionadas y que en comparación con las restantes explicaban poco a la OGP, lo que corrigió el problema de multicolinealidad y utilizando el método de White (1980) para eliminar los efectos de la heterocedasticidad.

Tabla 4.3 Estimación del modelo de regresión (OGP versus preguntas específicas) corrigiendo por multicolinealidad y por heterocedasticidad

Variable dependiente: OGP				
	Desviación			
Variable	Coefficiente	estándar	Estadístico t	Probabilidad
Constante	-0.141	0.016	-9.053	0.000
P1	0.151	0.016	9.320	0.000
P2	0.084	0.010	8.490	0.000
P4	0.089	0.017	5.381	0.000
P5	0.074	0.015	5.108	0.000
P7	0.058	0.022	2.666	0.008
P9	0.089	0.025	3.553	0.000
P11	0.175	0.022	7.950	0.000
P12	0.428	0.018	23.993	0.000
R-cuadrada		0.929	Estadístico F	4677.219
R-cuadrada ajustada		0.929	Probabilidad	0

Diagnóstico de Multicolinealidad (o Colinealidad)					
Variable	Tolerancia	VIF	Dimensión	Valor propio	Índice de Condición
P1	0.361	2.768	1	8.736	1.000
P2	0.598	1.673	2	0.078	10.612
P4	0.239	4.187	3	0.064	11.687
P5	0.328	3.048	4	0.038	15.223
P7	0.255	3.919	5	0.025	18.774
P9	0.265	3.774	6	0.020	20.909
P11	0.158	6.311	7	0.016	23.266
P12	0.188	5.326	8	0.013	25.939
			9	0.011	27.880

Como puede observarse el modelo no incluyó como regresores a las preguntas específicas: 3, 6, 8 y 10.¹¹ Pero, reportó una buena bondad de ajuste, explicó el 92.9% de la

¹¹ Para estimar el modelo se utilizó el método de paso a paso (Stepwise) con el fin de determinar las preguntas específicas que aportaron el mejor ajuste.

variabilidad de la información concentrada en la OGP, porcentaje idéntico al reportado por el modelo anterior.

Puede observarse que se ha erradicado el problema de multicolinealidad como lo demuestran los estadísticos de diagnóstico, y debido a que se corrigió por heterocedasticidad las pruebas de significancia estadística sobre los coeficientes y sobre el modelo total, son consideradas como confiables.

Todos los coeficientes estimados de las preguntas específicas fueron significativos, además afectaron positivamente a la pregunta 14, y para determinar las que tuvieron un efecto mayor al explicar la pregunta global, sólo basta con ordenarlas en forma descendente de acuerdo a la magnitud de sus coeficientes estimados, así se encontró que el orden fue: 12, 11, 1, 4, 9, 2, 5 y 7.

Esto significó, que para los alumnos fue importante que el profesor organizara e impartiera el curso de manera que favoreciera el aprendizaje, pregunta 12; mostrara una actitud comprometida y motivadora en la impartición y coordinación del curso, pregunta 11; manifestara dominio de la disciplina académica impartida, pregunta 1; fomentara una comunicación efectiva, se mantuviera al diálogo y creara un ambiente de confianza y respeto, pregunta 4; practicara y fomentara la misión del ITESM, pregunta 9; asistiera a las sesiones establecidas, fuera puntual y cumpliera con el servicio de asesoría, pregunta 2; evaluara de manera justa y objetiva, pregunta 5 y por último, que el profesor fomentara el aprendizaje individual y colaborativo, de conocimientos relevantes, conectados con situaciones reales y sustentado en información actualizada y confiable, pregunta 7.

En resumen, los resultados demostraron que fue posible explicar a la OGP en un 92.9%, mediante un modelo de regresión donde las variables explicativas fueron las preguntas específicas, resultando ser las preguntas 12, 11, 1, 4, 9, 2, 5 y 7 (ordenadas de mayor a menor en cuanto a su aportación), las que determinaron el modelo de regresión que mejor explicó a la pregunta global.

Sin embargo, falta aún verificar si los factores que resumen la información de las preguntas específicas y que son estimados mediante un análisis de factores, explican mejor a la OGP, así se procederá a realizar esta estimación.

4.1.2 EVALUACIÓN GLOBAL DEL PROFESOR MODELADA CON LOS FACTORES QUE RESUMEN LA INFORMACIÓN DE LAS PREGUNTAS ESPECÍFICAS

El análisis de factores es comúnmente utilizado para resumir la información obtenida en las preguntas específicas de la encuesta, dentro de un mínimo de factores que resultan de realizar una combinación lineal de las mismas. Estos factores pueden ser interpretados de acuerdo a la correlación que tengan con las preguntas para determinar qué desearon los alumnos de sus profesores.

Posteriormente, pueden utilizarse en un análisis de regresión con el fin de explicar la OGP, es decir, pueden ser utilizados como variables explicativas en lugar de utilizar todas las preguntas específicas, las que pueden estar muy correlacionadas entre sí.

El método utilizado para estimar las cargas de los factores fue el de componentes principales y para determinar cuantos estimar, se tomó como base la cantidad de valores propios mayores a 0.7 que resultaron de la matriz de correlaciones de las preguntas específicas de la encuesta, Dunteman (1989). Además, para realizar su interpretación se utilizó el método de rotación Varimax y para estimarlos el método de regresión. Por último, con el fin de identificar las preguntas de la encuesta que definieron los factores, se tomaron todas las cargas mayores a 0.6, ver Johnson y Wichern (1992) y Jobson (1992).¹²

Antes de presentar la información encontrada con este análisis, es conveniente explicar la tabla 4.4 donde se presentan los resultados. Se estimaron 2 factores que llevan el título de sin rotar y para darles una mejor interpretación, fue necesario rotarlos, es decir, hacer un cambio de ejes para poder distinguir las variables más correlacionadas con ellos.

Tabla 4.4 Estimación de los factores que resumen la información de las preguntas específicas

Pregunta	Sin rotar		Rotados		Comunalidades	Varianza específica
	Factor1	Factor2	Factor1	Factor2		
P1	0.821	0.056	0.683	0.460	0.678	0.322
P2	0.659	0.684	0.228	0.922	0.903	0.097
P3	0.83	0.352	0.542	0.721	0.813	0.187
P4	0.874	-0.222	0.867	0.246	0.813	0.187
P5	0.818	-0.162	0.789	0.270	0.696	0.304
P6	0.895	-0.231	0.891	0.248	0.855	0.145
P7	0.904	-0.101	0.833	0.366	0.828	0.172
P8	0.899	-0.097	0.827	0.366	0.818	0.182
P9	0.884	-0.009	0.770	0.436	0.782	0.218
P10	0.751	-0.012	0.656	0.366	0.564	0.436
P11	0.927	-0.002	0.804	0.463	0.86	0.14
P12	0.91	-0.065	0.820	0.400	0.833	0.167
Valores propios			8.693	0.748		
Proporción explicada			72.44%	6.23%		
Proporción acumulada			72.44%	78.67%		

En esta tabla es posible observar que el factor 1 explicó el 72.44% de la variabilidad de la información reportada por las preguntas específicas de la encuesta; en cambio, el factor 2 explicó el 6.23% y conjuntamente el 78.67%. Si bien, otro de los criterios que es utilizado para elegir la cantidad de factores, es incrementar su número hasta que la proporción

¹² Las cargas son los coeficientes de correlación entre las preguntas específicas y los factores. Así, si algunas preguntas están muy correlacionadas con un factor, significa que su información es muy similar, pero a su vez muy diferente de la que reportan las preguntas que resulten muy correlacionadas con otro factor. Estas correlaciones a su vez permiten interpretar lo que esta cuantificando el factor.

acumulada, esto es, el porcentaje total de la variabilidad de la información explicada por los factores sea de 80% ó 90%, no fue utilizado aquí debido a que cuando se incluyeron 3 factores el método de máxima verosimilitud y el de componentes principales arrojaron resultados diferentes, además éstos fueron difíciles de interpretar.¹³

Aunado a lo anterior, aparecen también en la tabla dos columnas más, una nombrada “Comunalidades” y otra “Varianza específica”. La primera se refiere a la variación de las preguntas específicas explicada por los factores 1 y 2 y la segunda, representa la variación que no fue explicada por los factores sino por el error; como puede observarse 8 de las 12 preguntas fueron explicadas por los factores con una proporción mayor al 78.67%, las preguntas fueron la 2, 11, 6, 12, 7, 8, 4 y 3 (ordenadas de mayor a menor explicación).

Con respecto a la interpretación de los factores, puede observarse que en el factor 1 todas las preguntas, con excepción de la 2 y la 3, tuvieron una correlación mayor a 0.6 (todas las correlaciones mayores a 0.6 fueron seleccionadas con color gris); a diferencia, en el factor 2 solamente las preguntas 2 y 3 excedieron este valor.

El factor 1 fue difícil de interpretar dado que estuvo muy correlacionado con la mayor parte de las preguntas, sin embargo si fue posible observar que las preguntas que estuvieron más correlacionadas con él fueron todas las subjetivas junto con dos objetivas (9 y 10), en el siguiente orden: la 6, 4, 7, 8, 12, 11, 5, 9, 1 y 10, lo cual permitió llamarlo el factor que mide la parte subjetiva del desempeño del profesor.

La elevada correlación de las preguntas con el factor 1 y el hecho de que este factor explicó la mayor parte de la variabilidad total de la información, pudo haberse debido a que la información que brindaron las preguntas fue muy similar, tan parecida que pudo explicarse con pocos factores.

En el caso particular del factor 2, este si pudo ser interpretado, y presentó solamente correlaciones altas (mayores a 0.6) con las preguntas 2 y 3, ambas objetivas, y dado lo que cuantifican, fue posible interpretarlo como el factor de cumplimiento en cuanto a la puntualidad del profesor, asistencia a clases, asesorías y al programa del curso. Pero, a diferencia del factor 1 tuvo menos importancia ya que explicó una variabilidad menor.

Los factores estimados, no sólo eliminan el problema de multicolinealidad cuando se utilizan como variables explicativas en lugar de las preguntas específicas, sino también permiten explicar la OGP considerando la información de todas las preguntas, ya que son una combinación lineal de éstas, pero dependiendo que también la resuman será su capacidad explicativa.

En la tabla 4.5 son presentadas las estimaciones del modelo de regresión que incluye a los factores estimados y se encontró que éstos explicaron a la OGP en una magnitud menor que las preguntas específicas, la R-cuadrada ajustada del modelo fue de 0.894.

¹³ Los resultados de la tabla 4.4 fueron similares a los reportados con el método de estimación de máxima verosimilitud.

Tabla 4.5 Estimación del modelo de regresión, OGP versus los factores estimados

Variable dependiente: P14. Resultados corregidos por heterocedasticidad				
Variable	Coefficiente	Desviación estándar	Estadístico t	Probabilidad
Constante	1.767	0.003	505.072	0.000
F1	0.470	0.005	101.059	0.000
F2	0.273	0.005	52.990	0.000
R-cuadrada		0.894	Estadístico F	12068.37
R-cuadrada ajustada		0.894	Probabilidad	0.00

Diagnóstico de Multicolinealidad (o Colinealidad)					
Variable	Tolerancia	VIF	Dimensión	Valor propio	Índice de condición
F1	1	1	1	1	1
F2	1	1	2	1	1
			3	1	1

El modelo estimado fue corregido por heterocedasticidad y dado que no presentó problemas de multicolinealidad, las pruebas de significancia de los coeficientes estimados fueron confiables.

Los coeficientes de este modelo fueron positivos, y puede observarse que el factor 1 fue el que tuvo más peso en la regresión, ya que su coeficiente estimado fue mayor al reportado por el factor 2.

Aplicando esta información a las preguntas, se encontró que el factor subjetivo, factor 1, fue el que aportó más información para explicar la pregunta global, el cual estuvo correlacionado con las preguntas específicas en el siguiente orden: 6, 4, 7, 8, 12, 11, 5, 9, 1, 10, 3 y 2. Esto indica que las preguntas que tuvieron más peso en el factor no fueron las que estuvieron más correlacionadas con la OGP, pregunta 11 y 12, por este motivo la bondad de ajuste no fue tan buena como la reportada por el modelo donde las variables explicativas fueron las preguntas específicas.

En resumen, se encontró que el modelo de regresión donde las variables explicativas fueron los factores estimados, no mostró un ajuste tan bueno como el reportado por el modelo donde lo fueron las preguntas específicas, esto se atribuyó a que la información resumida por los factores no estuvo tan correlacionada con la OGP como lo estuvieron individualmente las preguntas.

Hasta ahora, se ha analizado de que depende el indicador de evaluación del profesor, OGP, visto desde una perspectiva endógena al proceso de enseñanza-aprendizaje, es decir, se ha investigado como lo afecta la forma de enseñar del profesor, sin embargo aún no se ha

analizado como lo afectan características externas a este proceso, esto se hará a continuación.

4.2 RELACIÓN DE LA EVALUACIÓN GLOBAL DEL PROFESOR CON LAS VARIABLES QUE SE DESEA “NO” SEAN FUENTES DE EXPLICACIÓN

Existen variables externas al proceso de enseñanza-aprendizaje profesor-alumno que influyen en la Opinión Global Promedio, OGP, del profesor, es decir, en el indicador utilizado para su evaluación elaborado con base en las opiniones de los alumnos.

Estas variables hacen diferir las opiniones de los alumnos sobre el desempeño del profesor y se consideran errores sistemáticos porque si no se controlan, se volverán a repetir cada vez que se aplique la encuesta. Por esto, es común que existan grupos en donde se presenten evaluaciones extremas, por ejemplo donde exista la mayor parte de las evaluaciones de los alumnos en los valores de 1 y 2 de la escala y una o dos en el de 7.

Determinar la relación que tienen estas variables con el indicador, permitirá cuantificar la magnitud de su sesgo sobre las opiniones de los alumnos, por ejemplo, que tan significativo es el incremento reportado en el valor de 7 de la escala del patrón de respuestas de la pregunta global de la encuesta.

Además, el cuantificar la magnitud del sesgo ayudará a conocer la validez de las opiniones de los alumnos para evaluar el desempeño de los profesores, entre más pequeño sea, más validas serán éstas.

Así, se estimarán modelos de regresión donde se pretenderá explicar a la OGP con sólo variables que reflejen características del curso, grupo y/o del profesor que nada tienen que ver con el desempeño del profesor y con base en la bondad de ajuste de estos modelos se cuantificará el sesgo de estas variables.

Las variables consideradas en este análisis fueron las siguientes:

Cuantitativas

- El tamaño y la calificación promedio del grupo
- La edad y los años de experiencia del profesor en el Instituto

Cualitativas¹⁴

- La disciplina a la que perteneció el curso.

¹⁴ Para una explicación detallada sobre como fueron medidas las variables cualitativas (nominales) ver De la Garza (1995).

- El tipo de grupo: tradicional o rediseñado
- El género, tipo de contrato y grado máximo de estudios del profesor

El procedimiento para investigar si las variables mencionadas afectaron a la OGP, consistió en estimar un modelo base que incluye sólo las variables cuantitativas, para después de esto ir incluyendo a éste individualmente cada una de las características modeladas por las variables cualitativas, para terminar con un modelo donde se incluyen a las variables más significativas del total, sin considerar al modelo base.

Mas antes de comenzar las estimaciones es importante realizar un análisis descriptivo para estudiar la relación gráfica que tienen estas variables con el indicador utilizado para evaluar al profesor, la Opinión Global Promedio, OGP.

4.2.1 RELACIÓN GRÁFICA DE LA EVALUACIÓN GLOBAL DEL PROFESOR Y LAS VARIABLES QUE SE DESEA “NO” SEAN FUENTES DE EXPLICACIÓN

Se inicia primero realizando un análisis descriptivo de las variables que reflejan características del curso, luego del grupo y por último del profesor.

4.2.1.1 EFECTO DE LA DISCIPLINA QUE IMPARTE EL PROFESOR

Es factible que las disciplinas sean muy diferentes en cuanto a contenido y a dificultad, así como sus patrones de respuesta, los que pueden pasar desapercibidos cuando se juntan para formar el patrón global de profesional en la pregunta 14 con la cual se elabora la OGP del Campus.

En la tabla 4.6 se presentan las disciplinas que se consideraron en el análisis y en la tabla 4.7 se reporta la evaluación obtenida por las disciplinas en la pregunta global del encuesta, así que ese valor es la OGP de la disciplina, asimismo se muestra el número de grupos y el total de respuestas.

En total se consideraron 41 disciplinas y las que obtuvieron una OGP mayor a la global del Campus en el nivel de profesional, 1.78, fueron, de la disciplina peor evaluada a la mejor: EX, F, AA, CB, EC, Q, E, CV, SI, TF, TA, IS, MI, CS, MA, IN, CD e IG.

En cambio, las disciplinas que obtuvieron una OGP por debajo de 1.78 fueron, de la disciplina peor evaluada a la mejor: P, D, FZ, PA, AF, IQ, H, M, MT, CF, CO, NI, AR, RI, HI, IA, SM, MR, RH, OR, SA, RN y CC.

Tabla 4.6 Disciplinas consideradas en el nivel de profesional

Disciplina	Descripción	Disciplina	Descripción
AA	Agronegocios	IQ	Ingeniería Química
AF	Administración Financiera	IS	Ingeniería de Sistemas
AR	Arquitectura	M	Ingeniería Mecánica
CB	Computación Básica	MA	Matemáticas
CC	Ciencias de la Conducta	MI	Medios de Información
CD	Ciencias Administrativas	MR	Mecatrónica
CF	Contabilidad Financiera y Administración	MT	Mercadotecnia
CO	Comunicación	NI	Negocios Internacionales
CS	Ciencias Computacionales	OR	Organización
CV	Ingeniería Civil	P	Ciencia Política
D	Derecho	PA	Programa para Alumnos Sobresalientes
E	Ingeniería Eléctrica	Q	Química
EC	Economía	RH	Recursos Humanos
EX	Programa de Apoyo a la Misión	RI	Relaciones Internacionales
F	Física	RN	Recursos Naturales
FZ	Finanzas	SA	Sistemas Agropecuarios
H	Humanidades	SI	Sistemas de Información
HI	Idiomas	SM	Centro de Sistemas de Manufactura
IA	Centro de Inteligencia Artificial	TA	Tecnología de Alimentos
IG	Ingeniería Agrícola	TF	Térmica de Fluidos
IN	Ingeniería Industrial		

Si al observar la evaluación de las disciplinas se pudiera encontrar una relación negativa entre la evaluación del profesor y la dificultad de la disciplina que imparte, entonces este sería un factor de sesgo que las personas administrativas deberían de tomar en cuenta cuando determinan si un profesor se desempeña bien o mal al enseñar su curso.

Sin embargo, es difícil establecer tal relación ya que esto es muy subjetivo, por ejemplo, a criterio del tesista, en el anexo 3 se calcula un coeficiente de correlación de Spearman que captura la relación entre un ordenamiento subjetivo de la dificultad de las disciplinas y la evaluación real que obtuvieron, encontrándose un coeficiente de 0.2946, este valor indica que a mayor dificultad se obtiene una peor evaluación (un valor de la OGP mayor) lo que significa una relación negativa pero subjetiva entre la dificultad del curso y la evaluación del profesor, que puede diferir dependiendo del criterio de ordenamiento que se realice.

Tabla 4.7 Composición del total de grupos de profesional por disciplina

Disciplina	OGP	Respuestas	% del total	Grupos	% del total
EX	2.50	18	0.03%	1	0.03%
F	2.04	2810	4.18%	105	3.64%
AA	2.00	178	0.26%	17	0.59%
CB	1.99	3616	5.37%	126	4.37%
EC	1.96	3059	4.55%	118	4.09%
Q	1.93	1492	2.22%	57	1.98%
E	1.92	2021	3.00%	78	2.71%
CV	1.90	1021	1.52%	59	2.05%
SI	1.90	1462	2.17%	62	2.15%
TF	1.89	46	0.07%	4	0.14%
TA	1.87	252	0.37%	22	0.76%
IS	1.87	1489	2.21%	58	2.01%
MI	1.87	120	0.18%	6	0.21%
CS	1.82	1115	1.66%	45	1.56%
MA	1.82	6421	9.54%	216	7.49%
IN	1.81	2801	4.16%	99	3.44%
CD	1.81	1977	2.94%	91	3.16%
IG	1.79	71	0.11%	8	0.28%
P	1.77	532	0.79%	24	0.83%
D	1.77	3381	5.02%	158	5.48%
FZ	1.74	1329	1.98%	55	1.91%
PA	1.74	588	0.87%	39	1.35%
AF	1.74	363	0.54%	13	0.45%
IQ	1.74	1107	1.65%	42	1.46%
H	1.73	3860	5.74%	159	5.52%
M	1.72	3100	4.61%	116	4.02%
MT	1.72	2282	3.39%	118	4.09%
CF	1.71	3077	4.57%	114	3.96%
CO	1.71	3291	4.89%	165	5.73%
NI	1.70	788	1.17%	39	1.35%
AR	1.69	1839	2.73%	139	4.82%
RI	1.68	4002	5.95%	156	5.41%
HI	1.68	2441	3.63%	103	3.57%
IA	1.67	24	0.04%	6	0.21%
SM	1.67	39	0.06%	3	0.10%
MR	1.65	483	0.72%	26	0.90%
RH	1.53	1911	2.84%	78	2.71%
OR	1.53	1907	2.83%	87	3.02%
SA	1.47	172	0.26%	19	0.66%
RN	1.43	471	0.70%	34	1.18%
CC	1.37	331	0.49%	17	0.59%
Total	1.78	67287	100.00%	2882	100.00%

De esta forma, sólo resta conocer las disciplinas que fueron, en el periodo analizado, las que reportaron patrones de respuesta más afectados por evaluaciones en el valor de 7 de la escala.

En el gráfico 4.1 se presentan los patrones de respuesta de las evaluaciones de los alumnos en la pregunta 14 de la encuesta con la que es elaborada la OGP, en cada una de las disciplinas consideradas.

Este gráfico muestra que las disciplinas que reportaron patrones de respuesta decrecientes en todos los valores de la escala fueron (ordenadas de la peor a la mejor evaluación):

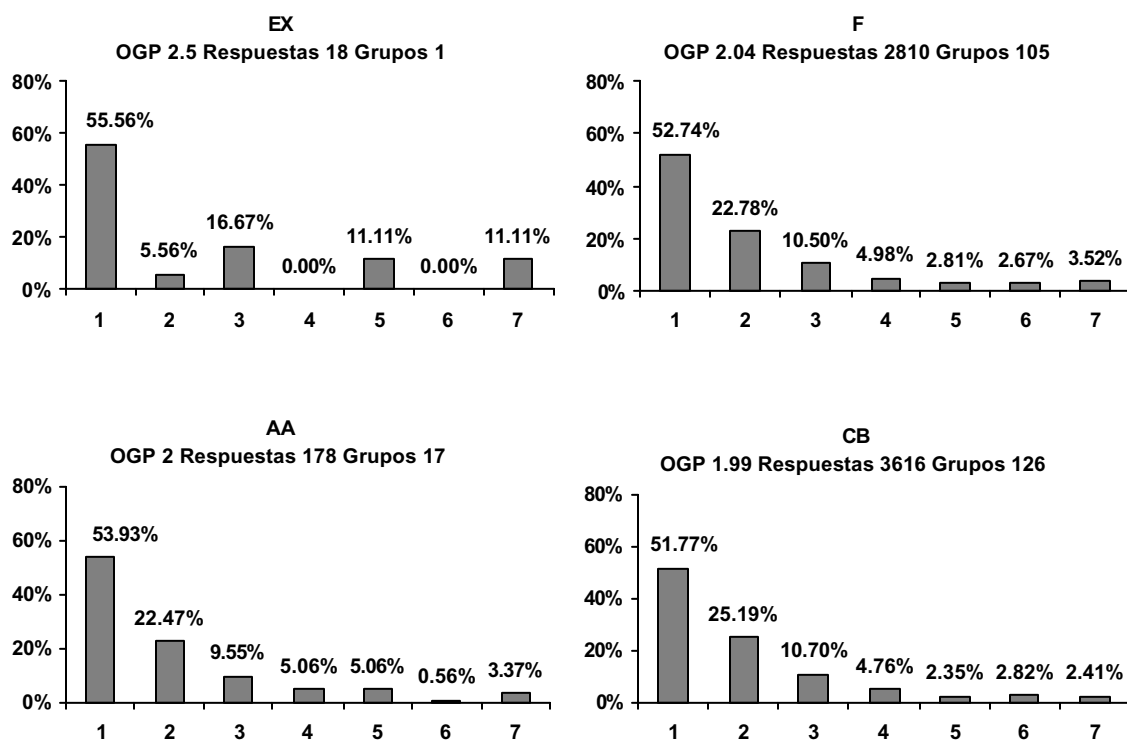
CV, TA, IS, CS, CD, PA, M, CO y OR.

Asimismo, informa que las disciplinas que reportaron patrones de respuesta decrecientes en los valores de 1 a 6 de la escala con un incremento en el valor de 7 fueron (ordenadas de la peor a la mejor evaluación):

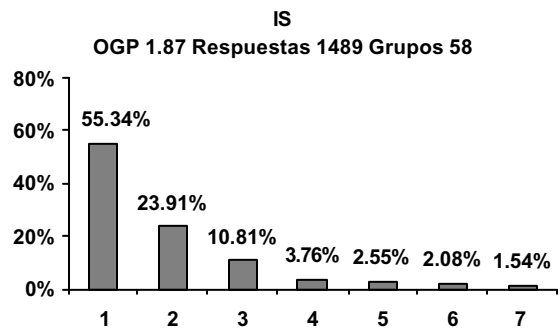
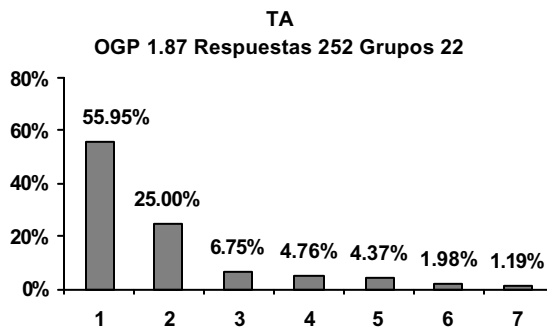
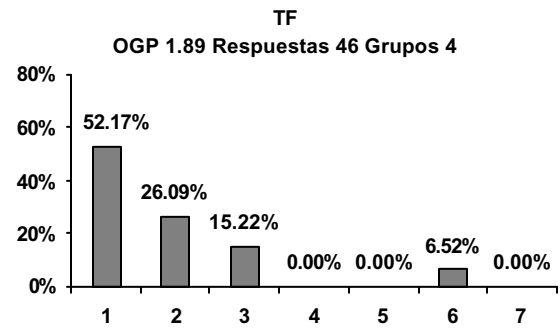
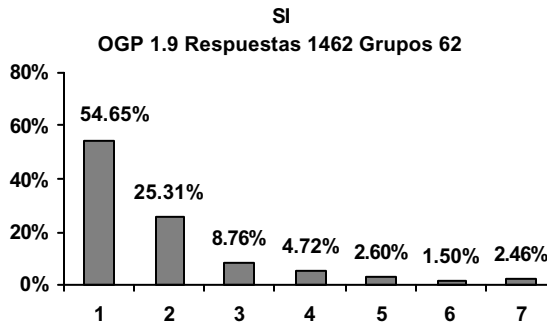
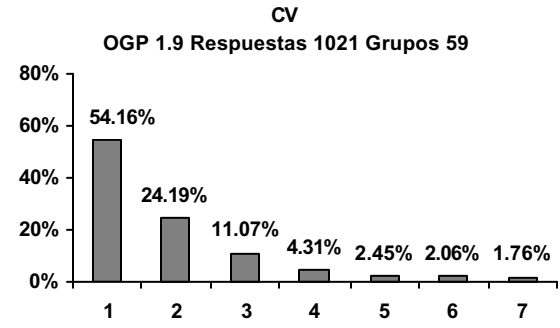
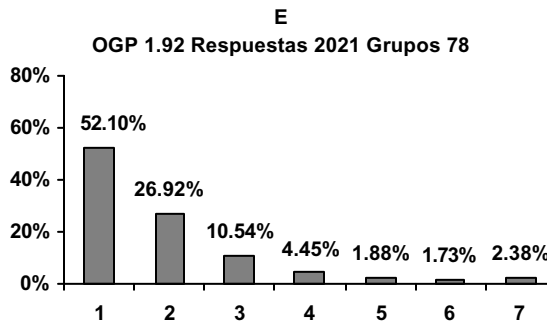
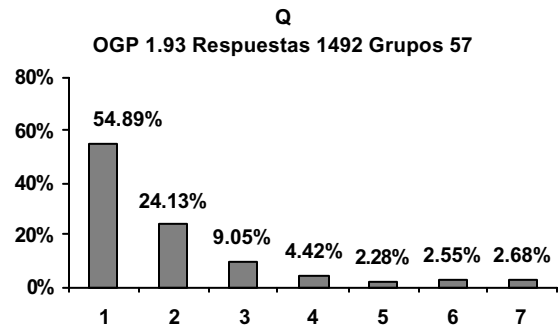
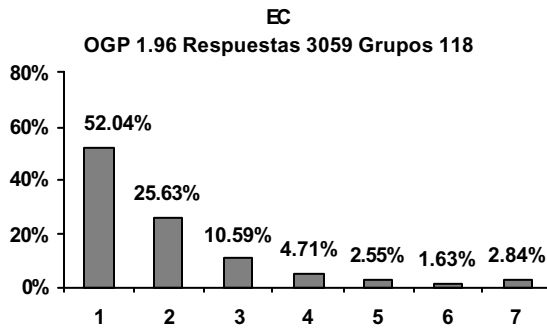
F, AA, EC, E, SI, MA, IN, P, D, FZ, IQ, H, MT, AR, RI, MR, RH, RN y CC.

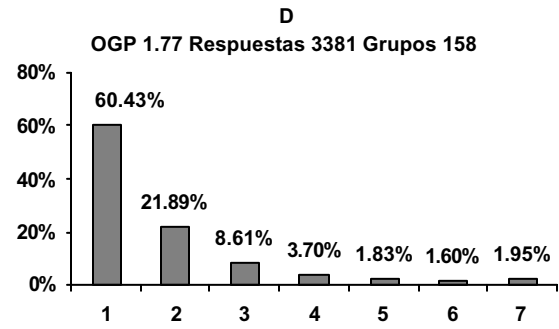
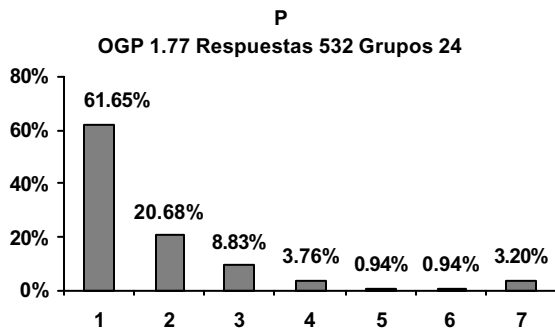
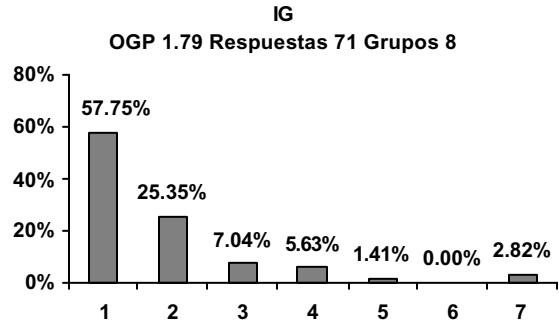
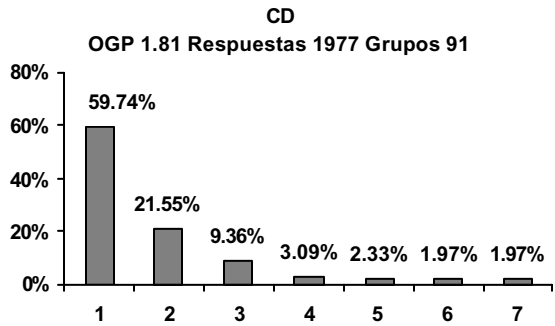
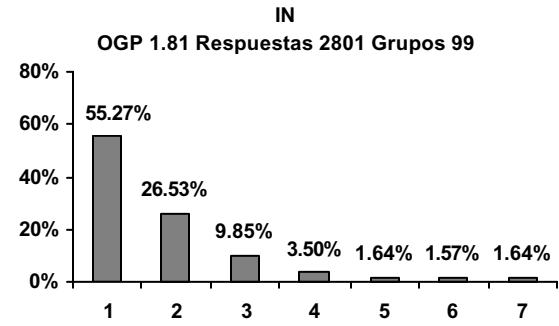
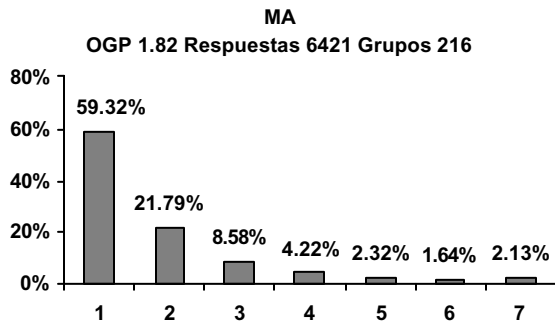
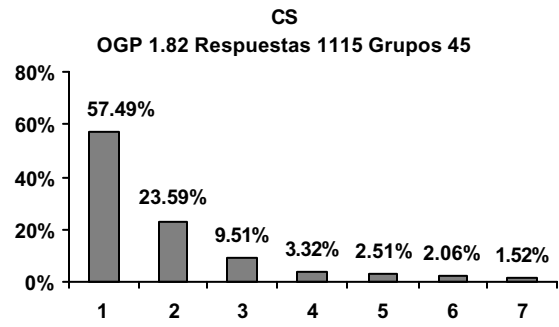
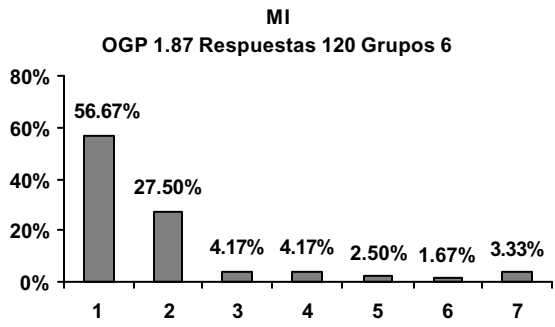
Y por último, señala que las disciplinas de CB, Q, NI, HI, SA, AF y CF reportaron patrones de respuesta decrecientes en los valores de la escala pero con incrementos ya sea en el valor de 6 o de 7 o en ambos.¹⁵

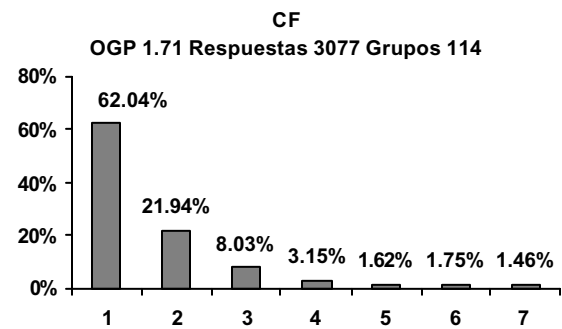
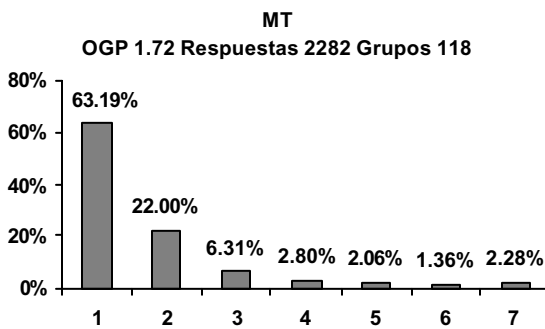
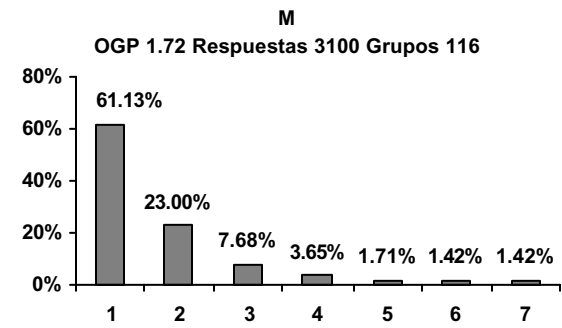
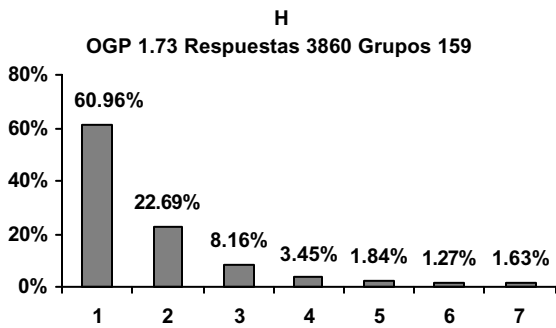
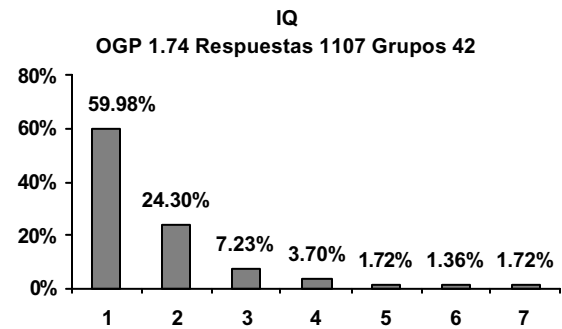
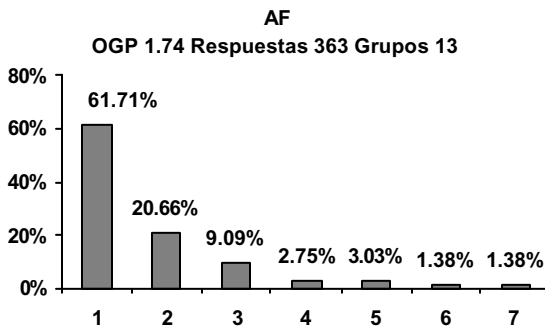
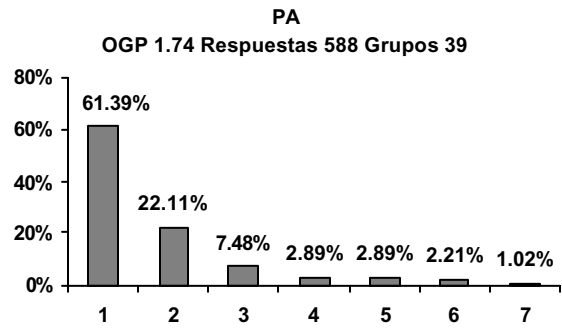
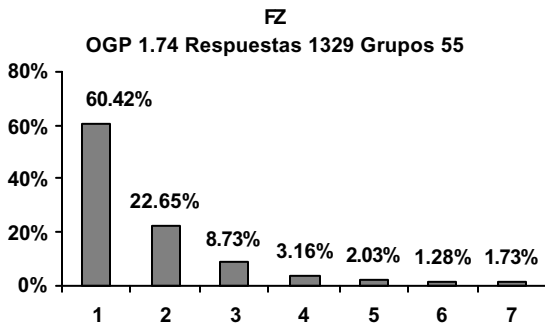
Gráfico 4.1 Patrones de respuesta de la pregunta global por disciplina

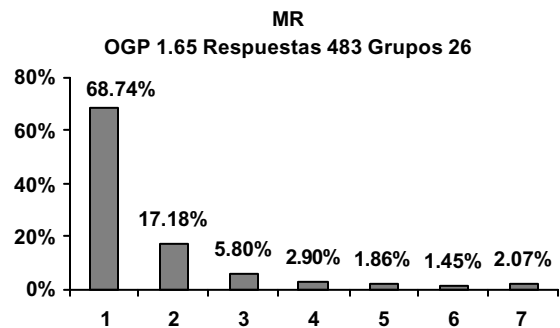
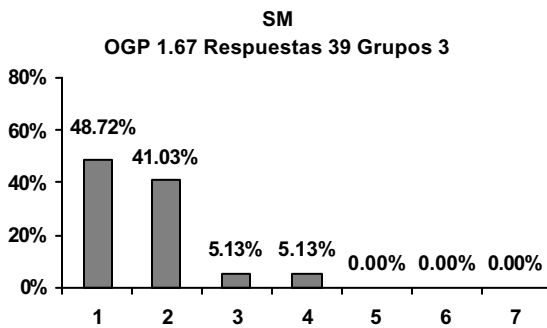
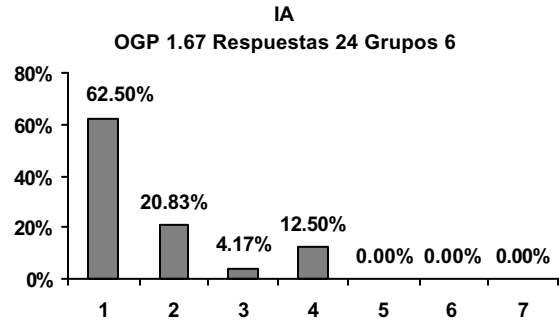
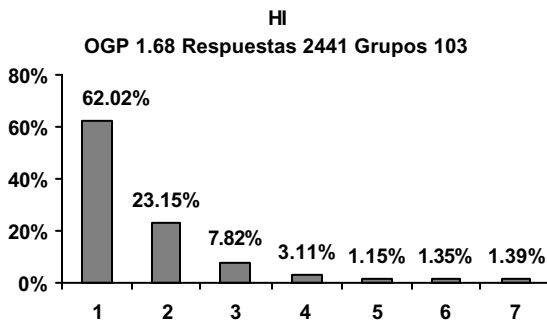
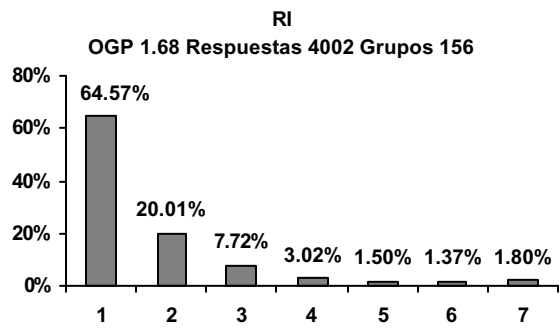
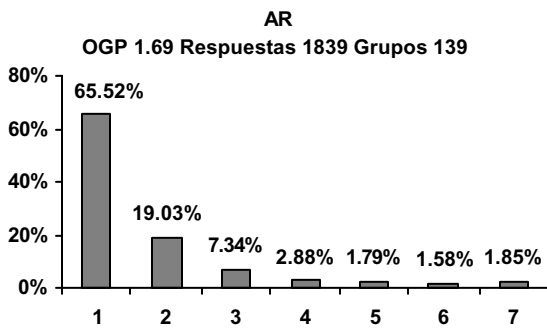
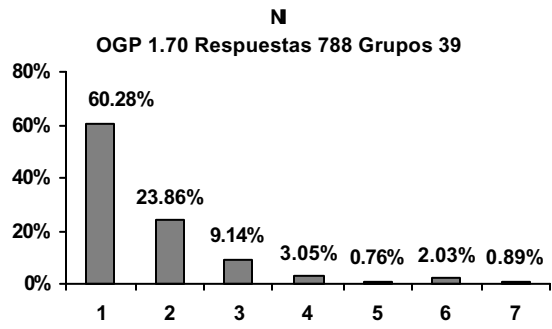
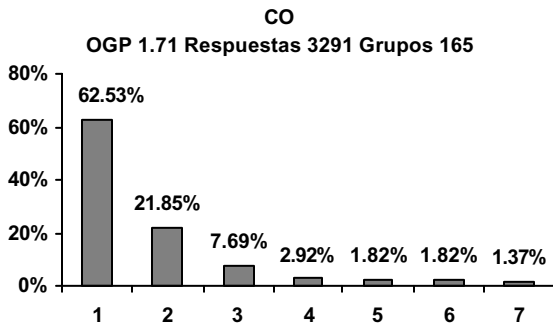


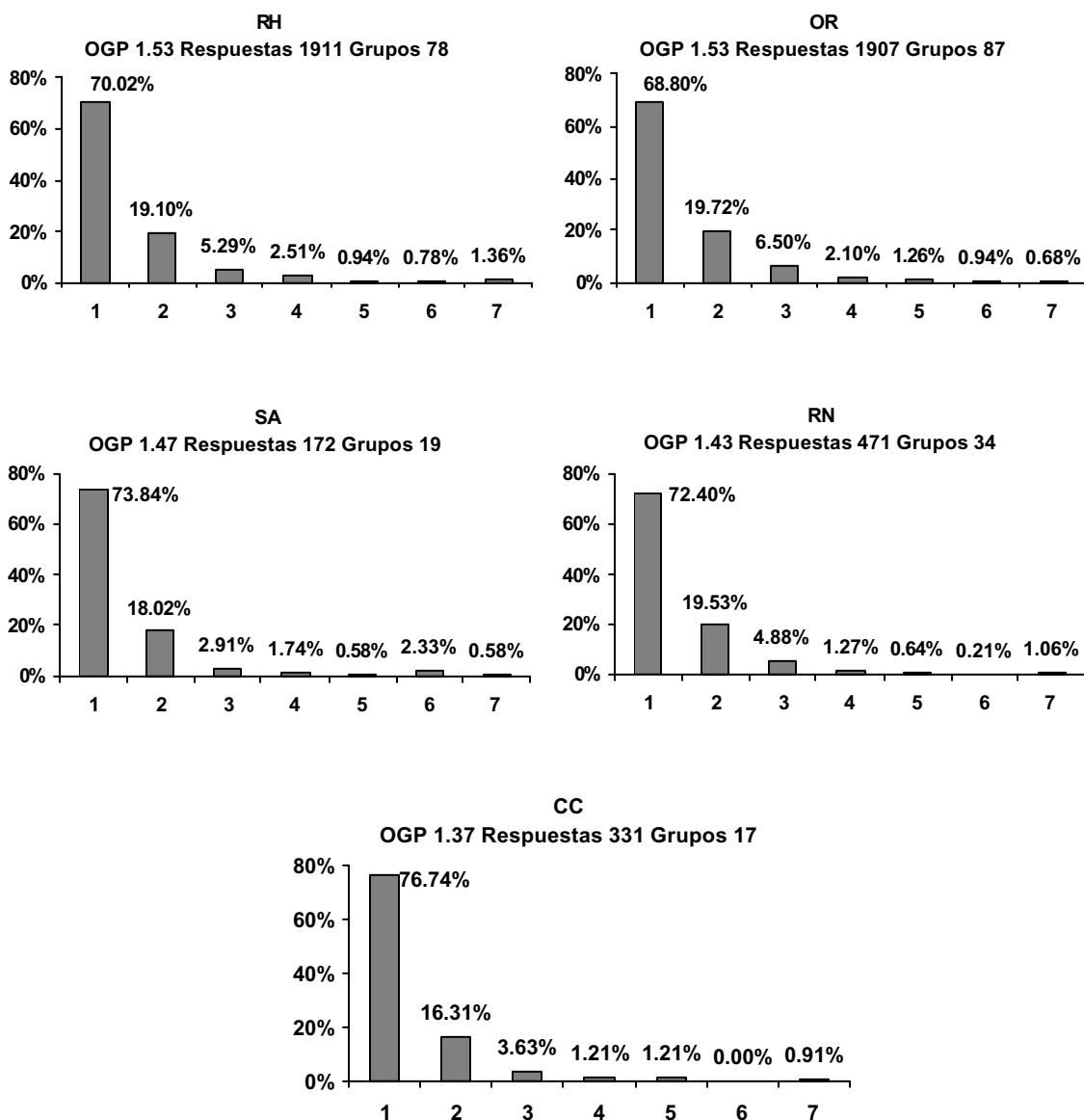
¹⁵ Las disciplinas con menos de 10 grupos (EX, TF, MI, IG, IA y SM) no se analizaron ya que su patrón de respuestas no fue considerado como representativo.











4.2.1.2 EFECTO DEL TIPO DE GRUPO (REDISEÑADO O TRADICIONAL)

El ITESM ha puesto mucho énfasis en rediseñar todos los cursos que ofrece, con el fin de utilizar más las técnicas didácticas en el proceso de enseñanza-aprendizaje. Esto puede hacer que existan diferencias entre el desempeño de los profesores que imparten grupos que siguen el esquema rediseñado y el tradicional, debido al apoyo de una técnica didáctica, lo que puede discernirse con la evaluación global obtenida y con el patrón global de respuestas de cada tipo de grupo.

En la tabla 4.8 se presenta la composición del total de grupos de acuerdo al tipo de grupo, rediseñado o tradicional. Es posible observar que predominaron los grupos rediseñados reportando a su vez una mejor evaluación que los tradicionales.

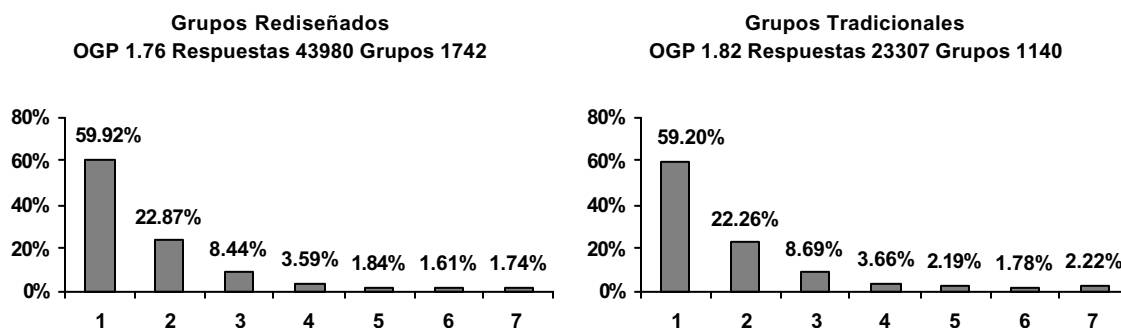
En el gráfico 4.2 se muestran los patrones de respuesta de los dos tipos de grupos, ambos son decrecientes en los valores de 1 a 6 de la escala reportando un incremento en el de 7.

Tabla 4.8 Composición del total de grupos por tipo de grupo (rediseñado o tradicional)

Tipo de grupo	OGP	Respuestas	% del total	Grupos	% del total
Rediseñado	1.76	43980	62.60%	1742	60.44%
Tradicional	1.82	23307	37.40%	1140	39.56%
Total	1.78	67287	100%	2882	100%

Según parece los patrones de respuesta son similares, pero el de los grupos tradicionales reportó un incremento mayor en el valor de 7 de la escala, esto pudo haber ocasionado que resultaran peor evaluados comparados con los rediseñados.

Gráfico 4.2 Patrones de respuesta de la pregunta global por tipo de grupo (tradicional o rediseñado)



Esta información indicó que el tipo de grupo tiene un efecto sobre la evaluación que hacen los alumnos sobre sus profesores, los grupos rediseñados se evalúan mejor que los tradicionales.

4.2.1.3 EFECTO DEL TAMAÑO DEL GRUPO

El tamaño del grupo puede tener una influencia en las evaluaciones que realizan los alumnos sobre sus profesores, esto es, profesores impartiendo grupos grandes pueden tener dificultad para obtener buenas evaluaciones a diferencia de los que imparten grupos chicos.

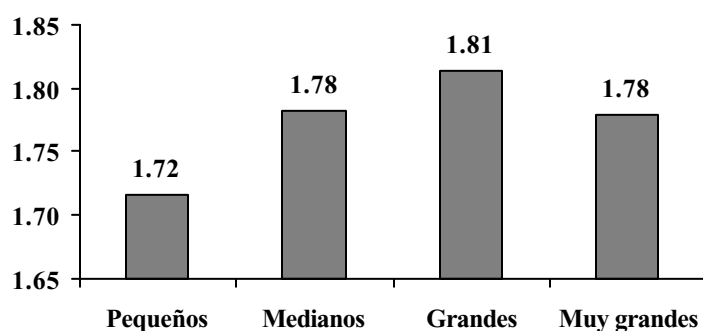
Por este motivo, para analizar este efecto se procedió a dividir el total de grupos conforme al número de alumnos inscritos, dando como resultado 4 tipos de grupos, como se muestra en la tabla 4.9.

Tabla 4.9 Composición del total de grupos por el tamaño del grupo

Tamaño del grupo	Alumnos	OGP	Respuestas	% del total	Grupos	% del total
Pequeño	3-20	1.72	6023	9%	629	22%
Mediano	21-40	1.78	45278	67%	1850	64%
Grande	41-60	1.81	12048	18%	353	12%
Muy grande	61-197	1.78	3938	6%	50	2%
Total		1.78	67287	100%	2882	100%

En esta tabla puede observarse que existieron más grupos de tamaño mediano (de 21 a 40 alumnos), seguidos por los pequeños (3-20), grandes (41-60) y muy grandes (61-197). Además, como muestra el gráfico 4.3 existe una relación curvilínea entre el tamaño del grupo y la evaluación global del profesor, esto es, a medida que se incrementa el tamaño de grupo empeora la evaluación del profesor (relación negativa), pero al pasar de grupos grandes a los muy grandes ésta mejora (relación positiva).

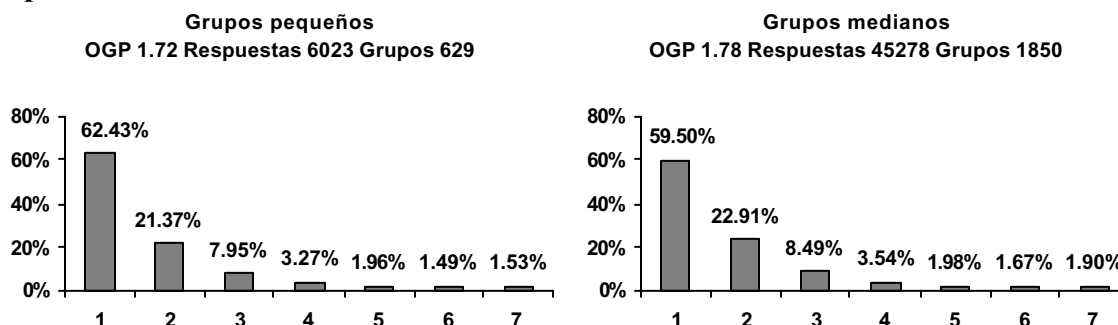
Gráfico 4.3 Evaluación en la pregunta global por tamaño de grupo

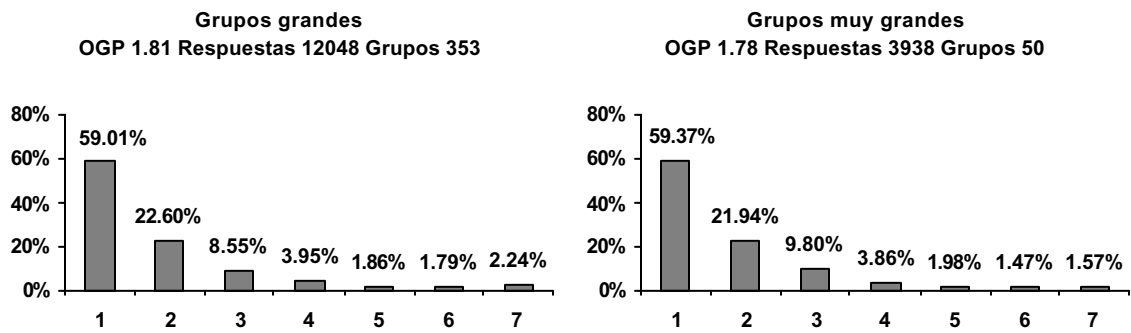


Este curioso resultado puede deberse a que los grupos muy grandes están conformados principalmente por cursos introductorios cursados por alumnos de primer semestre, quienes posiblemente no sean muy críticos al evaluar a sus maestros.

Al considerar en el gráfico 4.4 los patrones de respuesta de estos tipos de grupos, se observa que todos son decrecientes en los valores de la escala de 1 a 6 con un ligero incremento en el de 7, resultando ser el patrón de los grupos grandes el más afectado.

Gráfico 4.4 Patrones de respuesta de la pregunta global de acuerdo al tamaño del grupo.





4.2.1.4 EFECTO DEL RENDIMIENTO ACADÉMICO DE LOS ESTUDIANTES

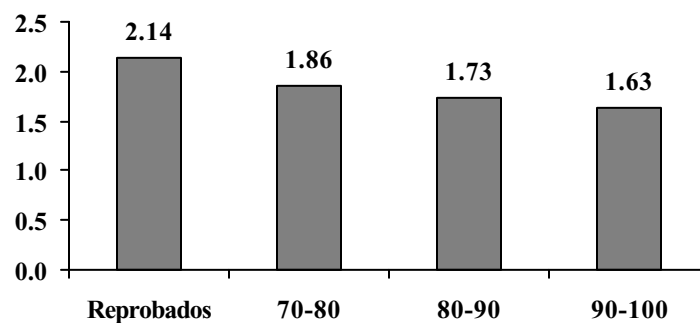
Es posible que las calificaciones obtenidas por los alumnos influyan en las evaluaciones que hacen sobre sus profesores, esto es, que exista una relación positiva entre la evaluación del profesor y la calificación que obtiene el alumno. Para analizar lo anterior, se utiliza la calificación promedio del grupo y se divide el total de grupos de acuerdo a esta variable, como es mostrado en la tabla 4.10.

Tabla 4.10 Composición del total de grupos de acuerdo a la calificación promedio del grupo

Calificación promedio	OGP	Respuestas	% del total	Grupos	% del total
Reprobados	2.14	3998	6%	174	6%
70 a menos de 80	1.86	21900	33%	898	31%
80 a menos de 90	1.73	29982	45%	1295	45%
90 a 100	1.63	11407	17%	515	18%
Total	1.78	67287	100%	2882	100%

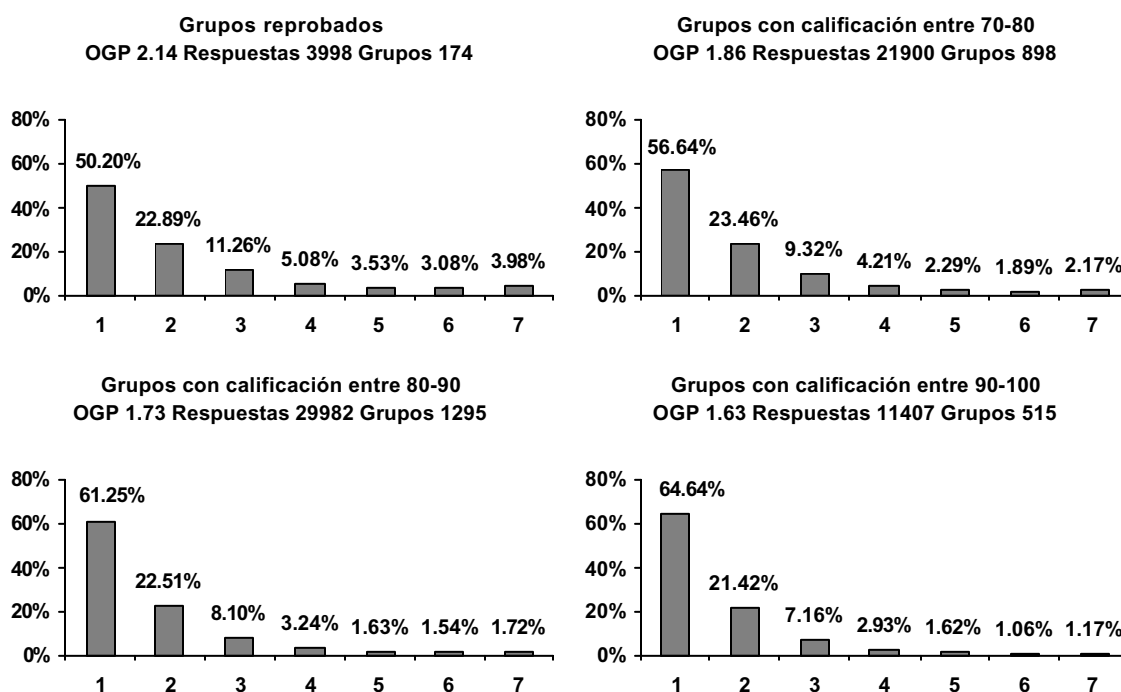
La tabla informa que predominaron los grupos con calificación promedio entre 80 y 90, seguidos por los de 70 a 80, 90 a 100 y los reprobados y de acuerdo al gráfico 4.5 se demuestra que existe una relación positiva entre la evaluación del profesor y la calificación obtenida por el grupo, ya que al aumentar la calificación mejora la evaluación del profesor.

Gráfico 4.5 Evaluación en la pregunta global de acuerdo a la calificación promedio obtenida por el grupo



Los patrones de respuesta de estos grupos son también decrecientes en los valores de 1 a 6 de la escala con un incremento en el valor de 7 como se muestra en el gráfico 4.6, resultando ser el de los grupos reprobados el más afectado por respuestas en 7, seguido por el de los grupos con calificación promedio entre 70-80, 80-90 y 90-100.

Gráfico 4.6 Patrones de respuesta de la pregunta global de acuerdo a la calificación promedio obtenida por el grupo



4.2.1.5 EFECTO DEL GÉNERO DEL PROFESOR

Identificar los patrones de respuesta de los grupos impartidos por mujeres y hombres, permitirá definir quien resultó mejor evaluado. En la tabla 4.11 aparece la composición del total de grupos divididos con base al género del profesor.

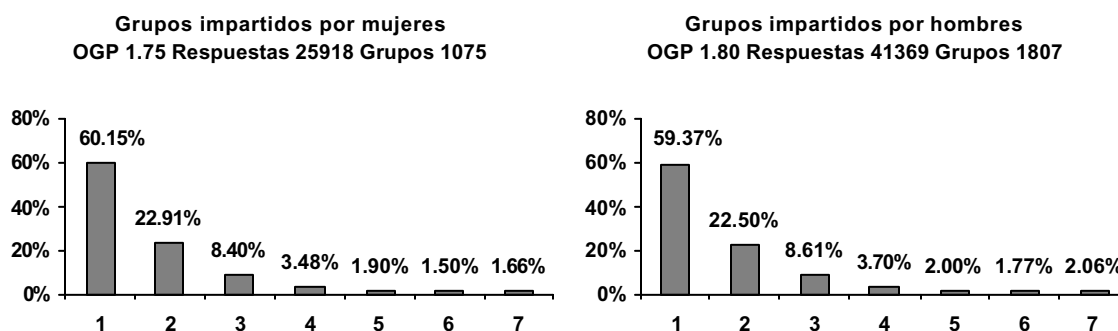
Tabla 4.11 Composición del total de grupos de acuerdo al género del profesor

Género	OGP	Respuestas	% del total	Grupos	% del total
Femenino	1.75	25918	39%	1075	37%
Masculino	1.80	41369	61%	1807	63%
Total	1.78	67287	100%	2882	100%

Predominaron los grupos impartidos por hombres pero resultaron peor evaluados que los impartidos por mujeres, posiblemente debido a que su patrón de respuestas fue el más afectado por evaluaciones en el valor de 7.

El gráfico 4.7 presenta los patrones de acuerdo al género del profesor, ambos son decrecientes en los valores de 1 a 6 de la escala con un incremento en el de 7, resultando ser el de los grupos impartidos por hombres el más afectado por este tipo de evaluaciones.

Gráfico 4.7 Patrones de respuesta de la pregunta global de acuerdo al género del profesor



Si bien esta información aporta evidencia para pensar que las mujeres se desempeñaron mejor que los hombres, es conveniente, para concluir lo anterior, investigar las características de los cursos que imparten y sobre todo si esta relación se mantiene a través del tiempo.

4.2.1.6 EFECTO DE LOS AÑOS DE EXPERIENCIA DEL PROFESOR EN EL INSTITUTO

Otra de las variables que influyen en el desempeño del profesor, es la experiencia como docente de éste dentro del Instituto (años de antigüedad). Cuando un profesor es de recién ingreso comienza a familiarizarse con el método de enseñanza que se sigue en el Instituto, esto puede repercutir en su desempeño. Dividir el total de grupos de acuerdo a esta característica, e identificar sus patrones de respuesta, permitirá investigar si existe una relación positiva entre la evaluación del profesor y los años de experiencia de éste en el Instituto.

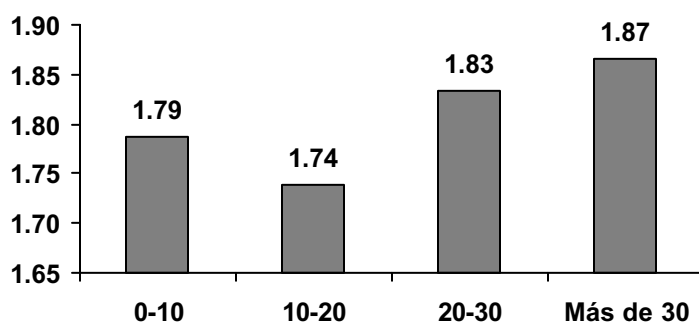
En la tabla 4.12, se presenta los intervalos de antigüedad elegidos, su longitud es de 10 años y van desde 0 a 10 hasta 30-40 años. La información señala que predominaron los grupos impartidos por profesores con 0 a menos de 10 años de experiencia dentro del ITESM, después los impartidos por profesores con 10 a menos de 20, de 20 a menos de 30 y por último los de más de 30 años.

Tabla 4.12 Composición del total de grupos de acuerdo a los años de experiencia del profesor en el Instituto

Años de experiencia	OGP	Respuestas	% del total	Grupos	% del total
0 a menos de 10	1.79	45347	67%	1970	68%
10 a menos de 20	1.74	15293	23%	644	22%
20 a menos de 30	1.83	3902	6%	161	6%
Más de 30	1.87	2745	4%	107	4%
Total	1.78	67287	100%	2882	100%

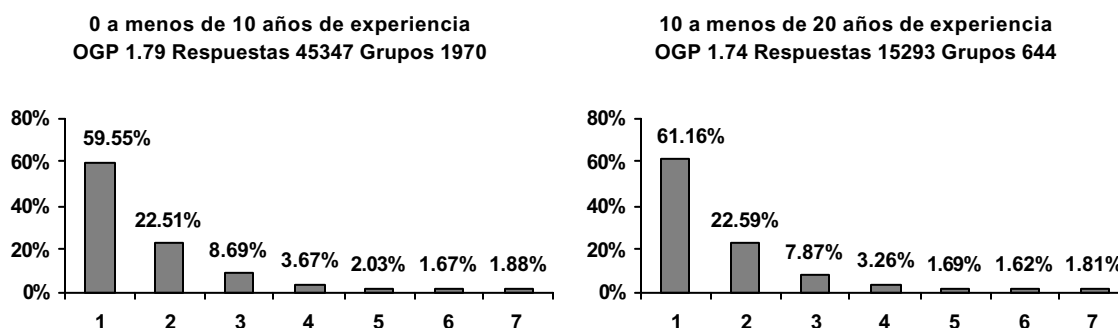
Asimismo, el gráfico 4.8 muestra que existe una relación curvilínea entre la evaluación del profesor y los años de experiencia de éste dentro del Instituto. La evaluación del profesor mejora al aumentar sus años de experiencia (relación positiva), pero sólo hasta cierto punto (10-20 años) porque después empeora (relación negativa).

Gráfico 4.8 Evaluación en la pregunta global de acuerdo a los años de experiencia del profesor dentro del Instituto

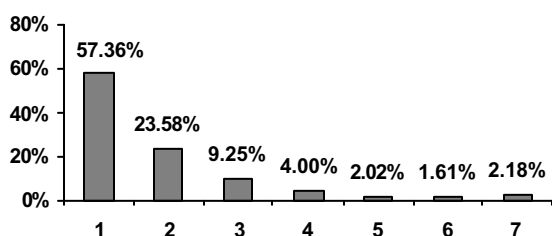


Al analizar los patrones de respuesta de cada una de las divisiones de grupos realizadas de acuerdo a los años de experiencia del profesor, se puede identificar patrones decrecientes en los valores de 1 a 6 de la escala con incrementos en el valor de 7, resultando ser el de los grupos impartidos por profesores con más de 30 años el más afectado por este tipo de evaluaciones, seguido por el de los grupos de 20 a 30, 0 a 10 y por último por el de 10 a 20.

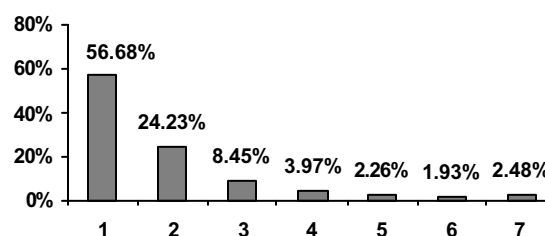
Gráfico 4.9 Patrones de respuesta de la pregunta global de acuerdo a los años de experiencia del profesor en el Instituto



20 a menos de 30 años de experiencia
OGP 1.83 Respuestas 3902 Grupos 161



Más de 30 años de experiencia
OGP 1.87 Respuestas 2745 Grupos 107



La información aquí presentada advierte que los años de experiencia del profesor en el Instituto o de antigüedad óptimos son de 20 años. Sin embargo, esta variable solo refleja la experiencia del profesor dentro del Instituto, pero no dice nada acerca de la que es adquirida en otras Instituciones o en empresas, la que pudiera guardar una relación más estrecha con su desempeño y que puede ser modelada mediante una variable como lo es la edad del profesor.

4.2.1.7 EFECTO DE LA EDAD DEL PROFESOR

La edad del profesor también puede influir en las evaluaciones que asignan los alumnos, ya que a diferencia de los años de antigüedad, permite capturar la experiencia que tiene el profesor no sólo en el ITESM si no fuera de él, en otras instituciones educativas o en empresas. Así, dividir el total de grupos de acuerdo a esta característica e identificar sus patrones de respuesta, permitirá indagar si existe una relación entre la edad del profesor y su evaluación.

La tabla 4.13 informa que imperaron los grupos impartidos por profesores cuya edad fue de 30 a menos de 40 años, seguidos por los de 40 a menos de 50, 50 a menos de 60, 23 a menos de 30, 60 a menos de 70 y por último los impartidos por profesores mayores de 70 años.

Tabla 4.13 Composición del total de grupos de acuerdo a la edad del profesor

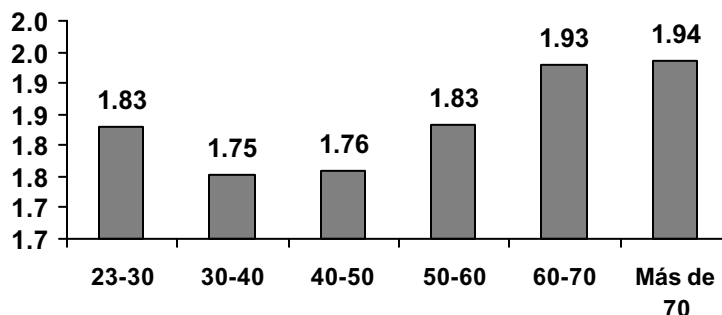
Edad del profesor	OGP	Respuestas	% del total	Grupos	% del total
23 a menos de 30	1.83	5868	9%	261	9%
30 a menos de 40	1.75	26187	39%	1098	38%
40 a menos de 50	1.76	21643	32%	917	32%
50 a menos de 60	1.83	10422	15%	463	16%
60 a menos de 70	1.93	2913	4%	127	4%
Más de 70	1.94	254	0%	16	1%
Total	1.78	67287	100%	2882	100%

La información aporta evidencia que señala una relación curvilínea entre la evaluación del profesor y la edad, a medida que se incrementa ésta última la evaluación del profesor

mejora (relación positiva) pero hasta cierto punto, entre los 30 y 40 años, porque después empeora (relación negativa), ver gráfico 4.10.

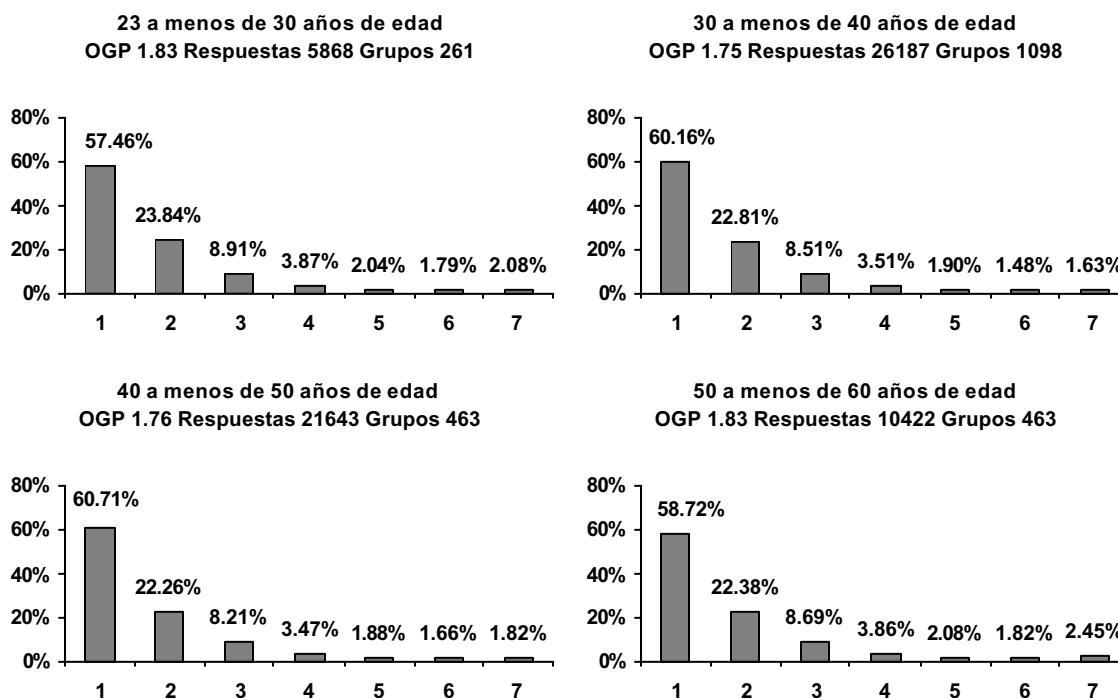
Todos los patrones de respuesta de estas divisiones de grupos fueron decrecientes en los valores de 1 a 6 de la escala y afectados por evaluaciones en el valor de 7, resultando los más afectados los que obtuvieron una peor OGP, ver gráfico 4.11.

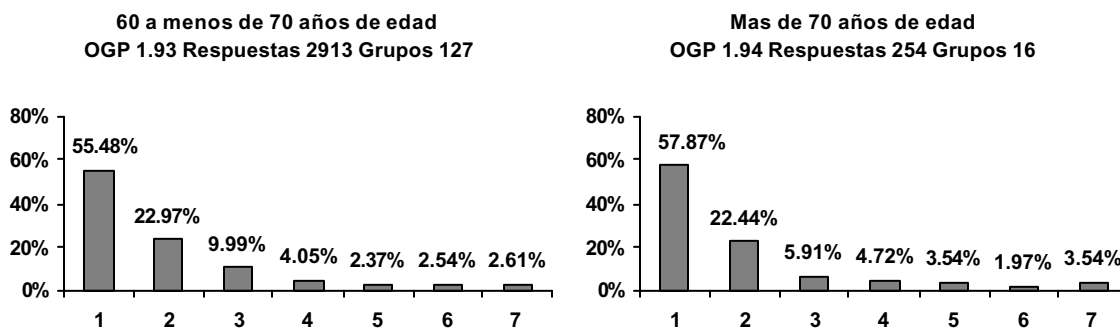
Gráfico 4.10 Evaluación en la pregunta global de acuerdo a la edad del profesor



Los resultados encontrados con esta variable sustentan los obtenidos con la de “años de experiencia del profesor en el Instituto”, la evaluación del profesor mejora con el tiempo debido posiblemente a la madurez que adquiere como persona y a la experiencia, pero esta relación solamente se mantiene hasta cierto punto.

Gráfico 4.11 Patrones de respuesta de la pregunta global de acuerdo a la edad del profesor





4.2.1.8 EFECTO DEL TIPO DE CONTRATO DEL PROFESOR

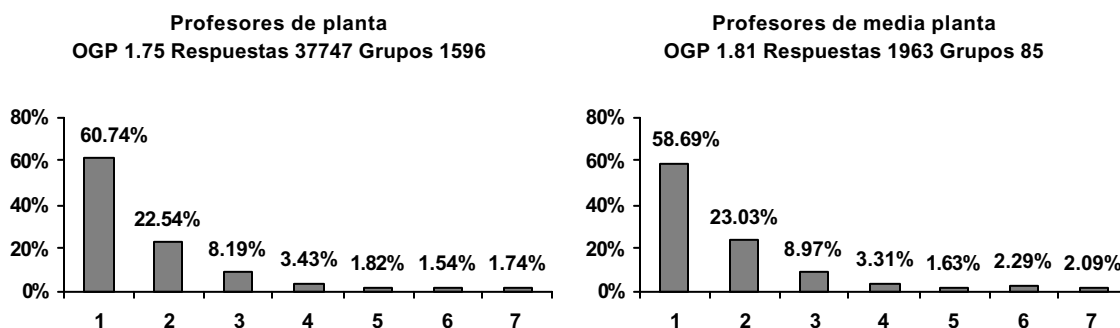
Existen los grupos impartidos por profesores de planta, media planta y auxiliares (éstos últimos son los profesores que no entran en las dos primeras categorías), por tanto podría separarse el total de grupos en estas tres divisiones y es factible pensar que tanto sus evaluaciones como sus patrones de respuesta son diferentes. En la tabla 4.14 se presentan los grupos divididos de acuerdo al tipo de contrato del profesor. La información muestra que los profesores de planta impartieron el mayor número de grupos, luego los profesores auxiliares y por último los de media planta.

Tabla 4.14 Composición del total de grupos de acuerdo al tipo de contrato del profesor

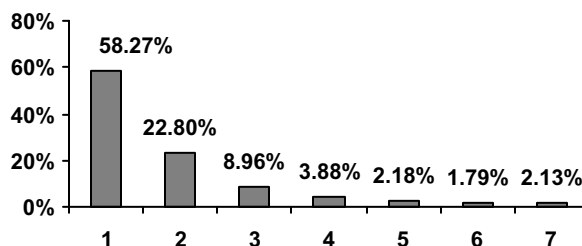
Tipo de contrato	OGP	Respuestas	% del total	Grupos	% del total
Planta	1.75	37747	56%	1596	55%
Media Planta	1.81	1963	3%	85	3%
Auxiliares	1.83	27577	41%	1201	42%
Total	1.78	67287	100%	2882	100%

También, puede cotejarse que los grupos impartidos por profesores de planta fueron los mejor evaluados, seguidos por los de media planta y por los auxiliares; los tres tipos de grupos reportaron patrones de respuesta decrecientes en los valores de 1 a 6 con un incremento en el de 7, lo que afectó las evaluaciones obtenidas ya que los grupos con los patrones más afectados por evaluaciones en 7 resultaron peor evaluados, ver gráfico 4.12.

Gráfico 4.12 Patrones de respuesta de la pregunta global de acuerdo al tipo de contrato del profesor



Profesores auxiliares
OGP 1.83 Respuestas 27577 Grupos 1201



4.2.1.9 EFECTO DEL GRADO MÁXIMO DE ESTUDIOS DEL PROFESOR

Por último, los grupos pueden separarse dependiendo del grado máximo de estudios del profesor, es decir, por los impartidos por profesores cuyo grado máximo es el título profesional, maestría o doctorado. Esta diferencia en cuanto a preparación, puede verse reflejada en la evaluación de los profesores y en sus patrones de respuesta.

En la tabla 4.15 se reporta el desglose de los grupos de acuerdo al grado máximo de estudios del profesor. Es posible observar que los profesores con Maestría impartieron la mayor parte de los grupos y después de ellos estuvieron los profesores con Doctorado y detrás de éstos, aparecieron los profesores con sólo título profesional.

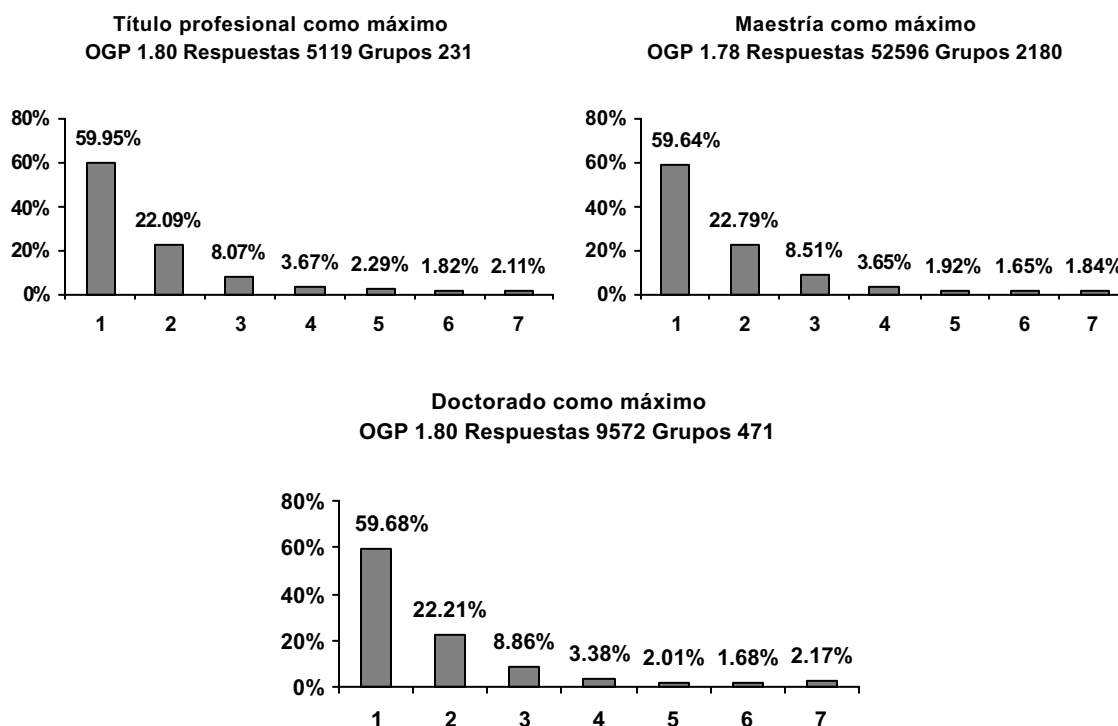
Tabla 4.15 Composición del total de grupos de acuerdo al grado máximo de estudios del profesor

Grado Máximo del profesor	OGP	Respuestas	% del total	Grupos	% del total
Título Profesional	1.80	5119	8%	231	8%
Maestría	1.78	52596	78%	2180	76%
Doctorado	1.80	9572	14%	471	16%
Total general	1.78	67287	100%	2882	100%

La evaluación obtenida, OGP, en estas divisiones de grupos aporta evidencia que indica que los grupos mejor evaluados fueron los impartidos por profesores con Maestría como máximo, apareciendo con una evaluación similar los grupos impartidos por profesores con título profesional como máximo y con doctorado.

Los patrones de respuesta de estas divisiones de grupo se reportan en el gráfico 4.13 y se observa que todos fueron decrecientes en los valores de 1 a 6 de la escala con un incremento en el de 7; asimismo, puede identificarse que el patrón de los grupos impartidos por profesores con grado de Maestría como máximo resultó el menos afectado por evaluaciones en el valor de 7 de la escala obteniendo la mejor evaluación.

Gráfico 4.13 Patrones de respuesta de la pregunta global de acuerdo al grado máximo de estudios del profesor



Estos resultados señalan que no existe una relación entre la evaluación que hacen los alumnos sobre el desempeño de sus profesores y el grado de preparación académica de éstos.

Hasta aquí, se ha investigado la posibilidad de que la OGP está sesgada por variables externas al proceso de enseñanza-aprendizaje, lo cual significaría que no es un indicador válido para evaluar el desempeño de los profesores. Sin embargo, estas relaciones sólo son descriptivas, así hasta que no se determine qué tan significativo es estadísticamente el sesgo encontrado, no será posible realizar una conclusión.

En la siguiente sección se abordarán de nuevo estos resultados utilizando herramientas estadísticas que permitan determinar significativamente si la OGP es válida para cuantificar el desempeño del profesor, es decir, si esta sesgada significativamente por variables externas al proceso de enseñanza-aprendizaje.

4.2.2 EVALUACIÓN GLOBAL DEL PROFESOR MODELADA CON LAS VARIABLES QUE SE DESEA “NO” SEAN FUENTES DE EXPLICACIÓN

Los resultados de las estimaciones de los modelos de regresión donde se modela la OGP con las variables externas al proceso de enseñanza-aprendizaje se reportan en la tabla 4.16, pero para comprender mejor las estimaciones es necesario explicar detalladamente la

información. Esta tabla presenta los modelos estimados en las columnas, en total 7, y las variables explicativas en las filas. No obstante, es conveniente aclarar que los modelos no incluyeron a todas las variables reportadas en las filas, las variables no incluidas en un modelo se señalaron en color gris además de reportar lo siguiente “-”.

También, la tabla incluye en las filas la R-cuadrada ajustada, la que determinará qué tanto explicaron las variables explicativas (las externas) a la OGP, y reporta un estadístico F junto con su probabilidad para verificar si el modelo fue significativo conjuntamente, además cada coeficiente estimado reportó una probabilidad para verificar si se rechaza o no la hipótesis nula de significancia estadística, un valor menor a 0.05 indicará que existe evidencia suficiente para rechazar la hipótesis de que el valor del coeficiente es igual a cero.

Es importante señalar que todos los modelos presentaron problemas de heterocedasticidad, por tal motivo se estimó la matriz consistente de varianzas y covarianzas de White (1980) para poder realizar pruebas válidas de significancia estadística sobre los coeficientes estimados de las variables explicativas. Adicionalmente, debido a que las variables no estuvieron correlacionadas, el problema de multicolinealidad no estuvo presente.¹⁶

Explicando los resultados encontrados, se observó en los modelos del 1 al 4 y en el 6, que todas las variables cuantitativas fueron significativas, en cambio en el modelo 5 sólo resultó significativa la variable de la calificación promedio y la del tamaño del grupo, al igual que en el modelo 7.¹⁷

En cuanto a los signos reportados por las variables cuantitativas, se encontraron resultados interesantes consistentes en todos los modelos.

El coeficiente de la variable “calificación promedio del grupo” resultó con signo negativo, lo que expresó una relación positiva y significativa con la OGP, es decir, se observó que a mayor calificación promedio obtenida por el grupo mejor fue la evaluación del profesor.

Esta relación debe de interpretarse con cautela ya que puede pensarse que los alumnos obtuvieron buenas calificaciones porque aprendieron muy bien dado que el profesor tuvo un buen desempeño, o porque el profesor fue muy complaciente.

¹⁶ Fue necesario eliminar 6 grupos que reportaban calificaciones promedio erróneas (de cero), por lo cual el total de observaciones utilizado para realizar las estimaciones de la tabla 5 fue de 2876 grupos.

¹⁷ En el modelo 7, se incluyeron los términos cuadráticos de las variables tamaño de grupo, antigüedad del profesor y edad, sin embargo el método de paso a paso utilizado para especificar el modelo, no las consideró.

Tabla 4.16 Estimación del modelo de regresión, OGP versus las variables externas al proceso de enseñanza-aprendizaje

Variable dependiente: OGP. Todas las estimaciones fueron corregidas por heterocedasticidad.							
Variables	Modelo 1 Coef. Prob.	Modelo 2 Coef. Prob.	Modelo 3 Coef. Prob.	Modelo 4 Coef. Prob.	Modelo 5 Coef. Prob.	Modelo 6 Coef. Prob.	Modelo 7 Coef. Prob.
Constante	2.647 0.000	2.509 0.000	2.769 0.000	2.692 0.000	2.721 0.000	2.655 0.000	2.864 0.000
Calificación del grupo	-0.013 0.000	-0.012 0.000	-0.014 0.000	-0.013 0.000	-0.013 0.000	-0.013 0.000	-0.013 0.000
Tamaño del grupo	0.003 0.000	0.002 0.002	0.003 0.000	0.003 0.000	0.003 0.000	0.003 0.000	0.003 0.000
Años de experiencia en el Instituto	-0.004 0.016	-0.005 0.004	-0.003 0.040	-0.004 0.006	-0.001 0.463	-0.004 0.014	-- --
Edad	0.003 0.013	0.004 0.002	0.003 0.032	0.003 0.025	0.003 0.053	0.003 0.023	-- --
Dummy_CB	-- --	0.190 0.002	-- --	-- --	-- --	-- --	0.233 0.000
Dummy_EC	-- --	0.170 0.002	-- --	-- --	-- --	-- --	0.157 0.003
Dummy_F	-- --	0.201 0.004	-- --	-- --	-- --	-- --	0.182 0.011
Dummy_HI	-- --	-0.150 0.000	-- --	-- --	-- --	-- --	-0.119 0.001
Dummy_IN	-- --	0.126 0.028	-- --	-- --	-- --	-- --	0.179 0.001
Dummy_OR	-- --	-0.126 0.006	-- --	-- --	-- --	-- --	-- --
Dummy_RH	-- --	-0.137 0.006	-- --	-- --	-- --	-- --	-- --
Dummy_Femenino	-- --	-- --	-- --	-0.067 0.002	-- --	-- --	-- --
Dummy_Rediseño	-- --	-- --	-0.103 0.000	-- --	-- --	-- --	-0.104 0.000
Dummy_Planta	-- --	-- --	-- --	-- --	-0.097 0.000	-- --	-0.121 0.000
Dummy_Media Planta	-- --	-- --	-- --	-- --	0.017 0.810	-- --	-- --
Dummy_Maestría	-- --	-- --	-- --	-- --	-- --	0.037 0.449	-- --
Dummy_Doctorado	-- --	-- --	-- --	-- --	-- --	-0.009 0.822	-- --
R-cuadrada Ajustada	0.043	0.059	0.050	0.046	0.049	0.043	0.072
Estadístico F de Fisher	33.36	17.50	31.35	28.59	25.48	22.66	23.20
Probabilidad	0.000	0.000	0.000	0.000	0.000	0.000	0.000

Estos resultados fueron equivalentes a los encontrados con el análisis descriptivo, en este análisis se aportó evidencia para pensar en la existencia de esta relación entre la OGP y el promedio del grupo, ya que los patrones de respuesta menos afectados por evaluaciones en el valor de 7 de la escala, fueron los de los grupos que obtuvieron buenas calificaciones.

Debido a que con este análisis no se puede inferir más allá de la existencia de dicha relación, sólo resta comentar que se debe de tener cuidado para que esta relación entre la OGP y la calificación promedio del grupo, no sea manejada mal intencionadamente.

En el caso particular del efecto del tamaño del grupo sobre la OGP del profesor, se observó que el coeficiente de esta variable fue positivo, lo que significa que se encontró una relación negativa entre el tamaño de grupo y la evaluación del profesor, más explícitamente, esto quiere decir que a medida que el tamaño de grupo se incrementa, la evaluación del profesor empeora debido posiblemente a que no tendrá el tiempo suficiente para atender a los alumnos. De este modo, es importante estudiar cual sería el tamaño de grupo óptimo en el que, tanto el profesor como el alumno se desempeñen bien.

Con lo que respecta a la relación de la OGP con la variable de los años de experiencia del profesor adquiridos en el ITESM, se advirtió que presentó un coeficiente con el signo esperado, negativo, esto implicó una relación positiva, esto es, entre más años de experiencia adquiera el profesor en el Instituto mejor será su desempeño.

La relación de la última variable cuantitativa considerada, la edad del profesor, con la OGP, fue negativa ya que el signo del coeficiente estimado fue positivo, es decir, a mayor edad peor evaluación. Esta relación fue contraria a la esperada, se consideró que esta variable podría cuantificar los años de experiencia del profesor no sólo los adquiridos en el ITESM si no también en otras Universidades o en empresas, así se esperaba que el coeficiente de esta variable fuera negativo.

En el análisis de las variables cualitativas, incluidas en el modelo base que sólo contemplaba a las cuantitativas, se pudo observar lo siguiente.

De las disciplinas que predominaron en profesional y en las Divisiones del Campus Monterrey, solamente resultaron significativas F, CB, EC, IN, OR, RH y HI, resultando las 4 primeras con signo positivo y las 3 restantes con signo negativo, modelo 2. Estas disciplinas se presentaron en orden descendente de acuerdo a la magnitud de su coeficiente, lo cual significa que de las disciplinas la peor evaluada fue F y la mejor HI. Esto es, los profesores que impartieron materias de la disciplina de F obtuvieron peores evaluaciones que los que impartieron materias de la disciplina HI, posiblemente debido a la dificultad del curso.

El modelo 3, indicó que la relación de la OGP con el rediseño fue negativa, esto es, los grupos rediseñados resultaron mejor evaluados que los grupos tradicionales.

Con respecto al género del profesor, modelo 4, el análisis de regresión indicó que las mujeres resultaron mejor evaluadas que los hombres, ya que el signo del coeficiente de esta variable fue negativo y significativo.

Analizando el efecto del tipo de contrato del profesor, modelo 5, se observó que de las dos variables incluidas, solamente la que representó a los profesores de planta fue significativa con un coeficiente negativo, en cambio la de los profesores de media planta fue no significativa y de signo positivo. Así, si la constante esta cuantificando bien el efecto de los profesores auxiliares, esta información indicaría que los profesores de planta obtuvieron mejores evaluaciones que los profesores de media planta y que los auxiliares.

En el modelo 6, donde se investigó la relación del grado máximo de estudios de los profesores con la OGP, se encontró que las dos variables incluidas para representar a los profesores con maestría como grado máximo y con doctorado, fueron no significativas, la primera con coeficiente con signo positivo y la segunda con uno negativo. Dados los coeficientes es posible identificar una relación positiva entre la evaluación del profesor y el grado máximo de estudios, sin embargo no significativa.

Por último, el modelo 7 el cual utilizó las variables más significativas sin importar si fueron o no cualitativas o cuantitativas, demostró significativamente que la evaluación del profesor depende positivamente de la calificación promedio del grupo y negativamente del tamaño del grupo; asimismo, que las disciplinas que resultaron peor evaluadas en profesional, en orden descendente, fueron: CB, F, IN y EC, en cambio la que resultó mejor evaluada fue: HI; además, que el efecto del rediseño fue importante, los grupos rediseñados resultaron mejor evaluados que los tradicionales, y que los profesores de planta resultaron mejor evaluados que los de media planta y que los auxiliares.¹⁸

Estos resultados indicaron que el indicador, OGP, utilizado para evaluar el desempeño de los profesores, estuvo sesgado por variables externas al proceso de enseñanza-aprendizaje, las que estuvieron representadas por características del curso, grupo y del profesor.

Si bien se encontró un sesgo significativo, éste fue muy pequeño como lo demuestra la R cuadrada ajustada reportada en todos los modelos presentados en la tabla 4.16, ésta estuvo entre 0.043 y 0.072, es decir, en un modelo de regresión donde la variable a explicar fue la OGP y las variables explicativas las variables externas al proceso de enseñanza-aprendizaje, se pudo explicar la variabilidad de la información de la OGP entre un 4.3% y un 7.2% resultando los modelos especificados significativos como lo demuestra el estadístico F de Fisher.

¹⁸ También se incluyeron los términos cuadráticos de las variables tamaño de grupo, antigüedad y edad del profesor, sin embargo, no fueron significativos.

4.3 RESUMEN DE LOS RESULTADOS ENCONTRADOS ACERCA DE LA VALIDEZ

En este capítulo se encontró evidencia acerca de la validez de las opiniones de los alumnos para evaluar el desempeño de los profesores.

El indicador utilizado para evaluar al profesor, la Opinión Global Promedio, OGP, elaborado a partir de las opiniones de los alumnos, fue explicado en mayor parte por variables internas al proceso de enseñanza-aprendizaje, las que se desean sean fuentes de explicación, y en menor magnitud por las variables externas, las cuales no tienen nada que ver con el desempeño del profesor y se desea no sean fuentes de explicación.

La variabilidad de la información aportada por la OGP fue explicada en un 92.9% por las preguntas específicas de la encuesta, utilizadas como variables internas, y por un 89.4% por dos factores estimados que resumieron en un 78.67% la información de estas preguntas; en cambio, fue explicada entre un 4.3% y un 7.2% por las variables externas modeladas con características del curso, grupo y/o del profesor.

Las preguntas específicas modelaron mejor a la OGP que los factores estimados siendo la 12, 11, 1, 4, 9, 2, 5 y 7 (ordenadas en orden descendente de acuerdo a su poder explicativo) las preguntas que determinaron el modelo.

De los factores estimados el primero fue el que tuvo más peso, resumió el 72.44% de la información de las preguntas específicas y se interpretó como el factor subjetivo porque estuvo correlacionado con casi todas las preguntas (con la 6, 4, 7, 8, 12, 11, 5, 9, 1 y 10, ordenadas de mayor a menor correlación), sobre todo con las subjetivas. Por otro lado, el segundo que se estimó, concentró el 6.23% de la información y sólo mostró correlaciones elevadas con dos preguntas (la 2 y 3), esto lo identificó como el factor cumplimiento por los aspectos que cuantificó (puntualidad, asistencia, cumplimiento con asesoría y con el programa del curso).

Si bien las variables externas explicaron en un pequeño porcentaje a la OGP, el sesgo que aportan a las evaluaciones de los alumnos no debe de ser pasado por alto porque fue significativo.

Las relaciones encontradas al modelar a la OGP con estas variables fueron las siguientes:

Efecto de la disciplina del curso

Las disciplinas que resultaron significativamente mejor evaluadas fueron las de HI, OR y RH, en cambio las peor evaluadas CB, EC y F, posiblemente debido a la dificultad de éstas últimas.

Efecto del tipo de grupo: rediseñado y tradicional

Se encontró tanto en el análisis descriptivo como en el de regresión, que los profesores impartiendo grupos rediseñados resultaron significativamente mejor evaluados que los que impartieron grupos tradicionales, entonces es factible que la técnicas didácticas ayuden a los profesores a tener un mejor desempeño en el salón de clases.

Efecto del tamaño del grupo

El análisis descriptivo indicó que el tamaño del grupo influyó en forma negativa sobre la evaluación del profesor, curiosamente esta relación fue curvilínea, esto es, se mantuvo solamente de los grupos pequeños (de 3 a 20 alumnos) a los grandes (de 41 a 60) pero después se modificó, fue positiva de los grupos grandes a los muy grandes (de 61 a 197).

Lo anterior se atribuyó a que los grupos muy grandes correspondieron a cursos introductorios, donde quizás los alumnos aún no tienen una visión madura para evaluar a sus profesores.

Sin embargo, en el análisis de regresión no fue posible capturar esta relación curvilínea (el término cuadrático incluido para capturarla no fue significativo), pero sí indicó que fue negativa y significativa.

Aún así, esta evidencia sugiere que es importante realizar una investigación para indagar el tamaño de grupo óptimo en donde interactúen eficientemente tanto los profesores como los alumnos.

Efecto del rendimiento académico de los estudiantes

Tanto el análisis descriptivo como el de regresión reportaron una relación positiva y significativa entre la evaluación del profesor y la calificación promedio del grupo, a mayor calificación mejor evaluación.

De esta forma, esta relación debe de estar presente al decidir si un profesor se desempeña bien o no, porque puede ser posible que la variable utilizada esté relacionada tanto con las expectativas de los alumnos sobre su calificación final del curso, como con la intención del profesor de asignar buenas calificaciones para obtener una buena evaluación.

Si bien con este análisis no se puede hacer inferencias sobre lo anterior, sólo resta comentar que se debe de tener cuidado para que esta relación entre la OGP y la calificación promedio del grupo sea manejada correctamente.

Efecto del género del profesor

Tanto el análisis descriptivo como el de regresión informaron que los grupos impartidos por mujeres resultaron significativamente mejor evaluados que los ofrecidos por hombres.

Con esta evidencia podría pensarse que las mujeres se desempeñaron mejor que los hombres, pero no necesariamente, esto sólo sería si las características del curso y del grupo fueran idénticas, y si esta relación se observara a través del tiempo.

Años de experiencia del profesor en el Instituto

La relación gráfica de esta variable indicó que la evaluación del profesor mejora al aumentar los años de experiencia (relación positiva), pero sólo hasta cierto punto (entre los 10 y 20 años) porque después empeora (relación negativa), esto señaló la presencia de una relación curvilínea entre estas dos variables.

En el análisis de regresión, también se advirtió una relación positiva y significativa, pero no fue posible capturar la relación curvilínea observada en el análisis gráfico, debido a que el término cuadrático incluido en el modelo estimado no fue significativo.

De esta manera, los años de experiencia del profesor dentro del Instituto son un factor importante que influye en la evaluación de los alumnos, lo que puede ser lógico ya que entre más años de experiencia se tengan más se conocerá el método de enseñanza.

Efecto de la edad del profesor

El análisis descriptivo también aportó evidencia que señala una relación curvilínea entre la evaluación del profesor y la edad, a medida que se incrementa ésta última la evaluación del profesor mejora (relación positiva) pero hasta cierto punto, entre los 30 y 40 años, porque después empeora (relación negativa).

En cambio, esta relación no fue encontrada en el análisis de regresión ya que el término cuadrático de esta variable no fue significativo, sólo se observó una relación negativa y significativa en la mayoría de los modelos estimados, esto brindó información diferente a la encontrada con la variable de años de experiencia, aquí el efecto del paso del tiempo si afectó en la evaluación del profesor.

Esto es, se pensaba que la variable de edad podría capturar los años de experiencia del profesor no solo dentro del Instituto, sino también en otras instituciones o en empresas, pero no fue así.

Efecto del tipo de contrato

Tanto en el análisis descriptivo como en el de regresión, se halló que los grupos impartidos por profesores de planta fueron los mejor evaluados, esto posiblemente porque estos profesores están dedicados de tiempo completo a la docencia y pueden invertir más tiempo en sus alumnos.

Efecto del grado máximo de estudios del profesor

Los resultados del análisis descriptivo indicaron que los grupos impartidos por profesores con Maestría resultaron mejor evaluados que los impartidos por profesores con Doctorado o con sólo título profesional (estos dos últimos obteniendo una evaluación similar), pero el análisis de regresión señaló que no existe una distinción significativa, se encontró una relación positiva entre la preparación del profesor y su evaluación pero no significativa.

En resumen, la información que resultó de este capítulo señala que las evaluaciones de los alumnos estuvieron relacionadas con variables internas al proceso de enseñanza-aprendizaje, como es lo deseado, pero también con variables externas lo cual no es deseado.

Sin embargo, al cuantificar la relación de la variables externas, fue posible determinar la magnitud en que las opiniones de los alumnos fueron sesgadas por estas variables, si bien el sesgo fue significativo no fue representativo, estuvo entre un 4.3% y un 7.2%, lo que permite concluir que las opiniones de los alumnos son válidas para evaluar el desempeño de los profesores.

CAPÍTULO 5

CONCLUSIONES Y RECOMENDACIONES

5.1 CONCLUSIONES

Los resultados de este estudio permitieron aportar evidencia para probar la confiabilidad y validez de la encuesta de opinión realizada a los alumnos para evaluar y retroalimentar el desempeño de los profesores del Campus Monterrey.

La actitud negativa hacia las evaluaciones de los alumnos por parte de los profesores, basada en experiencias propias o de compañeros, puede ser mitigada con la evidencia encontrada aquí, porque sugiere que las opiniones de los alumnos son confiables y válidas para evaluar y retroalimentar su desempeño.

Sin embargo, también sugiere que existen ciertos aspectos que deben de mejorarse, en la medida de lo posible, para incrementar la calidad de la información aportada por la encuesta y para controlar la validez de ésta.

5.1.1 CONFIABILIDAD

Para probar la confiabilidad de las opiniones de los alumnos obtenidas de la encuesta utilizada para retroalimentar y evaluar el desempeño del profesor, se calculó un coeficiente de consistencia interna llamado alfa de Cronbach.

Este coeficiente mide el grado de consistencia de las evaluaciones de los alumnos, de tal manera que entre menos variabilidad exista en las evaluaciones de los alumnos (mayor acuerdo) sobre el desempeño de su profesor, más confiable será la información que proporcione la encuesta.

Los resultados indicaron que este coeficiente fue en promedio de 0.891, lo que significó que el 89.1% de la varianza de las evaluaciones de los alumnos dependen de la varianza verdadera en el rasgo medido y el 10.9% depende de la varianza del error, por lo que se consideró que la información que aporta la encuesta es confiable para evaluar y retroalimentar el desempeño de los profesores.

Las diferencias en las evaluaciones de los alumnos se atribuyeron en gran parte a diferencias verdaderas en la percepción del desempeño del profesor, y no a diferencias ocasionadas por errores no sistemáticos que dependen de características como el estado de ánimo del alumno al contestar la encuesta, del tiempo disponible, del entorno, etc.

A su vez, la información señaló que los resultados de la encuesta serán más confiables a medida que se incremente el número de alumnos que la contesten. Al dividir el total de

grupos de acuerdo al número de alumnos que contestaron la encuesta, se encontró que el coeficiente de consistencia interna promedio de las divisiones realizadas se incrementaba al aumentarse el número de alumnos que la respondieron.

Al analizar las evaluaciones de los alumnos individualmente por cada una de las preguntas de la encuesta, se observó que las más consistentes para retroalimentar el desempeño del profesor, fueron las objetivas en donde el alumno tiene un parámetro base para evaluar al profesor. Dentro de la encuesta, a percepción del tésista, de las 13 preguntas analizadas solamente existieron 4 objetivas, lo que hace conveniente que el cuestionario actualmente utilizado sea modificado.

La variabilidad promedio de las evaluaciones que realizan los alumnos en las preguntas de la encuesta así como el promedio de estas evaluaciones, indican que los profesores son evaluados de una manera muy similar por sus alumnos en todas las preguntas, sobre todo con los valores de la escala cercanos a 1 y 2, esto es, sucede con mucha frecuencia que los alumnos asignan la misma evaluación al profesor en todas las preguntas.

Lo anterior es una cuestión que debe de mejorarse de la encuesta, porqué de que sirve que los resultados que reporte sean confiables, sí no aporta información relevante. Esto es, por ejemplo, sí en un grupo los alumnos evalúan al profesor en la pregunta 1 igual que en la pregunta 2 y así sucesivamente hasta la pregunta 14, de que sirve tener tantas preguntas en la encuesta, mejor que sólo se incluya la global, al fin y al cabo la información que aportan las específicas no es relevante.

En resumen, los resultados de la encuesta son confiables, esto es, existe consistencia en las evaluaciones de los alumnos sobre el desempeño de sus profesores, sin embargo es necesario que se mejore el cuestionario porqué los alumnos evalúan de una manera muy similar a sus profesores, sobre todo con los valores de 1 y 2 de la escala, lo que hace que las preguntas de la encuesta aporten información muy similar.

5.1.2 VALIDEZ

Al analizar los valores de la escala de likert de 1 a 7 (donde 1 es total acuerdo y 7 total desacuerdo) más utilizados para evaluar el desempeño de los profesores, se encontró que los patrones de respuesta de las preguntas de la encuesta (la representación gráfica de las evaluaciones de los alumnos en cada valor de la escala) fueron decrecientes en los valores de 1 a 6 de la escala con un incremento en el de 7.

Esto informó que se utilizaron más los valores de 1 y 2 de la escala y de acuerdo a la tendencia del patrón, se identificó el incremento en el valor de 7 como anormal.

Si bien en el análisis de la confiabilidad se cuantificó un tipo de error en las evaluaciones de los alumnos, el no sistemático, surgió la pregunta de qué tanto este error era afectado por el error sistemático, aquel que puede repetirse cada que se aplica la encuesta y que es atribuible a variables que nada tienen que ver con el desempeño del

profesor y que pudieran estar ocasionando el incremento anormal observado en el valor de 7 del patrón de respuestas de las preguntas de la encuesta.

Así, nació la necesidad de validar las opiniones de los alumnos para evaluar el desempeño de los profesores. Esto fue realizado mediante un modelo de regresión, donde se modeló la evaluación global del profesor con variables internas al proceso de enseñanza-aprendizaje, las que se desea sean fuentes de explicación porque miden aspectos específicos del profesor al enseñar, y por variables externas, las que no se desea que lo sean, porque nada tienen que ver con el desempeño del profesor.

De esta forma, se concluyó que las opiniones de los alumnos eran válidas porque la evaluación global del profesor estuvo más relacionada con las variables internas que con las externas.

Como variables internas se utilizaron las preguntas específicas de la encuesta y dos factores que resumieron la información proporcionada por éstas. El primero interpretado como el subjetivo, porque estuvo correlacionado con todas las preguntas subjetivas de la encuesta, resumió la mayor parte de la información; en cambio el segundo, concentró una menor proporción debido a que sólo estuvo correlacionado fuertemente con dos preguntas, las que permitieron interpretarlo como el factor de cumplimiento en cuanto a puntualidad, asistencia, asesorías y con el material del curso.

Las preguntas específicas de la encuesta explicaron en un 92.9% la evaluación global del profesor y los dos factores en un 89.4%, lo que se debió a que sólo resumieron el 78.67% de la información aportada por éstas.

Por otro lado, como variables externas se utilizaron ciertas características del curso (la disciplina), del grupo (el tipo, rediseñado o tradicional, el tamaño y la calificación promedio) y del profesor (género, años de experiencia en el Instituto, edad, tipo de contrato, grado máximo de estudios).

La magnitud del sesgo de estas variables sobre las evaluaciones de los alumnos fue significativa pero no representativa. Explicaron en un 4.3% y 7.2% la evaluación global del profesor.

Las relaciones encontradas en el análisis de regresión indicaron que:

- Las disciplinas mejor evaluadas fueron HI, OR y RH y las peor CB, EC y RH, posiblemente debido a la dificultad.
- Los profesores impartiendo grupos rediseñados obtuvieron mejores evaluaciones que los tradicionales, señalando que posiblemente el uso de las técnicas didácticas promovidas en el ITESM influyen en las evaluaciones de los alumnos.

- El tamaño del grupo afecta negativamente las evaluaciones de los alumnos, los profesores impartiendo grupos pequeños obtienen mejores evaluaciones que los grandes.
- La calificación promedio del grupo se relaciona en forma positiva con la evaluación del profesor, esto es, profesores con grupos que obtuvieron buenas calificaciones fueron mejor evaluados que sus contrapartes.
- Las profesoras obtuvieron mejores evaluaciones que los profesores.
- Los años de experiencia del profesor en el Instituto afectaron en forma positiva la evaluación del profesor.
- La edad del profesor presentó una relación negativa con las evaluaciones de los alumnos.
- Los profesores de planta obtuvieron mejores evaluaciones que los de media planta y los auxiliares.
- Por último, el efecto del grado máximo de estudios no fue significativo, pero indicó una relación positiva entre las evaluaciones de los alumnos y el nivel de preparación del profesor.

5.2 RECOMENDACIONES

Las recomendaciones de este trabajo están enfocadas en dos aspectos.

El primero, es mejorar el cuestionario, asegurándose que la información recolectada en la encuesta sea confiable pero además que cada pregunta aporte información diferente para retroalimentar al profesor, esto es, incluir preguntas que logren que los alumnos realicen evaluaciones similares dentro de ellas pero diferentes en una de otra.

El segundo, está relacionado con controlar el sesgo de las evaluaciones de los alumnos ocasionado por las variables externas consideradas. En el sentido de contemplar las relaciones de las variables a la hora de determinar el desempeño del profesor.

Con respecto al primer aspecto, las recomendaciones son las siguientes:

- Identificar todos los valores de la escala utilizada para evaluar al profesor. En el periodo analizado sólo estaban definidos el 1 y el 7, definir todos los valores puede lograr que los alumnos exploren más la escala ya que principalmente se utilizaron en mayor magnitud el valor 1 y el 2.
- Incluir más preguntas objetivas en la encuesta, actualmente a consideración del tesista, sólo existen 4. Esto porque se encontró que las preguntas objetivas aportaron

información más consistente que las subjetivas, esto es, la variabilidad en las evaluaciones de los alumnos fue menor, las opiniones fueron más congruentes.

Además, este tipo de preguntas pueden ser más amigables para los alumnos, ocasionando que la información que se recopile realmente mida los aspectos deseados, es decir, que el alumno asigne evaluaciones diferentes a su profesor en cada una de las preguntas, claro de acuerdo a su desempeño, y no como actualmente sucede que la evaluación del alumno generalmente es similar en todas las preguntas.

- Si no es posible encontrar preguntas objetivas, entonces brindarle al alumno un parámetro que utilice como base para realizar su evaluación, ya sea indicándole que evalúe a sus profesores conforme a su mejor profesor durante el semestre, o proporcionándole un perfil ideal establecido por estudiosos de la materia en donde se indique los aspectos que un profesor debe de cubrir para considerarse con buen desempeño.
- También, puede ayudar incluir en el cuestionario preguntas que los alumnos desean contestar y no las que las personas administrativas consideran convenientes incluir.
- Por último, que la encuesta sólo incluya aspectos relacionados con el profesor. El cuestionario analizado incluye preguntas sobre directivos y servicios lo que hace tedioso y tardado su respuesta. Entre más tiempo inviertan los alumnos a contestar la encuesta, menos confiable será la información que proporcionen, ya que por terminar pronto pueden asignar evaluaciones que realmente no reflejen la verdad sobre el desempeño del profesor.

Para esto es conveniente se divida la encuesta, en una parte que se incluyan solo aspectos del profesor y en otra, aplicada en un periodo diferente, se incluyan preguntas sobre los directivos y los servicios que se ofrecen en el Campus.

Por otra parte, en el caso particular del segundo aspecto:

Considerar las relaciones encontradas para calcular el percentil donde se ubica la OGP del profesor con respecto a la de todos los profesores del Campus. Esto es, por ejemplo, comparar el desempeño de los profesores que imparten grupos pequeños, grandes o elaborar el percentil contemplando sólo mujeres o hombres, asimismo para los que imparten grupos rediseñados o tradicionales, o para los profesores con muchos o pocos años en el Instituto, etc.

TRABAJOS FUTUROS

Durante el desarrollo del presente trabajo, se fueron identificando nuevas líneas de investigación que pueden ser realizadas en un futuro y que pueden aportar conocimiento valioso para el tema, las cuales se listan a continuación:

- Realizar un estudio en el que se analice si se encuentran los mismos resultados en los demás Campus del Sistema o se observan nuevos. Esto con la finalidad de determinar las necesidades específicas que cada uno posea.

Contar con la información que arroje un estudio de esta naturaleza permitirá identificar los esfuerzos de mejora acorde a los datos obtenidos.

- Efectuar una investigación sobre la forma en que las preguntas subjetivas de la encuesta sean más consistentes para evaluar o retroalimentar al profesor.
- Desarrollar un análisis para identificar la validez de la encuesta a través del tiempo.
- Elaborar un estudio en el que se determine la forma de minimizar, o desaparecer, el sesgo que presentan el indicador utilizado para evaluar a los profesores.

BIBLIOGRAFÍA

- Aiken L. R. (1996). Tests Psicológicos y Evaluación. Prentice Hall, octava edición.
- Anastasi, A. y Urbina, S. (1998). Test Psicológicos. Prentice Hall, séptima edición.
- Aubrecht, J. D. (1979). Are Student Ratings of Teacher Effectiveness Valid? (Individual Development Educational Assessment, IDEA, Paper 2). Manhattan, KS: Center for Faculty Evaluation and Development.
- Aubrecht, J. D. (1981). Reliability, Validity and Generalizability of Student Ratings of Instruction (IDEA, Paper 6). Manhattan, KS: Center for Faculty Evaluation and Development.
- Basow, A.S. (1995). Student Evaluations of College Professors: When Gender Matters. *Journal of Educational Psychology*, 87, 4, 656-665.
- Campus Ciudad de México, CCM, ITESM, (2001). Proyecto sobre el diseño y la utilización de un Programa de Evaluación del Desempeño Docente.
- Campus Monterrey, MTY, ITESM, (1998) Proposición de Adecuaciones al Proceso de Evaluación por alumnos, del Desempeño Docente de sus Profesores.
- Cashin W.E. (1988). Student Ratings of Teaching: A Summary of the Research. (IDEA, Paper 20). Manhattan, KS: Center for Faculty Evaluation and Development.
- Cashin W.E. (1990). Student Ratings of Teaching: Recommendations for USE. (IDEA, Paper 22). Manhattan, KS: Center for Faculty Evaluation and Development.
- Cashin W.E. (1995). Student Ratings of Teaching: The Research Revisited. (IDEA, Paper 32). Manhattan, KS: Center for Faculty Evaluation and Development.
- Cranton, P. A. And Smith, R. A. (1986). A new look at the effect of course characteristics on students ratings of instruction. *American Educational Research Journal*, 23, 117-128.
- Cranton, P. A. And Smith, R. A. (1990). Reconsidering the Unit of Analysis: A Model of Student Ratings of Instruction. *Journal of Educational Psychology*, 1990, 82, 2, 207-212.
- Dirección de Información Institucional, (2001). “Manual de ECOAS”, Instituto Tecnológico y de Estudios Superiores de Monterrey.
- Díaz Aguilar, E.M., (1997). Tesis. Evaluación de profesores: Sistema instrumentado mediante encuestas dirigidas a los estudiantes. Universidad Iberoamericana.

De la Garza G., Jorge (1995). Análisis de la información mercadológica a través de la estadística multivariante. Editorial Alambra Mexicana, primera edición

De Arce, R. (2001). Conceptos básicos sobre la heterocedasticidad en el modelo básico de regresión lineal tratamiento con E-views. Universidad Autónoma de Madrid. Departamento de Ciencias Económicas y Empresariales.
<http://www.uam.es/departamentos/economicas/econapli/pdf/heterocedasticidad.pdf>

Dunteman, G. H. (1989). Principal components analysis. Newbury Park : Sage.

Feldman K.A. (1976). Grades and college students' evaluations of their courses and teacher. *Research in Higher Education*, 4, 69-111.

Feldman K. A. (1978) Course characteristics and college student's ratings of their teachers-What we know and what we don't know. *Research in Higher Education*, 9, 199-244.

Feldman K. A. (1983) Seniority and experience of college teachers as related to evaluations they receive from students. *Research in Higher Education*, 18, 3-124.

Feldman K. A. (1984) Class size and college students' evaluations of teacher and courses: A closer look. *Research in Higher Education* 21, 45-116.

Freeman (1994). Student Evaluations of College Instructors: Effects of Type of Course Taught, Instructor Gender and Gender Role, and Student Gender. *Journal of Educational Psychology*, 86, 4, 627-630.

Greene, William H. (1991). Econometric Analysis. Maxwell Macmillan International Editions, primera edición.

Gigliotti R. J., y Buchtel, F.S. (1990). Attributional Bias and Course Evaluations. *Journal of Educational Psychology*, 82, 2, 341-351.

González V., T. (1996). Investigación etnográfica sobre maestros mejor evaluados Monterrey : ITESM. Centro para la Excelencia Académica.

Gujarati, D. N. (1995). Basic Econometrics. New York : McGraw-Hill, tercera edición.

Jerez Méndez, M. (1996). Multicolinealidad. Universidad Complutense de Madrid. Departamento de Economía Cuantitativa.
<http://www.ucm.es/info/ecocuan/ectr1/Ectr1-JCasals-Multicol.pdf>

Jobson, J.D. (1992). Applied Multivariate Data Analysis, Volume II: Categorical and Multivariate Methods, Springer-Verlang New York.

Johnson, R.A. y Wichern, D.W. (1992). Applied multivariate statistical analysis. Englewood Cliffs, N.J. : Prentice Hall, tercera edición.

Jude, G.G., W.E. Griffiths, R.C. Hill, H. Lütkepohl, and TC Lee. (1985). The Theory and Practice of Econometrics. New York : Wiley, segunda edición.

L'Hommedieu Randi, Menges Robert J., Brinko Kathleen T. (1990). Methodological Explanations for the Modest Effects of Feedback From Student Ratings. *Journal of Educational Psychology*, 82, 2, 232-241.

Maddala, G. S. (1992). *Introduction to econometrics*. New York : Macmillan, segunda edición.

Manual E-Views, Versión 2.1.

Marsh H.W. (1980). The influence of student, course, and instructor characteristics in the evaluations of university teaching. *American Educational Research Journal*, 17, 219-237.

Marsh, H.W. (1984). Students's evaluation of university teaching: Dimensionality, reliability, validity, potential biases, and utility. *Journal of Educational Psychology*, 76, 707-754.

Rectoría Zona Sur, ITESM, (2000). *ECOAS, ¿Qué pesa más? La metodología, la exigencia, la relación maestro-alumno*.

Rotem, A. (1978). The effects of feedback from students to university professors: An experimental study. *Research in Higher Education*, 9, 303-318.

Rueda B., M. y Díaz B. A., F (2000). *Evaluación de la docencia, perspectivas actuales*. Paidós Educador, primera edición.

Scriven, M. (1994). Using student ratings in teaching evaluation. *Evaluations perspectives*, 4: 4-7.

Stumpf, S. A., Freedman, R. D., y Aguanno, J.C. (1979). A path analysis of factors often found to be related to student ratings of teaching effectiveness. *Research in Higher Education*, 11, 111-123.

White, H. (1980). A Heteroskedasticity- Consistent Covariance Estimator and a Direct Test for Heteroskedasticity, *Econometrica*, 48, 817-838.

Wigington, H., Tollefson, Nona y Rodríguez, Edme. (1989). Students ratings of instructors revisited: interactions among class and instructor variables. *Research on Higer Education*, 30, 331-344.

ANEXO 1

DESCRIPCIÓN DE LOS ESTADÍSTICOS EMPLEADOS EN LOS MODELOS DE REGRESIÓN ESTIMADOS

Dentro de las tablas, se reportan los coeficientes estimados de las variables explicativas, las desviaciones estándar, los estadísticos t de Student y los valores de la probabilidad, estos últimos utilizados para hacer inferencias individuales sobre el valor de los coeficientes estimados. También, se reporta la R cuadrada y la R cuadrada ajustada por el número de variables del modelo, ambas utilizadas para determinar la magnitud en que las preguntas específicas explicaron a la OGP. Por último, se incluye una prueba conjunta de Fisher para probar si el modelo fue significativo.

A su vez, se incluye la prueba de White (1980), para verificar si se cumple o no el supuesto de varianza constante supuesto por MCO. Esta prueba consiste en estimar un modelo de regresión donde la variable dependiente son los residuales al cuadrado de MCO del modelo original y las variables independientes son las explicativas junto con sus términos cuadráticos.¹⁹ La prueba ofrece dos formas de contrastar la hipótesis nula sobre si los errores tienen varianza constante:

- Con el estadístico F con $(k-1)$ y $(n-k)$ grados de libertad, donde k es el número de coeficientes estimados, el que brinda una medida de la bondad de ajuste del modelo, es decir, si se confirma la validez conjunta de las variables utilizadas para explicar los errores al cuadrado, se concluiría que los errores no tienen varianza constante.
- Con el estadístico $\text{Observaciones} \cdot R\text{-cuadrada}$ el cual se distribuye como una Ji-cuadrada con “ $p-1$ ” grados de libertad, donde “ p ” es el número de variables explicativas, también se supone la hipótesis nula de varianza constante.

Adicionalmente, se aporta información para diagnosticar si el modelo está afectado por el problema de multicolinealidad (o colinealidad), es decir, si las variables explicativas están muy correlacionadas, de esta manera las pruebas presentadas para detectar este problema son:²⁰

- Tolerancia. Este estadístico es utilizado para determinar que tanto las variables explicativas están linealmente relacionadas con otras (multicolinealidad). Una variable con poca tolerancia contribuirá con poca información al modelo. Así, un valor cercano a cero, menor a 0.1, será considerado como un problema ya que este estadístico es calculado como $1-R_j\text{-cuadrada}$ del modelo de regresión donde la variable independiente “ j ” es explicada por todas las restantes variables del modelo.

¹⁹ Para una explicación detallada de esta prueba también puede verse a Greene (1991), Jude, G.G, W.E. Griffiths, R.C. Hill, H. Lütkepohl, and TC Lee. (1985). Asimismo, sobre como se realiza esta prueba en E Views a De Arce (2001) o el propio manual del E-Views.

²⁰ Ver Greene (1991) y Jerez (1996).

- VIF o Factor de Inflación de la Varianza. Este estadístico es similar al anterior, se elabora como $1/(1-R_j\text{-cuadrada})$, de esta manera si la variable independiente se encuentra muy bien explicada por las restantes variables del modelo (con un VIF mayor a 10, lo que es equivalente a una $R_j\text{-cuadrada}$ igual a 0.9) entonces significará que existen problemas de multicolinealidad.
- Valor propios. Este estadístico indica que si los valores propios son cercanos a cero, implicará que el determinante de la matriz $X'X$ utilizada por MCO para estimar los coeficientes del modelo, será cercano a cero, lo cual implica que esta matriz no tendrá una inversa única, lo cual brindará un número infinito de soluciones, es decir, el problema de multicolinealidad estará presente.
- Índice de condición. También informa si existe evidencia de multicolinealidad, este índice es la raíz cuadrada de la razón del valor propio más grande con su valor sucesivo, de esta manera un valor mayor a 30, indicará problemas serios de multicolinealidad.

ANEXO 2

METODOLOGÍA UTILIZADA PARA ERRADICAR EL PROBLEMA DE HETEROCEDASTICIDAD

Bajo el supuesto de varianza constante, los estimadores de Mínimos Cuadrados Ordinarios (MCO) tienen la siguiente matriz de varianzas-covarianzas:

$$\text{Var}(\hat{\beta}) = (X'X)^{-1} X' \Phi X (X'X)^{-1}$$

donde Φ es una matriz con $\phi_{ii} = \text{Var}(\epsilon_i)$. Cuando los errores son homocedásticos, Φ puede ser escrita como $\Phi = \sigma^2 I$. De esta manera, con este supuesto, la matriz de varianzas-covarianzas de los estimadores de MCO puede ser reducida, como:

$$\text{Var}(\hat{\beta}) = \sigma^2 (X'X)^{-1}$$

Así, definiendo los residuales como $e_i = y_i - x_i \hat{\beta}$, la matriz de varianzas-covarianzas de MCO puede ser estimada de la siguiente forma:

$$\text{MVCMCO} = \frac{\sum e_i^2}{N - K} (X'X)^{-1} = s^2 (X'X)^{-1}$$

Si los errores son heterocedásticos, los estimadores de MCO reportados en MVCMCO serán insesgados y las pruebas de significancia estadística usuales serán inconsistentes. Sin embargo, cuando Φ es conocida, entonces puede ser utilizada para corregir la heterocedasticidad. Desafortunadamente, ésta es raramente conocida, así es necesario un estimador consistente de Φ con el fin de utilizarlo para corregir el problema.

La matriz de varianzas-covarianzas consistente propuesta por White (1980) se basa en la idea de que los e_i^2 pueden ser utilizados para estimar los ϕ_{ii} , entonces la matriz sería de la siguiente forma:

$$\text{MVCCW} = \frac{N}{N - K} (X'X)^{-1} X' \hat{F} X (X'X)^{-1}$$

donde $\hat{F} = \text{diagonal}[e_i^2]$, es decir, una matriz diagonal donde los elementos de la diagonal son los e_i^2 , K es el número de coeficientes estimados y $N/(N-K)$ es una corrección por los grados de libertad. Así, el MVCCW es un estimador consistente de $\text{Var}(\hat{\beta})$ en la presencia de heterocedasticidad de una forma desconocida.

ANEXO 3

DESCRIPCIÓN SUBJETIVA DE LA DIFICULTAD DE LAS DISCIPLINAS

A continuación se presenta el ordenamiento que se le dio a las disciplinas; éste se basó en analizar una relación subjetiva entre la dificultad de la misma y la evaluación que obtiene.

Para determinar la dificultad, el tesista situó a las disciplinas de acuerdo a su criterio, considerando como más difíciles aquellas en las que se incluyeran más matemáticas.

Con la finalidad de materializar lo anterior, se estudió la relación tomando al ordenamiento propuesto como el Rango 1; a su vez, el Rango 2 se determinó en función de la OGP obtenida por la disciplina (de peor a mejor evaluación). Utilizando dichos datos se estimó el coeficiente de correlación de Spearman, el cual fue de 0.2946, lo que significa que ha mayor dificultad se obtiene una peor evaluación (una OGP mayor), esto implica una relación negativa pero subjetiva entre la dificultad del curso y la evaluación del profesor.

La siguiente tabla muestra el análisis de la correlación de Spearman para las disciplinas.

Disciplina	Rango 1	Rango 2	OGP
MA	1	15	1.82
IQ	2	24	1.74
Q	3	6	1.93
F	4	2	2.04
EC	5	27	1.72
M	6	26	1.72
E	7	7	1.92
IS	8	12	1.87
CS	9	14	1.82
CB	10	4	1.99
MR	11	36	1.65
AR	12	31	1.69
CV	13	8	1.90
IN	14	16	1.81
IG	15	18	1.79
FZ	16	21	1.74
CF	17	28	1.71
AF	18	23	1.74
CD	19	17	1.81
MT	20	5	1.96
AA	21	3	2.00

Disciplina	Rango 1	Rango 2	OGP
RN	22	40	1.43
SA	23	39	1.47
SM	24	35	1.67
TA	25	11	1.87
TF	26	10	1.89
NI	27	30	1.70
P	28	19	1.77
D	29	20	1.77
MI	30	13	1.87
RH	31	37	1.53
OR	32	38	1.53
CC	33	41	1.37
CO	34	29	1.71
H	35	25	1.73
RI	36	32	1.68
HI	37	33	1.68
SI	38	9	1.90
IA	39	34	1.67
PA	40	22	1.74
EX	41	1	2.50

Correlación de Spearman

La fórmula clásica suele expresarse como

$$r_{S(x,y)} = 1 - \frac{6 \cdot \sum d_i^2}{n \cdot (n^2 - 1)}$$

Donde,

$d_j = (\text{Rango } 1)_k - (\text{Rango } 2)_\lambda$ para $k = 1, 2, \dots, 41$; y para $\lambda = 1, 2, \dots, 41$.

$n = 41$.