

**Universidad de Puerto Rico en Cayey  
Oficina de Avalúo e Investigación Institucional**



**Instrumento Piloto de Evaluación Estudiantil de los  
Profesores: Indicadores de Confiabilidad y Validez**

**Preparado Por:  
Fernando Vázquez-Calle  
Investigador Auxiliar**

**Para:  
Dr. José N. Caraballo  
Decano de Asuntos Académicos**

**Remitido por:  
Prof. Irmannette Torres-Lugo  
Directora**

**Diciembre de 2007**

## TABLA DE CONTENIDO

|   |    |
|---|----|
| Introducción.....   | 1  |
| Trasfondo de la Evaluación Estudiantil de Profesores (EEP).....                 | 1  |
| Dimensionalidad de la EEP.....  | 4  |
| Modelo Integrador de Apodaca & Grad.....  | 6  |
| Objetivos del Estudio.....  | 7  |
| Métodos y Procedimientos.....   | 7  |
| El Instrumento Piloto de EEP de la UPR en Cayey.....                            | 7  |
| Administración del Formulario y Participantes.....                              | 7  |
| Métodos Estadísticos.....   | 8  |
| Evidencias de Confiabilidad y Validez.....                                      | 11 |
| Los Reactivos: Confiabilidad y Discriminación.....                              | 11 |
| Confiabilidad por Subescala y Escala Total.....                                 | 15 |
| Validez de Contenido.....   | 15 |
| Dimensiones del Instrumento de EEP: Análisis Factorial Exploratorio.....        | 16 |
| Validez de “Constructo”.....  | 18 |
| Estructuras Factoriales Alternativas.....                                       | 18 |
| Adecuación de las Estructuras Factoriales Estimadas.....                        | 18 |
| Medidas de la Bondad de Ajuste.....   | 20 |
| Resultados del Modelo de Segundo Orden.....                                     | 20 |
| Conclusiones e Implicaciones.....   | 21 |
| Conclusiones.....   | 21 |
| Implicaciones.....  | 22 |
| Estudios Adicionales.....   | 23 |
| Limitaciones.....   | 24 |
| Referencias.....  | 24 |
| Apéndice A: Formulario del Instrumento Piloto<br>de EEP de la UPR en Cayey..... | 27 |
| Apéndice B: Matriz de Correlaciones de los 32 Reactivos.....                    | 31 |

## INTRODUCCIÓN

### Trasfondo de la Evaluación Estudiantil

El uso de la Evaluación Estudiantil del Profesor (EEP) en la enseñanza superior, de acuerdo con Algozzine et al. (2005), se remonta a principios del siglo 20 con los trabajos precursores de Remmers (1927, 1928, 1930) y Remmers & Brandenburg (1927). En esos inicios, según Apodaca & Grad (2005), la EEP se empleó con propósitos formativos, es decir, como un mecanismo para identificar los aspectos fuertes y débiles de las experiencias de enseñanza-aprendizaje, a fin de retroalimentar el proceso educativo para mejorarlo. Sin embargo, los mismos instrumentos diseñados para propósitos formativos, posteriormente en varias universidades se comenzaron a utilizar para propósitos sumativos. Es decir, la administración comenzó a emplear los resultados de la Evaluación Estudiantil del Profesor en las decisiones de retención, ascenso y retribución.

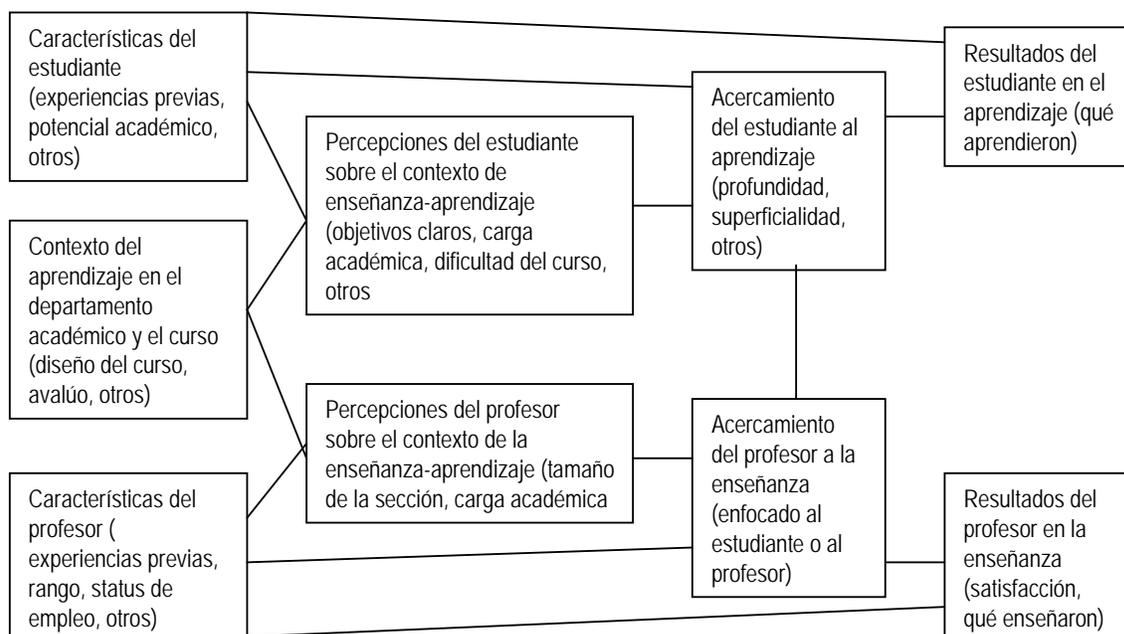
Aunque los propósitos formativos y sumativos de un instrumento de EEP pueden armonizarse, es importante tener conciencia de las contradicciones que hay en ellos. Para propósitos formativos es prácticamente imprescindible que el instrumento incluya aspectos (y mediciones de los mismos) detallados de la calidad del profesor en la enseñanza, puesto que de otra manera el mismo no sería útil como mecanismo de retroalimentación del proceso educativo. Pero, para propósitos sumativos, todo lo que necesita la administración universitaria es un puntaje global de la eficacia del docente en las tareas de la enseñanza. En la práctica ese único puntaje global se ha logrado sumando todos los reactivos del instrumento, o mediante promedios ponderados de las distintas dimensiones o subescalas incluidas en el instrumento de EEP. Sin embargo, si se desea lograr una retroalimentación eficaz, que le sirva sobre todo al profesor para elevar la calidad de las experiencias de enseñanza-aprendizaje, no basta con un puntaje global. Es necesario que el profesor reciba un informe detallado de los puntajes obtenidos, tanto por reactivo como por subescala.

Es abundante y diverso lo que se ha escrito en torno a la evaluación de la calidad de la enseñanza desde la óptica del estudiante. En Apodaca y Grad (2005) se citan las investigaciones de los siguientes autores: Marsh,

1991; Marsh & Hocevar, 1991; Abrami & D' Apollonia, 1991; Abrami et al., 1996; D' Apollonia & Abrami, 1997; Feldman, 1976; Cohen, 1981; Apodaca & Rodríguez, 1999; Apodaca & Grad, 2000 y Aleamoni, 1981. Uno de los temas más tratados en las investigaciones citadas son las dimensiones que componen el concepto de la calidad del profesor en la enseñanza, según la percepción del estudiante. Casi todas las investigaciones en torno al tema de la EEP se concentran en la segunda mitad del modelo de Enseñanza y Aprendizaje de Dunkin & Biddle (1974) y Biggs (1993), el cual se muestra en el Diagrama 1.

Hay bastante evidencia empírica que vincula positivamente la evaluación del estudiante en torno a la eficacia de la enseñanza, con medidas concretas del aprendizaje de los estudiantes. Por tanto, la EEP puede considerarse como un indicador válido de la calidad de la enseñanza (Cohen, 1981; Marsh, 1984 y Murray, 1980). Por ejemplo, Cohen (1981) encontró una correlación de 0.43 entre la evaluación general del estudiante y el puntaje medio en los exámenes finales. También, Feldman (1989), Marsh (1984), y Murray (1980), informaron una correlación positiva de la evaluación estudiantil con respecto a la enseñanza y la evaluación de ex-estudiantes, colegas y observadores externos. A la misma vez, los estudios mencionados hallaron una correlación débil entre la evaluación de los estudiantes e indicadores exógenos, tales como: tamaño de las secciones, rigurosidad del sistema de calificación, y el nivel del curso. Por otra parte, la mayoría de los estudios sobre el tema de la evaluación del estudiante con respecto a la calidad de la enseñanza, se basan en medidas más bien globales, tales como "entusiasmo, "claridad", "eficacia general", pero también, aunque menos, hay estudios (por ejemplo, Murray, 1983) basados en conductas más específicas en el salón de clases, como es llamar a los estudiantes por su nombre, escribir en la pizarra los objetivos de la clase, bosquejar las lecturas asignadas, entre otras. Las investigaciones han mostrado que las medidas más específicas de las destrezas del profesor están fuertemente correlacionadas con las evaluaciones generales que hacen los estudiantes en torno a las destrezas del profesor, evidenciándose así que la evaluación de la eficacia general es predecible por conductas específicas del profesor en el salón de clase (Cranton & Hilgartner, 1981; Murray, 1983). En resumen, hay evidencia de peso indicativa de que

**Diagrama 1. Modelo de Enseñanza y Aprendizaje Dunkie & Biddle (1974) y Biggs (1993)**



la evaluación estudiantil sobre la calidad de la enseñanza, refleja con razonable validez la eficacia de las destrezas del profesor, y que la misma está mucho menos influenciada por variables extrañas.

La EEP tiene fuertes críticos que señalan que su implementación en las instituciones de educación superior no está justificada por las siguientes razones (Aleamoni, 1981):

1. El estudiante carece de la madurez y el peritaje para hacer juicios válidos acerca del contenido del curso y del estilo del profesor.
2. La evaluación del estudiante después de todo es una mera percepción.
3. Los formularios de evaluación estudiantil carecen de confiabilidad y validez.
4. Hay variables extrañas que influyen en la evaluación estudiantil, entre otras: tamaño de las secciones, nivel del curso, si el curso es medular o electivo, sexo del estudiante y del profesor, edad del profesor, status de empleo del profesor (contrato, probatorio o permanente), rango académico y experiencia del profesor en la enseñanza.

A pesar de los detractores de la Evaluación Estudiantil de los Profesores, lo cierto es que por la posición única y privilegiada del estudiante como fuente directa de la información sobre el ambiente de aprendizaje, incluyendo las habilidades del profesor, convierte al estudiante en uno de los evaluadores lógicos de la calidad de la enseñanza. Por tal razón, todo parece indicar que la EEP en la educación superior llegó para quedarse.

#### **Dimensionalidad de la EEP**

Partiendo de que es poco probable que se erradique la EEP, y que por el contrario, su uso cada vez parece cobrar mayor importancia en la administración de las instituciones de educación superior, entonces resulta ineludible que los instrumentos de evaluación estén sustentados por los análisis de confiabilidad y validez de rigor. Uno de los elementos más importantes en la validación de un instrumento es lo que se denomina como validez de "constructo". La unidimensionalidad o multidimensionalidad de la estructura del "constructo" de las competencias del profesor, amerita un lugar central en la validación del instrumento. Este dilema de la unidimensionalidad versus la multidimensionalidad ha estado matizado

principalmente por los requerimientos contradictorios de los propósitos formativos y sumativos. Es obvio que para este último basta con una medida global de la calidad de la enseñanza, por lo que el enfoque unidimensional resulta más conveniente. Mientras que para el primero es indispensable una estructura multidimensional, puesto que es de la única forma que se puede dar retroalimentación útil al profesor.

Resulta lógico que cualquier evaluación de la eficacia del proceso de enseñanza debe estar guiada por las teorías de aprendizaje, pero también debe recordarse que no existe un modelo de aprendizaje aceptado unánimemente. Desde la perspectiva de la administración, Newble & Cannon (1995) han propuesto las siguientes dimensiones: organización, instrucción, evaluación, relaciones y dominio de los temas. Lowman (1984) sintetizó los aspectos anteriores en un modelo bidimensional, en el cual una de las dimensiones es la estimulación intelectual (entre otros: claridad, dominio de los temas y habilidad para modular la voz) y, la otra dimensión es el "rapport" (destrezas de relaciones interpersonales, motivación y aprendizaje independiente).

Desde la perspectiva del conductismo, Gagné & Briggs (1974) y Davis (1977) establecieron ocho (8) criterios en los cuales debe basarse la evaluación de las experiencias de aprendizaje:

1. El profesor establece el contexto de la experiencia de enseñanza-aprendizaje.
2. Uso de tecnologías apropiadas para facilitar la comunicación y ayudar al estudiante en el entendimiento y aprendizaje.
3. Estimulación de la participación activa del estudiante en la experiencia de aprendizaje.
4. El profesor explica los conceptos y principios desde distintos puntos de vista.
5. El profesor ofrece retroalimentación apropiada y específica a los estudiantes.
6. El diseño de la experiencia de aprendizaje se adapta a las diferencias individuales de los estudiantes.
7. Los temas son organizados para una comprensión efectiva de los mismos.
8. El profesor provee para que el estudiante pueda responder de distintas maneras a un mismo concepto, idea o principio.

Por otra parte, desde la perspectiva de la teoría del procesamiento de la información, centrada en la transmisión de información, Brown & Atkins (1988) establecieron las siguientes dimensiones:

1. Intenciones y planificación: preparación, información de las metas y planes de trabajo.
2. Transmisión:
  - a. Claridad en las explicaciones: organización (estructuración y secuencia adecuada de los temas, cantidad adecuada del contenido), estructuración de los cambios de temas (señales en la dirección de las lecturas, énfasis en los puntos clave, y la vinculación de los temas (introducción, cierre y transiciones en los temas), resumen al final de cada tema, pausas y énfasis apropiados, otros
  - b. Expresión: este es el principal medio para promover el interés del estudiante. Incluye el uso apropiado de la comunicación verbal y no verbal (gestos, ademanes, contacto visual, inflexiones en la voz), explicaciones con ejemplos, analogías y anécdotas, mostrar entusiasmo, dinamismo y pasión por los temas.
3. Receptor: mantener la atención del estudiante, facilitando la memorización (ejemplo: relacionando lo discutido con información conocida).
4. Salida: notas que toma el estudiante en la libreta, reacciones del estudiante y resultados del aprendizaje.

El enfoque anterior centrado sólo en la conducta del profesor, contrasta con el balance logrado por Marton & Säljö (1976), Entwistle & Ramsdend (1983) y Ramsdend (1992), en los siguientes principios de la enseñanza:

1. Claridad en los objetivos y desafíos intelectuales.
2. Calidad en las explicaciones y en la estimulación del interés del estudiante.
3. Promoción de la independencia, sentido de control y compromiso activo de los estudiantes con el proceso de aprendizaje.
4. Consideración y respeto hacia el estudiante y el aprendizaje de éste (incluyendo la flexibilidad en los métodos

de enseñanza y la disponibilidad para que el estudiante pueda consultar).

5. Aprendizaje desde el estudiante (atención a los comentarios y opiniones del estudiante; y análisis de los resultados del aprendizaje).
6. Retroalimentación apropiada al estudiante, así como también un sistema de calificación adecuado.

De acuerdo con Apodaca y Grad (2005), las categorías conceptuales discutidas anteriormente se pueden resumir en cinco (5) aspectos básicos de la calidad de las competencias en la enseñanza en el salón de clases.

1. Planificación y preparación de las lecciones (repasso de los temas, actualización, establecer metas, planes de trabajo, criterios de evaluación) e incluir una cantidad apropiada de material, así como un nivel de dificultad adecuado.
2. Destrezas de comunicación y habilidad para transmitir información efectivamente: organización, claridad en las explicaciones, estimulación del interés de los estudiantes (entusiasmo en las expresiones, uso de ejemplos y aplicaciones prácticas, pluralidad en los puntos de vistas sobre los temas) y la estimulación de la participación y el involucramiento del estudiante.
3. Uso de recursos didácticos y metodológicos adecuados al tema y el auditorio.
4. Interacción con los estudiantes: respuesta a los comentarios, disponibilidad para consultas, atención a la diversidad de la clase.
5. Retroalimentación apropiada para orientar el aprendizaje del estudiante.

La discusión anterior en torno al “constructo” de la calidad de la docencia, desde la perspectiva del estudiante, muestra una tremenda convergencia de las teorías y los resultados empíricos en relación con la estructura de factores de dicho “constructo”. Por supuesto, la conclusión anterior se refiere a los autores que en estudios empíricos han hallado una estructura multidimensional congruente con las teorías de aprendizaje (Glynn, Sauer & Wood, 2006; Marsh & Hocevar (1984); Prosser & Trigwell, 2006; Kember & Kwan, 2002. Hay otros autores que han desafiado la

multidimensionalidad de la evaluación estudiantil, sobresaliendo los trabajos de Abrami & D' Apollonia (1991, 1997). Estos últimos postulan que la percepción del estudiante en lo que concierne a la eficacia del profesor en la enseñanza, es unidimensional. Lo anterior no significa que ellos rechacen de plano que la eficacia de las destrezas y habilidades de los profesores en la docencia por naturaleza sea multidimensional. Lo que ocurre, según los autores mencionados, es que el estudiante se comporta más como un consumidor de un producto y tiende a simplificar su percepción del complejo proceso de enseñanza-aprendizaje, centrando su atención en el evaluado. Desde luego, este debate entre la unidimensionalidad versus la multidimensionalidad continúa abierto, aunque como se verá más adelante, ya hay enfoques en los que se comienzan a reconciliar las contradicciones inherentes.

### Modelo Integrador de Apodaca & Grad

Resulta interesante el enfoque integrador de Apodaca & Grad (2002, 2005), y Apodaca & Rodríguez (1999), en el contexto de universidades de habla española. En particular, el esfuerzo de los primeros, en el estudio realizado en la Universidad del País Vasco, España, titulado: “*The dimensionality of student ratings of teaching: integration of uni-and multidimensional models*”. En el mismo los autores hacen un esfuerzo tanto por el lado conceptual como empírico, por superar las contradicciones aparentemente antagónicas e irreconciliables entre la unidimensionalidad y multidimensionalidad de la estructura factorial de la calidad de la enseñanza, según es percibida por los estudiantes. Conceptualmente, los autores se inclinan por la teoría del procesamiento de información, en la cual se inserta la unidimensionalidad, como la más cercana a la manera en que los estudiantes evalúan las tareas docentes de sus profesores. Pero, sin dejar de reconocer, que el proceso educativo por naturaleza es multidimensional. Según los autores mencionados, lo que ocurre es que los estudiantes al evaluar a sus profesores concentran la atención en un grupo específico de reactivos indicativos de la satisfacción general con las experiencias del proceso de enseñanza, obviando otros reactivos, e inclusive otras subescalas de los instrumentos de evaluación.

Apodaca & Grad (2005), apoyados en la teoría del procesamiento de la información y en las

convergencias con las teorías de aprendizaje, postularon un modelo factorial integrando las dimensiones o subescalas de las competencias del profesor, con un factor general representativo de la evaluación general del estudiante. Específicamente, evaluaron las propiedades psicométricas de dos estructuras factoriales: una suponiendo que el “constructo” está subdividido en cinco subdimensiones (planificación, interacción con los estudiantes, comunicación y organización, avalúo y calificación de los estudiantes, y recursos didácticos y metodológicos; y la otra, suponiendo las cinco subdimensiones anteriores conjuntamente con un factor general que ejerce influencia sobre cada una de ellas.

Tanto el modelo integrador de los autores antes mencionados, como los modelos conceptuales fundamentados en las teorías de aprendizaje sirvieron de guías en el diseño del presente estudio.

### **OBJETIVOS DEL ESTUDIO**

El objetivo general de este estudio es evaluar la adecuación de los reactivos del instrumento piloto de Evaluación Estudiantil de los Profesores (EEP), así como también las características psicométricas de confiabilidad y validez. Los objetivos concretos son:

1. Evaluar la validez de contenido del instrumento piloto de EEP.
2. Estimar los índices de confiabilidad de Cronbach y de Mitades Spearman-Brown de las dimensiones de las competencias del profesor enseñando.
3. Evaluar los indicadores de confiabilidad y discriminación de cada uno de los reactivos del instrumento piloto de EEP.
4. Evaluar la validez de “constructo”, incluyendo la dimensionalidad, de la percepción del estudiante con respecto a la calidad del proceso de enseñanza-aprendizaje.

### **MÉTODOS Y PROCEDIMIENTOS**

#### **El Instrumento Piloto de EEP de la UPR en Cayey**

El instrumento piloto de EEP (en el Apéndice A puede verse una copia) consta de 32 reactivos divididos en tres partes: Competencias del

Profesor en la Enseñanza, Relaciones con los Estudiantes y Cumplimiento de Deberes. Como se desprende de su propio contenido y organización, el mismo se diseñó suponiendo que la percepción de calidad del estudiante en relación con las experiencias de enseñanza-aprendizaje, tiene una estructura de tres dimensiones. Las tres partes o dimensiones del instrumento, respectivamente, están compuestas por 18, 7 y 7 reactivos. Los reactivos de la Parte I se midieron a través de una escala bipolar, estableciendo dos polos absolutamente opuestos con alternativas fijas numeradas del I al IV. Mientras más cerca del I, más fuerte el grado de acuerdo con la premisa negativa formulada en un reactivo dado. Por el contrario, mientras más lejos del I y más cerca del IV, más de acuerdo con la premisa positiva. En el caso de las otras dos partes, se empleó una escala Likert de cuatro categorías, donde 0=nunca, 1=casi nunca, 2=a veces, 3=a menudo, y 4=siempre.

Para obtener un indicador global de un profesor evaluado, se suman los números correspondientes a las categorías de las respuestas dadas por todos los estudiantes de una sección en los 32 reactivos, y se divide entre el total de estudiantes que evaluaron al profesor en dicha sección. Esto arroja un puntaje promedio, cuyo campo de valores va de 18 a 128. Mientras más cerca de 128, mayor es la satisfacción de los estudiantes con las experiencias de enseñanza facilitadas por un profesor.

#### **Administración del Formulario y Participantes**

La muestra de diez y ocho (18) secciones se seleccionó a partir del archivo de horarios y cursos correspondiente al período lectivo de agosto a diciembre de 2007. A fin de asegurar una representación adecuada por rango académico, la plantilla de profesores se clasificó en: Instructor, Catedrático Auxiliar y Catedrático Asociado. También las secciones atendidas por profesor se clasificaron por departamento. La selección al azar de los profesores se hizo en dos etapas. En la primera, las unidades de muestreo se conformaron de acuerdo con la acumulación de estudiantes en las secciones atendidas por cada profesor. Es decir, el total de estudiantes-secciones por profesor. Mediante el procedimiento de selección sistemática y con probabilidad proporcional a la acumulación de estudiantes en las secciones por profesor, se seleccionaron al azar seis profesores por cada

uno de los tres rangos académicos, para un total de diez y ocho (18) profesores. En la segunda etapa, a base del total de estudiantes matriculados en las secciones escogidas en la etapa anterior, se seleccionó al azar una sección por profesor. Así se aseguró que en la muestra sólo hubiera una sección por profesor, a la vez que se logró una muestra con igual probabilidad de selección, por lo que no fue necesario ponderar los estimados.

Excluidas las secciones correspondientes a los profesores con rango de catedrático, la población objetivo quedó constituida por 670 secciones con aproximadamente 15,500 estudiantes-secciones y 160 profesores. No se consideró el personal docente con rango de catedrático porque éstos no son evaluados una vez reciben dicho rango. Se seleccionaron al azar diez y ochos (18) secciones, las cuales registraron una matrícula agregada de aproximadamente 415 estudiantes-secciones. De ese total de secciones seleccionadas, en diez y seis (16) de ellas se administró el formulario piloto de Evaluación Estudiantil de Profesores (EEP), arrojando un total de 351 formularios completados. Una de las secciones se descartó porque tan solo un estudiante llenó el formulario, manteniendo para el análisis quince (15) profesores y 350 formularios completados por estudiantes matriculados en la institución durante el primer semestre del año académico de 2007-08. Esto representó una tasa de participación de aproximadamente ochenta y cuatro (84) por ciento.

Para administrar el formulario de EEP se siguió un protocolo similar al que suele seguirse regularmente. Los profesores seleccionados fueron informados con anticipación de que personal de la Oficina de Avalúo e Investigación Institucional lo visitaría para distribuirle el formulario de EEP al grupo de estudiantes. Los propios estudiantes llenaron el formulario durante los primeros 15 minutos de la clase y lo devolvieron a la persona encargada. Antes de distribuir el formulario se le informaba al grupo de estudiantes que no tenían que escribir su nombre porque la información era anónima y confidencial. Los formularios se administraron durante el período del 8 al 17 de octubre de 2007.

## Métodos Estadísticos

### EVALUACIÓN DE LOS REACTIVOS

Cada uno de los treinta y dos (32) reactivos del instrumento piloto de EEP se examinó en vista de dos estadísticas: Correlación Reactivo-Total y el Alfa de Cronbach Corregida (eliminando el reactivo examinado). La primera consiste de la correlación bivariada de Pearson entre cada reactivo individual con la suma de los otros reactivos que compongan la escala objeto de estudio. Se espera que el signo algebraico de tal correlación sea positivo, indicando que, en promedio, los puntajes del reactivo y de la escala que presumiblemente mide éste, se mueven en la misma dirección. Debe recordarse que las respuestas a los reactivos son manifestaciones concretas de la idea o concepto que se pretende medir, el cual se vincula empíricamente con una escala, es decir, la suma de un grupo "representativo" de reactivos. Por tanto, lo que se espera es que la correlación entre el reactivo y la escala sea relativamente fuerte. Hay diferencias notables entre los autores en torno al punto crítico para considerar como aceptable la correlación Reactivo-Total estimada. En este estudio se adoptó la regla "práctica" sugerida por DeVellis, (1991), Spector (1992) y Kim & Mueller (1978), quienes consideran aceptable una correlación de por lo menos 0.35.

La segunda estadística hace referencia al coeficiente Alfa de Cronbach. En el contexto de la evaluación de los reactivos, el mismo se emplea para medir los cambios que ocurren en el nivel de confiabilidad de una escala si el reactivo no se toma en cuenta. Si un reactivo se obvia y el coeficiente Alfa de Cronbach de la escala aumenta significativamente, probablemente el mismo necesite revisarse, o inclusive en casos extremos, eliminarse.

Por otra parte, el instrumento de EEP se diseñó principalmente con un propósito sumativo, aunque con algunos cambios de enfoque podría utilizarse también con propósitos formativos. El mismo forma parte del conjunto de criterios de evaluación que la administración considera en las decisiones de reclutamiento, permanencia y ascensos. Por tanto, es muy importante que los reactivos de la EEP puedan discriminar entre los profesores que reciben puntajes altos de aquellos que reciben puntajes bajos. Es decir, que tengan la capacidad de diferenciar a los profesores que son "excelentes" enseñando de los que no lo son.

Para evaluar el poder de discriminación de cada reactivo se calcularon dos estadísticas, a saber: la correlación de Pearson Reactivo-Total y el contraste de las medias del grupo de profesores con puntajes más altos en la escala total de la EEP versus el grupo con los puntajes más bajos. La primera estadística tiene la ventaja de que en el cálculo de la misma se toman en cuenta las respuestas válidas de la muestra completa de estudiantes. Mientras que la segunda se basa en las respuestas de una porción de la muestra. Para realizar el contraste de las medias aludido anteriormente, se dividió la muestra en el 27 por ciento de los profesores (equivalente a cuatro profesores) con los puntajes más altos e igual porcentaje con los puntajes más bajos. Formados el Grupo Alto y el Grupo Bajo, a base de la estadística de t y un nivel de significación de 0.05, cada uno de los treinta y dos (32) reactivos se sometió a la prueba de hipótesis de que la diferencia entre las medias de ambos grupos era igual a cero. Es decir, en la hipótesis nula ( $H_0$ ) se planteó que no existían diferencias entre los grupos.

#### DIMENSIONALIDAD: ANÁLISIS FACTORIAL EXPLORATORIO

Tomando como marco de referencia los estudios citados, se analizó con cuál estructura factorial el patrón observado de covarianzas de los 32 reactivos del formulario de la EEP era más congruente. Para ello se utilizó el programa SPSS 15.0. La adecuación de cada una de las tres estructuras factoriales exploratorias se evaluó a partir de los siguientes criterios: 1) valor propio o raíz (mejor conocido en inglés como "eigenvalue") igual a uno o mayor (criterio de Kaiser, 1960), 2) porcentaje de la varianza que explica cada estructura factorial, 3) el diagrama de ladera ("screeplot"), 4) el peso de cada factor en los reactivos, y 4) la interpretación de los resultados en vista de los modelos conceptuales y la evidencia empírica revisados.

La evaluación de las estructuras factoriales se realizó mediante uno de los métodos de Análisis Factorial Exploratorio (AFE). En particular se empleó el método de Factorización del Eje Principal (Principal Axis Factoring), y para identificar con mayor claridad los reactivos que se relacionaban con cada uno de los factores de la estructura objeto de análisis, se empleó el algoritmo de rotación conocido como Quartimax. Una explicación breve y no técnica del AFE, es que el mismo consiste en hallar las variables observadas (en el caso del instrumento de EEP

son los 32 reactivos) que por su intercorrelación comparativamente alta, se infiere que están explicadas por factores no observados, los cuales representan conceptos o dimensiones abstractas no mesurables directamente. El siguiente dicho popular resume de manera simple el propósito de los modelos de Análisis Factorial: "*pájaros de una misma pluma vuelan juntos*". Por otra parte, el método de rotación Quartimax busca simplificar la interpretación de la estructura factorial hallada en el patrón de los datos observados, minimizando tanto el número de variables observadas por factor o dimensión, como el número de factores que explican las correlaciones entre las variables observadas.<sup>1</sup>

#### CONFIABILIDAD DE LAS ESCALAS: CRONBACH Y SPEARMAN-BROWN

La confiabilidad es una de las propiedades más importantes de una escala diseñada para medir un concepto en los campos de las ciencias sociales y la educación. De acuerdo con la teoría clásica de pruebas, las medidas recopiladas, que no son otra cosa que respuestas a los reactivos, tienen dos componentes: el puntaje verdadero (T) y el error (E). Conceptualmente, la confiabilidad es igual a la razón entre las varianzas de T y T+E. Por consiguiente, una escala se considera más confiable en la medida en que la razón anterior se acerca a uno (1), ya que implicaría que la varianza del error es cero.

Sin embargo, ya que el puntaje verdadero no se puede medir directamente, se han desarrollado múltiples medidas estadísticas de confiabilidad. El coeficiente Alfa de Cronbach es una de las estadísticas de confiabilidad más usadas. La misma se obtiene a través de la siguiente ecuación.

$$\alpha = [(k/(k-1))(1-(\sum \sigma_i^2 / \sigma_x^2))]$$

donde:

$\alpha$  es el coeficiente de Alfa estimado

k es el número de reactivos

$\sigma_i^2$  es la varianza del reactivo i (i=1,2, ..., k)

<sup>1</sup> Para una exposición técnica y estadístico-matemática del tema de Análisis Factorial, pueden consultarse las siguientes fuentes: 1) Tacq, J. (1997). *Multivariate analysis techniques: Social sciences research*. Beverly Hills and London: Sage Publications; 2) Kim, J. O. & Mueller, C. W. (1978). *Factor analysis*. Sage University Paper series on Quantitative Applications in the Social Sciences, series no. 07-014. Beverly Hills and London: Sage Publications.

$\sigma_x^2$  es la varianza de la escala total

Mientras más acerca de uno (1), mayor es la confiabilidad de los reactivos, puesto que significa que miden una proporción relativamente alta del factor o dimensión de interés. De acuerdo con Nunnally (1978) y Shultz & Whitney (2005) la gran mayoría de los autores concuerdan en que es aceptable un Alfa de Cronbach de por lo menos 0.70.

También la confiabilidad se puede medir en términos de la estabilidad de las medidas a través del tiempo. De ahí que uno de los índices de confiabilidad es la correlación del puntaje observado en una escala en dos o más momentos distintos. Sin embargo, en la práctica pocas veces se tiene la oportunidad de poder administrar una escala en más de una ocasión. Entre otras razones por las limitaciones en los recursos y por los problemas de disponibilidad de los sujetos en participar en más de una ocasión. La estabilidad de las medidas en el tiempo es una propiedad de particular importancia, ya que es deseable que en un intervalo de tiempo relativamente corto las respuestas a los reactivos de un mismo grupo de sujetos no varíen significativamente. La falta de estabilidad de las medidas de una escala es similar al problema que representaría un termómetro que arrojará temperaturas significativamente distintas en intervalos de tiempo cortos.

Ya que era poco probable poder administrar el instrumento de la EEP a los mismos estudiantes en dos ocasiones, se decidió emplear un índice que atendiera de algún modo este aspecto de la estabilidad de las medidas. A esos efectos se calculó el Índice de Confiabilidad Spearman-Brown de las Mitades tanto para los treinta y dos (32) reactivos como para cada una de las subescalas. El Índice de las Mitades consiste en dividir los reactivos aleatoriamente en dos grupos y estimar la correlación de Pearson entre los puntajes de ambos grupos. La confiabilidad, entonces, se evalúa en función del signo del coeficiente y del tamaño de éste. Si el signo es positivo quiere decir que los puntajes de ambos grupos se mueven en la misma dirección. Un signo negativo sería evidencia clara de un problema serio de confiabilidad de la escala. Ello es así porque estaría indicando que, en promedio, mientras los puntajes en uno de los grupos sube, en el otro bajan. Si se presupone que todos los reactivos miden un mismo concepto (la calidad de la enseñanza), un coeficiente negativo estaría

indicando que uno de los grupos está midiendo algo distinto. Por otra parte, al igual que con el Alfa de Cronbach, una correlación de por lo menos 0.70 se considera aceptable. Análogamente, un coeficiente relativamente pequeño sería indicativo de que el grupo de reactivos está midiendo principalmente un factor distinto al que se pretende medir.

#### VALIDEZ DE CONTENIDO

La validez de contenido se refiere a si las manifestaciones concretas del concepto objeto de estudio, contenidas en los 32 reactivos formulados, abarcan adecuadamente el espectro completo de las dimensiones del concepto de la calidad del proceso de enseñanza-aprendizaje. En otras palabras, se trata de cuán representativo es el grupo de reactivos desarrollados de todas las posibles manifestaciones de la calidad del profesor en la enseñanza.

Usualmente la validez de contenido se evidencia a través del juicio de un panel de expertos en lo que respecta a la representatividad de los reactivos. Se desconoce si este procedimiento se había efectuado como parte del desarrollo del instrumento de EEP, por lo que se decidió evaluar esta faceta de la validez comparando los aspectos que miden los 32 reactivos del instrumento mencionado, con las dimensiones o subescalas de instrumentos similares hallados en la literatura sobre el tema de la evaluación estudiantil de la enseñanza. A esos fines se revisaron los estudios de los siguientes autores: Apodaca, 2005; Gagné & Briggs, 1974; Brown & Atkins, 1998; Marton & Säljo, 1976; Apodaca & Grad, 2005; Jahangir, 1998; Riaz, 2000; Renaud & Murray, 2005 y Aleamoni, 1981.

#### VALIDEZ DE "CONSTRUCTO": ANÁLISIS FACTORIAL CONFIRMATORIO

Según Pérez-Gil, Chacón-Moscoso & Moreno-Rodríguez (2000), la validez de "constructo" es la más importante porque integra y unifica los demás tipos de validez (contenido, criterio, discriminante y concurrente). Los autores mencionados citan a Cronbach (1980) afirmando lo siguiente:

*"la meta final de la validación es la explicación y comprensión y, por tanto, esto nos lleva a considerar que toda validación es validación de "constructo"*

También, en Silva (1993) se halla la siguiente cita de Cronbach (1980):

*“Toda validación es una, y en un sentido todo es validación de “constructo”.*

En resumen, de acuerdo con los autores antes mencionados, las evidencias relativas a la dimensionalidad y a la validez de contenido, así como también a la confiabilidad, deben interpretarse como parte de la validación del “constructo” de la calidad de las experiencias de enseñanza desarrolladas por el profesor.

La validez del “constructo” antes definido se evaluó empíricamente mediante el enfoque deductivo del Análisis Factorial Confirmatorio (AFC). Dicho método forma parte de las técnicas desarrolladas en el campo de la econometría denominadas Modelos de Ecuaciones Estructurales (MEE). Los MEE combinan los métodos de análisis de regresión y de varianzas-covarianzas, para estimar los parámetros de un conjunto dado de relaciones lineales entre variables dependientes (endógenas) e independientes (exógenas).

En el contexto específico de la validación del instrumento de EEP, a base de la guía de los modelos conceptuales de la calidad percibida de la enseñanza a nivel universitario, se formularon tres estructuras factoriales distintas y se evaluó cuál de ellas se ajustaba mejor al patrón de covarianzas observado. La primera consistió de las tres dimensiones o subescalas literalmente establecidas en el propio formulario de EEP de la UPR en Cayey: Competencias del Profesor en la Enseñanza, Relaciones con los Estudiantes y Cumplimiento de Deberes. La segunda estructura probada consistió de cinco (5) dimensiones, a saber: 1) Planificación, Preparación y Dominio, 2) Interacción con los Estudiantes, 3) Comunicación y Organización, 4) Avalúo y Evaluación del Aprovechamiento Académico, y 5) Cumplimiento con el Horario. La tercera estructura consistió de las cinco (5) dimensiones de la estructura anterior, pero añadiendo una sexta dimensión vinculada con la evaluación general que los estudiantes hacen de las experiencias de enseñanza-aprendizaje. Las tres estructuras mencionadas, identificando los reactivos específicos vinculados con cada dimensión o subescala, se muestran en las Tablas 1 y 2.

En la nomenclatura del AFC las variables no observadas directamente se representan con círculos u óvalos. También, los residuales de cada una de las ecuaciones estructurales propuestas se representan en óvalos. Tanto las dimensiones como los residuales no se pueden medir directamente, pero entonces se miden indirectamente a través de los reactivos. Éstos, por su parte, se consideran variables observadas y se representan en cuadrados o rectángulos. Por último, las flechas en una dirección representan el efecto de una variable independiente o exógena sobre una variable dependiente o endógena. Cabe señalar que las estructuras factoriales estimadas son recursivas. Esto es, todos los pesos factoriales estimados son unidireccionales.

Las ecuaciones estructurales derivadas de los modelos factoriales propuestos se estimaron por el método de Máxima Verosimilitud (ML), el cual al igual que el método de los Cuadrados Mínimos Generalizados (CMG), suele generar estimados eficientes y estables con muestras relativamente grandes. La muestra de 350 estudiantes utilizada en este análisis, de acuerdo con varios autores se considera grande (Ullman, 2006; Shultz & Whitney, 2005; Bentler, 2005). Los estimados se realizaron con el programa AMOS 4.0 (Analysis of **MO**ment **Str**uctures).

## EVIDENCIAS DE CONFIABILIDAD Y VALIDEZ

### Los Reactivos: Confiabilidad y Discriminación

Las estadísticas de confiabilidad y de correlación Reactivo-Total se calcularon para las estructuras de tres (3) y cinco (5) subescalas. Las Tablas 1 y 2, muestran las estadísticas antes mencionadas, respectivamente, de las estructuras de 3 y 5 subescalas. Como se puede ver en las Tablas mencionadas, sólo el reactivo 2.1 registró un coeficiente de correlación Reactivo-Total por debajo del límite de 0.35 de la regla “práctica” en ambas estructuras factoriales. Esto quiere decir que dicho reactivo requiere alguna revisión en su fraseo y contenido. Además el reactivo 3.4, en la estructura de cinco factores, arrojó una correlación de 0.31 con respecto a la subescala Comunicación y Organización. Esta misma estadística de la correlación Reactivo-Total también se interpretó como evidencia de la capacidad de cada reactivo para discriminar los profesores con los puntajes más altos en cada subescala de aquellos con los puntajes más

**Tabla 1. Confiabilidad por Reactivo: Estructura de 3 Subescalas**

| Número del Reactivo                                      | Reactivos por Subescala   | Correlación Reactivo-Total | Cronbach Eliminando Reactivo |
|--|---|----------------------------|------------------------------|
| <b>Factor: Competencias del Profesor en la Enseñanza</b> |   |                            |                              |
| 1.1  | La forma en que distribuye el tiempo en cada clase...   | 0.6440                     | 0.9310                       |
| 1.2  | La organización que el profesor le da al curso es...  | 0.6850                     | 0.9300                       |
| 1.3  | Las explicaciones del profesor...   | 0.6760                     | 0.9300                       |
| 1.4  | Los materiales que utiliza el profesor (pizarra, bibliografía, recursos audiovisuales, etc.)...                             | 0.4710                     | 0.9350                       |
| 1.5  | Cuando los estudiantes preguntan, el profesor... El nivel de dificultad en la presentación del profesor es...               | 0.5990                     | 0.9320                       |
| 1.6  | es...   | 0.6690                     | 0.9300                       |
| 1.7  | El ambiente que el profesor vomenta en clase es...  | 0.5990                     | 0.9320                       |
| 1.8  | Durante la clase el profesor... La preparación que demuestra el profesor en su clase es...                                  | 0.5870                     | 0.9320                       |
| 1.9  | es...   | 0.6960                     | 0.9300                       |
| 1.10   | En general, las destrezas de enseñanza del profesor son...  | 0.8020                     | 0.9270                       |
| 1.11   | Los exámenes, asignaciones y tareas... Cuando la evaluación refleja que los estudiantes tienen dificultades, el profesor... | 0.7020                     | 0.9300                       |
| 1.12   | La cantidad de exámenes, asignaciones, tareas y otros trabajos...   | 0.5970                     | 0.9320                       |
| 1.13   | El tiempo asignado para cada examen, asignación o tarea...  | 0.6800                     | 0.9300                       |
| 1.14   | La relación entre la evaluación y el contenido del curso...   | 0.5910                     | 0.9320                       |
| 1.15   | El método de asignar calificaciones (notas) en el curso es...   | 0.6460                     | 0.9310                       |
| 1.16   | El profesor informa a los estudiantes sobre su progreso...  | 0.6990                     | 0.9300                       |
| 1.17   | En general, los métodos de evaluación del profesor son...   | 0.5310                     | 0.9350                       |
| 1.18   |   | 0.7800                     | 0.9280                       |
| <b>Factor: Relación con los Estudiantes</b>              |   |                            |                              |
| 2.1  | El profesor trata a los estudiantes con respeto y cortesía, y no como inferiores.   | 0.2450                     | 0.7940                       |
| 2.2  | Los estudiantes tienen libertad de diferir del profesor.  | 0.4470                     | 0.7630                       |
| 2.3  | El profesor estimula a los estudiantes a expresar sus opiniones o dudas.  | 0.4800                     | 0.7570                       |
| 2.4  | El profesor reconoce cuando los estudiantes no entienden el material.   | 0.5690                     | 0.7410                       |
| 2.5  | El profesor es justo e imparcial con los estudiantes.   | 0.5470                     | 0.7450                       |
| 2.6  | El profesor demuestra sensibilidad por las necesidades o circuntancias especiales de los estudiantes.                       | 0.6660                     | 0.7170                       |
| 2.7  | En general, el profesor establece una relación profesional y personal efectiva con sus estudiantes.                         | 0.5780                     | 0.7370                       |
| <b>Factor: Cumplimiento de Deberes</b>                   |   |                            |                              |
| 3.1  | Asiste puntualmente a la clase.   | 0.5440                     | 0.7100                       |
| 3.2  | Cumple con el horario establecido para la clase.  | 0.3900                     | 0.7430                       |
| 3.3  | Cumple con sus horas de oficina.  | 0.4790                     | 0.7250                       |
| 3.4  | Especifica los objetivos y normas del curso.  | 0.5890                     | 0.6990                       |
| 3.5  | Se comunica, de forma oral y escrita, con claridad. Repone el material de las clases a las que se ausenta.                  | 0.3550                     | 0.7540                       |
| 3.6  |   | 0.4600                     | 0.7300                       |
| 3.7  | Demuestra compromiso con su tarea.  | 0.5470                     | 0.7200                       |

**Tabla 2. Confiabilidad por Reactivo: Estructura de 5 Subescalas**

| Número del Reactivo   | Reactivos por Subescala  | Correlación Reactivo-Total | Cronbach Eliminando Reactivo |
|---|--|----------------------------|------------------------------|
| <b>Factor: Planificación, Preparación y Dominio</b>                                 |  |                            |                              |
| 1.1   | La forma en que distribuye el tiempo en cada clase...  | 0.6060                     | 0.8010                       |
| 1.4   | Los materiales que utiliza el profesor (pizarra, bibliografía, recursos audiovisuales, etc.)...        | 0.4840                     | 0.8340                       |
| 1.6   | El nivel de dificultad en la presentación del profesor es...   | 0.6190                     | 0.7970                       |
| 1.9   | La preparación que demuestra el profesor en su clase es...   | 0.6500                     | 0.7890                       |
| 1.10  | En general, las destrezas de enseñanza del profesor son...   | 0.7840                     | 0.7460                       |
| <b>Factor: Interacción con los Estudiantes</b>                                      |  |                            |                              |
| 1.7   | El ambiente que el profesor vomenta en clase es...   | 0.5619                     | 0.7137                       |
| 1.8   | Durante la clase el profesor...  | 0.6688                     | 0.6842                       |
| 1.12  | Cuando la evaluación refleja que los estudiantes tienen dificultades, el profesor...                   | 0.5395                     | 0.7193                       |
| 2.1   | El profesor trata a los estudiantes con respeto y cortesía, y no como inferiores.                      | 0.1969                     | 0.7976                       |
| 2.2   | Los estudiantes tienen libertad de diferir del profesor.   | 0.5353                     | 0.7205                       |
| 2.3   | El profesor estimula a los estudiantes a expresar sus opiniones o dudas.                               | 0.5414                     | 0.7191                       |
| <b>Factor: Comunicación y Organización</b>  |  |                            |                              |
| 1.2   | La organización que el profesor le da al curso es...   | 0.5580                     | 0.8470                       |
| 1.3   | Las explicaciones del profesor...  | 0.6690                     | 0.8370                       |
| 1.5   | Cuando los estudiantes preguntan, el profesor...   | 0.6580                     | 0.8390                       |
| 2.4   | El profesor reconoce cuando los estudiantes no entienden el material.                                  | 0.5930                     | 0.8460                       |
| 2.5   | El profesor es justo e imparcial con los estudiantes.  | 0.5350                     | 0.8490                       |
| 2.6   | El profesor demuestra sensibilidad por las necesidades o circunstancias especiales de los estudiantes. | 0.6560                     | 0.8380                       |
| 2.7   | En general, el profesor establece una relación profesional y personal efectiva con sus estudiantes.    | 0.5800                     | 0.8450                       |
| 3.4   | Especifica los objetivos y normas del curso.   | 0.3120                     | 0.8650                       |
| 3.5   | Se comunica, de forma oral y escrita, con claridad.  | 0.6100                     | 0.8430                       |
| 3.7   | Demuestra compromiso con su tarea.   | 0.5800                     | 0.8500                       |
| <b>Factor: Avalúo y Evaluación del Aprovechamiento Académico de los Estudiantes</b> |  |                            |                              |
| 1.11  | Los exámenes, asignaciones y tareas...   | 0.6590                     | 0.8670                       |
| 1.13  | La cantidad de exámenes, asignaciones, tareas y otros trabajos...                                      | 0.7250                     | 0.8580                       |
| 1.14  | El tiempo asignado para cada examen, asignación o tarea...   | 0.6420                     | 0.8690                       |
| 1.15  | La relación entre la evaluación y el contenido del curso...  | 0.6470                     | 0.8690                       |
| 1.16  | El método de asignar calificaciones (notas) en el curso es...  | 0.7260                     | 0.8580                       |
| 1.17  | El profesor informa a los estudiantes sobre su progreso...   | 0.5990                     | 0.8780                       |
| 1.18  | En general, los métodos de evaluación del profesor son...  | 0.7370                     | 0.8570                       |
| <b>Factor: Cumplimiento con el Horario</b>  |  |                            |                              |
| 3.1   | Asiste puntualmente a la clase.  | 0.5210                     | 0.5560                       |
| 3.2   | Cumple con el horario establecido para la clase.   | 0.4090                     | 0.6330                       |
| 3.3   | Cumple con sus horas de oficina.   | 0.4460                     | 0.6090                       |
| 3.6   | Repone el material de las clases a las que se ausenta.   | 0.4460                     | 0.6100                       |

**Tabla 3. Estadísticas Descriptivas de los Grupos de Profesores con Puntajes Más Bajos versus Puntajes más Altos en la Escala Total**

| Número del Reactivo | Grupo con Puntajes más Bajos |       |                     |                | Grupo con Puntajes más Altos |       |                     |                | Diferencia | t    | Valor-p |
|---------------------|------------------------------|-------|---------------------|----------------|------------------------------|-------|---------------------|----------------|------------|------|---------|
|                     | Total de Estudiantes (n)     | Media | Desviación Estándar | Error Estándar | Total de Estudiantes (n)     | Media | Desviación Estándar | Error Estándar |            |      |         |
| 1.1                 | 71                           | 3.10  | 0.7200              | 0.0855         | 89                           | 3.89  | 0.3200              | 0.0337         | 0.7890     | 9.27 | 0.000   |
| 1.2                 | 71                           | 3.15  | 0.7500              | 0.0889         | 89                           | 3.94  | 0.2300              | 0.0245         | 0.7889     | 9.39 | 0.000   |
| 1.3                 | 71                           | 2.92  | 1.0500              | 0.1200         | 89                           | 3.89  | 0.3200              | 0.0337         | 0.9721     | 8.26 | 0.000   |
| 1.4                 | 71                           | 3.12  | 0.8000              | 0.0945         | 89                           | 3.82  | 0.3900              | 0.0409         | 0.7005     | 6.86 | 0.000   |
| 1.5                 | 71                           | 3.27  | 0.9600              | 0.1100         | 89                           | 3.93  | 0.2500              | 0.0267         | 0.6650     | 6.30 | 0.000   |
| 1.6                 | 71                           | 3.24  | 0.8400              | 0.0992         | 89                           | 3.82  | 0.4100              | 0.0439         | 0.5808     | 5.73 | 0.000   |
| 1.7                 | 71                           | 3.42  | 0.8200              | 0.0976         | 89                           | 3.92  | 0.3100              | 0.0328         | 0.4988     | 5.28 | 0.000   |
| 1.8                 | 71                           | 3.45  | 0.7100              | 0.0846         | 89                           | 3.92  | 0.2700              | 0.0287         | 0.4706     | 5.74 | 0.000   |
| 1.9                 | 71                           | 3.13  | 1.0400              | 0.1200         | 89                           | 3.97  | 0.1800              | 0.0192         | 0.8395     | 7.47 | 0.000   |
| 1.10                | 71                           | 2.92  | 0.9400              | 0.1100         | 89                           | 3.85  | 0.3600              | 0.0376         | 0.9384     | 8.70 | 0.000   |
| 1.11                | 71                           | 3.06  | 0.7900              | 0.0938         | 89                           | 3.80  | 0.4600              | 0.0484         | 0.7414     | 7.43 | 0.000   |
| 1.12                | 71                           | 3.31  | 0.8000              | 0.0946         | 89                           | 3.86  | 0.3500              | 0.0366         | 0.5497     | 5.26 | 0.000   |
| 1.13                | 71                           | 3.20  | 0.9500              | 0.1100         | 89                           | 3.81  | 0.4200              | 0.0448         | 0.6118     | 5.44 | 0.000   |
| 1.14                | 71                           | 3.37  | 0.9300              | 0.1100         | 89                           | 3.83  | 0.4800              | 0.0511         | 0.4653     | 4.09 | 0.000   |
| 1.15                | 71                           | 3.27  | 0.7100              | 0.0846         | 89                           | 3.88  | 0.3300              | 0.0351         | 0.6088     | 6.49 | 0.000   |
| 1.16                | 71                           | 3.01  | 0.8600              | 0.1000         | 89                           | 3.79  | 0.4000              | 0.0429         | 0.7781     | 6.99 | 0.000   |
| 1.17                | 71                           | 2.88  | 1.0200              | 0.1200         | 89                           | 3.62  | 0.6800              | 0.072          | 0.7433     | 5.53 | 0.000   |
| 1.18                | 71                           | 2.99  | 0.8500              | 0.1000         | 89                           | 3.87  | 0.3700              | 0.0388         | 0.8778     | 8.81 | 0.000   |
| 2.1                 | 71                           | 3.77  | 0.6100              | 0.0729         | 89                           | 3.96  | 0.4200              | 0.0449         | 0.1804     | 2.19 | 0.030   |
| 2.2                 | 71                           | 3.41  | 0.8000              | 0.0953         | 89                           | 3.89  | 0.3800              | 0.0405         | 0.4792     | 4.97 | 0.000   |
| 2.3                 | 71                           | 3.56  | 0.6900              | 0.0821         | 89                           | 3.93  | 0.2500              | 0.0267         | 0.3692     | 4.67 | 0.000   |
| 2.4                 | 71                           | 3.10  | 1.0700              | 0.1300         | 89                           | 3.63  | 0.5900              | 0.0627         | 0.5306     | 3.77 | 0.000   |
| 2.5                 | 71                           | 3.48  | 0.6700              | 0.0798         | 89                           | 3.92  | 0.2700              | 0.0287         | 0.4425     | 4.94 | 0.000   |
| 2.6                 | 71                           | 3.17  | 0.9700              | 0.1200         | 89                           | 3.87  | 0.3700              | 0.0388         | 0.7018     | 5.91 | 0.000   |
| 2.7                 | 71                           | 3.38  | 0.9300              | 0.1100         | 89                           | 3.97  | 0.1800              | 0.0192         | 0.5860     | 5.80 | 0.000   |
| 3.1                 | 71                           | 3.35  | 0.9400              | 0.1100         | 89                           | 3.89  | 0.3800              | 0.0405         | 0.5355     | 4.88 | 0.000   |
| 3.2                 | 71                           | 3.72  | 0.5100              | 0.0608         | 89                           | 3.98  | 0.1500              | 0.0158         | 0.2592     | 4.54 | 0.000   |
| 3.3                 | 71                           | 3.57  | 0.6800              | 0.0807         | 89                           | 3.88  | 0.3500              | 0.0375         | 0.3116     | 3.06 | 0.003   |
| 3.4                 | 71                           | 3.44  | 0.9200              | 0.1100         | 89                           | 3.99  | 0.1100              | 0.0112         | 0.5521     | 5.03 | 0.000   |
| 3.5                 | 71                           | 3.35  | 0.9000              | 0.1100         | 89                           | 3.93  | 0.2900              | 0.0311         | 0.5805     | 5.74 | 0.000   |
| 3.6                 | 71                           | 3.39  | 0.8900              | 0.1100         | 89                           | 3.91  | 0.2900              | 0.0305         | 0.5157     | 4.03 | 0.000   |
| 3.7                 | 71                           | 3.65  | 0.6800              | 0.0805         | 89                           | 3.98  | 0.1500              | 0.0158         | 0.3296     | 4.46 | 0.000   |

bajos. De acuerdo con la correlación mencionada, los reactivos 2.1 y 3.4 tienen problemas para diferenciar los profesores que puntúan alto de los que puntúan bajo en las subescalas correspondientes.

La otra evidencia relacionada con el poder de discriminación de los reactivos consistió de la diferencia de la media de los puntajes en cada reactivo entre el Grupo Alto (27% de los profesores con los puntajes más altos en la escala total) y el Grupo Bajo (27% de los profesores con los puntajes más bajos). Todos los reactivos arrojaron un puntaje medio significativamente mayor en el Grupo Alto. La Tabla 3 muestra los resultados de estas diferencias y de las pruebas estadísticas. Las columnas penúltima y última de la mencionada tabla, reflejan el valor del estadístico t y la significación "observada" (Valor-p), respectivamente. En los 32 reactivos se rechazó la hipótesis de que el puntaje medio de los grupos Alto y Bajo era igual. El valor-p menor que 0.05 para cada reactivo es indicativo de que, en promedio, el Grupo Alto puntuó significativamente más alto que el Grupo Bajo.

#### CONFIABILIDAD POR SUBESCALA Y ESCALA TOTAL

La Tabla 4 muestra los índices de confiabilidad estimados para las estructuras de 3 y 5 subescalas. De las cinco (5) subescalas o dimensiones consideradas una sola arrojó problemas de confiabilidad. Como puede verse en la Tabla 4, exceptuando la subescala Cumplimiento con el Horario, las demás registraron coeficientes de confiabilidad, tanto en la versión de Cronbach como de Spearman-Brown, por encima de la regla práctica de 0.70. La subescala que no cumplió con el criterio de la regla "práctica", registró un coeficiente Cronbach de 0.67, y el de Spearman-Brown resultó ligeramente más pequeño (0.62).

Tabla 4. Confiabilidad por Subescalas y Escala Total

| Subdimensiones por Estructura                                       | Reactivos                                     | Total de Reactivos | Alfa de Cronbach | Spearman-Brown |
|---|---|--------------------|------------------|----------------|
| Estructura de Tres (3) Subdimensiones                               |   |                    |                  |                |
| Competencia del Profesor en la Enseñanza                            | 1.1 a 1.18                                    | 18                 | 0.9350           | 0.8840         |
| Relación con los Estudiantes  | 2.1 a 2.7                                     | 7                  | 0.7800           | 0.7660         |
| Cumplimiento de Deberes   | 3.1 a 3.7                                     | 7                  | 0.7560           | 0.6900         |
| Estructura de Tres (3) Subdimensiones y Dominio                     |   |                    |                  |                |
| Planificación, Preparación y Dominio                                | 1.1, 1.4, 1.6, 1.9, 1.10,                     | 5                  | 0.8290           | 0.8120         |
| Interacción con los Estudiantes                                     | 1.7, 1.8, 1.12, 2.1, 2.2, 2.3, 2.4, 2.5, 2.6, | 6                  | 0.7630           | 0.7330         |
| Comunicación y Organización   | 2.7, 3.4, 3.5, 3.7                            | 10                 | 0.8600           | 0.8290         |
| Avalúo y Evaluación del Aproximamiento Académico de los Estudiantes | 1.11, 1.13, 1.14, 1.15, 1.16, 1.17, 1.18      | 7                  | 0.8820           | 0.8530         |
| Cumplimiento con el Horario   | 3.1, 3.2, 3.3, 3.6                            | 4                  | 0.6710           | 0.6180         |
| <b>Escala Total</b>   | <b>1.1 a 1.18, 2.1 a 2.7, 3.1 a 3.7</b>       | <b>32</b>          | <b>0.9400</b>    | <b>0.8810</b>  |

#### Validez de Contenido

Una de las cosas que sobresalió de la revisión de los estudios citados es que existe bastante similitud entre los instrumentos de EEP en relación con la esencia que los reactivos miden. Las diferencias mayormente se encuentran en el fraseo de los mismos. De la revisión de los estudios mencionados, se preparó la siguiente lista de posibles aspectos de la calidad del proceso de enseñanza:

1. Planificación
2. Preparación
3. Dominio de los temas
4. Organización
5. Claridad
6. Interacción con los estudiantes
7. Dificultad del curso y carga académica
8. Recursos didácticos y metodológicos
9. Destrezas de comunicación
10. Organización
11. Entusiasmo e interés del profesor
12. Calificación y retroalimentación del estudiante
13. Manejo del tiempo (puntualidad y asistencia)

El análisis del contenido de los 32 reactivos que componen el instrumento de EEP arrojó evidencia indicativa de que todos los aspectos de la lista anterior, estaban representados por al menos un reactivo del instrumento mencionado.

Ahora bien, la clasificación de los 32 reactivos en las tres subescalas (Competencias de los Profesores, Relación con los Estudiantes y Cumplimiento de Deberes) que se desprenden de la propia faz del instrumento de EEP, muestra unas incongruencias que van más allá del análisis cuantitativo de las propiedades psicométricas. Mediante el análisis del contenido de los reactivos, se derivaron los siguientes hallazgos sobre el particular. En la Parte 1 del instrumento de referencia, en la cual los reactivos supuestamente miden las competencias del profesor en la enseñanza, los reactivos 7, 8 y 12 más bien miden la satisfacción de los estudiantes con los aspectos de la Relación con los Estudiantes. Asimismo, los reactivos 13 al 18 de esa misma Parte 1, miden la percepción de los estudiantes con respecto al sistema de avalúo y evaluación del aprovechamiento académico, el cual tendría más sentido lógico que fuera una dimensión o subescala por sí misma. También, en la Parte 3 de Cumplimiento de Deberes, los

reactivos 4, 5 y 7 fundamentalmente miden aspectos de organización, claridad y entusiasmo en la comunicación del profesor.

Los demás reactivos están dirigidos a medir otros aspectos de las competencias del profesor, pero por su contenido pudieran organizarse en subdimensiones más específicas, entre otras: planificación, preparación, dominio, comunicación y organización.

#### **DIMENSIONES DEL INSTRUMENTO DE EEP: ANÁLISIS FACTORIAL EXPLORATORIO**

La Tabla 5 muestra los Pesos Factoriales y el porcentaje de la variación de los reactivos explicado por el correspondiente factor común o subescala. Los resultados del Análisis Factorial Exploratorio (AFE) sugieren que la estructura de correlaciones observadas entre los reactivos del instrumento de EEP (véase en el Apéndice B la matriz de correlaciones de los 32 reactivos) se explica mejor suponiendo una estructura de cinco (5) factores comunes o dimensiones. Aunque seis factores arrojaron un “eigenvalue” (valor propio) mayor que 1 (criterio de Kaiser), tomando en cuenta la interpretabilidad de cada factor y el principio de la parsimonia, se decidió escoger la solución de cinco (5) factores. Mediante el examen de los pesos de cada factor en el grupo de reactivos correspondiente, se infirió que los factores o dimensiones vinculadas con las subescalas eran:

**Factor 1:**  
Avalúo y Evaluación del Aprovechamiento Académico del Estudiante

**Factor 2:**  
Planificación, Preparación y Dominio

**Factor 3:**  
Interacción con los Estudiantes

**Factor 4:**  
Destrezas de Comunicación y Organización

**Factor 5:**  
Cumplimiento con el Horario

Es importante apuntar que tanto la solución de factores determinada como la selección de los nombres de éstos, se basaron en el juicio que se hizo de los resultados de la Matriz Rotada de Pesos Factoriales. Es bien común en el AFE encontrar reactivos cuyo contenido no es congruente con el factor en el cual registraron el

peso factorial mayor. Por ejemplo, en el Factor 1, denominado Evaluación del Progreso Académico del Estudiante, se encuentran los reactivos 1.1 y 1.6, cuyo contenido no está íntimamente vinculado con tal dimensión o subescala; pero los siete reactivos restantes sí lo están, aparte de que tienen los pesos factoriales más altos. Este es el tipo de juicio que se hizo para inferir las dimensiones a partir de los resultados cuantitativos del AFE.

Tanto la estadística KMO (kaiser-Meyer-Okin) como la prueba de Bartlett arrojaron resultados indicativos de que el tamaño de muestra es adecuado para efectuar el Análisis Factorial Exploratorio. Se estimó un KMO igual a 0.931, lo cual cumple con la regla “práctica” de que sea por lo menos 0.5 (Kaiser, 1974). Por otra parte, la prueba de hipótesis de Bartlett resultó altamente significativa, rechazándose así que las correlaciones entre las variables observadas (reactivos) fueran distintas de cero meramente por las fluctuaciones aleatorias del muestreo.

Hubo seis reactivos (1.4, 2.1, 2.2, 3.2, 3.3, y 3.6) con un coeficiente de “communality” menor que 0.32, que es la regla “práctica” (Costello & Osborne, 2005). El “communality” mide el porcentaje de la variación de un reactivo que se explica por el conjunto de factores comunes o dimensiones. Por tanto, un valor por debajo del límite inferior de la regla “práctica” quiere decir que el reactivo en cuestión tiene una relación muy pobre con los otros reactivos, por lo que es altamente probable que no esté midiendo ninguna de las subescalas halladas en la estructura factorial estimada. De los seis reactivos con valores de “communality” significativamente pequeños, sólo el reactivo 2.1 registró un peso factorial menor que 0.32 (regla “práctica” citada por Costello & Osborne, 2005: en Tabachnick & Fidell, 2001) en todos los factores comunes hallados. Este resultado abona evidencia adicional al problema señalado de este reactivo, cuyo contenido es el siguiente: *“El profesor trata a los estudiantes con respeto y cortesía, y no como inferiores”*. Es muy posible que el problema de este reactivo consista principalmente en que el contenido del mismo incluye tres asuntos distintos a la vez: respeto, cortesía e inferioridad. Cuando se incluye más de un asunto en una premisa, existe un alto riesgo de que las respuestas se alejen del objetivo de estandarización. Debido a que la premisa tiene asuntos múltiples y hasta cierto punto contradictorios, es de esperarse que

Tabla 5. Matriz Rotada de Pesos Factoriales

| Reactivo             | Dimensiones del Concepto de Calidad de la Enseñanza |               |               |               |               | "Communality" |
|----------------------|---|---------------|---------------|---------------|---------------|---------------|
|                      | 1   | 2             | 3             | 4             | 5             |               |
| 1.1                  | <u>0.4140</u>                                       | 0.3410        | 0.1060        | 0.3360        | 0.2310        | 0.4660        |
| 1.2                  | 0.3760  | <u>0.5150</u> | 0.2160        | 0.0790        | 0.3160        | 0.5600        |
| 1.3                  | 0.1910  | <u>0.6630</u> | 0.2560        | 0.3100        | 0.0670        | 0.6420        |
| 1.4                  | <u>0.2760</u>                                       | <u>0.3530</u> | 0.1270        | 0.2000        | -0.0340       | 0.2580        |
| 1.5                  | 0.0640  | <u>0.5510</u> | <u>0.4360</u> | 0.3220        | 0.0250        | 0.6020        |
| 1.6                  | <u>0.5040</u>                                       | 0.3530        | 0.2130        | 0.2280        | 0.1450        | 0.4970        |
| 1.7                  | 0.2930  | 0.2960        | <u>0.5810</u> | 0.1280        | 0.1050        | 0.5390        |
| 1.8                  | 0.2900  | 0.1940        | <u>0.7410</u> | 0.1450        | 0.0650        | 0.6970        |
| 1.9                  | 0.2760  | <u>0.7270</u> | 0.2150        | 0.2460        | 0.0080        | 0.7120        |
| 1.10                 | 0.4340  | <u>0.6330</u> | 0.2540        | 0.2390        | 0.1590        | 0.7360        |
| 1.11                 | <u>0.5170</u>                                       | 0.3970        | 0.1780        | 0.2010        | 0.2120        | 0.5410        |
| 1.12                 | 0.2580  | 0.2400        | <u>0.4110</u> | <u>0.4650</u> | 0.0250        | 0.5100        |
| 1.13                 | <u>0.6610</u>                                       | 0.2130        | 0.1890        | 0.1710        | 0.2690        | 0.6200        |
| 1.14                 | <u>0.6390</u>                                       | 0.1980        | 0.0360        | 0.1490        | 0.2090        | 0.5150        |
| 1.15                 | <u>0.4900</u>                                       | 0.1990        | 0.1660        | 0.3200        | 0.3260        | 0.5160        |
| 1.16                 | <u>0.5860</u>                                       | 0.2190        | 0.1550        | 0.3880        | 0.1910        | 0.6030        |
| 1.17                 | <u>0.5180</u>                                       | 0.0590        | 0.1190        | 0.1590        | 0.3750        | 0.4510        |
| 1.18                 | <u>0.5170</u>                                       | 0.3550        | 0.2900        | 0.2850        | 0.3070        | 0.6520        |
| 2.1                  | 0.0260  | -0.1090       | 0.1700        | 0.1970        | 0.2510        | 0.1430        |
| 2.2                  | -0.0480   | 0.1010        | <u>0.5090</u> | 0.2550        | 0.1510        | 0.3600        |
| 2.3                  | 0.0020  | 0.1420        | <u>0.6050</u> | 0.2170        | 0.0720        | 0.4390        |
| 2.4                  | 0.1370  | 0.3060        | 0.2810        | <u>0.5500</u> | -0.0390       | 0.4950        |
| 2.5                  | 0.2500  | 0.1280        | 0.1280        | <u>0.6540</u> | 0.1060        | 0.5340        |
| 2.6                  | 0.0120  | 0.2980        | 0.3870        | <u>0.5660</u> | 0.1380        | 0.5780        |
| 2.7                  | 0.1980  | 0.1880        | 0.2650        | <u>0.5180</u> | 0.1550        | 0.4370        |
| 3.1                  | 0.1330  | 0.0600        | 0.0440        | -0.0050       | <u>0.6720</u> | 0.4750        |
| 3.2                  | 0.0190  | 0.0620        | 0.0060        | 0.0890        | <u>0.4990</u> | 0.2610        |
| 3.3                  | 0.2440  | 0.0320        | 0.1320        | 0.1060        | <u>0.5110</u> | 0.3500        |
| 3.4                  | 0.2470  | 0.1800        | 0.0170        | 0.0570        | <u>0.6210</u> | 0.4830        |
| 3.5                  | 0.0860  | <u>0.5040</u> | 0.2610        | 0.2470        | 0.2130        | 0.4350        |
| 3.6                  | 0.2280  | 0.0210        | 0.2770        | 0.0410        | <u>0.4870</u> | 0.3670        |
| 3.7                  | 0.0930  | 0.3840        | 0.2020        | 0.2730        | <u>0.4410</u> | 0.4650        |
| Eigenvalue           | 11.6880   | 2.5640        | 1.7430        | 1.2810        | 1.1530        |               |
| % Varianza           | 36.5260   | 8.0130        | 5.4460        | 4.0030        | 3.6040        |               |
| % Varianza Acumulada | 36.5260   | 44.5390       | 49.9850       | 53.9870       | 57.5910       |               |

un grupo de estudiantes esté centrando su atención en la cortesía, otro grupo en el respeto y otro en el trato de inferioridad. Así que cada cual responde de acuerdo con lo que enfoque en la premisa.

### Validez de “Constructo”

#### ESTRUCTURAS FACTORIALES ALTERNATIVAS

Mediante el método de Análisis Factorial Confirmatorio (AFC) se estimaron tres estructuras factoriales distintas, definidas en la parte de Métodos y Procedimientos. Dichas estructuras representan explicaciones conceptuales alternas del “constructo” Calidad de la Enseñanza, según es percibida por el estudiante. La primera estructura postula que el mencionado “constructo” está dimensionado por tres factores, mientras que las otras dos plantean que el mismo se puede explicar en función de cinco factores. Aunque estas dos últimas estructuras comparten las mismas subdimensiones, una de ellas se diferencia porque postula un factor adicional, representativo de la satisfacción general del estudiante con respecto a las experiencias de enseñanza-aprendizaje. Este tipo de estructura técnicamente se denomina como un modelo factorial de segundo orden, porque las cinco subdimensiones se consideran factores primarios y están supeditados al factor general, denominado de segundo orden.

Los tres modelos propuestos se evaluaron a base de los criterios que se explican a continuación, con el fin de seleccionar el que mejor se ajusta al patrón de correlaciones observado.

#### ADECUACIÓN DE LAS ESTRUCTURAS FACTORIALES ESTIMADAS

Los modelos estimados se evaluaron a base de los siguientes criterios: 1) identificación y tamaño de muestra, 2) ausencia de estimados “ofensivos”, 3) signo algebraico y tamaño de los coeficientes estimados, y 4) medidas de la bondad de ajuste. Ninguno de los tres modelos propuestos arrojó evidencia de un problema de identificación, pues todos resultaron estar sobre-identificados. Como puede comprobarse en la Tabla 6, los grados de libertad de los tres modelos factoriales son mayores que cero por bastante. Técnicamente los grados de libertad (gl) se refieren al número de puntos muestrales disponibles para estimar los parámetros de un modelo dado. Cuando dicho número es menor

**Tabla 6. Estadísticas de Ajuste de las Estructuras Factoriales Propuestas**

| Estadística             | 3 Factores | 5 Factores | 5 Factores y un Factor General |
|-------------------------|------------|------------|--------------------------------|
| Ji-Cuadrada             | 2,151      | 2,561      | 1,612                          |
| Grados de Libertad (gl) | 464        | 464        | 459                            |
| RMSEA                   | 0.102      | 0.114      | 0.085                          |
| Ji-Cuadrada/gl          | 4.635      | 5.520      | 3.512                          |
| TLI                     | 0.958      | 0.947      | 0.971                          |
| NFI                     | 0.953      | 0.944      | 0.965                          |
| CFI                     | 0.963      | 0.954      | 0.975                          |
| AIC                     | 2,343      | 2,753      | 1,814                          |
| ECVI                    | 6.715      | 7.889      | 5.198                          |

que el número de parámetros a estimarse, se dice que el modelo está subidentificado, y por consiguiente, el número de soluciones es infinito. Si hay cero grados de libertad el modelo está exactamente identificado y, sólo existe una solución.

El tamaño de la muestra se refiere al total de sujetos seleccionados al azar que completaron el instrumento. Trescientos cincuenta (350) estudiantes completaron el instrumento piloto de EEP. Aunque hay bastante disparidad en las reglas sugeridas en la literatura revisada de AFC, sí parece haber una convergencia en que el número de observaciones por parámetro a estimarse no debería ser menor que cinco (Benter & Chou, 1987; Shultz & Whitney, 2005). Los tres modelos se estimaron con aproximadamente cuatro (4) estudiantes por parámetro.

Usualmente los estimados “ofensivos” están asociados a problemas con los datos recopilados. Bien sea porque el tamaño de la muestra no es suficientemente “grande”, o porque la distribución de los datos se aparta de la normal más allá de lo razonable. Entiéndase por estimados “ofensivos” varianzas negativas, pesos factoriales estandarizados mayores que uno (1), matrices de covarianzas no positivas y errores estándar o muy pequeños o muy grandes. Ninguno de los modelos propuestos arrojó resultados que pudieran considerarse “ofensivos”.

Todos los pesos factoriales estimados, como puede verse en la Tabla 7, reflejaron un signo positivo, lo cual es congruente con los modelos conceptuales de la Calidad de la Enseñanza. Tampoco se observó que hubiera algún peso factorial desmedidamente grande o pequeño.

**Tabla 7. Coeficientes Estimados de las Estructuras Factoriales Propuestas**

| Reactivo                                | 3 Factores |                |       |                         | 5 Factores |                |       |                         | 5 Factores y Un Factor General |                |       |                         |
|---|------------|----------------|-------|-------------------------|------------|----------------|-------|-------------------------|--------------------------------|----------------|-------|-------------------------|
|   | b          | Error Estándar | t     | R <sup>2</sup> Múltiple | b          | Error Estándar | t     | R <sup>2</sup> Múltiple | b                              | Error Estándar | t     | R <sup>2</sup> Múltiple |
| 1.1                                     | 0.663      | 0.064          | 13.33 | 0.439                   | 0.632      |                |       | 0.400                   | 0.665                          |                |       | 0.443                   |
| 1.2                                     | 0.714      | 0.052          | 14.63 | 0.510                   | 0.602      |                |       | 0.363                   | 0.685                          |                |       | 0.469                   |
| 1.3                                     | 0.717      | 0.066          | 14.72 | 0.514                   | 0.740      | 0.145          | 10.70 | 0.547                   | 0.759                          | 0.108          | 13.00 | 0.576                   |
| 1.4                                     | 0.498      | 0.065          | 9.54  | 0.248                   | 0.519      | 0.093          | 8.55  | 0.269                   | 0.516                          | 0.084          | 8.89  | 0.267                   |
| 1.5                                     | 0.633      | 0.057          | 12.62 | 0.401                   | 0.725      | 0.120          | 10.56 | 0.526                   | 0.708                          | 0.089          | 12.21 | 0.502                   |
| 1.6                                     | 0.688      | 0.061          | 13.97 | 0.474                   | 0.663      | 0.095          | 10.51 | 0.440                   | 0.687                          | 0.086          | 11.52 | 0.472                   |
| 1.7                                     | 0.623      | 0.053          | 12.36 | 0.388                   | 0.681      |                |       | 0.464                   | 0.718                          |                |       | 0.516                   |
| 1.8                                     | 0.603      | 0.054          | 11.91 | 0.364                   | 0.831      | 0.102          | 12.04 | 0.690                   | 0.786                          | 0.084          | 13.16 | 0.617                   |
| 1.9                                     | 0.741      | 0.057          | 15.36 | 0.55                    | 0.768      | 0.095          | 11.77 | 0.590                   | 0.782                          | 0.084          | 12.87 | 0.612                   |
| 1.10                                    | 0.840      | 0.062          | 18.20 | 0.706                   | 0.919      | 0.119          | 12.71 | 0.844                   | 0.876                          | 0.100          | 14.09 | 0.768                   |
| 1.11                                    | 0.723      | 0.060          | 14.86 | 0.522                   | 0.718      |                |       | 0.516                   | 0.733                          |                |       | 0.537                   |
| 1.12                                    | 0.609      | 0.058          | 12.02 | 0.371                   | 0.602      | 0.098          | 9.65  | 0.362                   | 0.667                          | 0.088          | 11.34 | 0.445                   |
| 1.13                                    | 0.694      | 0.069          | 14.11 | 0.481                   | 0.776      | 0.090          | 13.63 | 0.602                   | 0.756                          | 0.084          | 13.88 | 0.572                   |
| 1.14                                    | 0.602      | 0.072          | 11.89 | 0.363                   | 0.685      | 0.090          | 12.06 | 0.469                   | 0.659                          | 0.085          | 12.03 | 0.435                   |
| 1.15                                    | 0.656      | 0.054          | 13.16 | 0.431                   | 0.694      | 0.069          | 12.21 | 0.481                   | 0.700                          | 0.065          | 12.78 | 0.789                   |
| 1.16                                    | 0.705      | 0.068          | 14.38 | 0.497                   | 0.780      | 0.089          | 13.69 | 0.608                   | 0.770                          | 0.084          | 14.13 | 0.592                   |
| 1.17                                    | 0.546      | 0.083          | 10.61 | 0.298                   | 0.640      | 0.103          | 11.38 | 0.416                   | 0.623                          | 0.098          | 11.34 | 0.388                   |
| 1.18                                    | 0.801      |                |       | 0.641                   | 0.787      | 0.080          | 13.82 | 0.619                   | 0.821                          | 0.075          | 15.13 | 0.674                   |
| 2.1                                     | 0.257      | 0.072          | 4.30  | 0.066                   | 0.187      | 0.079          | 3.17  | 0.035                   | 0.202                          | 0.074          | 3.50  | 0.041                   |
| 2.2                                     | 0.476      | 0.086          | 7.71  | 0.227                   | 0.539      | 0.096          | 8.76  | 0.290                   | 0.527                          | 0.087          | 9.04  | 0.278                   |
| 2.3                                     | 0.546      | 0.082          | 8.71  | 0.298                   | 0.675      | 0.093          | 10.66 | 0.456                   | 0.627                          | 0.082          | 10.70 | 0.393                   |
| 2.4                                     | 0.661      | 0.118          | 10.23 | 0.437                   | 0.649      | 0.151          | 9.76  | 0.422                   | 0.617                          | 0.115          | 10.72 | 0.380                   |
| 2.5                                     | 0.630      | 0.084          | 9.83  | 0.397                   | 0.568      | 0.105          | 8.81  | 0.323                   | 0.565                          | 0.082          | 9.86  | 0.319                   |
| 2.6                                     | 0.784      | 0.107          | 11.49 | 0.615                   | 0.710      | 0.133          | 10.40 | 0.504                   | 0.657                          | 0.099          | 11.38 | 0.432                   |
| 2.7                                     | 0.672      |                |       | 0.451                   | 0.619      | 0.122          | 9.43  | 0.384                   | 0.609                          | 0.094          | 10.60 | 0.371                   |
| 3.1                                     | 0.633      | 0.186          | 8.79  | 0.401                   | 0.705      |                |       | 0.497                   | 0.633                          |                |       | 0.400                   |
| 3.2                                     | 0.457      | 0.129          | 6.88  | 0.209                   | 0.533      | 0.083          | 6.85  | 0.283                   | 0.490                          | 0.086          | 6.73  | 0.242                   |
| 3.3                                     | 0.554      | 0.151          | 8.00  | 0.307                   | 0.546      | 0.009          | 6.94  | 0.299                   | 0.598                          | 0.104          | 7.65  | 0.357                   |
| 3.4                                     | 0.704      | 0.180          | 9.36  | 0.496                   | 0.344      | 0.095          | 5.74  | 0.118                   | 0.390                          | 0.078          | 6.90  | 0.152                   |
| 3.5                                     | 0.446      | 0.162          | 6.74  | 0.199                   | 0.655      | 0.107          | 9.81  | 0.427                   | 0.635                          | 0.081          | 11.02 | 0.403                   |
| 3.6                                     | 0.525      | 0.168          | 7.67  | 0.275                   | 0.547      | 0.106          | 6.94  | 0.299                   | 0.606                          | 0.118          | 7.70  | 0.367                   |
| 3.7                                     | 0.613      |                |       | 0.376                   | 0.603      | 0.070          | 9.24  | 0.364                   | 0.594                          | 0.054          | 10.35 | 0.353                   |
| Planificación, Preparación y Dominio    |            |                |       |                         |            |                |       |                         | 0.952                          | 0.106          | 10.86 | 0.906                   |
| Interacción con los Estudiantes         |            |                |       |                         |            |                |       |                         | 0.810                          | 0.083          | 10.38 | 0.656                   |
| Comunicación y Organización             |            |                |       |                         |            |                |       |                         | 0.965                          |                |       | 0.932                   |
| Avalúo y Evaluación del Aprovechamiento |            |                |       |                         |            |                |       |                         | 0.867                          | 0.099          | 11.22 | 0.751                   |
| Cumplimiento con el Horario             |            |                |       |                         |            |                |       |                         | 0.527                          | 0.079          | 6.63  | 0.278                   |

Nota: Los parámetros se estimaron por el método de máxima verosimilitud y los valores omitidos se imputaron por el método EM (Expectation Maximization), disponible en el programa AMOS, el cual se utilizó para estimar los modelos factoriales propuestos.

#### MEDIDAS DE LA BONDAD DE AJUSTE

La bondad de ajuste de los modelos factoriales propuestos se evaluó a través de las siguientes estadísticas clasificadas por tipo<sup>2</sup>:

##### Ajuste General:

1. Ji-Cuadrada ( $\chi^2$ )
2. Ji-Cuadrada/gl ( $\chi^2$ /gl)

##### Ajuste Incremental:

3. TLI
4. NFI
5. CFI

##### Parsimonia:

6. RMSEA
7. AIC
8. ECVI.

A la prueba de hipótesis asociada a  $\chi^2$  se le dio poca importancia, porque según Schultz & Whitney (2005), Martens (2005) y Ullman (2006), la significación estadística es extremadamente sensible al tamaño de la muestra. Estos autores afirman que en muestras grandes (algunos consideran una muestra grande si es mayor de 150 y otros si es mayor de 200) la prueba de hipótesis basada en el estadístico  $\chi^2$  tiene muy poca validez. No obstante, el valor estimado de  $\chi^2$  sirvió para evaluar el ajuste de los modelos propuestos en términos relativos. En el caso del AFC, contrario a la mayoría de las pruebas de hipótesis, mientras menor sea el valor de  $\chi^2$ , mejor es el ajuste.<sup>3</sup> La segunda medida de ajuste general, según las reglas “prácticas”, se considera aceptable si está entre 2 y 4, pero si es mayor que 5 se considera totalmente inaceptable (Byrne, 2006, Bentler, 2001). También esta medida cuando es menor que 1, se interpreta como evidencia de un sobre-ajuste en el modelo.

<sup>2</sup> Para una discusión más amplia y detallada de las medidas de ajuste consúltense las siguientes fuentes: 1) Byrne, B. (2006). *Structural equation modeling with AMOS*, 2) Shultz, K. S. & Whitney, D. J. (2005). *Measurement theory in action: Case studies and exercises*. 3) Ullman, J. B. (2006). *Structural equation modeling: Reviewing the basics and moving forward*. *Journal of Personality Assessment* vol. 87 no.1.

<sup>3</sup> La prueba estadística de  $\chi^2$  en el contexto del AFC, es indicativa de un buen ajuste cuando se acepta la hipótesis nula. Es decir, cuando el valor-p (probabilidad de rechazar la hipótesis nula dado que es cierta) es mayor que 0.05, indicando que no existe diferencia significativa entre la matriz de covarianzas observada y la matriz estimada a base del modelo propuesto. El valor-p disminuye a medida que el valor del estadístico de  $\chi^2$  aumenta. De ahí que mientras menor es  $\chi^2$  mejor es el ajuste.

En el grupo incremental o comparativo, los índices considerados en este estudio (TLI, NFI y CFI) son aceptables cuando el valor es mayor que 0.90 (Byrne, 2006; Shultz & Whitney, 2005; Ullman, 2006; Pérez-Gil, 2001; Bentler, 2001). El tercer grupo, en el cual están las medidas de parsimonia, además de tomar en cuenta el ajuste general, también consideran la complejidad del modelo. De ahí el término de la parsimonia. La primera medida de este grupo, de acuerdo con Byrne, 2006: en MacCallum, et al., 1996, se considera aceptable e indicativa de un buen ajuste, si está entre 0.04 y 0.08, y muy pobre si es mayor que 0.10. Las últimas dos medidas de parsimonia, AIC y ECVI, son particularmente útiles para comparar dos o más modelos alternativos. Para las últimas dos medidas no hay un campo de valores definido para determinar si el ajuste es aceptable o no. Sin embargo, está establecido que mientras menor es el valor, mejor es el ajuste del modelo.

Como puede comprobarse en la Tabla 6, el modelo de 5 Factores y Un Factor General superó a los otros dos modelos en todas las estadísticas de ajuste. Registró la  $\chi^2$  más baja, y también fue el único que arrojó una Ji-Cuadrada Corregida ( $\chi^2$ /gl) dentro de los límites indicativos de un ajuste adecuado. Aunque los tres modelos propuestos exhibieron un desempeño bastante parecido en cuanto a las medidas de tipo incremental, ya que todos arrojaron valores aceptables (mayores que 0.90), el último modelo registró valores ligeramente superiores a los otros dos. En lo que respecta a las medidas de parsimonia, el tercer modelo superó por bastante a los otros, puesto que, registró valores más pequeños en las tres medidas de parsimonia.

Basado en los valores estimados de los índices de ajuste discutidos anteriormente, se concluye que el modelo de 5 Factores y Un Factor General, registró un ajuste a los datos recopilados superior a los otros dos modelos alternativos. Así las cosas, en la próxima sección se presentan los hallazgos específicos de la validez de “constructo” a la luz del modelo factorial de segundo orden seleccionado como el de mejor ajuste.

#### RESULTADOS DEL MODELO DE SEGUNDO ORDEN

Las columnas de la Tabla 7 correspondientes a la estructura de 5 Factores y Un Factor General tienen la información necesaria para analizar dicha estructura en detalle. Todos los reactivos, así como los cinco factores de primer orden

arrojaron un valor de  $t$  mayor que 2, indicativo de que el peso factorial estimado es significativamente mayor que cero (a un nivel de alfa igual a 0.05). Resalta de estos resultados que los reactivos 2.1 y 3.4, arrojaron una  $R^2$  bastante baja (0.04 y 0.15, respectivamente), indicativo de que contribuyen relativamente poco a medir el “constructo” antes mencionado. Por otra parte, la subdimensión Cumplimiento con el Horario reflejó una  $R^2$  relativamente baja (0.278), lo cual quiere decir que el factor general de segundo orden, explica alrededor del 28 por ciento de las variaciones de la subdimensión mencionada. Mientras que el restante 72 por ciento es explicada por el componente de error. Las razones de la débil influencia del factor general sobre la subdimensión Cumplimiento con el Horario, no son parte de este estudio. Sin embargo, parece ser que el estudiante centra su atención principalmente en las destrezas del profesor que están más íntimamente relacionadas con su aprendizaje.

El Diagrama 2 muestra el peso factorial estandarizado de cada uno de los factores de primer orden. Como se puede ver en dicho diagrama, el factor general de segundo orden registró efectos por encima de 0.70 en todos los factores de primer orden o subdimensiones, excepto en el de Cumplimiento con el Horario. También se desprende de ese diagrama que el peso factorial es mayor en unas dimensiones que en otras. Por ejemplo, los dos factores de primer

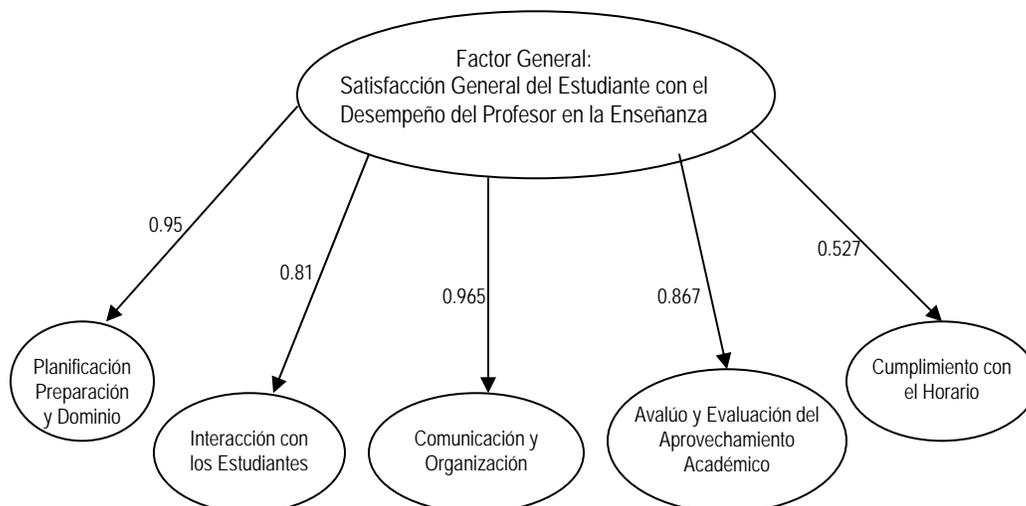
orden con el peso factorial mayor fueron: Comunicación & Organización y el de Planificación, Preparación & Dominio. Luego le siguieron los pesos factoriales de Avalúo & Evaluación del Aprovechamiento Académico y la subdimensión de Interacción con los Estudiantes.

## CONCLUSIONES E IMPLICACIONES

### Conclusiones

El conjunto de 32 reactivos del instrumento de la EEP cubre adecuadamente el espectro de posibles aspectos y subdimensiones del “constructo” de la Calidad del Profesor en la Enseñanza, según es percibida por el estudiante. Sin embargo, de acuerdo con la evidencia recopilada, los reactivos 2.1 y 3.4 (respectivamente, los reactivos #1 de la Parte 1 y #4 de la Parte 2 del Formulario de EEP) tienen problemas tanto de confiabilidad como de capacidad para discriminar los profesores con puntajes más altos y, por ende, de mayor calidad en las experiencias de enseñanza-aprendizaje, de aquellos profesores con los puntajes más bajos. Los problemas de ambos reactivos pudieran estar relacionados con el contenido de múltiples asuntos en sus premisas. Por ejemplo, el reactivo 2.1 tiene tres ideas un tanto distintas (respeto, cortesía e inferioridad), y el 3.4 habla de objetivos y normas, que no necesariamente son sinónimos.

Diagrama 2. Pesos Factoriales de los Factores de 1er. Orden



El análisis de confiabilidad de la estructura de 3 factores, arrojó evidencia indicativa de que la subdimensión de Cumplimiento de Deberes, compuesta por los reactivos 3.1 al 3.7 de la parte 3 del formulario de EEP, tiene problemas de confiabilidad, porque falló en rebasar la regla “práctica” de 0.70 en el índice de Spearman-Brown. En la estructura de 5 factores, los 7 reactivos antes mencionados, se reestructuraron en función de su contenido. Los reactivos 3.1, 3.2, 3.3 y 3.6 se reagruparon para formar la subdimensión que se denominó como Cumplimiento con el Horario. Dicha subdimensión arrojó evidencia contundente de falta de confiabilidad, ya que los dos índices de confiabilidad (Cronbach y Spearman-Brown) se quedaron por debajo de 0.70.

Por otra parte, la revisión de las investigaciones citadas en este estudio, conjuntamente con el análisis del contenido específico de los reactivos del instrumento piloto de EEP, permitió inferir que además de la estructura de 3 factores que se desprende literalmente de la organización del propio formulario, era altamente probable que el “constructo” de Calidad de la Enseñanza se subdividiera en 5 dimensiones: Planificación, Preparación y Dominio, Interacción con los Estudiantes, Comunicación y Organización, Avalúo y Evaluación del Aprovechamiento, y Cumplimiento con el Horario. Esta estructura de 5 dimensiones se configuró descomponiendo la estructura original de 3 dimensiones en 5 dimensiones más específicas.

Tanto el Análisis Factorial Exploratorio como Confirmatorio evidenciaron la multidimensionalidad del “constructo” de la Calidad del Profesor en la Enseñanza. Asimismo, el análisis comparativo de las tres estructuras factoriales probadas (3 factores, 5 factores y 5 factores de primer orden más un factor de segundo orden), arrojó evidencia empírica que apoya la conclusión de que el “constructo” de la Calidad de la Enseñanza es más compatible con una estructura de 5 subdimensiones y un fuerte factor general, que representa la visión unidimensional e integrada que tiene el estudiante al evaluar las experiencias de enseñanza-aprendizaje en el salón de clase. Apodaca & Grad (2005), Mazelan et al. (1991) y Marks (1997) argumentan que esa óptica unidimensional de los estudiantes universitarios obedece a la tendencia de éstos a enjuiciar la calidad de los profesores más bien como consumidores de un producto.

Aunque la evidencia empírica apoya la estructura factorial de segundo orden, el papel que desempeñan las cinco subdimensiones en la percepción general del estudiante no es ni siquiera aproximadamente uniforme. Las subdimensiones de Planificación, Preparación & Dominio y Comunicación & Organización, desempeñan un papel central en la percepción general del estudiante. Mientras que la Interacción y el sistema de Avalúo y Evaluación desempeñan un papel de menor importancia que las dos subdimensiones antes mencionadas. La subdimensión Cumplimiento del Horario, que mide aspectos de la puntualidad y la asistencia del profesor, demostró un rol marginal en la percepción general del estudiante con respecto a la calidad del profesor en la enseñanza. Estos resultados tienen unas implicaciones importantes para la administración universitaria en lo que respecta a las funciones sumativa y formativa del proceso de evaluación estudiantil de profesores.

En conclusión, salvo los reactivos 2.1 y 3.4, la evidencia empírica recopilada en este estudio apoya la validez del “constructo” de las percepciones del estudiante en torno a la calidad de las experiencias de enseñanza-aprendizaje desarrolladas por el profesor. Esto implica también el apoyo a la validez de los puntajes que se calculen a partir de los reactivos de la EEP para propósitos sumativo y formativo.

### **Implicaciones**

Las conclusiones esbozadas anteriormente tienen unas implicaciones, que se enumeran a continuación.

1. En vista de los problemas de confiabilidad y de capacidad de discriminación, los reactivos 2.1 y 3.4 deben revisarse o eliminarse del formulario de EEP.
2. Ya que tanto la evidencia empírica como teórica apoya una estructura multidimensional de la calidad de la enseñanza, se recomienda que el formulario se reorganice de tal manera que sea compatible con la estructura validada de 5 subescalas.
3. Dada la estructura jerárquica validada de 5 factores de primer orden fuertemente influidos por un único factor general de segundo orden, se recomienda que cualquier puntaje global que se defina para propósitos sumativos, tome en cuenta mediante un promedio ponderado, el peso

- diferenciado que registraron los factores de primer orden.
4. Como resultado de la evidencia empírica y teórica en apoyo de la multidimensionalidad del “constructo” de la calidad de la enseñanza, se recomienda que, además del propósito sumativo de la EEP, se institucionalice el uso de los resultados para propósitos formativos. Para ello es necesario que se prepare un reporte estándar que además del puntaje global, incluya puntajes para las cinco subdimensiones y por reactivo. Sólo así los resultados de la EEP pueden tener una verdadera utilidad práctica como mecanismo para retroalimentar el proceso de enseñanza, en aras de mejorarlo continuamente. Aun más, se recomienda que el reporte también incluya los puntajes del período corriente de evaluación y de los dos anteriores. Así el profesor evaluado puede tener un cuadro claro de cómo va mejorando. Igualmente, sería bueno que el reporte incluyera medidas estadísticas de la posición relativa de los puntajes global, por subescala y reactivo.
  5. El autor de este estudio comparte algunas de las críticas a la EEP, discutidas por Aleamoni, 1981 y resumidas en la página 4 de este estudio. Por las limitaciones que tiene el estudiante como “experto” observador de todos los aspectos del proceso de enseñanza, no es aconsejable que la EEP sea el criterio más importante en el proceso evaluativo de las competencias en la enseñanza, ni mucho menos la única fuente de información de dicho proceso.

### Estudios Adicionales

Obviamente el presente estudio tan sólo debe marcar el inicio de una serie de investigaciones sobre la EEP y sus usos en el contexto particular de la UPR en Cayey. Sin pretender ser exhaustivo, a continuación se enumeran algunas de las investigaciones que ayudarían a comprender mejor la temática de la calidad del proceso de enseñanza, desde la perspectiva del estudiante.

1. Replicar el análisis de confiabilidad y validez con una muestra mayor y con las modificaciones sugeridas en relación con el formulario y los reactivos.

2. Hacer un estudio de la validez predictiva o de criterio, en el cual se relacione una medida (o varias) del aprovechamiento del estudiante con las opiniones de éste en torno a la calidad de las experiencias de enseñanza facilitadas por el profesor.
3. Estudio de la validez predictiva de las 5 subdimensiones específicas validadas.
4. Múltiples estudios para evaluar el efecto interventor que pudieran tener las siguientes variables sobre la evaluación estudiantil de los profesores.

#### Profesor:

- Sexo
- Edad
- Años de experiencia enseñando
- Rango académico
- Nombramiento (permanente o contrato)

#### Estudiante:

- Sexo
- Escuela superior de procedencia
- Trasfondo socioeconómico
- Generación universitaria
- Créditos aprobados

5. Un estudio para contestar la interrogante de si existe un problema de “inflación” en las notas y su efecto en la percepción del estudiante en torno a las competencias del profesor en la enseñanza.

### Limitaciones

Debido a la compleja estructura multidimensional y jerárquica del “constructo” objeto de análisis, así como también al número relativamente grande de reactivos, la muestra de 350 estudiantes no resultó suficientemente grande para el Análisis Factorial Confirmatorio, ya que ni siquiera alcanzó el mínimo recomendado de cinco observaciones por parámetro estimado (Hair, 1999). Una muestra mayor pudiera reducir el error estándar de los estimados, así como también el riesgo de inestabilidad en los mismos. Dado que la estructura validada de 5 dimensiones y un factor general de segundo orden, ronda los 100 parámetros, se necesitaría una muestra de por lo menos 1,000 estudiantes (10 observaciones por

parámetro), los cuales se podrían obtener seleccionando al azar alrededor de 50 secciones.

## REFERENCIAS

- Abrami, P. C. (1989). How should we use student ratings to evaluate teaching? *Research in Higher Education* vol. 30.
- Abrami, P. C. & D'Apollonia, S. (1991). Multidimensional student evaluations of teaching effectiveness generalizability: comentado en Marsh (1991). *Journal of Educational Psychology* vol. 83.
- Abrami, P. C., D'Apollonia, S. & Rosenfield, S. (1996). The dimensionality of student ratings of instruction: What we know and what we do not en: J. C. Smart (Ed.) *Higher education: Handbook of theory and research* vol. 11, New York, Agathon Press.
- Algozzine, B., Beattie, H, Bray, M., Flowers, C., Gretes, J., Howley, L. Mohanty, G. & Spooner, F. (2005). Student evaluation college teaching: A practice in search of principles. *College Teaching*, vol. 52 no. 4.
- Aleamoni, L. M. (1981). Student evaluation of teacher performance. Consultado el 28 de noviembre de 2007 en: <http://ericae.net/edo/ED289887.htm>.
- Apodaca, P. & Rodríguez, M. (1999). La opinión de los alumnos en la evaluación de la calidad docente: posibilidades, limitaciones y estructura dimensional de sus indicadores, en: J. Vidal (Coord.) *Indicadores en la universidad: Información y decisiones* (Madrid, Consejo de Universidades, Ministerio de Educación y Cultural).
- Apodaca, P. & Grad, H. (2005). The dimensionality of student ratings of teaching: integration of uni-and multidimensional models. *Studies in Higher Education* vol. 30 no. 6.
- Arbuckle, J. L. & Wothke, W. (1999). *AMOS 4.0 User's Guide*. SmallWaters Corporation.
- Backhoff, E., Larrazolo, N. & Rosas, M. (2000). Nivel de dificultad y poder de discriminación del Examen de Habilidades y Conocimientos Básicos (EXHCOBA). *Revista Electrónica de Investigación Educativa*, 2(1). Consultado el 28 de noviembre de 2007 en: <http://redie.uabc.mx/vol2no1/contenido-backhorr.html>.
- Benter, P. M. & Chou, C. P. (1987). Practical issues in structural modeling. *Sociological Methods and Research* vol. 16 no. 1.
- Brown, G & Atkins, M. (1988). *Effective teaching in higher education* (London, Routledge).
- Byrne, B. M. (2006). *Structural Equation Modeling with AMOS: Basic concepts, applications, and programming*. Beverly Hills and London: Sage Publications.
- Cheung, D. (2000). Evidence of a single second-order factor in student ratings of teaching effectiveness. *Structural Equation Modeling* vol. 7 no. 3: Lawrence Erlbaum Associates, Inc.
- Clouder, L. (1998). Getting the right answers: Student evaluation as reflection of intellectual development? *Teaching in Higher Education*, vol. 3 no. 2.
- Cohen, P. A. (1981). Student ratings of instruction and student achievement: a meta-analysis of multisection validity studies, *Review of Educational Research*, 51.
- Costello, A. B. & Osborne, J. W. (2005). Best practices in exploratory factor analysis: Four recommendations for getting the most from your analysis. *Practical Assessment, Research and Evaluation* vol. 10 no. 7.
- Cranton, P. A. & Hilgartner, W. (1981). The relationship between student ratings and instructor behavior: Implications for improving teaching. *Canadian Journal of Higher Education* vol. 11.
- Cronbach, L. J. (1980). Validity on parole: How can we go straight. *New Directions for Testing and Measurement* vol. 5.
- D' Apollonia, S. & Abrami, P. C. (1997). Navigating student ratings of instruction. *American Psychology* vol. 52.
- Davis, R. H. (1977). Learning by design, en: J. A. Centra (Ed.) *Renewing and evaluating teaching*. San Francisco Jossey-Bass.
- DeVellis, R. F. (1991). *Scale development: Theory and applications*. Newbury Park, CA: Sage Publications.
- Entwistle, N. J. & Ramsden, P. (1983). *Understanding student learning* (London, Croom Helm).
- Feldman, K. A. (1989). Instructional effectiveness of college teachers themselves, current and former students, colleagues, administrator, and external (neutral) observers, *Research in Higher Education* vol. 30.
- Gagné, R. & Briggs, L. J. (1974). *Principles of instructional design*. New York, Holt Rinehart and Winston.
- Glynn, J. G., Sauer, P. L. & Wood, G. R. (2006). Dimensionality of student evaluation of teaching scale: A ten-year review. *Academy*

- of Educational Leadership Journal vol. 10 no. 2.
- Hair, J. F., Anderson, R., Tathan, R. L. & Black, W. C. (1999). *Multivariable data analysis*, 3<sup>rd</sup> edition, MacMillan, Inc. Publishing.
- Jahangir, S. F. (1988). Students' evaluation of their teachers' performance. *Pakistan Psychological Studies* vol. 6.
- Kaiser, H. F. (1958). The varimax criterion for analytic rotation in factor analysis. *Psychometrika* vol. 23.
- Kember, D., Lenny, D. & Kwan, K. (2002). Does the use of student feedback questionnaires improve the overall quality of teaching? *Assessment and Evaluation in Higher Education* vol. 27.
- Kim, J. O. & Mueller, C. W. (1978). *Factor analysis*. Sages University Paper series of Quantitative Applications in the Social Sciences, series no. 07-014. Beverly Hills and London: Sage Publications.
- Lowman, J. (1984). *Mastering the techniques of teaching*. San Francisco Jossey-Bass.
- MacCallum, R. C., Brown, M. W. & Sugawara, H. M. (1996). Power analysis and determination of sample size for covariance structure modeling. *Psychological Methods* vol. 1.
- Marks, A. (1997). Me-me-me monsters in our midst. *Times Higher Educational Supplement*.
- Marsh, H. H. & Hocevar, D. (1984). The multidimensionality of students' evaluation of teaching effectiveness. *Teaching & Teacher Education* vol. 7.
- Mazelan, P., Brannigan, C., Green, D., Tormay, P. & O'Shea, J. (1991). Using measure of student satisfaction: the implications of a user-led strategy of quality assurance in higher education, *Broadcast* vol. 18.
- Marsh, H. W. (1984). Students' evaluations of university teaching: Dimensionality, reliability, validity, potential biases, and utility. *Journal of Educational Psychological* vol. 76.
- Marsh, H. W. & Hocevar, D. (1991). The multidimensionality of students' evaluation of teaching effectiveness: The generality of factor structure across academia discipline, instruction level, and course level. *Teaching & Teacher Education* vol. 7.
- Marsh, H. W. & Roche, L. A. (1993). The use of students' evaluations and an individually structured intervention to enhance university teaching effectiveness. *American Educational Research Journal* vol. 30.
- Marton, F. & Säljö, R. (1976). On qualitative differences in learning, outcome, and process. *British Journal of Educational Psychology* vol. 46.
- Martens, M. P. (2005). The use of structural equation modeling in counseling psychology research. *Counseling Psychologist* vol. 33.
- McPherson, M. A. (2006). Determinants of how students evaluate teachers. *Journal of Economic Education*, vol. 37 no. 1.
- Murray, H. G. (1980). *Evaluating University Teaching: A review of research*, Ontario Confederation University Faculty Associations, Toronto.
- Murray, H. G. (1983). Low-inference classroom teaching behaviors and student ratings of college teaching effectiveness. *Journal of Educational Psychology* vol. 75.
- Murray, H. G. (1994). The impact of formative and summative evaluation of teaching in North American Universities. *Assessment and Evaluation in Higher Education* vol. 9 no. 2.
- Newble, D. & Cannon, R (1995). *A handbook for teachers in universities & colleges: A guide to improving teaching methods* (London, Kogan Page).
- Nunnally, J. (1978). *Psychometric theory*, 2da. Edición. Nueva Cork: McGraw-Hill.
- Pérez-Gil, J. A., Chacón-Moscoso, S. & Moreno-Rodríguez, R. (2000). Validez de constructo: El uso de análisis factorial exploratorio-confirmatorio para obtener evidencias de validez. *Psicothema* vol. 12 no. 2.
- Prosser, M. & Trigwell, K. (2006). Confirmatory factor analysis of the approaches to teaching inventory. *British Journal of Educational Psychology* vol. 6.
- Ramsden, P. (1992). *Learning to teach in higher education* (London, Routledge).
- Remmers, H. H. (1927). The Purdue Rating Scale for Instructor. *Educational Administration and Supervision* vol. 13.
- Remmers, H. H. & Brandenburg (1927). Experimental data on the Purdue Rating Scale for Instructors. *Educational Administration and Supervision*. vol. 6.
- Renaud, R. D. & Murray, H. G. (2005). Factorial validity of student ratings of instruction. *Research of Higher Education* vol. 46 no. 8.
- Riaz, M. N. (2000). Student evaluation of the university teaching quality: Analysis of a teacher's rating scale for a sample of university students. *Pakistan Journal of Psychological Research* vol. 14 no. 3.
- Rubina, H. & Pervez, S. (2004). Development and validation of teachers' job performance

- scale. Pakistan Journal of Psychological Research vol. 19 nos. 3-4.
- Scriven, M. (1995). Student ratings offer useful input to teacher evaluations. Practical Assessment, Research & Evaluation vol. 4 no. 7. Consultado el 28 de noviembre de 2007 en: <http://PAREonline.net/getvn.asp>.
- Shultz, K. S. & Whitney, D. J. (2005). Measurement theory in action: Case studies and exercises. Beverly Hills and London: Sage Publications.
- Silva, F. (1993). Psychometric foundations and behavioral assessment. Newbury Park, CA: Sage Publications.
- Sinclair, C. (1984). Students and lecturers-are we so different? Educational Change and Development vol. 15 no.2.
- Spector, P. E. (1992). Summated rating scales: An introduction. Sages University Paper series of Quantitative Applications in the Social Sciences, series no. 07-024. Beverly Hills and London: Sage Publications.
- Tabachnick, B. G. & Fidell, L. S. (2001). Using multivariable statistics. Boston: Allyn and Bacon.
- Ullman, J. B. (2006). Structural equation modeling: Reviewing the basics and moving forward. Journal of Personality Assessment vol. 87 no. 1.

## **APÉNDICE A: FORMULARIO DE EVALUACIÓN ESTUDIANTIL DE PROFESORES**







## APÉNDICE B: MATRIZ DE CORRELACIONES DE LOS 32 REACTIVOS

Tabla 1. Matriz de Correlaciones Bivariadas entre los Reactivos del Instrumento de Evaluación Estudiantil de Profesores

| Reactivos | 1.1   | 1.2   | 1.3   | 1.4   | 1.5   | 1.6   | 1.7   | 1.8   | 1.9   | 1.1   | 1.11  | 1.12  | 1.13  | 1.14  | 1.15  | 1.16  | 1.17  | 1.18  | 2.1   | 2.2   | 2.3   | 2.4   | 2.5   | 2.6   | 2.7   | 3.1   | 3.2   | 3.3   | 3.4   | 3.5   | 3.6   | 3.7 |  |  |  |  |  |
|-----------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-----|--|--|--|--|--|
| 1.1       | 1     |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |     |  |  |  |  |  |
| 1.2       | 0.504 | 1     |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |     |  |  |  |  |  |
| 1.3       | 0.507 | 0.540 | 1     |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |     |  |  |  |  |  |
| 1.4       | 0.404 | 0.436 | 0.390 | 1     |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |     |  |  |  |  |  |
| 1.5       | 0.345 | 0.437 | 0.615 | 0.314 | 1     |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |     |  |  |  |  |  |
| 1.6       | 0.497 | 0.422 | 0.433 | 0.323 | 0.359 | 1     |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |     |  |  |  |  |  |
| 1.7       | 0.292 | 0.428 | 0.449 | 0.236 | 0.509 | 0.457 | 1     |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |     |  |  |  |  |  |
| 1.8       | 0.353 | 0.398 | 0.409 | 0.315 | 0.524 | 0.392 | 0.586 | 1     |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |     |  |  |  |  |  |
| 1.9       | 0.407 | 0.517 | 0.661 | 0.322 | 0.642 | 0.535 | 0.466 | 0.412 | 1     |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |     |  |  |  |  |  |
| 1.1       | 0.570 | 0.653 | 0.667 | 0.470 | 0.539 | 0.592 | 0.496 | 0.475 | 0.718 | 1     |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |     |  |  |  |  |  |
| 1.11      | 0.465 | 0.549 | 0.486 | 0.356 | 0.379 | 0.525 | 0.406 | 0.393 | 0.550 | 0.562 | 1     |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |     |  |  |  |  |  |
| 1.12      | 0.379 | 0.315 | 0.498 | 0.271 | 0.551 | 0.445 | 0.427 | 0.474 | 0.446 | 0.460 | 0.442 | 1     |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |     |  |  |  |  |  |
| 1.13      | 0.443 | 0.533 | 0.356 | 0.285 | 0.293 | 0.510 | 0.385 | 0.409 | 0.439 | 0.546 | 0.581 | 0.376 | 1     |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |     |  |  |  |  |  |
| 1.14      | 0.461 | 0.406 | 0.299 | 0.227 | 0.203 | 0.486 | 0.332 | 0.298 | 0.371 | 0.460 | 0.471 | 0.268 | 0.602 | 1     |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |     |  |  |  |  |  |
| 1.15      | 0.478 | 0.381 | 0.375 | 0.282 | 0.354 | 0.475 | 0.334 | 0.423 | 0.396 | 0.525 | 0.561 | 0.456 | 0.510 | 0.469 | 1     |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |     |  |  |  |  |  |
| 1.16      | 0.505 | 0.462 | 0.477 | 0.298 | 0.379 | 0.479 | 0.404 | 0.386 | 0.425 | 0.513 | 0.541 | 0.456 | 0.595 | 0.539 | 0.542 | 1     |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |     |  |  |  |  |  |
| 1.17      | 0.338 | 0.364 | 0.246 | 0.194 | 0.177 | 0.382 | 0.331 | 0.266 | 0.250 | 0.389 | 0.393 | 0.356 | 0.475 | 0.431 | 0.425 | 0.520 | 1     |       |       |       |       |       |       |       |       |       |       |       |       |       |       |     |  |  |  |  |  |
| 1.18      | 0.507 | 0.559 | 0.510 | 0.336 | 0.437 | 0.526 | 0.557 | 0.486 | 0.542 | 0.673 | 0.571 | 0.456 | 0.594 | 0.482 | 0.537 | 0.626 | 0.573 | 1     |       |       |       |       |       |       |       |       |       |       |       |       |       |     |  |  |  |  |  |
| 2.1       | 0.116 | 0.147 | 0.076 | 0.036 | 0.113 | 0.032 | 0.125 | 0.131 | 0.033 | 0.099 | 0.119 | 0.119 | 0.171 | 0.011 | 0.138 | 0.142 | 0.139 | 0.148 | 1     |       |       |       |       |       |       |       |       |       |       |       |       |     |  |  |  |  |  |
| 2.2       | 0.224 | 0.236 | 0.320 | 0.117 | 0.371 | 0.181 | 0.310 | 0.417 | 0.229 | 0.267 | 0.218 | 0.419 | 0.188 | 0.085 | 0.184 | 0.172 | 0.156 | 0.270 | 0.288 | 1     |       |       |       |       |       |       |       |       |       |       |       |     |  |  |  |  |  |
| 2.3       | 0.197 | 0.233 | 0.303 | 0.158 | 0.338 | 0.237 | 0.435 | 0.578 | 0.238 | 0.329 | 0.224 | 0.377 | 0.194 | 0.099 | 0.253 | 0.257 | 0.122 | 0.290 | 0.030 | 0.382 | 1     |       |       |       |       |       |       |       |       |       |       |     |  |  |  |  |  |
| 2.4       | 0.330 | 0.282 | 0.432 | 0.288 | 0.499 | 0.390 | 0.282 | 0.375 | 0.480 | 0.426 | 0.338 | 0.520 | 0.312 | 0.223 | 0.349 | 0.396 | 0.239 | 0.391 | 0.097 | 0.248 | 0.431 | 1     |       |       |       |       |       |       |       |       |       |     |  |  |  |  |  |
| 2.5       | 0.396 | 0.276 | 0.372 | 0.247 | 0.361 | 0.372 | 0.294 | 0.317 | 0.384 | 0.412 | 0.331 | 0.421 | 0.348 | 0.307 | 0.432 | 0.460 | 0.287 | 0.457 | 0.229 | 0.205 | 0.238 | 0.465 | 1     |       |       |       |       |       |       |       |       |     |  |  |  |  |  |
| 2.6       | 0.385 | 0.331 | 0.477 | 0.307 | 0.477 | 0.371 | 0.423 | 0.411 | 0.421 | 0.438 | 0.322 | 0.449 | 0.296 | 0.204 | 0.328 | 0.410 | 0.154 | 0.411 | 0.152 | 0.401 | 0.460 | 0.465 | 0.490 | 1     |       |       |       |       |       |       |       |     |  |  |  |  |  |
| 2.7       | 0.412 | 0.333 | 0.361 | 0.300 | 0.348 | 0.335 | 0.430 | 0.378 | 0.369 | 0.443 | 0.350 | 0.436 | 0.350 | 0.309 | 0.355 | 0.384 | 0.253 | 0.468 | 0.199 | 0.284 | 0.237 | 0.442 | 0.469 | 0.525 | 1     |       |       |       |       |       |       |     |  |  |  |  |  |
| 3.1       | 0.175 | 0.321 | 0.113 | 0.063 | 0.036 | 0.210 | 0.149 | 0.132 | 0.106 | 0.228 | 0.252 | 0.123 | 0.269 | 0.192 | 0.330 | 0.206 | 0.364 | 0.314 | 0.141 | 0.073 | 0.127 | 0.049 | 0.135 | 0.114 | 0.130 | 1     |       |       |       |       |       |     |  |  |  |  |  |
| 3.2       | 0.224 | 0.250 | 0.106 | 0.030 | 0.124 | 0.066 | 0.071 | 0.067 | 0.057 | 0.175 | 0.199 | 0.118 | 0.101 | 0.167 | 0.194 | 0.166 | 0.197 | 0.241 | 0.140 | 0.116 | 0.085 | 0.041 | 0.145 | 0.084 | 0.153 | 0.422 | 1     |       |       |       |       |     |  |  |  |  |  |
| 3.3       | 0.361 | 0.314 | 0.212 | 0.086 | 0.180 | 0.250 | 0.214 | 0.228 | 0.141 | 0.285 | 0.229 | 0.169 | 0.406 | 0.308 | 0.352 | 0.298 | 0.347 | 0.272 | 0.221 | 0.211 | 0.093 | 0.092 | 0.184 | 0.202 | 0.245 | 0.366 | 0.242 | 1     |       |       |       |     |  |  |  |  |  |
| 3.4       | 0.307 | 0.340 | 0.225 | 0.087 | 0.178 | 0.316 | 0.166 | 0.195 | 0.213 | 0.280 | 0.337 | 0.138 | 0.392 | 0.346 | 0.487 | 0.348 | 0.420 | 0.371 | 0.122 | 0.090 | 0.142 | 0.107 | 0.158 | 0.191 | 0.162 | 0.463 | 0.287 | 0.399 | 1     |       |       |     |  |  |  |  |  |
| 3.5       | 0.356 | 0.402 | 0.479 | 0.265 | 0.471 | 0.364 | 0.394 | 0.354 | 0.521 | 0.486 | 0.355 | 0.338 | 0.358 | 0.287 | 0.308 | 0.345 | 0.222 | 0.426 | 0.082 | 0.273 | 0.319 | 0.450 | 0.245 | 0.436 | 0.355 | 0.189 | 0.089 | 0.185 | 0.351 | 1     |       |     |  |  |  |  |  |
| 3.6       | 0.288 | 0.337 | 0.150 | 0.090 | 0.162 | 0.363 | 0.319 | 0.304 | 0.171 | 0.274 | 0.274 | 0.168 | 0.353 | 0.249 | 0.297 | 0.242 | 0.316 | 0.379 | 0.288 | 0.194 | 0.177 | 0.119 | 0.162 | 0.224 | 0.246 | 0.360 | 0.256 | 0.380 | 0.307 | 0.194 | 1     |     |  |  |  |  |  |
| 3.7       | 0.331 | 0.388 | 0.413 | 0.148 | 0.396 | 0.371 | 0.367 | 0.305 | 0.426 | 0.431 | 0.401 | 0.292 | 0.352 | 0.313 | 0.376 | 0.355 | 0.309 | 0.510 | 0.043 | 0.255 | 0.292 | 0.307 | 0.339 | 0.471 | 0.386 | 0.306 | 0.250 | 0.289 | 0.453 | 0.481 | 0.324 | 1   |  |  |  |  |  |