

12 de septiembre de 2006.

Dr. Jorge I. Vélez Arocho, Rector
Presidente del Senado Académico
Recinto Universitario de Mayagüez

Dra. Doris Ramírez-Soto, Presidenta
Comité Institucional para la Elaboración de Instrumentos de Evaluación del Personal
Docente (CIEPD)

Miembros del Comité
Dra. Nidia S. López- Orientación
Prof. Jeannette Valentín-Biblioteca General
Prof. María Medina- Administración de Empresas
Dra. Darnyd Ortiz- Artes y Ciencias
Dr. Ángel Custodio- Ciencias Agrícolas
Prof. Baldomero Lloréns- Ingeniería
Dr. Andrés Collazo-Consultor

Informe sobre el “Cuestionario de Opinión Estudiantil” (COE)

I. Trasfondo

El Senado Académico, en su reunión del 15 de noviembre de 2005 consideró y aprobó el Informe del Comité Institucional para la Elaboración de Instrumentos de Evaluación del Personal Docente (CIEPD) sobre el documento “Módulo de la Evaluación de la Enseñanza por los Estudiantes”. El CIEPD debía evaluar las sugerencias y recomendaciones presentadas en dicha reunión y realizar un estudio piloto que permitiera establecer las propiedades psicométricas del documento. La Certificación Número 05-56 del 20 de enero de 2006 del Senado Académico incluye el cuestionario aprobado por el CIEPD conocido como “Cuestionario de Opinión Estudiantil” (Anejo 1). La Oficina de Investigación Institucional y Planificación (Dra. Janet Bonilla e Inv. Irmannette Lugo) generó el banco de datos del estudio piloto. El análisis e interpretación de los datos incluido en este informe fue realizado por el Dr. Andrés Collazo, consultor del CIEPD.

II. Resultados del estudio piloto del *Cuestionario de Opinión Estudiantil* (COE)

Durante el primer semestre del año académico 2005-2006, se llevó a cabo el estudio piloto con el propósito de evaluar las propiedades psicométricas de las puntuaciones obtenidas con el *Cuestionario de Opinión Estudiantil* (COE) en dos versiones: papel y electrónica. Este estudio se llevó a cabo con grupos de estudiantes cuyos profesores y profesoras accedieron, voluntariamente, a que el instrumento se administrara en dos de sus secciones de un mismo curso. En una de las secciones se administró la versión en

papel y en la otra se administró la versión electrónica (generada por el Centro de Cómputos RUM). Se observó, sin embargo, excepciones a este procedimiento que posiblemente contribuyeron a que el grupo de estudiantes que contestó una versión fuera distinto al grupo que contestó la otra versión (ver nota al calce de la Tabla 1).

En el estudio participaron 34 secciones, 17 por cada versión del instrumento. Además, participaron cinco secciones adicionales de las cuales 33 estudiantes contestaron solamente la versión electrónica. Estas secciones se eliminaron de los análisis que aquí se presentan. El total de secciones y la cantidad de estudiantes participantes se resume en la Tabla 1. Las tasas de respuesta correspondientes a las versiones en papel y electrónica fueron 65% y 60%, respectivamente.

Tabla 1: *Distribución de los participantes del estudio por facultad y versión del COE*

Facultad	Núm. Secciones	Cantidad estudiantes	
		Versión papel	Versión electrónica
Administración de Empresas	8	63	55
Artes y Ciencias	15 ^a	132	137
Ciencias Agrícolas	4	45	49
Ingeniería	7 ^b	71	44
Total	34	311	285

Nota. ^a Hubo una sección adicional de un curso que contestó la versión electrónica. ^b Una sólo sección de un curso participó; algunos estudiantes contestaron la versión en papel y otros contestaron la versión electrónica.

La evaluación de las características psicométricas de las puntuaciones obtenidas con el COE partió de varias hipótesis, que se fundamentaron en la premisa de que las puntuaciones válidas que se obtienen con un instrumento se manifiestan en forma predecible.

Hipótesis #1: Las puntuaciones totales del COE se distribuirán en forma similar en las versiones en papel y electrónica.

Mediante esta hipótesis se deseaba determinar si el formato del COE no afectaba la distribución de los resultados. Con este fin, se analizaron las estadísticas de distribución de frecuencias. Para determinar las puntuaciones totales, se excluyeron de todos los análisis los ítems 10, 11 y 12 por dos razones: (1) las respuestas a dichos ítems se midieron con una escala binaria (sí o no), lo que creaba grandes complicaciones para los análisis estadísticos y (2) el contenido de dichos ítems se centra en el cumplimiento de deberes del profesor/a y no en su desempeño en el curso. Además, se excluyó el ítem

38 (*Está disponible durante sus horas de oficina para atender las dudas o preguntas de los estudiantes*), pues una proporción alta de los estudiantes no lo contestó o indicó que no aplica. Por último, el ítem 37 se excluyó de los análisis debido a que, por error, no se incluyó en la versión electrónica (*Facilita el aprendizaje mediante el uso de métodos de enseñanza tales como ...*).

La puntuación total en el COE se definió como la media (promedio aritmético) de las respuestas a los ítems 13 al 36 (24 ítems). La Tabla 2 resume las estadísticas descriptivas de las puntuaciones totales obtenidas. Se encontró una diferencia estadísticamente significativa en las medias (promedios) de las puntuaciones de las dos versiones, esto es, una diferencia que no puede atribuirse a la casualidad ($t = 3.48$, $p < .01$). No obstante, la magnitud estandarizada de dicha diferencia ($d = .31$) se considera baja, de acuerdo con el criterio establecido por Cohen (1977). Por otro lado, no hubo diferencia estadísticamente significativa en la variabilidad de las puntuaciones obtenidas con las dos versiones. Es posible que la pequeña diferencia entre las medias se deba a la desigualdad de los grupos comparados, como consecuencia de la forma en que finalmente se constituyeron. La desigualdad no pudo corroborarse por la falta de datos sobre las características de los estudiantes que contestaron el COE. Se concluye que la hipótesis #1 se sostuvo sólo parcialmente.

Las puntuación media para cada versión fue alta (máximo posible era 5), hallazgo que puede deberse a dos posibles razones. Primero, los profesores participantes fueron voluntarios, lo que posiblemente contribuyó a que fueran profesores excelentes en su desempeño. Segundo, se ha observado una tendencia generalizada a la indulgencia por parte de los estudiantes al hacer los *ratings* de sus profesores (Hoyd y Lee, 2002).

Tabla 2: *Estadísticas descriptivas de las puntuaciones obtenidas con las dos versiones del COE*

Versión	Puntuación media	Desviación típica
Papel	4.48	.49
Electrónica	4.64	.54

Nota. La puntuación media es el promedio aritmético. La desviación típica o estándar es una medida de la variación entre las puntuaciones.

Hipótesis #2: *Las puntuaciones obtenidas con los ítems del COE se agruparán en varios conglomerados que representarán las dimensiones teóricas en que se fundamentó la construcción del instrumento.*

Esta hipótesis partió de la literatura relacionada con instrumentos para determinar las opiniones de los estudiantes sobre la labor de los profesores (por ejemplo: Arreola, 2000; Cashin, 1989; Marsh y Dunkin, 1992). En dicha literatura se documenta la multidimensionalidad de las puntuaciones obtenidas con dichos instrumentos en universidades de Estados Unidos. Para determinar la dimensionalidad del COE, se llevó a cabo un análisis factorial con las puntuaciones obtenidas con cada una de las versiones del instrumento. En dicho análisis, se encontró que en ambas versiones del COE las puntuaciones de los ítems se agruparon en un solo factor interpretable. Esto se observó tanto al utilizar el método factorial del eje principal como el método de los componentes principales. Las correlaciones de los ítems con dicho factor variaron desde .42 hasta .80 en la versión en papel y desde .55 hasta .81 en la versión electrónica. Aparentemente los estudiantes no consideraron dimensiones distintas al contestar el COE. Toda vez que la muestra utilizada en el estudio fue por disponibilidad, se debe analizar nuevamente la dimensionalidad de las puntuaciones del COE al administrarlo al universo de los estudiantes del Recinto. En resumen, la hipótesis #2 no se sostuvo. Por esta razón, todos los análisis posteriores se hicieron con las puntuaciones totales del COE, partiendo de la presunción de que las mismas eran unidimensionales.

Hipótesis #3: *Las puntuaciones totales del COE tendrán una correlación positiva y al menos moderada con la nota esperada por los estudiantes en el curso, la nota asignada por los estudiantes al profesor del curso y la disposición de los estudiantes para tomar otro curso con el/la profesor/a, de haber la oportunidad.*

Esta hipótesis parte de la premisa de que las puntuaciones totales obtenidas con el COE deben corresponder al aprovechamiento logrado por los estudiantes en el curso (relación entre el aprendizaje de los estudiantes y la labor del profesor/a), la nota asignada al profesor/a y la disposición de los estudiantes para tomar cursos con el profesor/a en el futuro, de haber la oportunidad. Para determinar la relación de estas variables con las puntuaciones del COE, se hizo un análisis de correlación y regresión, cuyos resultados se resumen en la Tabla 3.

Cada variable estudiada se correlacionó, por sí sola, con las puntuaciones del COE (r_{xy}). La correlación con la nota asignada al profesor fue alta en ambas versiones (.74 y .79 para las versiones en papel y electrónica, respectivamente); la correlación con la intención de tomar otro curso con el/la profesor/a fue moderada (.65 y .67 para las versiones en papel y electrónica, respectivamente); y la correlación con la nota esperada en el curso fue baja (.26 y .40 para las versiones en papel y electrónica, respectivamente). Las tres variables combinadas correlacionaron altamente con las puntuaciones del COE ($R = .75$ y $R = .81$ para las versiones en papel y electrónica, respectivamente). No sabemos si las diferencias observadas en las versiones se deban a posibles diferencias entre los estudiantes que conformaron los grupos comparados o al efecto del formato del instrumento. En general, los resultados del análisis de correlación y regresión sostuvieron la hipótesis #3.

Tabla 3: Resultados del análisis de correlación y regresión para la hipótesis # 3

Variable	Versión en papel		Versión electrónica	
	r_{xy}	β	r_{xy}	β
Nota esperada en el curso	.26	-.01	.40	.10
Nota asignada al profesor	.74	.58	.79	.62
Intención de tomar otro curso con el/la profesor/a	.65	.22	.67	.18
	R = .75, R ² = .57, p < .001		R = .81, R ² = .65, p < .001	

Nota. r_{xy} = correlación bivariada entre la variable correspondiente y las puntuaciones del COE; β = coeficiente estandarizado de regresión. Este coeficiente indica la importancia relativa de una variable, dentro de la combinación de la cuatro variables, al predecir las puntuaciones del COE. R = correlación entre la combinación de las tres variables con las puntuaciones del COE.

Hipótesis #4: *Las puntuaciones del COE tendrán una correlación insignificante con ciertas variables que no están bajo el control total del profesor tales como las percepciones de los estudiantes sobre su motivación para tomar el curso, sus hábitos de estudio, su esfuerzo en el curso y la dificultad del curso.*

Esta hipótesis parte de la premisa de que si las puntuaciones que se obtienen con el COE son válidas, aquellas variables que son ajenas a la labor del profesor/a y sobre las cuales el/ella no ejerce control no se relacionarán con dichas puntuaciones. Para someter a prueba la hipótesis #4, se efectuó un análisis de correlación y regresión cuyos resultados se resumen en la Tabla 4. Las correlaciones entre cada variable considerada y las puntuaciones del COE (r_{xy}) fueron insignificantes o bajas en ambas versiones. La combinación de las cuatro variables correlacionó en una magnitud baja con las puntuaciones del COE en la versión en papel (R = .30) y en una magnitud moderadamente baja en la versión electrónica (R = .43). En la versión en papel, sólo el 9% de la variación en las puntuaciones del COE se relaciona con las cuatro variables mientras que esa cifra fue de 18% en la versión electrónica. En la versión en papel, fue la dificultad percibida del curso la variable que más contribuyó a dicha relación; en la versión electrónica, fue la motivación para tomar el curso.

Tabla 4: *Resultados del análisis de correlación y regresión para la hipótesis 4*

Variable	Versión en papel		Versión electrónica	
	r_{xy}	β	r_{xy}	β
Motivación para el curso	.21	.18	.37	.37
Hábitos de estudio	.01	-.08	-.02	-.14
Dificultad del curso	-.16	-.24	-.15	-.22
Esfuerzo en el curso	.08	.18	.08	.14
	R = .30, R ² = .09, p < .001		R = .43, R ² = .18, p < .001	

Nota. r_{xy} = correlación bivariada entre la variable correspondiente y las puntuaciones del COE; β = coeficiente estandarizado de regresión. Este coeficiente indica la importancia relativa de una variable, dentro de la combinación de las cuatro variables, al predecir las puntuaciones del COE. R = correlación entre la combinación de las cuatro variables con las puntuaciones del COE.

Al introducir la nota esperada en el curso como variable control, se encontró que la combinación de las cuatro variables añadía un 5% en la explicación de la variación en las puntuaciones de la versión en papel y un 13% en la versión electrónica por encima de la variación explicada por la nota esperada. Cabe señalar que la asociación entre las puntuaciones del COE con la motivación para el curso, los hábitos de estudio, la dificultad del curso y el esfuerzo en el mismo depende tanto de la labor del profesor/a como de factores que no están bajo el control de éste/ésta. Por lo tanto, se debe hacer un análisis para determinar la parte de la variación que se puede explicar por la ejecución del profesor/a y desarrollar fórmulas para ajustar por la porción de la variación explicada por los factores ajenos al profesor/a, según el procedimiento sugerido por Hoyd y Lee (2002).

Hipótesis #5: Las puntuaciones del COE tendrán una confiabilidad alta.

La confiabilidad de las puntuaciones del COE se estimó mediante el coeficiente de correlación intraclase (CCI). Este es un estimado de confiabilidad recomendado cuando se llevan a cabo *ratings* sobre algo (Cashin, 1995; Shrout, 1998, Shrout & Fleiss, 1979). En este estudio, el CCI indica la consistencia o acuerdo entre los *ratings* dados a la labor del profesor/a por los estudiantes en una clase en particular. En la versión en papel, se encontró que los CCI variaron desde .58 hasta .95, con un CCI mediano de .86; el 71% de los CCI del centro de la distribución tuvieron valores desde .82 hasta .96. En la versión electrónica, los CCI variaron desde .46 hasta .99, con un CCI mediano de .88; el 71% de los CCI del centro de la distribución variaron desde .75 hasta .99. La mayoría de los CCI en la versión en papel (70%) y la versión electrónica (65%) fueron de .82 o

más. De acuerdo con la rúbrica presentada por Shrout (1998) esos valores son altos. Dos de los 17 CCI de la versión en papel y cinco de los 17 CCI de la versión electrónica fueron de .66 o menos. Casi todos estos valores fueron obtenidos con grupos muy pequeños de estudiantes (menos de 10). Cashin (1995) recomienda que estos valores deben interpretarse con cautela ya que el valor del CCI depende del número de personas que hace los *ratings*.

Referencias

- Arreola, R. A. (2000). *Developing a comprehensive faculty evaluation system*. Williston, VT: Anker Publishing.
- Cashin, W. E. (1989). *Defining and evaluating college teaching. IDEA Paper No. 21*. Manhattan, KS: Kansas State University, Center for Faculty Evaluation & Development.
- Cashin, W. E. (1995). *Student ratings of teaching: The research revisited. IDEA Paper No. 32*. Manhattan, KS: Kansas State University, Center for Faculty Evaluation and Development.
- Cohen, J. (1977). *Statistical power analysis for the behavioral sciences*. NY: Academic Press.
- Hoyd, D. P., & Lee, E. J. (2002). *Basic data for the Revised IDEA System. Technical Report No. 12*. The IDEA Center, Manhattan, KS.
- Marsh, H. W., & Dunkin, M. (1992). Students' evaluations of university teaching: A multidimensional perspective. *Higher Education: Handbook on theory and research* (Vol. 8, pp 143-234). NY: Agathon.
- Marsh, H. W., & Roche, L. A. (2000). Effects of grade leniency and low workload on students' evaluations of teaching: Popular myths, bias, validity or innocent bystanders? *Journal of Educational Psychology, 92*, 202-228.
- Shrout, P. E. (1998). Measurement reliability and agreement in psychiatry. *Statistical Methods in Medical Research, 7*, 301-317.
- Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin, 86*, 420-428.

III. Recomendaciones

1. Implantar el COE (versión 2, anejo 2) a partir del Primer Semestre del Año Académico 2006-2007. El estudio piloto demostró que su validez y confiabilidad es alta. La versión 2 incluye los cambios recomendados por miembros de este Senado Académico y aprobados por el CIEPD. Además, se recomiendan las siguientes modificaciones.

- a. Modificar la redacción de las siguientes preguntas:

COE-versión 1	COE-versión 2
28. Logra los objetivos establecidos en el prontuario del curso.	29. Contribuye a que los/as estudiantes logren los objetivos del curso.
30. Anuncia la fecha de exámenes parciales o de pruebas cortas con una semana o más de anticipación.	31. Anuncia la fecha de exámenes parciales con una semana o más de anticipación.

32. Redacta los exámenes o pruebas cortas con claridad.	33. Redacta con claridad los exámenes u otras actividades de evaluación (pruebas cortas, presentaciones, simulaciones, representación dramática, diarios reflexivos, representación de roles, otras).
---	---

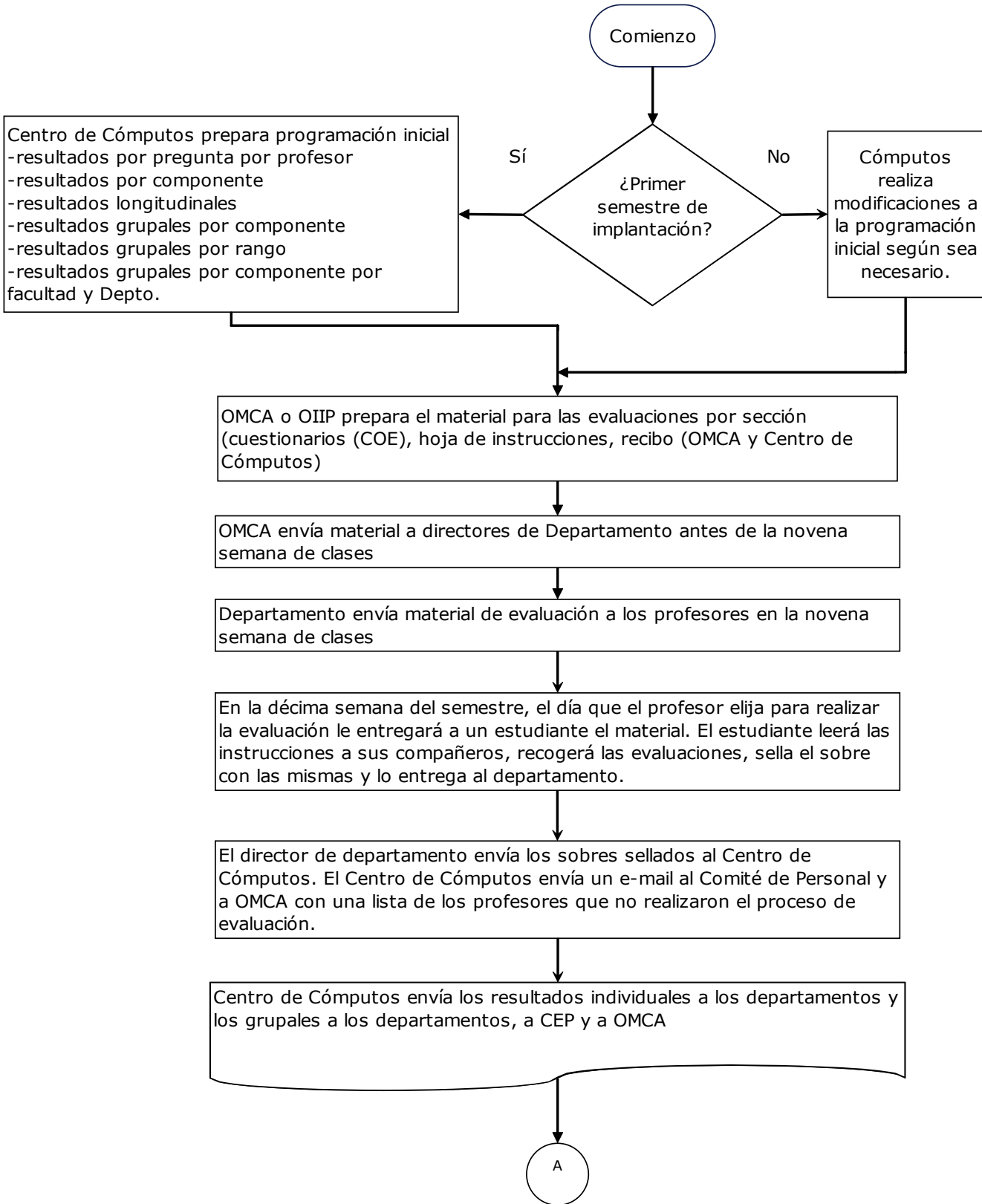
b. Añadir las siguientes preguntas:

¿Qué información aporta la pregunta añadida?	COE-versión 2
Conocer sobre la ejecutoria del profesor/a de acuerdo al nivel del programa académico.	1. Nivel del programa: subgraduado o graduado
Conocer sobre la disponibilidad del profesor/a para atender sus estudiantes.	39. Hace posible que los estudiantes se comuniquen con el/ella fuera de horas de clase.

2. Elevar a la Junta Administrativa (JA) del RUM la recomendación de sustituir el Módulo H "Módulo para la Evaluación de la Enseñanza por los Estudiantes" de la Certificación 86-87-476 por el Cuestionario de Opinión Estudiantil.
3. Implantar la versión en papel del COE manteniendo como indicadores de validez todas las preguntas en la parte de información del/la estudiante. A pesar de que las preguntas en torno a motivación, hábitos de estudio, haber trabajado más fuerte y dificultad del curso no tuvieron efecto significativo en las puntuaciones del COE se deben incluir, por lo menos el primer año de su implantación, para estudiar su validez por la totalidad de población del RUM. El estudio del COE estará a cargo de la Oficina de Investigación Institucional y Planificación del RUM.
4. Considerar la implantación del COE versión electrónica en años académicos subsiguientes al 2006-07 dado que el estudio piloto demostró su validez y confiabilidad. Además, una evaluación del COE electrónico por los estudiantes que usaron esta versión sugiere una opinión positiva sobre el mismo. Este estudio fue realizado por la Oficina de Investigación Institucional y Planificación a petición del CIEPD (anexo 3).
5. Utilizar la información obtenida en los ítems 11, 12, 13 y 40 del COE al evaluar el componente de Cumplimiento de Deberes en el Área de Docencia General según aparece en el documento "Propuesta para establecer un Nuevo Sistema de Evaluación para el Personal Docente del RUM" www.uprm.edu/evaluaciondocente. Considerar los demás ítems en el COE en la evaluación del Área de Enseñanza.
6. Estandarizar el procedimiento de la administración del COE. La estandarización del procedimiento involucra además de la uniformidad del cuestionario utilizado la consistencia en el procedimiento de administración del mismo. La estandarización del procedimiento de administración de una

medida de evaluación repercute en la confiabilidad y validez de la misma. Se recomienda el siguiente procedimiento de estandarización en la administración del COE en papel.

Proceso del COE en Papel



Proceso del COE en Papel (cont.)

