

Estudios para la Evaluación de las Propiedades de  
Medición del Módulo para la Evaluación de la  
Enseñanza por los Estudiantes:  
Pasado, Presente y Futuro

Recinto Universitario de Mayagüez

3 de marzo de 2004



## Objetivo General

- Evaluar las Propiedades de Medición del Instrumento de Evaluación de la Enseñanza por los y las Estudiantes
  - CIEQ (Course/Instructor Evaluation Questionnaire)
  - Lawrence M. Alemoni (1975)



## Descripción del Instrumento

Contiene 33 reactivos que se agrupan en siete sub-escalas

- Actitud (1,6,15,21)
- Método (2,7,12,18)
- Contenido (4,8,13,16)
- Interés (5,10,17,19)
- Instructor (3,9,11,14,20)
- Destrezas de Diseño (22-28)
- Administración Curso (29-33)

Se contesta utilizando una escala likert

- Totalmente Acuerdo (4); Parcialmente Acuerdo (3); Parcialmente Desacuerdo (2); Totalmente Desacuerdo(1)



## Estudio I (Fase 1)

Se llevó a cabo un estudio piloto en una muestra a conveniencia

- Evaluar la claridad de las instrucciones
- y de los reactivos del instrumento
- Evaluar el tiempo de administración del instrumento
- Ensayar el procedimiento de administración



## Estudio I (Fase I 2000)

### Resultados (Cualitativo)

- Varios reactivos se identificaron como confusos y el tiempo de los verbos causó dificultad para entender los mismos.
  - Se modificaron los reactivos #30 y #33 para aclarar la frase: tarea o actividad rutinaria y la palabra sistemática.
  - Se modificó el tiempo de los verbos en los reactivos (presente) ya que el curso está aún corriendo cuando se hace la evaluación.
- Muchos estudiantes no leyeron las las instrucciones.
  - Se recomendó leer en voz alta las instrucciones del instrumento



## Estudio I (Fase I)

### Resultados (Cuantitativo)

- Se encontró que un porcentaje alto de los/as estudiantes no contestó dos reactivos.
  - #33 (sistemática) y #39 (disponible hrs. oficina)
- Se encontró una confiabilidad interna de .93.
  - No se identificaron reactivos que disminuyan la confiabilidad interna del instrumento
- Se encontró tres reactivos con una correlación menor
  - 14 (material curso muy difícil); 30 (actividad rutinaria) y 39 (disponible hrs. oficina).



# Estudio I (Fase II - 2000)

- 1) Evaluar la confiabilidad y validez del instrumento utilizando una muestra representativa de las secciones del clase del RUM
  - Confiabilidad interna
    - grado de intercorrelación entre los reactivos
  - Validez grado en que el instrumento mide el constructo o concepto para el cual fue desarrollado.
    - construcción lógica

# Marco de Muestreo

- Secciones de clases de enero a mayo de 2000 (n=2,987).
- Estratificado por Facultad
- Selección aleatoria

- ADM	196	7%	70	24.4%
- ACiencias	1,537	51%	510	62.5%
- CAgrícolas	113	4%	40	5 %
- Ingeniería	621	21%	210	10%
- Div.Extensión	66	2%	20	2.6%
- Otros Prog.	34	1%	10	
- EGraduada	420	14%	140	

215-S

# Características de la Muestra

- Género
  - Varones - 54.4% Mujeres - 44.6%
- Año Estudio
  - Primero 22.1%; Segundo 22.5%; Cuarto ó más 35.9%
- Promedio
  - 4:00 a 3:00 (42.6%); 2.99 a 2.00 (53.7%)
- Cursos Evaluados
  - Requisito - 80.6%
- Tipo Calificación
  - Nota - 96.6%
- Calificación Esperada por el/a estudiante
  - A - 42.8%
  - B - 36.7%
  - No Contestó - 20.5%

# Evaluación General

	E	MB	B	R	D	MD
			%			
Contenido Curso	26	<b>36.3</b>	25.4	10.5	1.3	6
Profesor/a	<b>40.3</b>	27.8	19.7	9.2	2.0	.7
Curso General	22.6	<b>35.7</b>	27.1	11.9	1.7	.5

## Escala de Evaluación

E=Excelente; MB= Muy Bueno; B = Bueno; R= Regular

D=Deficiente; MD= Muy Deficiente

# Resultados: Confiabilidad Interna

<b>Dimensión</b>	<b>Cronbach Alfa</b>
• Escala Total (1-21)	.89
• Escala Total (1-33)	.93
• Actitud (1,6,15,21)	.80
• Método (2,7,12,18)	.04⇐
• Contenido (4,8,13,16)	.58
• Interés (5,10,17,19)	.77
• Instructor (3,9,11,14,20)	.68
• Destrezas de Diseño (22-28)	.85
• Administración Curso (29-33)	.76

No se identificaron reactivos que disminuyen dramáticamente la confiabilidad interna del instrumento

## Resultados: Validez/Análisis de Factores

Resultaron: 5 factores (55.26%)

- Factor 1 (16.98%)
  - 1,2,5,6,8,10,16,21,22,23,24
- Factor 2 (13.45%)
  - 4,5,7,9,12,13,15,17,18, 19,20
- Factor 3 (10.29%)
  - 14,28,29,30,31,32,33
- Factor 4 (9.41%)
  - 22,23,24,25,26,27
- Factor 5 (5.13%)
  - 3,4,5,11,21,28

# Resultados

- Reactivos que mayor # no contestó
  - 29 n=243/ Profesor está en horas oficina
  - 33 n=230/ Profesor explica sistema de notas
  - Ultimos reactivos
- 
- Reactivos/ correlación menor
  - 4 (material muy difícil); 20 (actividad rutinaria) y 29 (disponible hrs. oficina); 32 (trabajos devueltos prontitud).



## Conclusiones a partir de las Hipótesis de Trabajo

- Un índice de confiabilidad interna entre moderadamente alto y alto. (Cronbach Alfa)
- Un índice de confiabilidad interna entre moderadamente alto y alto para las sub-escalas (Cronbach Alfa).
- Los reactivos del instrumento se agrupan en factores que responden al modelo teórico del cual parte el instrumento (Análisis Factores).



# Recomendaciones Generales

- Revisar reactivos correlación menor con otros
- Revisar índices bajos de CI en algunas SE
- Verificar proceso codificación (invierte escala de respuesta de algunos reactivos)
- Tomar decisiones
  - evaluación por reactivo; evaluación por dimensión (unas necesitan revisarse ej. método)
  - puntuación total no debe incluir reactivo general (duplicidad)
  - Hoja Electrónica cónsona a formato cuestionario



## Estudio II (2002)

- Se llevaron a cabo una serie de estudios con el objetivo de comparar las propiedades de medición del instrumento en desarrollo con el Módulo H
  - Confiabilidad interna
  - Validez (construcción lógica)
    - Análisis en las evaluaciones de secciones de curso provistas por el Centro de Cómputos (año 2002)



# Resultados Estudio II

- Administración de Empresas
- Artes y Ciencias
- Ciencias Agrícolas



## Conclusiones y Recomendaciones

- Ambos instrumentos tienen fortalezas y debilidades.
- Debe utilizarse el mismo instrumento para la evaluación de todos/as los y las profesores/as
- Debe haber consistencia en el Proceso de administración y corrección del instrumento
  - Protocolo de administración
  - Elaborarse criterio de corrección y puntos de corte del instrumento (puntuación total y puntuación por sub-escalas)



## Resultados/Recomendaciones

- Ambos instrumentos tienen fortalezas y debilidades

### **Recomendaciones al Nuevo Instrumento**

- Evaluar la interpretación de la puntuación del instrumento a partir de la escala (Acuerdo).
- Modificar reactivos que requieran la alternativa de NO APLICA, en caso de que todos los reactivos midan conductas esperadas.
- Revisar los reactivos con doble dimensión o constructo.
- Eliminar o disminuir repetición en los reactivos.
- Revisar que los reactivos sean apropiados para la escala de respuesta.
- Revisar todos los reactivos, para uso de acción de personal, en función de la evaluación de la ejecución



# Resultados/Recomendaciones

## Recomendaciones al Módulo H

- Usar escala que indique ocurrencia (frecuencia).
- Usar escala donde el cero indique “ausencia/nunca” de ocurrencia de lo que se evalúa
- Eliminar la alternativa No Aplica en caso de que los reactivos midan conductas esperadas.
- Revisar los reactivos con doble dimensión (constructos) (2,7,18,19 y 30, entre otros).
- Eliminar o disminuir repetición en los reactivos
- Revisar que los reactivos sean apropiados para una escala ordinal de respuesta.
- Revisar los reactivos en su redacción para que aluda a una conducta observable (escala ocurrencia)
- Verificar la consistencia entre el Perfil del Profesor/a y los contenidos en los reactivos (Tecnología en el salón de clase y La investigación en el salón de clase).



# Recomendación General

- Uniformidad en el Instrumento Utilizado
- Consistencia en el Proceso de Administración
  - Protocolo de administración
- Elaborar criterio de corrección del instrumento
- Determinar una puntuación total y una puntuación por sub-escalas



## Objetivos y Resultados del Estudio III

### Estudio Piloto en muestras a conveniencia

#### 1) Evaluar la claridad de las instrucciones del instrumento (n=25).

- Se utilizó una hoja que recopila esta información.

### Resultados

- Instrucciones no indican ¿cómo constestar el instrumento? Las nuevas instrucciones indican ¿qué y cómo?
- Preocupa el aspecto de confidencialidad mencionado en las instrucciones. Se añadió el período en que se entregan las evaluaciones a los/as profesores/as.



## Continuación...

- En las instrucciones no se motiva al estudiante para que participe seriamente en la evaluación. Se incluye un párrafo sobre la importancia e implicaciones del proceso de evaluación.

En resumen se desarrolló el documento **Procedimiento** para la Administración del Cuestionario de Evaluación de la Enseñanza por los/as Estudiantes. El documento atiende: las instrucciones al instrumento, pertinencia del proceso de evaluación y el aspecto de la confidencialidad de las respuestas, entre otros aspectos.



## Objetivos y Resultados del Estudio III

- 2) Evaluar la claridad de los reactivos del instrumento (muestra estudiantil  $n=27$  y expertos en aspectos de medición  $n=4$ ).
- Se usó una hoja que recopila esta información.

### Resultados

- Todos los reactivos comienzan con una acción (más que menos observable que se puede traducir en frecuencia para que corresponda a la escala)
- Se modificaron reactivos; Se eliminaron reactivos
- Se añadieron reactivos (Preg. abiertas ¿tiempo?); Se creó un patrón de salto
- Se desarrolló una instrucción específica
- Se encontró que un porcentaje alto de los/as estudiantes no contestó el reactivo: El profesor está disponible en las horas de oficina



## Objetivos y Resultados del Estudio III

3) Evaluar el tiempo de administración del instrumento (n=24).

### Resultados

Tomó entre siete y 20 minutos completar el cuestionario. Se solicitó al estudiante indicar en el reverso de la hoja de respuesta hora de comienzo y hora en que terminó de contestar. La variabilidad en tiempo de respuesta estuvo, entre otros aspectos, mediada porque **No** todos los estudiantes comenzaron al mismo tiempo. Se realizó la evaluación al principio de la clase y algunos estudiantes llegaron tarde.

- Sesgo a sociado a la clase del día
- Rapidez por salir



## Objetivos y Resultados del Estudio III

- 4) Ensayar el procedimiento de administración (n=17)
- Logística y recursos en la implantación
  - Se utilizó un estudiante del curso para la administración incorporando al procedimiento elementos de control de calidad del proceso.
  - El estudiante que administró, así como sus compañeros, señaló que las directrices verbales no son adecuadas. Proponen el desarrollo de instrucciones para la persona que administra el cuestionario.

- Resultados

- Se incorporó al documento de procedimiento las instrucciones a la persona que administra. Esto atiende a su vez la uniformidad del proceso.



## Objetivos y Resultados del Estudio III

5) Evaluar la confiabilidad interna y la validez del instrumento (n=30)

### Resultados en torno a la confiabilidad

- Se encontró una confiabilidad interna (grado de intercorrelación entre los reactivos) de .91
  - No se identificaron reactivos que disminuyan la confiabilidad interna del instrumento
- Se encontró que los siguientes reactivos correlacionan, en promedio, menos con el resto de ellos:
  - Entrega el prontuario durante la primera semana y disponible hrs. oficina ¿naturaleza de la pregunta, escala, tamaño de la muestra?



## Objetivos y Resultados del Estudio III

### Resultados en torno a la validez

(construcción lógica). Es el grado en que el instrumento mide el constructo o concepto para el cual fue desarrollado.

- Los reactivos del instrumento se agrupan en factores que responden al modelo teórico del cual parte el instrumento (Análisis Factores)
- Los reactivos fueron asignados a las categorías del modelo del cual parten (Expertos n=4).

# ¿Los Reactivos se Agruparon en las Categorías Identificadas?

## Calidad de la Enseñanza

- Dominio de la materia
- Métodos y técnicas de enseñanza
- Métodos y técnicas de evaluación
- Habilidad para comunicarse

## Componentes de la Enseñanza

- Dominio de la materia
- Destrezas de presentación
- Diseño del curso
- Administración del curso

**Mayor acuerdo**  
sub-escalas

Se recomienda una definición de las categorías (no a partir de los mismos reactivos)

# Escala de Respuesta del Instrumento

## 1) Asignación de Valor: Reactivos 1 al 31

Siempre	( A )	4
Casi Siempre	( B )	3
A veces	( C )	2
Casi Nunca	( D )	1
Nunca	( E )	0
No Aplica	( F )	Missing

(No se invierte la escala de ningún reactivo)

## 2) Asignación de Valor: Reactivos 32 al 35

Sí (A) = 1 No (B) = 0

## 3) Otros reactivos

36 y 37 (0 a 4 puntos)

38 al 40 preguntas abiertas

¿Punto de Corte?

Descripción del  
Instrumento

¿Pregunta  
abiertas info.  
Profesor/a?



## El instrumento en revisión...

- Evaluación por reactivo
- Evaluación por dimensión (sub-escalas)
- Evaluación para la puntuación total
  - ¿no debe incluir reactivo general?
- Provee información para para acción de personal, además de otros propósitos
- Consistencia en el Proceso de Administración
  - Protocolo de administración



## Trabajamos en...

- Evaluar la interpretación de la puntuación del instrumento a partir de la escala ¿Escala?
  - Elaborar un criterio de evaluación de la enseñanza a partir de aspectos teóricos (modelo de Arreola)
- Revisar que los reactivos son apropiados para la escala de respuesta (dos escalas).
- El uso de la alternativa No Aplica
- El uso de patrones de salto
- El uso de lenguaje inclusivo (el/la)
- Perfil de Profesor (tecnología en el salón de clase se incluyó pero la investigación en el salón de clase no)
- Descripción del instrumento
- Instrucciones al/la profesor/a que enseña el curso
- Instrucciones al personal del departamento



## Trabajamos en...


- Informar a la facultad en torno a los estudios realizados y solicitar retroalimentación

Modulo de Evaluación de la Enseñanza:  
Pasado, Presente y Futurto



## Aspectos a Considerar ...

- Análisis e interpretación de los resultados del proceso de evaluación
  - **Tendencias**
- Elaborar un criterio de evaluación de la enseñanza a partir de aspectos teóricos (modelo de Arreola)
- Propósito del instrumento
  - Evaluación Sumativa
  - Evaluación Formativa
    - Modelo de Arreola “A comprehensive Faculty Evaluation System”

- 
- EP : Exemplary Professional Performance
    - Consistently exceeds accepted standards of professional performance (JUSTIFICATION MUST BE INCLUDED)
  - HP : High Professional Performance
    - Frequently exceeds accepted standards of professional performance
  - SP : Standard Professional Performance
    - Consistently meets accepted standards of professional performance
  - MP : Minimal Performance
    - Does not consistently meet accepted standards of professional performance (JUSTIFICATION MUST BE INCLUDED)
  - UP : Unsatisfactory Performance
    - Does not meet minimal standards of professional performance (JUSTIFICATION MUST BE INCLUDED)