

Universidad de Costa Rica.

Vicerretería de Investigación

Instituto de Investigación En Educación

**FACTORES DE SESGO ASOCIADOS A LA
EVALUACIÓN DOCENTE UNIVERSITARIA**

Nº 724-A9-118

Raziel Acevedo Álvarez

Año 2009-2010

INDICE

INFORME FINAL DEL PROYECTO DE INVESTIGACIÓN.....	3
1.- Datos generales del proyecto.....	3
1.1. Código del proyecto:	3
1.2. Título del proyecto:.....	3
1.3. Unidad de pertenencia:	3
1.4. Unidad de adscripción.....	3
1.5. Programa al que pertenece.....	3
1.6. Nombre del investigador y carga académica.....	3
1.7. Duración del proyecto.....	3
1.8. Objetivos del proyecto.....	3
11.8. Objetivo General.....	4
Determinar el grado asociación entre los resultados aportados por los cuestionarios de evaluación docente y las variables del estudiante, profesor y la clase como: género del estudiante, del profesor, tamaño de la clase, edad del estudiante, del profesor, entre otros, para observar el grado de validez de estos instrumentos.....	4
1.8.2. Objetivos específicos	4
2.1. ACTIVIDADES Y PORCENTAJE DE LOGRO DESARROLADAS DURANTE LA INVESTIGACIÓN	4
2.1. Fase 1.....	5
2.2. Fase 2.....	8
2.3. Fase 3 y 4.....	11
2.4. Fase 5.....	13
2.2. LOS MÉTODOS O LOS PROCEDIMIENTOS UTILIZADOS. LAS DIFICULTADES Y LA MANERA DE CONFRONTARLA.....	13
3. RESULTADOS FINALES	13
4. REFERENCIAS.....	15

INFORME FINAL DEL PROYECTO DE INVESTIGACIÓN

1.- Datos generales del proyecto

1.1. Código del proyecto:

N° 720-A7-054

1.2. Título del proyecto:

**FACTORES DE SESGO ASOCIADOS A LA EVALUACIÓN DOCENTE
UNIVERSITARIA**

1.3. Unidad de pertenencia:

Sede de Guanacaste

1.4. Unidad de adscripción

Instituto de Investigaciones Educativas

1.5. Programa al que pertenece

Programa de Investigación en el cambio, desarrollo y gestión de la Educación Superior No. 724-A5-900

1.6. Nombre del investigador y carga académica

Dr. Raziél Acevedo Álvarez T.C.; T.C. con un $\frac{1}{4}$ de tiempo adicional para el desarrollo del proyecto.

M.Sc. Lorena Kikut, Colaboradora C.E.A.

1.7. Duración del proyecto

01-01-2009 hasta 31-12-2010.

1.8. Objetivos del proyecto

11.8. Objetivo General

Determinar el grado asociación entre los resultados aportados por los cuestionarios de evaluación docente y las variables del estudiante, profesor y la clase como: género del estudiante, del profesor, tamaño de la clase, edad del estudiante, del profesor, entre otros, para observar el grado de validez de estos instrumentos.

1.8.2. Objetivos específicos

Seleccionar las variables que identifican a los estudiantes, profesores y cursos, para construir una base de datos que reúna información relevante sobre el tema en estudio, de manera que permita su caracterización.

Elaborar un Modelo Jerárquico Lineal con los datos suministrados por el CEA, para proponer una estructura con dos.

Determinar el grado asociación entre los resultados aportados por los cuestionarios de evaluación docente y las variables del estudiante, profesor y la clase como: género del estudiante, del profesor, tamaño de la clase, edad del estudiante, del profesor, entre otros, para observar el grado de validez de estos instrumentos.

2.1. ACTIVIDADES Y PORCENTAJE DE LOGRO DESARROLLADAS DURANTE LA INVESTIGACIÓN

Una investigación de esta naturaleza dependía enteramente de dos elementos:

- 1) Interactuar con los investigadores del C.E.A. para conseguir una base de datos reales, sobre algún instrumento reciente aplicado a los profesores de nuestra institución. Este paso se logro en un 100% pues los encargados del C.E.A. gustosamente y ágilmente facilitaron una gran cantidad de información de aplicaciones recientes en las diferentes facultades de la U.C.R. sobre el nuevo instrumento de evaluación del docente universitario.
- 2) Solicitar al Centro de Informática el paquete estadístico MIWin para analizar las bases de datos por medio de los modelos jerárquicos lineales. Tal actividad no se cumplió con ellos, pero los investigadores de la Facultad de Educación, de la Universidad compútense de Madrid, nos facilitaron las computadoras para poder concluir este proyecto, con el

paquete estadístico que necesitábamos, además, pudimos interactuar con otros investigadores sobre las estimaciones de otros proyectos similares al nuestro.

Habiendo conseguido en principio, uno de los dos componentes, nos dimos a la tarea de realizar las siguientes actividades:

Para hacer tangibles los dos objetivos específicos se plasmaron las siguientes metas:

Caracterizar 10 variables del estudiante, 10 del profesor y 10 de la clase para iniciar la formación del modelo nulo.

Una base de datos numérica y otra bibliográfica que brinde contenido teórico y soporte cuantitativo a la investigación.

Un modelo jerárquico en dos niveles.

Con esas metas propuestas se atendieron las fases 1 y 2 de la investigación, anotadas en el cronograma de trabajo, que integraba las siguientes actividades:

2.1. Fase 1

a.1. Revisión de la bibliografía, para determinar los factores más importantes que afectan estas evaluaciones. En la revisión, fueron base los siguientes artículos:

Abrami, P.C.; D'apollonia, S. & Cohen, P. "Validity of Student Ratings of Instruction: What We Know and What We Do Not". En *Journal of Educational Psychology*, 82 (2), pp.219-231. 1990.

Abrami, P.S.; D'apollonia, S. & Rosenfield. "The Dimensionality Of Student Ratings of Instruction: What We Know and What We do Not". En R.P. PRRY & J.C. SMART (Eds.) *Effective Teaching in Higher Education: Research and Practice*, pp.321-367. New York: Agathon Press. 1997.

Aleamori, L.M. "Student Rating myths versus research facts from 1924 to 1998". En *Journal of Personnel Evaluation in Education*, 13 (2), pp. 153-166.1999.

Álvarez, V.; García, E. Y Flores, J. "Características de la Docencia Mejor Evaluada por los Alumnos En Las Diferentes Áreas de Enseñanza Universitaria". En *Revista Española de Pedagogía*, N° 214, pp. 445-464. 1999.

Bentler, P.M. "Multivariate Analysis with Latent Variables: Causal Modeling". En *Annual Review of Psychology*, 31, pp. 419-456. 1990.

Cajide, J. "Análisis Factorial Confirmatorio de las Características de Calidad Docente Universitario (solución LISREL)". En *Bordón*, 43 (4), pp. 389-405. 1994.

Cohen, R.J. y Swerdlik, M.E. *Psychological Testing And Assessment: An Introduction To Tests And Measurement* (4th ed.). Mountain View, CA.: Mayfield. 1999.

De Miguel, M. "La Evaluación Del Profesorado Universitario. Criterios y Propuestas Para Mejorar La Función Docente". En *Revista de Educación*, N° 315, pp. 67-83. 1998.

Escudero, T. "Los Estudiantes como Evaluadores de la Docencia y de los Profesores: Nuestra Experiencia". En *Revista Interuniversitaria De Formación Del Profesorado*. N° 34, Enero/ Abril. pp. 69-86. 1999.

Feldman, K.A. "Research Productivity and Scholarly Accomplishment of College Teacher as Related to Their Instructional Effectiveness: A Review and Exploration". En *Research in Higher Education*, 26 (3), pp. 227-298. 1987.

Feldman, K.A. "Identifying Exemplary Teachers and Teaching: Evidence from Student Ratings". En R.P. PERRY & J.C. SMART (eds.), *Effective teaching in Higher Education*:

Fernández, J.; Mateo, M. & Muñiz, J. "Is There Relationship Between Class Size and Student Ratings of Teacher Quality?" En *Educational And Psychological Measurement*, 58 (August), 596-604. 1998.

García Ramos, J.M. "El Análisis de Estructuras de Covarianza en El Estudio De La Competencia Docente Del Profesor Universitario". En *Revista De Investigación Educativa*, 16 (1), Madrid. pp. 155-184. 1998.

Green, Sbt.; Akey, T.M.; Fleming, K.K.; Hershberger, S.C. Y Marquis, J.G. "Effects of the number of scale points on chi-square fit indices in confirmatory factor analysis". En *Structural Equation Modeling*, 4 (2), pp. 108-120. 1997.

Hair, J.; Anderson, R.; Tatham, Black, W. *Análisis Multivariante*. España: Prentice Hall. 1999.

Kwan, K.P. "How Fair Are Rating in Assessings The Teaching Performance Of University Teacher?". En *Assessment & Evaluation In Higher Education*, 24 (June), 181-195.1999.

Marsh, H. & Bailey, M. "Multidimensional Students' Evaluations of Teaching Effectiveness". En *Journal of Higher Education*, 64 (1), pp. 1-18. 1993.

Marsh, H. & Roche, L. R. "Making Students' Evaluation of Teaching Effectiveness Effective: The Critical Issues of Validity, Bias, and Utility". En *American Psychologist*, Vol. 52, 11, 1187-1197. 1997.

Marsh, H. W., & Dunkin, M. J. "Students' evaluations of university teaching: A multidimensional perspective". En J.C. SMART (ed.) *Higher Education: Handbook of Theory and Research*, (Vol. 8, pp. 143-233). New York: Agathon Press. 1992.

Marsh, H.W. y Roche, L.A. "Effects of grading leniency and low workload on students' evaluation of teaching: popular myth, bias, validity or innocent bystanders?". En *Journal of Educational Psychology*, 92 (1), pp. 202-228. 2000.

Murray, H.G. "Does Evaluation of Teaching Lead to Improvement of Teaching?". En *International Journal for Academic Development*, 2 (1), pp. 20-41. 1997.

Ramsden, P. *Learning of Teach in Higher Education*. London: Routledge. 1992.

Tejedor, F. y Montero, L. "Indicadores de la Calidad Docente para La Evaluación del Profesor Universitario". En *Revista Española De Pedagogía*, año XLVIII, N° 186, mayo-agosto, pp. 259-279. 1990.

Ting, K.F. "A Multilevel Perspective On Student Ratings of Instruction: Lessons From the Chinese Experience". En *Research in Higher Education*. Vol. 41, 5, pp.637-653. 2001.

Young, P.; Delli, D. y Johnson, L. "Student Evaluation of faculty: Effects of Purpose on Pattern". En *Journal of Personnel Evaluation in Education*, 13 (2), pp. 179-190. 1999.

a.2. Elaboración de una base de datos escrita sobre los autores más destacados.

Se construyó una base de citas bibliográficas con el programa FILE MAKER, para hacer accesible al información encontrada y tenerla a mano para desarrollar el marco teórico del documento. La base integra diferentes componentes como: nombre del autor, bibliografía, indicadores de la lectura escogida, tema principal, tema secundario, resumen y página de dónde se tomó la lectura. Esto permite un orden en la lectura y utilización de las fuentes documentales.

a.3. Coordinar con el CEA y otras unidades académicas para obtener sus bases de datos.

El C.E.A. aplicó durante el 2009 su instrumento de evaluación docente, a aquellas unidades académicas que lo solicitaron. Con estos datos se integró una base que conforma una estructura jerárquica.

Con los datos de esta base se conocen las variables anotadas en el objetivo uno, como son: edad del estudiante, del profesor, tipo de curso, unidad de pertenencia, género, años de estudio, años de servicio, unidad de pertenencia, carrera, rango en régimen académico, jornada laboral, curso obligatorio u optativo, cantidad de estudiantes, horas de trabajo en clase. Todas estas variables identifican a los estudiantes, profesores y cursos

Con los datos a mano, organizamos e integramos cada una de las bases de datos facilitadas por el C.E.A., para ajustar una estructura propia de los modelos jerárquicos lineales, la cual es requiere de una organización rigurosa, que implica: eliminar los datos en blanco o respuestas sin contestar, ajustar todos los valores de cada una de las variables utilizadas a cero. Este trabajo minucioso le dedicamos una gran cantidad de horas, aún así, una vez que trabajamos en ella, fue necesario depurarla nuevamente. El trabajo de organización se logró en un 100% y con ello se cumple el objetivo número uno y las metas, de seleccionar las variables del estudiante, profesor y clase. En el trabajo del primer objetivo, fue indispensable el trabajo, apoyo y experiencia, de la señora M.Sc. Lorena Kikut, del C.E.A., quién también, dedicó muchas horas a la organización de la base de datos, que sería utilizada en la investigación.

2.2. Fase 2

De esta fase se ordena la base para el modelado jerárquico lineal. El ordenamiento de la base integra información del estudiante, el profesor y la clase, la cual es no está integrada encuentra en diferentes

Ahora bien, para cumplir totalmente con el objetivo número dos, haremos un modelos con una estructura en dos niveles, el primero los estudiantes y en el segundo lo compone el profesor y la clase. El modelo se puede representar por medio de dos ecuaciones matemáticas.

$$Y_{ij} = \beta_{0j} + e_{ij}$$
$$\beta_{0j} = \beta_{00} + \mu_{0j}$$

La ecuación de primer nivel es:

$$Y_{ij} = \beta_{0j} + e_{ij}$$

Simboliza:

Y_{ij} es la variable dependiente observada en el nivel inferior o micro nivel, o sea, la unidad mínima definida en la investigación como la evaluación del estudiante, i (determinada por el instrumento del CEA) agrupado por una clase j . Notemos que desde el inicio la representación de los dos niveles; i para el estudiante y j la escuela de pertenencia.

β_{0j} corresponde al intercepto, el punto de partida de la recta en el plano. Para nuestra investigación, representa la media general de la evaluación del profesor dentro de la clase j . El valor del rendimiento medio de todas las clases de una muestra determinada puede ser un número negativo o positivo.

e_{ij} constituye el residuo o varianza residual, cuya media es cero y tiene una varianza σ_e^2 . Representa las diferencias en la evaluación de los estudiantes, entre el valor estimado por la regresión y el valor real. En su modelo más simple asume la varianza de error al azar, es la misma para todos los estudiantes.

Como vemos la ecuación hace referencia directa a los componentes descritos en el nivel uno de la ecuación de regresión, o sea, los estudiantes.

Analizada la ecuación de nivel uno, ingresamos a la de nivel dos:

$$\beta_{0j} = \beta_{00} + \mu_{0j}$$

Es la ecuación de nivel macro referente a la información o características de las clases y profesores.

β_{00} representa el intercepto del nivel dos, la gran media general de todas las clases j incluidas en el estudio.

μ_{0j} es el residuo o varianza residual del nivel dos, la desviación del valor estimado para la clase de su valor real. Tiene media cero y varianza $\sigma_{\mu_0}^2$. Cuando el puntaje de la desviación es alto indica que existe una marcada diferencia entre las clases.

Como observamos, el segundo nivel hace referencia directa a las clases y ofrece estimaciones de su media, con el respectivo residuo.

La información contenida en las ecuaciones, se condensa en una sola integrando los elementos de ambas, de esa forma se obtiene, una representación para las dos ecuaciones. La misma quedaría elaborada de la siguiente forma:

$$Y_{ij} = \beta_{00} + u_{0j} + e_{ij}$$

Donde:

Y_{ij} corresponde a la variable dependiente. El rendimiento del sujeto i en la escuela j .

β_{00} identifica la gran media general del rendimiento entre las clases.

u_{0j} es el residuo o varianza residual de segundo nivel (la clase) y representa la desviación entre la media entre las clases; las diferencias de estimación entre la predicción y el valor real.

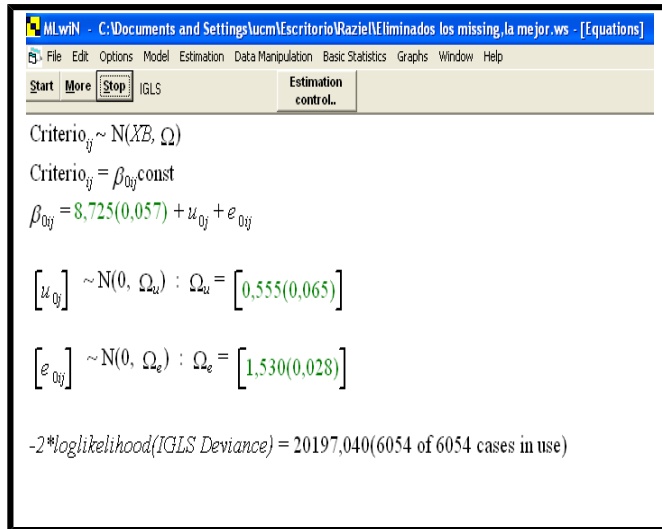
e_{ij} hace referencia a la varianza residual del primer nivel: los estudiantes.

El primero de los coeficientes el β_{00} es llamado la parte fija del modelo. Los otros dos coeficientes, u_{0j} y e_{ij} representan las varianzas residuales denominada la parte aleatoria del modelo. Como se deduce, la ecuación de regresión del modelo nulo ofrece información sobre dos elementos, la parte fija y la aleatoria en la estructura jerárquica de los datos.

En la parte aleatoria del modelo sólo están las varianzas de los residuales. Los residuales no son parámetros del modelo, pero sus varianzas sí. Los residuales son las diferencias entre el valor que toma la variable dependiente para una unidad y lo que el modelo predice para esa misma unidad (sujeto o escuela). Hay tantos residuales de nivel uno como sujetos y tantos residuales de nivel dos como clases, pero esos no son parámetros, los parámetros son sus varianzas.

Con la base de datos organizada, se procedió a desarrollar el segundo objetivo Elaborar un Modelo Jerárquico Lineal con los datos suministrados por el CEA, para proponer una estructura con dos niveles. El mismo quedo de la siguiente manera:

Modelo Nulo



El modelo se construye en dos niveles, y ofrece la suficiente varianza para que se desarrollen modelos más complejos, dado que $8.725/0.057= 153.07$ lo cual es un valor mayor a 2. Este número es considerado el mínimo para que exista la suficiente varianza entre los datos. Observando el segundo valor o sea $0.555/0.065= 8.53$, ello muestra claramente un valor mayor a 2. Finalmente, se divide el tercer dato o sea el primer nivel, $1.530 / 0.028 = 54.64$. El resultado muestra un valor mayor a 2. Todos los resultados de las divisiones entre los parámetros son superiores a dos, lo cual indica valores de la varianza muy significativos, ello demuestra que el modelo nulo es significativo y se puede proceder a desarrollar modelos más complejos, por lo tanto, con estos datos se cumple en 100% el objetivo dos.

2.3. Fase 3 y 4

Para este apartado se considera incluir las variables del profesor y clase. No obstante, ninguna de ellas alcanza los niveles de significación relevantes como para ser incluidas en el modelo, lo que indica que el instrumento no se vé afectado por estas variables.

Una vez excluidas del modelo estas variables, se procede a incluir aquellas variables del estudiante que fueron significativas con los paquetes tradicionales, como género, edad y situación laboral.

M.L.N. - C:\Documents and Settings\lucm\Escritorio\Waziel\Eliminados los missing, la mejor.ws

File Edit Options Model Estimation Data Manipulation Basic Statistics Graphs Window Help

Start More Stop IGLS Estimation control.

Equations

$$\text{Criterio}_{ij} \sim N(\mathcal{X}\beta, \Omega)$$

$$\text{Criterio}_{ij} = \beta_{0ij}\text{const} + 0,063(0,034)\text{sex-est}_{ij} + 0,013(0,006)\text{edad-est}_{ij} + 0,002(0,030)\text{sit-labor}_{ij}$$

$$\beta_{0ij} = 8,348(0,138) + u_{0ij} + e_{0ij}$$

$$\begin{bmatrix} u_{0ij} \end{bmatrix} \sim N(0, \Omega_u) : \Omega_u = \begin{bmatrix} 0,541(0,063) \end{bmatrix}$$

$$\begin{bmatrix} e_{0ij} \end{bmatrix} \sim N(0, \Omega_e) : \Omega_e = \begin{bmatrix} 1,529(0,028) \end{bmatrix}$$

$-2*\log\text{likelihood(IGLS Deviance)} = 20190,680(6054 \text{ of } 6054 \text{ cases in use})$

De todas las variables incluidas sexo, edad, situación laboral, la única que es significativa es la edad del estudiante. Por tanto se eliminan las variables y solamente se introduce esta en el modelo.

Introducción de la Variable edad del estudiante

M.L.N. - C:\Documents and Settings\lucm\Escritorio\Waziel\Eliminados los missing, la mejor.ws

File Edit Options Model Estimation Data Manipulation Basic Statistics Graphs Window Help

Start More Stop IGLS Estimation control.

Equations

$$\text{Criterio}_{ij} \sim N(\mathcal{X}\beta, \Omega)$$

$$\text{Criterio}_{ij} = \beta_{0ij}\text{const} + \beta_{1j}\text{edad-est}_{ij}$$

$$\beta_{0ij} = 8,404(0,153) + u_{0ij} + e_{0ij}$$

$$\beta_{1j} = 0,015(0,007) + u_{1j}$$

$$\begin{bmatrix} u_{0ij} \\ u_{1j} \end{bmatrix} \sim N(0, \Omega_u) : \Omega_u = \begin{bmatrix} 1,229(0,377) \\ -0,031(0,015) \quad 0,001(0,001) \end{bmatrix}$$

$$\begin{bmatrix} e_{0ij} \end{bmatrix} \sim N(0, \Omega_e) : \Omega_e = \begin{bmatrix} 1,518(0,028) \end{bmatrix}$$

$-2*\log\text{likelihood(IGLS Deviance)} = 20184,350(6054 \text{ of } 6054 \text{ cases in use})$

El modelo indica que la variable edad del estudiante es significativa, pero con un valor muy pequeño, que señala que la edad de los estudiantes afecta cuando el estudiante tiene alrededor de 18 años. Aunque el valor es 0.015 y la covarianza -0.031 expresan que esta asociación actúa como estabilizador, dado que cuando los estudiantes menores valoran muy alto a sus profesores, los estudiantes de mayor edad lo hacen un poco bajo y viceversa, cuando los estudiantes menores bajan la calificación del profesor, los estudiantes de mayor

edad incrementan su valoración del profesor. No obstante el incremento de 0.015 es tan pequeño, puede no ser tomado en cuenta.

2.4. Fase 5

Elaborados y analizados los modelos jerárquicos lineales y aplicados al instrumento de evaluación de la docencia universitaria en la UCR, se puede expresar que la robustez del instrumento utilizado por el CEA es altamente fiable y válido, lo cual demuestra que no se ve afectado por factores de sesgo y contexto de la docencia universitaria, por tanto, con las fases 3, 4 y 5 se cumple al 100% el objetivo número tres.

2.2. LOS MÉTODOS O LOS PROCEDIMIENTOS UTILIZADOS. LAS DIFICULTADES Y LA MANERA DE CONFRONTARLA

Como método de trabajo se fundamentó en la lectura de las bases de datos bibliográficas sobre el tema. Inicialmente hicimos búsqueda dentro del país y encontramos muy poca, por tanto, ampliamos nuestro interés al ámbito internacional, donde logramos muchísimos documentos pero en idioma inglés, que no es problema para nosotros.

Posteriormente reconstruyó una base de datos bibliográficos con FILE MAKER y con los datos empíricos aportados por las aplicaciones reunidas por el C.E.A. se analizaron utilizando la metodología de modelos jerárquicos lineales, los cuales fueron facilitados por la Universidad Complutense de Madrid.

Respecto a las dificultades encontradas en el desarrollo de la investigación, únicamente podemos mencionar una: la falta de paquetes informáticos de última generación para el análisis de datos, es uno de los problemas más difíciles, pues hemos estado expresando lo mismo durante varios años. Este hecho es lamentable en una Universidad dedicada al desarrollo de la investigación y que no cuente con estas herramientas de apoyo. Las Universidades actuales dejaron hace muchos años las herramientas de análisis tradicionales porque se encontró que producían resultados erróneos, por tanto, no es lo mismo trabajar con los métodos tradicionales que con los emergentes.

Finalmente, el cuarto objetivo no se cumplió porque se necesita mayor tiempo para redactar el artículo que se puede extraer de toda esta discusión científica.

3. RESULTADOS FINALES

Uno de los aportes más significativos es la demostración de la validez del instrumento utilizado por el C.E.A, por medio de los modelos jerárquicos lineales. Una metodología de estadística de avanzada, que permite una profundidad de análisis muy superior a los métodos tradicionales. De esa forma:

1. A las Unidades académicas les permitirá contar con instrumentos válidos, que no se vean afectados por factores ajenos a la docencia universitaria. Lo cual les posibilitará la construcción de nuevos instrumentos o tener toda la confianza en los que tienen.
2. Los investigadores tendrán material relevante para análisis de primera mano
3. Los Profesores y administrativos podrán confiar plenamente en su evaluación que les suministrará información verás, la cual permitirá mejorar su acción.
4. La institución en general al poder contar con mejores instrumentos de evaluación de la docencia universitaria.
5. Las unidades académicas que evalúan constantemente a sus docentes, podrán observar el comportamiento de estos y la opinión de sus estudiantes, con instrumentos muy consistentes y validos.

Además, participamos en una pasantía de investigación, en el departamento de Métodos de Investigación, diagnóstico y Evaluación de la Universidad Complutense de Madrid, lo cual nos permitió laborar con otros investigadores e investigadoras que trabajan sobre el tema y se pueden considerar los siguientes beneficios:

1.- Brindar a sus profesores y administrativos la confianza y la certeza de que sus instrumentos de evaluación de la docencia universitaria son válidos y miden el modelo de profesor (a) propuesto por el Consejo Universitario, dado que el análisis de los datos fundamentados en los modelos jerárquicos lineales, ofrece un nivel de profundidad que ninguna técnica tradicional de estadística ofrece.

2.- La institución se muestra ante la comunidad nacional e internacional, como una entidad preocupada no solo por la calidad universitaria, sino también, por la calidad de sus instrumentos de medición, posibilitando la oportunidad de réplica en otras instituciones de educación superior nacional e internacional, dado, que pocas unidades académicas utilizan instrumentos de evaluación, analizados en profundidad con los modelos jerárquicos lineales.

Como recomendación final, es sumamente importante considerar dentro del I.N.I.E. talleres de apoyo a los investigadores, profesores, estudiantes y público en general, acerca de las herramientas cuantitativas emergentes para el análisis de datos, debido a que existe un vacío muy grande sobre este tema. Además, es recomendable adquirir para el laboratorio del I.N.I.E. estos paquetes informáticos, a fin de que los investigadores puedan practicar y mejorar los trabajos de investigación cuantitativa.

4. REFERENCIAS

Abalde, E.; De Salvador, X; González Carbanach, R. Y Muñoz C.; J.M. (1995). Análisis de la Evaluación de la Docencia Universitaria por los (as) alumnos (as) en la Universidad de la Coruña (1993-1994). **En Estudios de Investigación Educativa en Intervención Pedagógica**, pp. (289-292) Valencia: AIDIPE.

Abrami, P. C., Dickens, W. J., Perry, R. P., & Leventhal, L. (1980). Do teacher standards for assigning grades affect student evaluations of instruction. **Journal of Educational Psychology**, **72**, 107-118.

Abrami, P.C. (1989a). How Should We Use Student Ratings to Evaluate Teaching?. **Research in Higher Education**, **30** (2), 221-227.

Abrami, P.C.; Cohen, P.A. & D'apollonia, S. (1988). Implementation Problems in Meta-Analysis. **Review of Educational Research**, **58**, pp. 151-179.

Abrami, P.C.; D'apollonia, S. & Cohen, P. (1990). .Validity of Student Ratings of Instruction. What We Know and What We Do Not. **Journal of Educational Psychology**, **82** (2), pp.219-231.

Abrami.C. y D'apollonia, S. (1990b). .The dimensionality of ratings and their use in personnel decisions. M. Theall Y J. Franklins (Eds.), **In Student Ratings of Instruction. Issues for Improving Practice** (pp. 97-111). New Directions for Teaching and Learning.

Abrami, P.S.; D'apollonia, S. & Rosenfield. (1997). .The Dimensionality Of Student Ratings of Instruction. What We Know and What We do Not.. R.P. PRRY & J.C. SMART (Eds.) **In Effective Teaching in Higher Education. Research and Practice**, pp.321-367. New York. Agathon Press.

Acevedo Alvarez, R. & Rodríguez, N. M. (2006). Factores de sesgo asociados a la validez de la evaluación docente universitaria: un modelo jerárquico lineal. **Archivos Analíticos de Políticas Educativas**, **14** (34). Recuperado [27 marzo 2007] de <http://epaa.asu.edu/epaa/v14n34/>

Acevedo, R. y Fernández, M. J. (2004): "La percepción de los estudiantes universitarios en la medida de la competencia docente: validación de una escala". En **Revista de Educación**. N° 28 (2): 145-166. San José: Editorial de la Universidad de Costa Rica.

Albanese, M.A. (1991). .The validity of lecturer ratings by student and trained observers. **Academic Medicine**, **66** (1), pp. 26-28.

Aleamori, L. (1978). .Development and factorial validation of the Arizona Course/ Instructor Evaluation Questionnaire. **Educational and Psychological Measurement**, **38**, 1063-1067.

Aleamori, L. (1981). .Student ratings of instruction.. In J. MILLAMN (ed.), **Handbook of Teacher Evaluation**, pp.110-145. Newbury Park, CA. Sage.

Aleamori, L.M. (1999). .Student Rating myths versus research facts from 1924 to 1998.. **Journal of Personnel Evaluation in Education**, **13** (2), pp. 153-166.

Aleamori, L.M. y Hexner, P.Z. (1980). .A review of the research on student evaluation and a report on the effect of different sets of instructions on student course and instructor evaluation. **Instructional Science**, **9**, pp. 67-84.

Aleamori, L.M. y Yimer, M. (1974). Graduating Senior Ratings' Relationship to Colleague Rating Student Rating. **Research Productivity and Academic Rank in Rating Instructional Effectiveness (Research Report N°352)**. Urbana. University of Illinois, Office of Instructional Resources, Measurement and Research Division.

Anderson, K.H. y Siegfried, J.J. (1997). Gender differences in rating the teaching economics. **Eastern Economic Journal**, **23** (3), pp. 347-357.

Arubayi, E. (1987). .Improvement of Instruction and Teacher Effectiveness. Are Student Ratings Reliable and Valid?. **Higher Education**, **16**, pp. 267-288.

Barnes, L.B. y Barnes, M.W. (1993). .Academic discipline and generalizability of student evaluations of instruction. **Research in Higher Education**, **34** (2), pp. 135-149.

Bendig, A.W. (1953). .Relation of level of course achievement of student, instructor and course ratings in introductory psychology. **Educational and Psychological Measurement**, **13**, pp. 437-488.

Braskamp, L.A. & Ory, J.C. (1994). **Assessing Faculty Work.ancing Individual and Institutional Performance**. San Francisco. Jossey-Bass.

Braskamp, L.A.; Ory, J.C. & Pieper, D.M. (1981). Student Written Comments. Dimensions of Instructional Quality. **Journal of Educational Psychology**, **73** (1), pp. 75-70.

Broder, J.M. & Dorfman, J.H. (1994). .Determinants of Teaching Quality; What's Important to Students?. **Research in Higher Education**, **35** (2), pp. 235-249.

Burdsal, C.A. & Bardo, J.W. (1986). Measuring Student's Perceptions of Teaching. Dimensions of Evaluation. **Educational And Psychological Measurement**, **56**, pp. 63-79.

Campbell, D.T. y Fiske, D.W. (1959). Convergent and discriminant validation by the multitrait – multimethod matrix. **Psychological Bulletin**, **56**, pp. 81-105.

Carson, B.H. (1999). Bad News in the Service of Good Teaching. Students Remember Ineffective Professors. **Journal On Excellence In College Teaching**, **19** (1), pp.91-105.

Casey, R.J.; Gentile, P. y Bigger, S.W. (1997). Teaching appraisal in higher education. an Australian perspective. **Higher Education**, **34** (3), 459-482.

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.

Cashin, W.E. (1988). **Student Ratings of teaching. A summary of the research. IDEA. Paper No 20.** Manhattan, KS. Kansas State University, Center Faculty Evaluation and Development.

CHACKO, T.I. (1983): "Student ratings of instruction: a function of grading standards". En **Educational Research Quaterly**, 83 (1), pp. 19-25.

Centra, J. (1977). Student Ratings of Instruction and Their Relationship to Student Learning. **American Educational Research Journal**, 14, pp. 17-24.

Centra, J.A. (1972). **The Utility of Student Ratings for Instructional Improvement.** Princeton, NJ. Educational Testing Services.

Centra, J.A. (1979). **Determining Faculty Effectiveness.** San Francisco: Jossey – Bass.

Centra, J.A. (1993). **Reflective Faculty Evaluation. hancing Teaching and Determining Faculty Effectiveness.** San Francisco: Jossey-Bass.

Cheng, Y. y Hoshower, L.B. (1998). Assessing student motivations to participate Teaching evaluations. an application of expectancy theory. **Issues in Accounting Education**, **13** (3), pp. 531-549.

Coffey, M. & Gibbs, G. (2001). The Evaluation of the Student Evaluation of Educational Quality Questionnaire (SEEG) in U.K. Higher Education. **Assessment & Evaluation in Higher Education**, **Vol 26**, (1), pp. 89-93.

Cohen, P. A. (1981). Student ratings of instruction and student achievement. A meta-analysis of multisection validity studies. **Review of Educational Research**, **51**, 281-309.

Cohen, P.A. (1983). A selective review of the validity of student ratings of teaching. **Journal of Higher Education**, **54**, 448-458.

Cohen, P.A. (1989). Do grades influence students' evaluation of clinical courses?. **Journal of Dental Education**, **53** (4), pp. 238-240.

Costin, F. (1968). **Survey of Opinions About Lectures**. University of Illinois: Department of Psychology.

Costin, F.; Greenough, W.T. y Menges, R.J. (1971). Student Ratings Of College Teaching. Reliability, Validity and Usefulness. **Review Of Educational Research**, **41**, 511-535.

Craton, P. y Smith, R. A. (1990). Reconsidering the Unit of Analysis. A Model of Student Ratings of Instruction. **Journal of Educational Psychology**, **82** (2) pp. 207-212.

Cruse, D. B. (1987). Student Evaluation and the University Professor. **Higher Education**, **15** (6), pp. 723-737.

D'apollonia, S. & Abrami, P. C. (1997). Navigating Student Ratings of Instruction. **American Psychologist**, **51** (11), pp. 1198-1208

De Neve, H. M. F. y Janssen, P. J. (1982). Validity of Student Evaluation of Instruction. **Higher Education**, **11** (5), pp. 543-552.

Dickinson, D. J. (1990). The relationship between ratings of teacher performance and student learning. **Contemporary Educational Psychology**, **15**, pp. 142-151.

Dowell, D. A. , & Neal, J. A. (1982). A selective view of the validity of student ratings of teaching. **Journal of Higher Education**, **53**, 51-62.

Doyle, K. O. (1975). **Student Evaluation of Instruction**. Lexington, MA: Lexington Books.

Drews, D. R. ; Burroughs, W. J. Y Nokovich, D. A. (1987). Teacher self ratings as a validity criterion for student evaluation. **Teaching of Psychology**, **14** (1), pp. 129-143.

Drucker, A. J. Y Remmers, H. H. (1950). Do Alumni and Students Differ in Their Attitudes Toward Instructors?. **Purdue University Studies in Higher Education**, **70**, pp. 62-64.

Drucker, A. J. Y Remmers, H. H. (1951). Do Alumni and Students Differ in Their Attitudes Toward Instructors?. **Journal of Educational Psychology**, **42**, pp. 129-143.

Erdle, S.; Murray, H.G. & Rushton, J.P (1985): "Personality, Classroom Behavior and Student Ratings of College Teaching Effectiveness: A Path Analysis". En **Journal of Educational Psychology**, **77** (4), pp. 394-407.

Feldman, K. A. (1977). Consistency and Variability among College Students in Rating Their Teachers and Courses. **Research in Higher Education**, **6** (2), pp. 223-274.

Feldman, K. A. (1978). Course characteristics and college students' ratings of their teachers. What are know and what we don't know. **Research in Higher Education**, **9** (2), pp. 199-242.

Feldman, K. A. (1979). The Significance of Cirunstances for College Students ratings of Their Teachers and Courses. **Research in Higher Education**, **10** (2), pp. 149-172.

Feldman, K. A. (1984). Class size and college students' evaluation of teachers and courses. a closer look. **Research in Higher Education**, **21** (11), pp. 45-116.

Feldman, K. A. (1989a). Instructional effectiveness of college teachers as judged by teachers themselves, current and formed student, colleagues, administrators and external (neutral) observers. **Research in Higher Education**, **30** (2), pp. 137-194.

Feldman, K. A. (1989b). The association between student ratings of specific instructional dimensions and student achievement. refining and extending the synthesis of data from multisection validity studies. **Research in Higher Education, 30**, pp. 583-645.

Feldman, K. A. (1997). Identifying Exemplary Teachers and Teaching. Evidence from Student Ratings. R. P. En PERRY & J. C. SMART (eds.), **Effective teaching in Higher Education. Research and Practice**, . (pp. 368-395). Bronx, N. Y: Agathon.

Fernández, J. ; Mateo, M. & Muñiz, J. (1998). Is There Relationship Between Class Size and Student Ratings of Teacher Quality?. **Educational And Psychological Measurement, 58** (August), 596-604

Franklin, J. & Theall, M. (1989). **Who read ratings. Knowledge, attitude, and practice of users of student ratings of instruction**. Paper presented at the annual meeting of the American Educational Research Association, San Francisco.

Franklin, J. & Theall, M. (1989). **Who read ratings. Knowledge, attitude, and practice of users of student ratings of instruction**. Paper presented at the annual meeting of the American Educational Research Association, San Francisco.

Frey, P. W. (1973). Student Ratings of Teaching. Validity of Several Rating Factors. **Science, 182**, pp. 83-85

García Ramos, J. M. (1999^a). Análisis multirasgo- multimétodo en la validación de instrumentos de medida para la evaluación de la calidad docente en instituciones universitarias. **Revista Española de Pedagogía, 214**, pp. 417-444.

García Ramos, J. M. (1997). Valoración De La Competencia Docente Del Profesor Universitario. Una Aproximación Empírica. **Revista Complutense De Educación, 8** (2), pp. 81-108

García Ramos, J. M. y Congosto Luna, E. (1996). **Un Modelo de Evaluación Institucional en la Universidad**. Salamanca: Studia Pedagógica

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

Gillmore, G. (1973). **Estimates of Reliability Coefficients for Items and Subscales of the Illinois Courses Evaluation Questionnaire**. (Research Report N°341). Urbana: University of Illinois, Office of Instructional Resources, Measurement, and Research Division.

Gillmore, G. M. ; Kane, M. T. & Maccarato, R. W. (1978). The Generalizability of Student Ratings of Instruction. Estimation of the Teacher And Course Components. **Journal of Educational Measurement**, **15**, pp. 1-13.

Gilmore, G. (1984). Student ratings as a factor in faculty employment decisions and periodic review. **Journal of College and University Law**, **10**, 557-576 .

Goddard, R; Hoy, W & Woolfolk, A. (2001). Collective Teacher Efficacy. Its Meaning, Measure, and Impact on Student Achievement. **Collective Teacher Efficacy**. Manuscrito sometido para publicación. 1-40.

Goldman, L. (1993). On erosion of education and the eroding foundation of teacher education (or why we should nor take student evaluation of faculty seriously). **Teacher Quarterly**, **20** (2), pp. 57-64.

Greenwald, A. G. (1997b). Validity Concern and Usefulness of Student Ratings Of Instruction. **American Psychologist** , **51** (11), 1182-1186.

Greenwald, A. G. Y Gillmore, G. M. (1997). Grading Leniency is a Removable Contaminant of Student Ratings. **American Psychologist** , **51** (11), 1209-1217.

Guthrie, E. R. (1954). **The Evaluation of Teaching**. A Progress Report. Seattle: University of Washington.

Harrison, P. D. ; Ryan, J. M. Y Moore, P. S. (1996). College studen's self-insight and common implicit theories in the ratings of teaching effectiveness. **Journal of Educational Psychology**, **88** (4), pp. 775-782.

Hativa, N. (1996). University instructors' rating profiles. Stability over time, and disciplinary differences. **Research in Higher Education**, **37** (3), pp. 341-365.

Hativa, N. y Raviv, A. (1993). Using a single score for summative teacher evaluation by students. **Research in Higher Education**, **34** (5), pp. 625-646.

Hepworth, D. Y Oviatt, B. E. (1985). Using student course evaluations. findings, issues and recommendations. **Journal of Social Work Education**, **21** (3), pp. 105-112.

Hilton, H. (1993). Realibility and Validity of Student Evaluation. Testing Models versus Survey Research Model, p. s. **Political Science & Politics**, **26**, pp. 562-569.

Hogan, T. P. (1973). Similarity of student ratings across instructors, courses and time. **Research in Higher Education, 1**, pp. 149-154.

Holmes, D. S. (1972). Effects of grades and disconfirmed grade expectancies on students' evaluations of their instructor. **Journal of Educational Psychology, 63**, 130-133.

Howard, G. S. , Conway, C. G, & Maxwell, S. E. (1985). Construct validity of measures of college teaching effectiveness. **Journal of Educational Psychology, 77**, 187-196.

Howard, G. S. Y Maxwell, S. E. (1980). Correlation between student satisfaction and grades. a case of mistaken causation?. **Journal of Educational Philosophy, 72** (4), pp. 810-820.

Howell, A. J. Y Symbaluk, D. (2001). Published Student Ratings of Instruction. Revealing and Reconciling the Views of Student and Faculty. **Journal of Educational Psychology, (4)**, 790-796

Jornet, J. M. ; González Such, J. Y Pérez Carbonell, A. (1995) . Evaluación De la Actividad Universitaria. G. Quintás (eds.), En **Reforma y Evaluación de la Universidad**. Valencia: Servei de Publicacions de la Universitat de Valencia. pp. 189-244.

Koblitz, . N. (1990). Are student ratings unfair to women?. **Newsletter of the Association for Women in Mathematics, 20**, pp. 17-19.

Koermer, C. D. Y Petelle, J. L. (1991). Expectancy violation and student rating of instruction. **Communication Quarterly, 39** (4), pp. 341-350.

Kolitch, E. & Dean, A. V. (1999). Student Ratings of Instruction en the U. S. A. Hidden Assumptions and Missing Conceptions About . Good. Teaching. **Studies in Higher Education, Vol. 24**, (1), pp. 27-42.

Koon, J. Y Murray, H. G. (1995). Using multiple outcomes to validate student ratings of overall teacher effectiveness. **Journal of Higher Education. 66** (1), pp. 61-81.

Lin, W. Y. ; Watkins, D. Y Meng, Q. M. (1995). Student's Evaluation of University teaching. A China perspective. **Higher Education and Research y Development, 14** (1), pp. 61-74.

López Feal, R. (1986). **Construcción de Instrumentos de Medida en Ciencias Conductuales y Sociales**. Barcelona: Alamex.

Mahmound, M. M. (1991). Descriptive models of student decision behaviour in evaluation of higher education. **Assessment and Evaluation in Higher Education**, **16** (2), pp. 133-148.

Marques, T. E. ; Lane, D. M. Y Dorfman, P. W. (1979). Toward the development of a system for instructional evaluation. Is there consensus regarding what constitutes effective teaching?. **Journal of Educational Psychology**, **71**, pp. 840-849.

Marsh, H. & Roche, L. R. (1997). Making Students' Evaluation of Teaching Effectiveness Effective. The Critical Issues of Validity, Bias, and Utility. **American Psychologist**, **Vol. 52**, 11, 1187-1197.

Marsh, H. (1982). SEEQ. a Reliable, Valid, and Useful Instrument for Collecting Students Evaluation of University Teaching. **British Journal of Psychology**, **52**, pp. 77-95.

Marsh H. (1984). Student's evaluation of university teaching; dimensionality, reliability, validity, potential biases ad utility. **Journal of Educational Psychology**, **76** (5), 707-754.

Marsh, H. (1987a). Students' evaluation of university teaching; Research findings, methodological issues, and directions for future research. **International Journal of Educational Research**, **11**, pp. 253-288.

Marsh, H. (1987b). Student Evaluations of Teaching. M. J. DUNKINS (eds.), **The International Encyclopedia of Teaching and Teacher Evaluation**. (pp. 181-187) Oxford: Pergamon Press

Marsh, H. y Bailey, M. (1993). Multidimensional Students' Evaluations of Teaching Effectiveness. **Journal of Higher Education**, **64** (1), pp. 1-18.

Marsh, H. A. y Overall, J. U. (1979b). **Validity of student's evaluation of teaching. A comparison with self evaluations by teaching assistants, undergraduate faculty, and graduate faculty**. Paper presented Annual Meeting of the American Educational Research Association, San Francisco.

Marsh, H. W. (1977). The validity of students' evaluations of instructors independently nominated as best and worst teacher by graduating senior. **American Educational Research Journal**, **14**, pp. 441-447.

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

Marsh, H. W. (1992b). **A Longitudinal Perspective of Student's Evaluations of University Teaching. Ratings of The Same Teacher over a 13 Year Period.** Documento presentado en Annual Meeting of the American Educational Research Association (p. 18) San Francisco, Ca. Abril.

Marsh, H. W. (2001). Distinguishing between good (useful) and bad workloads on Students' Evaluations of teaching. **American Educational Research, 38** (1), pp. 183-212.

Marsh, H. W. y Dunkin, M. J. (1992). Students' Evaluation of University Teaching. A Multidimensional Perspective. J. SMART (ed.) **Higher Education. Handbook of Theory and Research.** (pp. 143-223) New York: Agathon.

Marsh, H. W. y Hocevar, D. (1985). Application of confirmatory factor analysis to the study of self-concept. First and higher order factor models and their invariance across groups. **Psychological Bulletin, 97**, pp. 562-582.

Marsh, H. W. y Overall, J. U. (1981). The relative influence of course level, course type, and instructor on students' evaluations of college teaching. **American Educational Research, 18**, pp. 103-112.

Marsh, H. W. y Roche, L. (1992). **The Use of Student's Evaluations of University Teaching To Improve Teaching Effectiveness.** Canberra: Australian Government Publishing Services.

Marsh, H. W. y Roche, L. A. (1994). The Use of Student Evaluations of University Teaching in Different Settings. The Applicability Paradigm. **Australian Journal of Education, 36** (3), pp. 278-300.

Marsh, H. W. y Roche, L. A. (2000). Effects of grading leniency and low workload on students' evaluation of teaching. popular myth, bias, validity or innocent bystanders?. **Journal of Educational Psychology, 92** (1), pp. 202-228.

Marsh, H. W. ; Balla, J. R. y Mcdonald, R. P. (1988). Goodness of Fit Indices in Confirmatory Factor Analysis. The Effects of Sample Size. **Psychological Bulletin, 103**, pp. 391-410.

Marsh, H. W. ; Hau, K. T. ; Chung, C. M. & Siu, T. L. P. (1997). Students' s Evaluations of University Teaching. Chinese Version of The Student's Evaluations of Educational Quality Instrument. **Journal of Educational Psychology, 89** (3), pp. 568-572.

Marsh, H. & Roche, L. R. (1997b). Making Students's Evaluation of Teaching Effectiveness Effective. The Critical Issues of Validity, Bias, and Utility. **American Psychologist, Vol. 52** (11), pp. 1187-1197.

Marsh, H. W. ; Touron, J. y Wheeler, B. (1985). Students' s Evaluations of University Instructor. The Applicability of American Instrument in a Spanish Setting. **Teaching and Teacher Education, 1**, pp. 123-138.

Mccallum, L. W. (1984). A meta-analysis of course evaluation data and its use in the tenure decision. **Research in Higher Education, 21**, 150-158.

McKeachie, W. J. (1979). Student ratings of faculty. A reprise. **Academe, 65**, pp. 384-397.

McKeachie, W. J. (1987). Instructional evaluation. current issues and possible improvement. **Journal of Higher Education, 58** (3), pp. 344-350.

McKeachie, W. J. (1990). Research on College Teaching. The Historical Background. **Journal of Educational Psychology, 82** (2), pp. 189-200.

McKeachie, W. J. (1997). Student Ratings. The Validity of Use. **American Psychologist, 52** (11), pp. 1218-1225.

McKeachie, W. J. ; LIN, Y. G. y MENDELSON, C. N. (1978). A small study assessing teacher effectiveness. Does learning last?. **Contemporary Educational Psychology, 3**, pp. 352-357.

Meeth, L. R. (1976). **The stateless art of teaching evaluation. report on teaching. Vol. 2, Change, 8**, pp. 3-5.

Menges, R. J. (1991). The real world of teaching improvement. A faculty perspective. M THEALL & FRANKLIN (Eds.) **Effective Practices for Improving Teaching, New Directions for Teaching and Learning, Vol. 48**, (pp. 21-37). San Francisco: Jossey-Bass.

Miller, R. I. (1987). **Evaluation Faculty for Promotion and Tenure**. San Francisco: Jossey-Bass.

Miller, S. (1984). Student rating scales for tenure and promotion. **Improving College and University Teaching, 32** (2), pp. 87-90

Millman, J. (1981). **Handbook of Teacher Evaluation**. Beverly Hills, CA: Sage.

Monroe, C. y Borzi, M. G. (1989). Methodological issues regarding student evaluation of teacher. A pilot study. *ACA Bulletin*, 70, pp. 73-79.

Moses, I. (1986). Self and Student evaluation of academic staff. **Assessment and Evaluation in Higher Education**, pp. 76-78.

Muñíz, J. ; García, A. y Virgos, J. M. (1991). Escala de la Universidad de Oviedo para la evaluación del profesorado. **Psicothema**, 3 (2), pp. 269-281.

Murray, H. G.; Rushton, P. & Paunonen, S. V. (1990). Teacher Personality Traits and Student Instructional ratings in Six Types of University Courses. **Journal of Educational Psychology**, 82 (2), pp. 250-261.

Nimmer, J. G. y Stone, E. F. (1991). Effects of grading practices and time of rating on student ratings of faculty performance and student learning. **Research in Higher Education**, 32 (2), pp. 195-215.

Noser, T. C. ; Manakyan, H. y Tanner, J. R. (1996). Research productivity and perceived teaching effectiveness. A survey of economic faculty. **Research in Higher Education**, 37 (3), pp. 299-321.

O'connell, D. Q. y Dickenson, D. J. (1993). Student ratings of instruction as a function of testing conditions and perceptions amount learned. **Journal of Research and Development in Education**, 27 (1), pp. 18-23.

Overall, J. U. & Marsh, H. W. (1980). Students' s Evaluations of Instruction. A Longitudinal Study of Their Stability. **Journal of Educational Psychology**, 72, pp. 321-325.

Palchik, N. S. (1988). Student assessment of teaching effectiveness in a multi-instructor course for multidisciplinary health professional student. **Evaluation and the Health Professions**, 11 (1), pp. 55-73.

Powell, R. W. (1977). Grades, learning, and student evaluation of instruction. **Research in Higher Education**, 7, 193-205.

Prave, R. S. y Bavril, G. L. (1993). Instructor rating. controlling for bias from initial student interest. **Journal of Educational for Business**, 68 (2), pp. 362-366.

Prosser, M. & Trigwell, K. (1990). How will Future Academic be Evaluated? Using Student Study Strategies to **Check the Validity of Student Evaluations of Teaching Courses**. G. MULLIS (Ed)

Ramsden, P. (1991). A performance indicator of teaching quality in higher education. the course experience Questionnaire. **Studies in Higher Education, 16**, pp. 129-150.

Rindermann, H. & Schofield, N. (2001). Generalizability of Multidimensional Student Rating of University Instruction Across Courses and Teacher. **Research in Higher Education, Vol. 42**, 4, pp. 377-400.

Rodin, M. & Rodin, B (1972). Student Evaluation of Teachers. **Science, 177**, pp. 1164-1166.

Rutland, P. (1990). Some considerations regarding teaching evaluations. **Political Science Teacher, 3**, pp. 1-2.

Ryan, J. M. & Harrison, P. D. (1995). The Relationship Between Individual Instructional Characteristic and the Overall Assessment of Teaching Effectiveness Across Different Instructional Context. **Research In Higher Education, Vol. 36**, No 5, pp. 577-594

Ryans, D. G. (1960) **Characteristics of Teachers**. Washington, D. C.: American Council on Education.

Salvador, L. (1990). **Los Docentes Universitarios Exitosos Desde La Perspectiva del Alumno. Su Caracterización Psicopedagógica**. Tesis Doctoral. Universidad de Salamanca. España.

Saroyan, A. & Amundsen, Ch. (2001). Evaluating University teaching. Time to Take Stock. **Assessment & Evaluation in Higher Education, Vol 26**, (4), pp. 341-353

Seldin, P. (1984). **Changing Practices in Faculty Evaluation. A Critical assessment and Recommendations for Improvement**. San Francisco: Josser – Bass.

Seldin, P. (1993a). The use ad abuse of students ratings of professors. **The Chronicle of Higher Education**, p. 40

Seldin, P. (1993b). **Successful Use Of Teaching Portafolios**. Bolton, MA: Anker Publishing Co.

Shadish, W. (1998). **Some Evaluation Questions. Practical Assessment, Research & Evaluation, 6 (3)**. Recuperado el 14 de enero de 2001 de <http://ericae.net/pare/getvn.asp?v=6&n=3>].

Snyder, C. R. , & Clair, M. (1976). Effects of expected and obtained grades on teacher evaluation and attribution of performance. **Journal of Educational Psychology, 68**, 75-82.

Spencer, P. A. Y Flyr, M. L. (1992). **The Formal Evaluation As Ac Impetus To Classroom Change. Myths or Reality?**. Research/ Technical Report, Riverside, CA.

Tagomori, H. & Bishop, L. (1995). Student Evaluation of Teaching. Flaw Instruments. Thought and Action. **The National Education Association Higher Education Journal, 11**, pp. 63-78.

Tatro, C. N. (1995). Gender effect on student evaluations of faculty. **Journal of Research and Development in Education, 28 (3)**, pp. 169-173.

Tejedor, F. y Montero, L. (1990). Indicadores de la Calidad Docente para La Evaluación del Profesor Universitario. **Revista Española De Pedagogía, año XLVIII**, N° 186, mayo-agosto, pp. 259-279.

Timpson, W. W. Y Andrew, D. (1997). Rethinking student evaluation and the improvement of teaching. instrument for change at the University of Queensland. **Studies in Higher Education, 22 (1)**, pp. 55-65.

Ting, K. F. (2001). A Multilevel Perspective On Student Ratings of Instruction. Lessons From the Chinese Experience. **Research in Higher Education. Vol. 41, 5**, pp. 637-653.

Trigwell, K. & Prosser, M. (1996). Changing Approaches to Teaching. A Relational Perspective. **Studies in Higher Education, 21**, pp. 275-284.

Vandewalle (1997). Development and validation of a work domain goal orientation instrument. **Educational and Psychological Measurement, 57 (6)**, pp. 995-1015.

Vasta, R. , & Sarmiento, R. F. (1979). Liberal grading improves evaluations but not performance. **Journal of Educational Psychology, 71**, 207-211.

Villa, A. Y Morales, P. (1993). **La Evaluación Del Profesor. Una Visión De Los Principales Problemas Y Enfoques De Diversos Contextos.** Vitoria: Departamento de Educación, Universidades e Investigación. Gobierno Vasco.

Vu, T. R. ; Marrito, D. J. ; Stratos, G. A. Y Litzelman, D. K. (1997). Prioritizing areas for faculty development of clinical teachers by using student evaluations for evidence- based decisions. **Academic Medicine, 72** (10), pp. 57-59.

Wachtel, H. K. (1998). Student Evaluation of College Teaching Effectiveness. A Brief Review. **Assessment & Evaluation In Higher Education, 23** (June), 191-211.

Weinbach, R. W. (1988). Manipulation of student evaluations. No laughing Mater. *Journal of Social Work Education, 24* (1), pp. 37-34.

Wigington, H. ; Tollefson, N. Y Rodríguez, E. (1989). Student's ratings of instructor visited. Interactions among class and instructor variables. **Research in Higher Education, 30** (3), pp. 331-334.

Wilkerson, D. ; Rogers, M. A. Y Maughan, R. (2000). Validation of Student, Principal, and Self Ratings in 360 Feedback for Teacher Evaluation. **Journal of Personnel Evaluation in Education, 14** (2), pp. 179-192

Wilson, K. L.; Lizzio, A. & Ramsden, P. (1997). The Development , Validation, and Application of the Course Experience Questionnaire. **Studies in Higher Education, 22** (1) , pp. 33-52.

Worthington, A. G. , & Wong, P. T. P. (1979). Effects of earned and assigned grades on student evaluations of an instructor. **Journal of Educational Psychology, 71** , 764-775.

Worthington, A. C. (2002). The impact of student perception and characteristics on teaching evaluation. a case study in finance education. **Assessment & Evaluation in Higher Education, 27** (1), pp. 49-64.

Zoller, U. (1992). Faculty Teaching Performance Evaluation in Higher Science Education. Issues and Implications. **Science Education, 76** (6), pp. 673-684.