

DRA. SONIA SALAS B.  
DR. LEE HUMPHREYS  
PAMELA MESSER

## EVALUACION DE LA INSTRUCCION SUPERIOR A TRAVES DE LOS ALUMNOS: ¿RIESGOS O BENEFICIOS ?

### *Resumen*

*El presente artículo discute el uso de las evaluaciones de los alumnos en la Enseñanza Superior.*

*Se analizan tres tipos de variables en torno al tema:*

- 1. Factores relacionados con la validez y confiabilidad de los instrumentos de medición utilizados.*
- 2. Factores vinculados con las características personales y psicológicas de los alumnos y de los docentes.*
- 3. Variables relacionadas en el proceso mismo de enseñanza.*

*El artículo analiza una gran variedad de fuentes de evaluación de la enseñanza; entre éstas se destacan: la autoevaluación, la evaluación realizada por comités, jefes de departamentos y decanos, la opinión de los colegas, diversas observaciones en la sala de clases y por supuesto la evaluación sistemática por parte de los estudiantes.*

*La revisión de la bibliografía denota la presencia de gran controversia y discrepancia entre los investigadores. Sin embargo emerge el supuesto que debiera utilizarse múltiples fuentes evaluativas convenientemente validadas, considerando la naturaleza crítica de los juicios implícitos; de esta forma se puede lograr una evaluación instruccional que corresponda en forma fidedigna con los esfuerzos profesionales de una institución superior.*

## **Abstract**

*The article focusses upon the use of student evaluations on teaching instruction.*

*Three broad areas are discussed:*

- 1) factors concerning the validity and reliability of measurement instruments;*
- 2) factors related to personal and psychological characteristics of both instructors and students, and*
- 3) variables related to the teaching process itself.*

*A wide variety of teaching evaluation sources are reviewed. Among these, self-evaluation, comitee evaluation, dean and chairman evaluation, colleagues' opinion, classroom visits, and of course, systematic student ratings are most prevalent.*

*Research findings have produced a great deal of controversy among researchers. Careful thought to the proper balance of as many sources of evidence and perceptions as possible in the light of the critical nature of the judgments to be based on them, can lead to proper consideration of instruction as a vital component in the professional efforts of faculty.*

Las decisiones más difíciles que enfrenta la educación superior se relacionan con la selección, promoción o despido de los integrantes de la planta docente y administrativa. Ninguna otra decisión afecta de manera tan vital la calidad de la institución y de los departamentos insertos en ella. Lo que es más crucial aún, es que estas determinaciones definen en esencia la carrera profesional y la vida misma de los colegas con quienes se trabaja en forma estrecha y a quienes con frecuencia se le considera como amigos. Estas decisiones exigen juicios que son complejíssimos, las más de las veces difíciles de formular y que deben basarse en criterios informativos multidimensionales apoyados y evaluados con criterios claros. A pesar de todas estas exigencias, a veces los juicios deben emitirse enfrentando ambigüedades como también diversos grados de incertidumbre. Por lo mismo, este tipo de determinaciones de carácter tan complejo, no debe sólo reducirse a un tipo de numerología moderna que pretende evaluar basándose solamente en un conjunto de dígitos, seguidos de decimales arbitrariamente seleccionados.

En el momento actual, en que se está asignando una gran importancia a la "rentabilidad" que muestran los individuos dentro y fuera de una institución, a las demandas evidentes de valores en término de dineros gastados y en un contexto de competencia permanente, la presión por evidencias cuantificables claras y rápidas se hace manifiesta; no es de extrañar entonces, que exista cautela ante cualquier información que no esté basada en criterios cuantificables.

Casi la totalidad de los planteamientos y criterios en relación a la propiedad de los cargos y la promoción dentro de ellos, sostiene que la docencia e investigación constituyen componentes esenciales de los deberes profesionales de todo miembro académico; se dice que se busca la excelencia académica y que la calidad de los esfuerzos al interior de la sala de clases, junto a la productividad investigativa, publicaciones y actividades de carácter creativo ayudaran a tomar decisiones de incidencia promocional; sin embargo muchos de los docentes piensan, y con justa razón, que al momento de tomar las decisiones, los aspectos que se toman en cuenta (por los colegas y autoridades universitarias), son sólo aquellos relacionados con las investigaciones, publicaciones, actividades creativas las que son productivas y rentables económicamente.

Hay algo de verdad en esta percepción: la situación pareciera no deberse a criterios torcidos por parte de la institución superior o de sus líderes, sino más bien al hecho que las actividades de investigación, publicación, y esfuerzo por la creatividad parecen determinarse con criterios cuantificables en términos de dinero, están mejor definidos y por tanto son defendibles aún ante inspecciones externas. En contraste, la calidad de la docencia pareciera notoriamente difícil de definir, los criterios son menos claros y las evidencias se toman ambiguas y con frecuencia dudosas.

Es cierto que la evaluación sistemática de la docencia superior por parte de los estudiantes está aumentando en frecuencia de uso (especialmente en las facultades de ciencias sociales); sin embargo, este hecho no logra disipar las dudas al respecto (Seldin, 1984). Para algunos, este énfasis en el aporte evaluativo de los estudiantes generaría problemas de carácter ético. Existirían serias interrogantes acerca de lo que la evaluación de los estudiantes podría decir, acerca de su potencialidad, limitaciones, relación con la calidad de la docencia y por último, con el aprendizaje del alumno. Estas interrogantes presentan un carácter complejo y debieran constituir el centro de interés como también lo debieran ser las implicaciones éticas subyacentes. Se sabe que la evaluación de la docencia por parte del estudiante se puede convertir en un simple "concurso de popularidad", que la "actuación dramática" puede obtener las mejores calificaciones, que estas sólo pueden medir el valor de

"entretenimiento" de algunos docentes o cursos, que pueden llegar a ser fácilmente manipuladas por los docentes y que ofrecen la tentación de desarrollar un curriculum "relajado" disminuyendo los criterios de exigencia. Otras opiniones vertidas en forma más seria, han sugerido que los diseños utilizados en el proceso de evaluación de la docencia serían deficientes ya que piden información a los estudiantes que no están en condiciones de proporcionar; también se afirma que las evaluaciones no se realizan en forma sistemática utilizándose sin normas definidas lo cual impediría comparaciones válidas. Finalmente, se sugiere que debiera medirse cuidadosamente la relación entre la utilización de los criterios evaluativos con propósitos de optimizar la enseñanza y/o toma de decisiones de tipo personal.

¿Hasta que nivel son exactas estas percepciones? ¿Qué investigaciones se han realizado acerca del uso y calidad de la evaluación por parte de los estudiantes? ¿Cómo debiera utilizarse la evaluación a fin de aumentar su potencial y evitar los abusos? Esta y otras interrogantes son las que se revisarán a continuación.

## EL USO DE LA EVALUACION ESTUDIANTIL EN LA DOCENCIA SUPERIOR

El interés creciente relacionado con la evaluación de la docencia ha despertado un creciente interés de parte de los profesores e investigadores los que han tratado de determinar la multiplicidad de factores que estarían influyendo en este tipo de evaluaciones. Las investigaciones realizadas se han orientado principalmente hacia tres áreas: a) factores relacionados con la validez y confiabilidad de los instrumentos de medición; b) factores relacionados con las características personales y psicológicas tanto del profesor como del alumno; y c) variables relacionadas con el proceso mismo de enseñanza.

### Confiabilidad y Validez de la Evaluación Estudiantil

El aumento en el uso de la evaluación estudiantil en universidades e institutos ha generado una gran preocupación en torno a la validez y confiabilidad de las medidas utilizadas; a su vez las investigaciones realizadas en torno al tema, han provocado una serie de controversias y discrepancias entre los expertos. Un grupo postula la idea que las evaluaciones de los

alumnos constituyen medios válidos y útiles que ayudan a los profesores a optimizar su docencia (Marsh, 1982). Centra (1979) afirma que la confiabilidad de la evaluación de los estudiantes se ha demostrado en forma consistente, utilizando una variedad de procedimientos; entre estos se destacan: el estudio del grado de acuerdo de la evaluación del alumno "en la clase"; la aplicación de coeficientes de correlación dentro del curso a fin de estimar el grado de homogeneidad de la evaluación; y por último, determinando los puntajes básicos observados en cursos escindidos aleatoriamente por mitades los que a su vez se correlacionan con otros cursos. Centra encontró coeficientes de confiabilidad cercanos a .70 para cada uno de 30 ítemes al aplicar la fórmula de correlación dentro del curso (con 10 evaluadores) y coeficientes de .78 al evaluar la efectividad de los instructores. Inversamente, Zelby (1977) sostiene que el sistema de evaluación sólo estimula a los profesores a bajar los estándares ("hagan pruebas fáciles hasta el momento de la evaluación, permitiendo que cada estudiante obtenga una nota suficientemente alta").

Lamberth y Kostascki (1981) compararon la auto evaluación de los profesores-asistentes y la evaluación de estos por parte de los estudiantes en dos oportunidades (períodos de primavera y otoño). El instrumento que constaba de 28 ítemes, tenía como objetivo medir la efectividad de las clases como también la relación profesor-alumno. Se obtuvo una correlación alta (.71 y .69 para los dos períodos respectivamente) entre la evaluación de los alumnos y los profesores asistentes, sustentando con ello la validez de la evaluación por parte de los alumnos. Un año más tarde, sin embargo, Bonge (1982) discrepó de las conclusiones de Lamberth y Kostascki argumentando en el sentido que las correlaciones obtenidas eran artificialmente altas (véase la respuesta de Lamberth y Kostascki, 1982). De todas maneras, McKeachie (citado por Seldin, 1984) llegó a concluir que la mayoría de los estudiantes han demostrado altos niveles de estabilidad y consistencia, respaldando el hecho que este tipo de evaluaciones poseería un grado alto de confiabilidad.

Cohen (1982), utilizó la técnica conocida como meta-análisis en un intento por sintetizar los resultados de las investigaciones sobre la relación entre la evaluación y el rendimiento de los estudiantes en cursos de psicología; a través de esta técnica el autor seleccionó dieciseis estudios que incluían quince variables diferentes. Los resultados de este meta-análisis arrojaron un promedio del coeficiente de correlación igual a .44 entre la evaluación y el rendimiento de los estudiantes. Este coeficiente promedio puede considerarse con un carácter moderadamente alto, a pesar del hecho que las correlaciones variaron de curso en curso. Cohen piensa que este estudio argumenta a favor de la validez de la evaluación de los estudiantes como una estrategia de medición de los rendimientos y eficiencia mostrado por los

estudiantes de psicología. En otras palabras, "aquellos profesores cuyos estudiantes presentan buenos resultados en términos de su rendimiento en el curso, tienden a recibir mejores evaluaciones de su docencia comparados con aquellos docentes cuyos alumnos no les va bien". Estos resultados sugieren que algunos aspectos de la enseñanza parecen más relacionados con el aprendizaje que otros (ej. experiencia del profesor  $r=.65$  para un académico de jornada completa). Por otra parte, los factores ligados a variables como simpatía e interacción obtuvieron correlaciones positivas más bajas (.29 y .16 respectivamente). Cabe hacer notar que existen discrepancias en cuanto a los resultados usando este tipo de análisis, debido a las diferencias de los diseños y métodos utilizados, tamaños muestrales y variedad de los instrumentos utilizados. De acuerdo a Seldin (1984), aunque el uso de formas de evaluación para medir la efectividad de la docencia está aumentando rápidamente en las universidades norteamericanas, esfuerzos mínimos se habrían dedicado para validar estos instrumentos.

## 2. Características del profesor y del alumno

Se han formulado dos preguntas básicas con respecto a las características del profesor y del alumno:

- a) ¿En qué medida la evaluación de los alumnos es influida por las características del profesor?
- b) ¿En qué medida las características propias del alumno influyen en su evaluación?

Una serie de investigaciones se han orientado hacia las variables que pudieran tener influencia en la evaluación de los estudiantes. Los aspectos más comunmente descritos por la literatura se relacionan con el sexo, experiencia académica, edad, carga y productividad académica. De acuerdo a Centra (1979) las características del profesor generalmente no muestran ninguna relación con la evaluación de los estudiantes a excepción de la variable "años de experiencia docente". Feldman (1983) ha descrito un tipo de relación curvilínea en relación a los años de experiencia; en este sentido, este autor observó que la evaluación de los alumnos tiende a ser mejor cuando la experiencia del docente se encuentra dentro de un período de tiempo determinado (9 a 15 años), después del cual parece declinar. Las evaluaciones más bajas las recibieron aquellos docentes con uno o dos años de experiencia. Otras características del profesor (sexo, jerarquía, carga académica y productividad en investigación) no mostraron relaciones significativas con respecto a las evaluaciones (Centra, 1979).

Los estudios tanto de Centra (1979) como de Murray (1983) parecieran indicar que los docentes tienden a obtener calificaciones positivas cuando evalúan a sus estudiantes en forma justa presentando las clases en un contexto lógico a la vez que denotando "interés y entusiasmo". Murray a su vez señala haber encontrado diferencias significativas en relación con ciertas conductas exhibidas por los profesores (entusiasmo y simpatía). Estos resultados se obtuvieron a partir de la visita que hicieron "observadores entrenados" a tres conjuntos de clases que se agruparon en tres niveles (alto, medio y bajo) de acuerdo a la evaluación de los estudiantes en torno a la efectividad de la enseñanza.

La literatura también ha descrito varias características del estudiante que pueden influir en el proceso de evaluación del docente. Entre las variables que se han estudiado con más frecuencia se mencionan: curso de la carrera, sexo, edad, promedio de notas acumulado, habilidad académica y estado emocional. En relación a estos factores, Centra y Creech (1976), observaron la existencia de relaciones insignificantes entre la evaluación de los estudiantes acerca de la efectividad del docente y las variables: sexo, promedio de notas acumulado, curso de la carrera, habilidad académica y edad; sin embargo, un análisis más profundo indica que las características arriba mencionadas, podrían tener un efecto significativo, dependiendo de las diversas circunstancias que se observen. Small et al. (1982) concluyeron que de hecho existiría un escaso número de investigaciones acerca de los factores potenciales que pudieran influir la evaluación de los profesores por parte de los estudiantes. Así, los autores investigaron la relación existente entre el estado emocional del estudiante y sus evaluaciones del docente, observándose que aquellos alumnos que mostraban síntomas de ansiedad, hostilidad o depresión tendían a evaluar al profesor en forma significativamente más negativa. Asimismo, es interesante hacer notar que las variables emocionales y evaluativas evidenciaron una correlación más alta al administrarse las mediciones en forma simultánea.

Recientemente, ha aparecido en la literatura una serie de estudios relacionados con el efecto que tendría sobre la evaluación, las expectativas de nota por parte de los alumnos; los resultados sin embargo no logran llegar a conclusiones definidas. Mientras algunos autores (South et al., 1979); Orpen, 1980; Doyle y Whitely, 1974; Marsh et al., 1975; Frey, 1973) afirman que las notas "sí" influyen la evaluación del docente, otros especialistas (Eagle, 1977; Centra, 1979; and Martin, 1979) enfatizan la idea que no existiría una asociación definida entre la evaluación (positiva o negativa) y las expectativas de notas que los estudiantes esperan obtener. Rodin y Rodin (1972), encontraron una correlación negativa entre las calificaciones de los alumnos y la evaluación de los profesores ayudantes de postgrado. Al parecer, mientras

algunos estudios han mostrado correlaciones bajas entre las variables evaluación y calificaciones, otros trabajos han señalado la ausencia de relación; incluso, un número escaso de investigaciones, ha llegado a observar correlaciones inversas. Centra (1979) por ejemplo, señaló que la correlación promedio entre la evaluación y las notas de los alumnos, alcanzó un .20 coeficiente que puede ser considerado relativamente bajo (4% de la varianza explicada); Ducette y Kenny (1982) sin embargo, encontraron una relación significativa entre las notas esperadas y tres dimensiones del proceso de enseñanza: efectividad del curso, del profesor y grado de dificultad de la asignatura; estos tres factores en conjunto logran explicar un 54% de la variación total (donde la variable "efectividad del profesor" por sí sola explica el 41%). Schevrich et., al. (1983) describe otras cuatro variables que pudieran producir discrepancias en los resultados: unidad de observación, análisis sobresimplificados, estrechez de enfoque y sesgo debido a la ausencia de respuestas.

¿Cómo puede explicarse estos resultados incongruentes? La única respuesta que surge, deriva del hecho de considerar la evaluación como un proceso multidimensional y donde algunos autores (Ducette y Kenny, 1982) recomiendan la construcción cuidadosa de los instrumentos de evaluación, considerando una gama amplia de dimensiones a medir.

### 3. Variables relacionadas con el proceso de enseñanza

Una serie de variables de tipo administrativa se han estudiado con el fin de determinar el grado de relación que presenta el proceso de enseñanza y la evaluación de la docencia por parte de los estudiantes: entre los más frecuentemente mencionados se destacan los análisis sobre tamaño del curso, material instructivo, contenidos y requisitos de la cátedra.

Overall y Marshall (1982) como también Centra (1979) determinaron que los cursos pequeños (menos de 30 o 35 alumnos) y los cursos grandes (sobre 100 alumnos) eran mejor evaluados; por otra parte, aquellos cursos entre 35 y 100 estudiantes eran los que tendían a recibir evaluaciones más bajas por parte de los alumnos. Centra (1979) además planteó que tanto la interacción profesor-alumno como los años de experiencia del docente, constituirían factores que podían incidir en la evaluación. Ambos estudios señalaron que las asignaturas de ciencias sociales y naturales tendían a recibir mejores evaluaciones que las asignaturas propiamente de humanidades. Asimismo, los cursos de la especialidad y los ramos electivos tendían a recibir evaluaciones levemente superiores comparados con aquellos cursos tomados solamente para cumplir con los requisitos.



Robert Sommer (1981) utilizó un enfoque ideográfico (estudio individual) a fin de estudiar el proceso de evaluación por parte de los estudiantes a lo largo del tiempo. Al revisar su colección de veinte años de evaluaciones de sus estudiantes, concluyó que los alumnos podían discriminar entre su propio desempeño como profesor y la dificultad de los contenidos de la clase. Así, aquellos segmentos del curso que aparecían obvios y triviales eran aquellos que resultaban menos preferidos. Este autor recomienda el uso de un enfoque con métodos múltiples a fin de evaluar la calidad de la docencia.

La tabla siguiente(1) presenta algunos principios (Centra, 1983; Seldin, 1979) que debieran tomarse en cuenta al utilizar la evaluación por parte de los alumnos. En términos globales estas recomendaciones plantean que la evaluación de la docencia por parte de los alumnos puede proporcionar información útil siempre y cuando sea diseñada, puesta a prueba y administrada en forma cuidadosa. También se deja en claro que este tipo de información proporciona evidencia limitada sobre un fenómeno muy complejo debiendo modificarse a la luz de otros tipos de información. El asignar escasa o demasiada confianza a este tipo de evaluación, especialmente si se ha utilizado instrumentos evaluativos que evidencian un diseño deficiente, que no han sido probados ni administrados en forma sistemática, resulta anti ético especialmente si se considera la naturaleza crítica de las decisiones a tomar sobre esta base.

**TABLA 1**  
**Principios y sugerencias**

Centra (1983)	Seldin (1984)
1) Al tomar decisiones prefiera evaluaciones globales	1) Diseñe un folleto explicando la utilización de la evaluación.
2) Use fuentes de información múltiples (triángulo evaluativo: alumno - pares - situación de aprendizaje)	2) Estandarice la administración de las evaluaciones.
3) Use evaluaciones longitudinales (múltiples conjuntos evaluativos).	3) Use la evaluación en forma juiciosa (bajando los niveles de ansiedad).
4) Use procedimientos de administración estandarizados.	4) Al tomar decisiones use evaluaciones de fuentes variadas.
5) Use un mínimo de 15 o más alumnos evaluadores (propósito de confiabilidad y generalización de resultados).	5) Al evaluar la docencia use fuentes informativas variadas.
6) Tome en consideración las características del curso (ej.: tamaño).	6) Debe manejarse con criterio las diferencias poco significativas.
7) Proporcione retroalimentación diagnóstica a fin de mejorar la docencia.	
8) Utilice información comparable (normas nacionales o locales).	
9) No se extralimite en el uso de los cuestionarios evaluativos (se recomienda un equilibrio adecuado).	

Los principios reseñados más arriba sugieren la necesidad de rechazar aquellas áreas donde se cree que no hay suficiente evidencia o que la información presenta un grado de confiabilidad inaceptable. De igual forma se advierte la necesidad de equilibrar el proceso evaluativo con otro tipo de evidencias.

Fink et. al., (1983) recomienda el considerar tres conjuntos de variables al evaluar la enseñanza; estas variables se relacionan con: el uso, las fuentes de información posibles y los componentes del proceso instruccional.

**Uso:** Se debe dejar en claro (tanto a los evaluadores como a quien es evaluado) la utilización que se dará a la evaluación; este aspecto debe considerarse al diseñar los instrumentos como también al planificarse su administración. Aquellas evaluaciones diseñadas para mejorar la docencia no deben usarse para tomar decisiones con respecto al personal docente. Asimismo, cabe señalar que la optimización del proceso evaluativo requiere de un seguimiento permanente.

**Fuentes de información.** Los estudiantes son sólo "una" fuente potencial de información; los estudiantes ya graduados, los pares y colegas, los jefes de departamento y decanos, pueden proporcionar información valiosa y combinada; más aún, el docente mismo puede proporcionar una perspectiva importante acerca de su nivel de docencia.

Cada fuente de información tiene potencialidades y limitaciones; cada una proporciona un marco referencial limitado y con posibles distorsiones y sesgos; sin embargo tomadas en conjunto, pueden proveer una base sólida que asegure decisiones justas y solventes.

**Componentes del proceso de docencia.** Las actividades del docente en la sala de clases constituyen tan solo un "ingrediente" de la enseñanza. Además de las actividades mismas, debe tomarse en cuenta el nivel de conocimientos, el manejo de las competencias necesarias, como asimismo los esfuerzos desplegados para el desarrollo, renovación o redireccionalidad de estas. Otros componentes importantes de la enseñanza son:

- \* las decisiones previas acerca de los textos a usar, el programa del curso, las lecturas exigidas, resúmenes, etc.
- \* las decisiones acerca del equilibrio que debe mantener la actividad docente (clases teóricas y prácticas; integración del trabajo en laboratorio o en terreno).
- \* las decisiones acerca de la naturaleza, tipo y frecuencia de los trabajos y controles.

- \* las decisiones en relación a los criterios de diagnóstico, evaluación y comunicación de los resultados.

Asimismo, debe prestarse atención a la "ejecución" de las decisiones consideradas anteriormente; entre estas se destaca:

- \* calidad de las clases teóricas
- \* calidad y pertinencia de las discusiones en clases
- \* integración de las lecturas personales a la clase misma
- \* comunicación efectiva de los objetivos del curso y los criterios de evaluación.
- \* diseño de las pruebas y controles
- \* interacción con los alumnos y disposición a solucionar conflictos

A continuación debe considerarse la calidad de los logros del alumno (trabajos, ejercicios, exámenes, informes de laboratorio, proyectos) lo que puede facilitar una evaluación flexible de la enseñanza al estar concientes de la complejidad de la relación entre la enseñanza y el aprendizaje (Seldin, 1980).

Finalmente debiera considerarse lo que podría llamarse "contexto de la enseñanza":

- \* tamaño del curso
- \* duración de la clase
- \* obligatoriedad del curso
- \* implementación física
- \* carga académica del profesor (semestral o anual)
- \* otras demandas del docente (investigación, actividades subvencionadas, prestación de servicios, etc.)
- \* factores personales

Es evidente que por cada componente de la docencia debiera incluirse una o más fuentes informativas. Los colegas y los directores de departamento, como también los decanos están en mejores condiciones de evaluar el conocimiento de un profesor que lo que pueden estar los alumnos de pregrado, siendo el profesor quien tiene los medios para informar de aquellas actividades que realiza para desarrollar ese conocimiento. Los colegas en el mismo programa o departamento están mejor calificados para evaluar la propiedad de los objetivos del curso, la selección de guías y textos, el perfil del curso, los requisitos y modelos de evaluación; a su vez los alumnos pueden informar si los objetivos del curso son claros y si perciben que ellos se cumplen a medida

que el curso avanza; también pueden proporcionar información acerca de la efectividad y claridad de la presentación del material del docente. Las autoridades en cambio pueden informar acerca del contexto en que se realiza el curso, como también acerca de otras actividades que ocupan el tiempo del profesor. Los colegas y alumnos de los últimos niveles pueden informar mejor acerca de la relación del curso con respecto al programa global, a la vinculación con otras asignaturas y con posibles trabajos a futuro.

Lo expuesto constituyen solo sugerencias. La idea básica apunta a que más allá del diseño de los instrumentos de evaluación, de la definición clara de los criterios y de la administración sistemática de ellos, (todos factores importantes) es esencial el tomar conciencia de la potencialidad y limitaciones de la gama de fuentes informativas que posee la docencia y del uso que puede dársele al proceso evaluativo. Se sabe que el calibrar las fuentes potenciales de información exige tiempo y meditación; sin embargo vale la pena el esfuerzo si con ello se logra una base sólida para formular juicios al respecto.

Con respecto al uso de la "auto evaluación" puede afirmarse que en general ha sido desestimada por los docentes. Al respecto, Elrich (1983; Cross, Centra, 1973; Clark y Blackburn, 1973) postulan que los profesores consistentemente se sobreestiman al utilizar instrumentos autoevaluativos. Los aportes de las investigaciones de los autores señalados más arriba indican que, existe una confiabilidad mayor en los criterios de medición cuando los mismos instrumentos se utilizan por los estudiantes y por los docentes, debido a que las discrepancias entre ambos tienden a ser más claramente discernibles. Marsh (1982) por ejemplo analizó los docentes de 329 cursos usando el mismo instrumento de evaluación (35 ítems) para los docentes (auto evaluación) y los alumnos. Los resultados indicaron que la evaluación de los estudiantes se correlacionaba significativamente con la autoevaluación del desempeño docente ( $r=.46$ ).

En 1984 Seldin concluyó a partir de sus estudios, que la autoevaluación podrían ser un medio importante para mejorar el desempeño del docente, ya sea por medio de la autosensibilización o a través de la identificación de las discrepancias que tuvieran lugar entre la evaluación del profesor y la del alumno. A partir de estas conclusiones se advierte la conveniencia del uso responsable de la autoevaluación junto a otras fuentes de información que pudieran aumentar la confianza y seguridad del docente en el proceso por el cual se llega a tomar decisiones de jerarquización o promoción.

El emitir juicios sobre materias complejas (y la docencia sin duda es una de las más complejas) no resulta tarea fácil, a raíz de lo cual los juicios emitidos no pueden cuantificarse en forma rápida; los intentos por hacerlo y

cuantificarlo constituyen en general esfuerzos disfrazados para entregar la responsabilidad a otros. Los estudiantes no toman decisiones acerca de la promoción y estabilidad de los profesores en sus cargos, como tampoco lo deciden los pares, las autoridades y el mismo docente. Cada uno de ellos sin embargo, pueden ofrecer evidencias, percepciones y perspectivas colectivas al momento de emitir y revisar aquellas decisiones. Confiar en un solo tipo de evidencia (generalmente la evaluación de los estudiantes) no sólo es ilegítimo sino que también es falto de ética.

De igual forma resulta poco ético, rechazar los esfuerzos docentes al tomar decisiones sobre su personal, aduciendo que la docencia es demasiado compleja de evaluar significativamente; una reflexión cuidadosa, tomando en cuenta tantas fuentes de información como sea posible (hasta llegar a un equilibrio pertinente), puede llevar a concluir el que una "docencia adecuada", constituye un componente vital de los esfuerzos profesionales en una facultad.

## BIBLIOGRAFIA

- Bonge, Dennis. "Using TA Ratings to Validate Student Evaluations: A Reply to Lambert and Kosteski". *Teaching of Psychology*, 9(2): 102, 1982.
- Centra, John. "Self-Ratings of College Teachers: A Comparison with Student Ratings". *Journal of Educational Measurement*, 10(4): 287-295, 1973
- Centra, John. *Determining Faculty Effectiveness*. San Francisco: Jossey-Bass Publishers, 1979.
- Centra, John. "The Fair Use of Student Ratings". *Proceedings of the Ninth International Conference on Improving University Teaching*. Dublin, Ireland. July 6-9, 1983.
- Centra, J.A. and F.R. Creech. "The Relationship Between Student, Teachers, and Course Characteristics and Student Ratings of Teacher Effectiveness". *Project Report 76-1*. Princeton, N.J.: Educational Testing Service, 1976.
- Clark, Mary Jo and Robert T Blackburn. "Faculty Performance Under Stress". *Faculty Effectiveness as Evaluated by Students*, Alan Sockloff, editor. Philadelphia: Temple University, 1973
- Cohen, Peter "Validity of Student Ratings in Psychology Courses: A Research Synthesis". *Teaching of Psychology*, 9(2), 1982

- Cross, K. Patricia. "Not Can But Will College Teaching Be Improved?", *New Directions for Higher Education: Reviewing and Evaluating Teaching*, John Centra, editor. San Francisco: Jossey-Bass, 1977.
- Doyle, K.O. and S.E. Whitely. "Student Ratings as Criteria for Effective Teaching", *American Educational Research Journal*, 11: 259-274, 1974.
- Ducette, Joseph and Jane Kenney. "Do Grading Standards Affect Student Evaluations of Teaching? Some New Evidence on an Old Questions", *Journal of Educational Psychology*, 74(3): 308-315, 1982.
- Eagle, N. "Studies in the Reliability and Validity of a Community College Student Rating of Instruction Questionnaire", *Community/Junior College Research Quarterly*, 1:303-314, 1977.
- Elrick, Mei-fei. "Methods of Evaluating Teaching", *Proceedings of the Ninth International Conference for Improving University Teaching*, Dublin, Ireland. July 6-9, 1983.
- Feldman, K. A. "Seniority and Experience of College Teachers as Related to Evaluations They Receive from Students", *Research in Higher Education*, 18(3): 124, 1983.
- Fincher, Cameron. "The Purpose and Functions of Faculty Evaluations. "Conference on Evaluation of Teaching Effectiveness, Georgia Center for Continuing Education, Sept. 8-9, 1982.
- Fink, Dee, et. al., *Instructional Improvement and Teacher Evaluation*. (The Faculty Senate Committee, University of Oklahoma, 1983).
- Frey, P. W. "Student Rating of Teaching: Validity of Several Ratings Factors", *Science*, 182: 83-85, 1973.
- Lambert, John and Debra Kostas. "Student Evaluations: An Assessment of Validity", *Teaching of Psychology*, 8(1): 8-11, 1981.
- Lambert, John and Debra Kostas. "Using TA Ratings to Validate Student Evaluations: The Important Issues", *Teaching of Psychology*, 9(2): 102-103, 1982.
- Marsh, Herbert. "Validity of Students' Evaluations of College Teaching: A Multi-Trait-Multi-Method Analysis", *Journal of Educational Psychology*, 74(2): 264-279, 1982.
- Marsh, H. W., H. Fleiner, and C. S. Thomas. "Validity and Usefulness of Student Evaluations of Instructional Quality", *Journal of Educational Psychology*, 67:833-839, 1975.

- Martin, R.R. "Correlation Between Instructor Evaluation and Student Test Scores", *Journal of Chemical Education*, 56: 461-462, 1979.
- Murray, H. G. "Low Inference Classroom Teaching Behaviors and Students Ratings of College Teaching Effectiveness", *Journal of Educational Psychology*, 75: 138 - 149, 1983.
- Orpen, C. "Student Evaluation of Lecturers as an Indicator of Instructional Quality: A Validity Study", *Journal of Educational Reserach*, 74: 5-7, 1980.
- Overall, Jesse and Herbert Marsh. "Students' Evaluations of Teaching: An Update", *Research Currents/American Association for Higher Education Bulletin*, 9 - 12, 1982.
- Rodin, M. and B. Rodin. "Student Evaluations of Teachers", *Science*, 177: 1164 - 1165, 1972.
- Schevrich, Venice, et. al. "Expected Grades vs. Specific Evaluations of the Teacher as Preditors of Students Overall Evaluation of the Teacher", *Research on Higher Education*, 19: 159 - 173, 1983.
- Seldin, Peter. *Successful Faculty Programs*. Crugers, N. Y.: Coventry Press, 1980.
- Seldin, Peter. *Changing Practices in Faculty Evaluation*. San Francisco: Jossey-Bass, 1984.
- Small, Arnold, et. al. "Effect of Emotional State on Student Ratings of Instructors", *Teaching of Psychology*, 9, 1982.
- Sommer, Robert. "Twenty Years of Teaching Evaluations: One Instructor's Experience", *Teaching of Psychology*, 8: 223 - 226, 1981.
- South, J. D., et. al. "Faculty Role Orientation and Perceived Teaching Effectiveness: The Case of an Intervening Variable", *Research in Higher Education*, 11: 99 - 110, 1979.
- Zelby, L. W. "Good Teaching: A Problem in Education", *Social Science*, 133 - 138, Summer, 1977.