

# APROXIMACIÓN PRAGMÁTICA A LA EVALUACIÓN DE LA UNIVERSIDAD

*por*

*Tomás Escudero Escorza*

Instituto de Ciencias de la Educación  
Universidad de Zaragoza

El presente trabajo pretende ser una aportación útil a esa interesante empresa de búsqueda de vías y procedimientos metodológicos eficaces para evaluar adecuadamente el quehacer de nuestras universidades, y de acercamiento progresivo a planteamientos evaluadores vigentes en otros sistemas universitarios. El objetivo de análisis es el rendimiento institucional, con sus diversas facetas y perspectivas de observación, pero somos conscientes de que no todos los terrenos de este rendimiento son igualmente explorables en el momento presente.

Estas reflexiones intentan partir desde los fundamentos y planteamientos de la investigación básica, antes de revisar los resultados empíricos más contrastados y los análisis crítico-teóricos con orientación práctica, para llegar finalmente al intento de sugerir propuestas viables de intervención evaluadora en nuestro sistema universitario, dentro del marco necesario del rigor científico.

## 1. RAZONES PARA LA EVALUACIÓN INSTITUCIONAL

Como ya señalaba hace años HOUSE (1973), algunos profesores universitarios siguen sin ver plenamente justificada la evaluación en su terreno profesional, cuando no existe en el caso de otros profesionales. Esta queja tiene alguna justificación, pero no responde totalmente a la realidad de los hechos, ya que todos los servicios públicos están permanentemente sujetos a una presión social evaluadora de muy diverso tipo e intensidad. Lo que ocurre es que la acción universitaria tiene unas especiales características en cuanto a dificultades definitorias, está en constante interacción con un colectivo crítico, es difícilmente controlable, genera expectativas especiales, etc., con lo que la presión social sobre su evaluación se hace más explícita que en otros sectores, aunque en la práctica, la acción evaluadora real no

sea, objetivamente, ni más sistemática, ni más rigurosa que la sufrida por profesionales de otros servicios públicos.

Desde la perspectiva científico-técnica parece incuestionable la importancia y la necesidad de la evaluación institucional como elemento básico para asegurar la “salud” del sistema, puesto que “una buena organización se caracteriza por poseer estructuras y procedimientos para detectar e identificar la existencia de problemas, describir las posibles soluciones, aplicarlas y medir su eficacia”. En definitiva, “un organismo sano es aquel que tiene capacidad y, de hecho, se autoevalúa permanentemente” (ESCUADERO, 1980: 74).

Poniéndose en la plataforma del planificador y del responsable universitario, desde la que se plantea como objetivo máximo la optimización del sistema en el manejo de los recursos legales, personales y materiales que ponen a su disposición los responsables políticos y sociales y diversas instituciones, se llega rápidamente a la conclusión de que la evaluación institucional es un elemento imprescindible en la organización universitaria, puesto que la búsqueda permanente de la optimización no solamente es imposible sin la evaluación sistemática, sino que, en rigor, ambos conceptos son difícilmente separables.

Tomando finalmente la perspectiva del profesor universitario, no nos queda otro remedio que reconocer que cada día es mayor la presión externa por la evaluación que, como apuntaba hace años HASTINGS (1973), se articula a través de estudiantes, el ciudadano en general, las instituciones y agencias financiadoras, los cuerpos y órganos legisladores, las administraciones, etc... Esta presión evaluadora, no siempre razonable, es difícil de contrarrestar, pero se puede y se debe orientar hacia el terreno en el que sea productiva para la universidad y sus miembros.

Si desde un punto de vista estratégico resulta prácticamente imposible mantener una posición contraria a la evaluación institucional, entendemos que, tácticamente, es poco inteligente ser beligerante ante la evaluación en sí, no solamente porque está regulada en la LRU, sino porque se va en contra de la dinámica de los sistemas universitarios más desarrollados (OS, W. VAN et al., 1987; BALL y HALWACHI, 1987; MILLER, 1988). Con esto no queremos decir que hay que estar a favor de cualquier evaluación, porque entendemos que hay que estar a favor, lógicamente, de la evaluación que sirve a la mejora y a la optimización del sistema. Si la evaluación no sirve para este fin, no tiene sentido llevarla a cabo, pues es costosa y se corren riesgos innecesarios.

## **2. CONTEXTUALIZAR DEBIDAMENTE LA EVALUACIÓN**

Las razones y circunstancias que acabamos de señalar nos empujan a reivindicar de nuevo como posturas más razonables, la superación de los debates centrados únicamente en la conveniencia o no de la evaluación, y el avance en la búsqueda de mejoras conceptuales y metodológicas en dicha evaluación (ESCUADERO, 1986).

La primera y, probablemente, principal tarea en este momento es la de vencer

reticencias y darle credibilidad a la evaluación institucional. Para ello, consideramos absolutamente imprescindible que la propia universidad, como institución, se adelante a la presión externa y tome un papel activo en el proceso evaluador. Este planteamiento ha sido tradicionalmente defendido por especialistas como GOLDSCHMID (1978) y BORICH (1985), quien dice que la única solución contra el miedo a la evaluación es la auto-evaluación. En esta misma línea se mueven DRENTH et al. (1986) cuando señalan que los problemas de la evaluación externa pierden relevancia ante un sistema elaborado de evaluación interna. Lógicamente, la institución que se auto-evalúa se adelanta a la presión evaluadora externa, conduciendo el proceso hacia los terrenos más eficaces para la institución y sus objetivos y poniendo, seguramente, freno a las experiencias de evaluación espontánea que de cuando en cuando se llevan a cabo en las universidades y que comportan muchos más riesgos que promesas de mejora (GOLDSCHMID, 1978; ESCUDERO, 1987).

Por otra parte, los profesores universitarios parecen ubicar sus resistencias y dudas mucho más en los propósitos y en la utilización de la evaluación que en el hecho de ser evaluados (ARUBAYI, 1987; RODRÍGUEZ ESPINAR, 1987; JORNET et al., 1987; "EL PAÍS EDUCACIÓN" de 29 de diciembre de 1987). Los profesores son reticentes, sobre todo, a la evaluación que no les reporta beneficios para su actividad profesional, a la evaluación que enfatiza la "advertencia", el "acoso público", pero que no sirve para mejorar el estado de las cosas.

Lógicamente, por su propia naturaleza, la auto-evaluación va a reforzar los aspectos formativos de la evaluación, lo que no es solamente aconsejable en este momento desde el punto de vista táctico, sino que es lo mejor para la universidad, para los estudiantes y para el sistema en conjunto. Además, la evaluación así planteada, no tiene otro remedio que desembocar en programas de intervención y perfeccionamiento, sobre todo en el terreno didáctico, algo que parece absolutamente necesario en estos momentos. La implicación institucional en un proceso de auto-evaluación expandirá el proceso fuera de los límites demasiado estrechos en los que ahora se ubican la mayoría de las experiencias de evaluación, esto es, la evaluación de los profesores a través de cuestionarios a los alumnos. Sin duda, los propios profesores exigirán la evaluación y mejora de otras muchas facetas de la vida universitaria.

En síntesis, creemos que un adecuado tratamiento del proceso evaluador en nuestras universidades nos debe conducir a una total implicación y una mayor iniciativa de las propias instituciones, al énfasis del enfoque formativo y al análisis del mayor número de aspectos y funciones, y todo esto, a través de un modelo de auto-evaluación permanente.

### **3. ENFOQUES EN LA EVALUACIÓN DEL RENDIMIENTO INSTITUCIONAL**

Partiendo de que la valoración completa del rendimiento de cualquier institu-

ción supone sopesar el valor de lo conseguido a la vista del punto de partida, de los recursos empleados y del proceso seguido, al plantearnos la evaluación del rendimiento institucional universitario debemos preguntarnos con LINDSAY (1982: 177), “con qué efectividad consigue la institución sus objetivos y con qué eficiencia usa sus recursos en el proceso”. Tras esta pregunta, probablemente el problema más grave se centra en la definición precisa de los objetivos institucionales.

Los conceptos de “rendimiento” y “efectividad” son con frecuencia tratados como sinónimos, sin embargo, tal como indica YORKE (1987), son sustancialmente diferentes. El “rendimiento” implica su contraste con unos criterios de referencia, esto es, con los objetivos previstos para la organización. Obviamente, cualquier proceso evaluador medianamente riguroso debe contemplar al menos la apreciación de la dimensión de “efectividad”, aunque en algunos casos y momentos concretos sea dificultoso superar el nivel del simple análisis descriptivo de los resultados.

La apreciación de la “eficiencia” supone el establecer la relación de los productos conseguidos con los recursos de entrada puestos a disposición de la institución. En suma, podemos decir que una institución es “efectiva” cuando alcanza sus objetivos y que mejora su “eficiencia” a medida que los consigue con menos recursos.

Los objetivos y límites de este trabajo nos obligan a no extendernos mucho más en este asunto, que el lector puede estudiar con mayor profundidad en un buen número de artículos más especializados en el tema (LINDSAY 1981 y 1982; YORKE, 1987; BALL y HALWACHI, 1987). Sin embargo, sí que conviene que hagamos hincapié sobre la importancia que tienen estas precisiones conceptuales en determinados momentos y en que no podemos olvidarnos de ninguna de las dimensiones al plantear la evaluación, aunque no podamos acercarnos a todas en todos los casos.

Todos los especialistas coinciden en que es muy difícil evaluar la eficiencia de la universidad y que, en determinados aspectos, es muy discutible que se pueda medir con rigor y que sea útil hacerlo; la evaluación del rendimiento institucional universitario tiene que apoyarse en gran parte en apreciaciones y análisis cualitativos. YORKE (1987: 17) va un poco más lejos y dice que “la valoración del rendimiento institucional es esencialmente un ejercicio cualitativo (aunque supone el uso de datos cuantitativos)”, ya que los juicios valorativos son parte sustancial del proceso en cuanto se analizan los primeros datos. En concreto, este autor cita el ejemplo del porcentaje de abandonos, cuya valoración positiva o negativa dependerá, entre otras cosas, del modelo de acceso y de los requerimientos de selección. Ahora bien, lo anterior no implica que los evaluadores se puedan olvidar del concepto de “eficiencia”, porque se puede caer en situaciones confusas como la que CROMBAG (1978) señala alrededor del concepto de “calidad de la educación”. Este autor dice que no se puede hablar de calidad de la educación analizando solamente los productos educativos, esto es, moviéndose solamente en el dominio de la efectividad, puesto que entonces de lo que se habla es de “calidad de los

graduados”, y no de “calidad de la educación” que es sinónimo de eficiencia. La calidad de los graduados no es razón suficiente para asegurar que existe una acción educativa de alta calidad. Con una selección rigurosa de alumnos, podemos asegurar la salida de buenos graduados, apoyados en una acción didáctica muy poco eficiente, esto es, en una acción educativa mediocre.

A pesar de lo que acabamos de señalar, es lógico que la práctica evaluadora se centre en el dominio de la efectividad, por las dificultades de cuantificación de muchos productos y entradas y porque como indica LINDSAY (1982), las estrategias de auto-evaluación de carácter formativo —que son las que con mayor frecuencia se usan en las universidades (ESCUADERO, 1986), y que hemos defendido anteriormente como las más viables y útiles para aplicar en este momento en nuestro país—, enfatizan la valoración de la efectividad. La mayor parte de los trabajos que hemos podido ver en revistas especializadas, relativos al análisis de costos (dimensión de eficiencia), terminan estableciendo hipótesis que, a nuestro entender, simplifican excesivamente la realidad.

Dado el estado de cosas y sin perder de vista la necesidad de la eficiencia, ya sería un paso trascendental el plantearse la evaluación en los términos que señala CAMERÓN (1984), esto es, a la búsqueda de la efectividad definida como expresión de ausencia de ineffectividad, de forma que se plantee como objetivo básico de la evaluación la identificación de los puntos del rendimiento que han descendido por debajo de niveles aceptables.

#### **4. LOS OBJETIVOS, LAS INDEFINICIONES Y EL PLURALISMO DE LA INSTITUCIÓN UNIVERSITARIA**

Las grandes dificultades teóricas y prácticas que conlleva la evaluación de las instituciones universitarias surgen fundamentalmente de la propia naturaleza de estas organizaciones. El pluralismo ecológico y de objetivos, así como la ambigüedad y diversidad en sus prioridades de actuación, hacen que sea extremadamente complejo el establecimiento de índices y criterios válidos de evaluación. Los problemas derivados de la “ambigüedad de metas”, así como otros derivados de las contradicciones internas propias de los sistemas de valores preponderantes entre los diferentes sectores universitarios, llevaron a COHEN y MARCH (1974) a acuñar la frase de que las universidades como organizaciones son una “anarquía organizada”. MOSES (1986) y BALL y HALWACHI (1987) profundizan en estas ideas antes de sugerir esquemas e indicadores para la evaluación institucional.

Puestos a medir, por ejemplo, la efectividad institucional, tenemos que plantearnos obligatoriamente la multiplicidad de funciones, docencia, investigación, otros servicios externos y otros servicios internos, y, lógicamente, en cada una de estas funciones deberemos establecer objetivos, indicadores y criterios definitorios de la institución y de su calidad. En este segundo nivel van a empezar a surgir las discrepancias de enfoque y la diversidad entre instituciones: énfasis docente, énfasis

sis investigador, orientación cultural, orientación profesionalizadora, enseñanza pre-graduada, enseñanza post-graduada, investigación básica, investigación aplicada, etc.

Además de estos debates clásicos, que producen balanceos en la orientación de las universidades a lo largo del tiempo y diferencias entre unas universidades y otras, existen otros problemas latentes que todavía dificultan más el proceso. Tal es el caso del pretendido antagonismo entre la libertad académica y cualquier tipo de control de la actividad docente e investigadora del profesorado, o las habituales discrepancias en el énfasis de objetivos entre los diversos sectores de la comunidad universitaria, que se acentúan si incluimos también en el análisis a las fuerzas sociales y a la propia Administración. Asimismo, el sentido de estas funciones, sobre todo de la referente a los servicios externos, es permanentemente cambiante y, al mismo tiempo, está muy condicionado por el contexto propio de cada universidad.

Este panorama funcional debe enmarcarse a la hora de la evaluación en una multiplicidad de niveles (DIPPERHOFER-STIEM, 1986), diferentes entre universidades y con elementos muy variados dentro de cada universidad; nos estamos refiriendo a individuos, cursos, disciplinas, carreras, departamentos, centros, etc. Asimismo, la evaluación no puede ser ajena al hecho de que la institución se apoya fundamentalmente en lo que FERNÁNDEZ et al. (1987) denomina como estratos de la evaluación, esto es, los alumnos, los profesores, y el resto de elementos estructurales y organizativos.

En estas circunstancias y en virtud de esta naturaleza plural, se comprende perfectamente que ROTEM y GLASMAN (1977:77) digan que la “universidad es una institución inconsistente; no es una comunidad sino varias; y sus límites son difusos. La universidad es el lugar de confrontación de ideas, valores y expectativas, mantenidas tanto por sus comunidades internas (estudiantes, profesores y administradores) como por numerosos segmentos de la sociedad (graduados, legisladores y ciudadanos)... El pluralismo tiene virtudes obvias. La universidad es una ciudad de variedad infinita, que ofrece un rango amplio de elecciones para los estudiantes y los profesores y numerosos servicios a la sociedad. Cuando hay que tomar decisiones sobre qué metas y qué intereses deben considerarse prioritarios, la magnitud de la complejidad es inmensa”.

## **5. INDICADORES DEL RENDIMIENTO INSTITUCIONAL UNIVERSITARIO**

La búsqueda de indicadores del rendimiento es una de las parcelas que más esfuerzo ha requerido y requiere entre los especialistas en evaluación de la universidad, aunque, lógicamente, por la complejidad global expuesta, por las dificultades técnicas existentes, por los muchos problemas pendientes de resolución y porque los intereses con los que se plantean las evaluaciones son muchas veces direcciona-

les, no es fácil encontrar en la literatura especializada trabajos de evaluación del rendimiento institucional de manera global, en todas sus facetas, niveles y circunstancias (FERNÁNDEZ et al., 1987). En particular, según el interés desde el que se plantea la evaluación, encontramos que la perspectiva social nos obliga al análisis de las características profesionales de los graduados y su distribución cuantitativa, la perspectiva del estudiante centra su análisis de calidad en problemas como abandonos, satisfacción académica, niveles de aprendizaje, etc., la preocupación del administrador y el político es el uso eficiente de recursos, la planificación de tareas y el control del rendimiento, etc. (OS, W. VAN et al., 1987).

Un problema adicional en la evaluación del rendimiento institucional es la diversidad en la naturaleza de muchos de los denominados indicadores y criterios de la evaluación. Se utilizan de manera bastante indiscriminada indicadores que son verdaderos productos o resultados, características que no pasan de ser aspectos estructurales o descriptivos de la institución y otros que son juicios de colectivos o individuos implicados en el sistema. La complejidad del problema obliga a que esto sea así, pero es necesario evitar la confusión de conceptos y deja claro que, en rigor, los verdaderos criterios e indicadores del rendimiento son los logros de la institución (del tipo que sean) y es con éstos con los que podemos acercarnos a la valoración de la efectividad y de la eficiencia de las que antes hemos hablado.

Esta precisión de conceptos no está en contra de que en la evaluación institucional en general, e incluso en la evaluación del rendimiento, se analicen y valoren otro tipo de indicadores que no sólo sirven para describir adecuadamente a la institución, sino que son condicionantes de su funcionamiento y, en definitiva, de sus realizaciones. En algunos contextos de la evaluación, pueden ser de mayor interés ciertos índices indirectos que determinados productos. Así por ejemplo, para algún tipo de valoración de un departamento universitario, puede que resulte de más interés el análisis de la comunicación interna entre sus miembros, que el dato del número de artículos publicados en el último quinquenio, por mucho que sea un índice directo, esto es, un producto institucional. Los indicadores indirectos toman una relevancia mucho mayor dentro del enfoque formativo que dentro del sumativo, puesto que están mucho más cercanos al terreno de la intervención para la mejora. Por ello, no hay duda de que en el enfoque de evaluación defendido en este trabajo —básicamente auto evaluación de carácter formativo—, no solamente tenemos que considerar también indicadores indirectos, sino que algunos de ellos son fundamentales para la evaluación.

En el proceso de definición de indicadores existen algunos intentos globalizadores pero en la mayoría de los casos se toma una perspectiva de análisis restringida. DE LA ORDEN (1987), por ejemplo, centrándose en la valoración de la “eficacia del profesorado universitario”, destaca los indicadores siguientes: medidas de aprendizaje de los estudiantes, juicio de los estudiantes, juicio de los colegas, juicio del director del departamento, número y calidad de tesis y tesinas dirigidas, proyectos de investigación financiados por entidades oficiales y privadas, número y calidad de publicaciones, número y calidad de ponencias y comunicaciones en reuniones

científicas y congresos nacionales y extranjeros y patentes aceptadas en los registros correspondientes. Como puede verse, este listado recoge indicadores de muy diversa naturaleza, tal como hemos comentado anteriormente.

Un informe del gobierno británico (GREEN PAPER, 1985) ofrece una serie de indicadores de rendimiento relacionados con distintos objetivos. Así por ejemplo, para la evaluación de la producción de profesionales cualificados por parte de las universidades, se sugieren algunos indicadores como los siguientes: costos por estudiante, graduados en diversas áreas, costos por graduado, calificaciones de entrada de los estudiantes y éxitos en el mercado de empleo.

Otro informe británico (JARRAT et al., 1985) ofrece una veintena de indicadores del rendimiento institucional universitario, divididos en tres bloques. Estos indicadores tienen una orientación muy global y aunque hay que elaborarlos más, nos parece un listado interesante. En concreto, los indicadores son los siguientes:

A. *Indicadores del rendimiento interno*: mercado de aspirantes (estudiantes nuevos), tasas de graduación y tipos de titulaciones, atractivo para estudios de post-gradado, tasas de éxito en los estudios de post-gradado, atracción de fondos de investigación, calidad de la enseñanza.

B. *Indicadores del rendimiento externo*: aceptación de los graduados en el mercado de trabajo, primer destino de los graduados, reputación externa, publicaciones y citas, patentes, inventos y asesoramientos, premios, distinciones, etc., ponencias en conferencias, congresos,...

C. *Indicadores del rendimiento en el funcionamiento*: costos de las unidades, proporciones estudiante/profesor, tamaño de las clases, opciones curriculares, carga del profesorado, disponibilidad de fondos bibliográficos, disponibilidades informáticas y de cálculo.

La simple lectura de estos indicadores, aun después de ser traducidos de manera no literal en todos los casos, nos muestra claramente que están condicionados por el sistema universitario al que hacen referencia. Es obvio que algunos de los índices anteriores no son tan relevantes en nuestro sistema universitario como en el británico y, sin embargo, podrían incluirse otros que sí lo son. Parece claro, por tanto, que es imprescindible contextualizar, modificar y adaptar a la institución los indicadores de rendimiento.

También con orientación de globalidad como en el caso anterior, pero tomando una perspectiva que entendemos más ligada a los aspectos organizativos, YORKE (1987) sugiere lo que él denomina once áreas interrelacionadas de rendimiento institucional, construidas a partir de los enfoques de evaluación de la efectividad de la organización de ANTIA (1976), CAMERÓN (1978) y LINDSAY (1981). Estas once áreas son las siguientes: consecución de las metas oficiales, adquisición de recursos, calidad y desarrollo del profesorado, currícula, calidad y desarrollo de los estudiantes, clima del "campus", investigación y asesoramiento, gestión institucional, relaciones con instituciones externas, habilidad para explorar el entorno, reputación institucional. Estas áreas, tal como están expresadas, no son utilizables por el evaluador y requieren un desarrollo en verdaderos índices de rendimiento.



Los indicadores deben buscarse con el apoyo de la literatura especializada, pero contextualizándolos al tipo de evaluación que se está haciendo y al tipo de institución universitaria que se está evaluando. Así por ejemplo, no tendría sentido evaluar una universidad organizada siguiendo el modelo de centros y carreras, utilizando un conjunto de indicadores de rendimiento institucional elaborados para la evaluación de una universidad organizada según el modelo departamental.

### 5.1. La infraestructura de datos

La utilización de muchos de los índices de rendimiento institucional no solamente viene condicionada por los problemas definitorios que hemos señalado, sino que en muchos casos existen problemas graves de infraestructura informativa, resultado muy difícil y costosa la obtención de datos rigurosos. En determinadas ocasiones, los excesivos costos de la evaluación rigurosa, nos llevan a cuestionarnos si tiene sentido la realización de algunos estudios, porque su utilidad puede ser mucho menor que la magnitud del esfuerzo necesario. Un ejemplo evidente es el estudio del fracaso y éxito escolar universitario. En primer lugar nos encontramos con el habitual problema de la diversidad de definiciones operacionales de los conceptos de fracaso y éxito escolar, que afecta a todos los niveles educativos, pero que se acentúa en la universidad, por tratarse de ciclos terminales, con enfoques académico-profesionales varios, con multitud de contenidos y exigencias curriculares y con una amplia gama de niveles de dificultad en función de los tipos de estudios y carreras. En segundo lugar, lo más probable es que nos encontremos con carencias en los datos básicos necesarios, bien porque existen lagunas, bien porque los datos disponibles sean de dudosa fiabilidad.

En los últimos años está cambiando rápidamente la infraestructura de datos en nuestras universidades, al introducirse con mayor o menor celeridad sistemas informáticos para el control de matrículas, expedientes, etc. Hasta ahora, un problema grave es la no existencia de datos fidedignos en muchas universidades sobre nuevos ingresos (matriculados) año a año en las distintas carreras. Cuando hablamos del control de nuevos ingresos nos estamos refiriendo a un control nominal, no solamente de número, puesto que si nos apoyamos exclusivamente en los números de alumnos, no se controlan las dobles matrículas, muchos de los abandonos son simples cambios de carrera, cambios de universidad, o abandonos temporales. Lógicamente, el control absoluto de este problema requiere una informatización completa en las universidades españolas, con una interconexión absoluta entre todas ellas, lo que nos lleva a plantearnos si las contrapartidas justificarían el esfuerzo de todo tipo que habría que realizar. Probablemente, en el momento presente sería más que suficiente la informatización en cada universidad y la obtención de índices rigurosos de éxito y fracaso de cada carrera y cada universidad. En la actualidad, con los problemas citados, tenemos sospechas fundadas de que, por ejemplo, los índices promedios de abandono que a menudo se indican para

nuestro sistema universitario, e incluso para determinadas universidades, son algo más elevados que la realidad; probablemente existe cierto inflamiento de los matrículados, debido seguramente a las dobles matrículas, y de los abandonos, debido seguramente a los traslados y cambios de carrera contabilizados como abandonos.

Este problema de la fiabilidad de los datos, que como ya hemos señalado será menor en el futuro, se da en todos los sistemas educativos, incluso en los más desarrollados, por lo que no resulta fácil encontrar estudios completos de éxito y fracaso universitario en las revistas especializadas extranjeras (ESCUADERO, 1985). En el sistema norteamericano por ejemplo, existen problemas similares a los que hemos expuesto para nuestro país, tal como describen recientemente Le COMPTE y GOEBEL (1987) en un magnífico artículo sobre el registro de abandonos escolares en los Estados Unidos.

## 6. EL BINOMIO DOCENCIA-INVESTIGACIÓN

La naturaleza multidimensional de las universidades les conduce con bastante frecuencia a situaciones en las que se produce competencia, o incluso confrontación, entre funciones, enfoques, etc., introduciendo problemas adicionales a los que hemos venido reseñando con anterioridad. Un ejemplo típico es el que hace referencia al binomio docencia-investigación.

La coexistencia de la docencia y la investigación en la universidad occidental actual, es mucho menos armoniosa de lo que pudiera desprenderse de las definiciones institucionales y, por supuesto, de lo deseable. Esta falta de armonía afecta decisivamente a la evaluación institucional, puesto que se produce fundamentalmente entre la estructura organizativa de la institución y su sistema de recompensas. BECHER y KOGAN (1980), por ejemplo, hablan de que los valores atribuidos a las universidades no concuerdan necesariamente con las prácticas y los criterios de selección. Aunque el profesorado es contratado para enseñar, la selección se basa en su capacidad y rendimiento en la investigación y, además, el sistema de recompensas favorece al investigador. Tal como indica MOSES (1986), la competencia docente no supone promoción y el profesorado percibe que existe una clara discrepancia entre lo que las universidades dicen que son sus metas y el sistema de recompensas a sus miembros por el trabajo hacia tales metas. Las recompensas son desiguales según las metas; se prima mucho más la investigación. Lógicamente, ante esta situación, el profesorado se orienta hacia actividades con mayor recompensa, a igualdad de esfuerzo.

Con independencia de las declaraciones de principios, no existe duda de que las universidades son "burocracias profesionales" dominadas básicamente, en lo que se refiere al proceso de evaluación, por los criterios de los grupos académico-profesionales (HIND et al., 1974), que dan una clara preponderancia de la labor investigadora a la hora de juzgar la actividad y el rendimiento de un profesor.

Por todo lo anterior, el profesorado universitario acepta con naturalidad la

evaluación de su tarea por colegas, dentro del sistema de valores y de recompensas descrito anteriormente. Por el contrario, el profesorado no acepta siempre de buen grado la evaluación que pueden defender los responsables administrativos del sistema y los estudiantes. Ante las pocas perspectivas promocionales que promete este tipo de evaluación, parte del profesorado la ve más como un peligro para su autonomía profesional que como una oportunidad para la mejora. En todo caso, quizás, la única manera de equilibrar el interés del profesorado entre la docencia y la investigación pasa por equilibrar el sistema de recompensas como producto de todo proceso de evaluación que se realice.

Todo lo que acabamos de comentar no hace sino dificultar todavía más la valoración docente e investigadora de los profesores universitarios, algo de por sí complejo y delicado, aunque se caiga en el artificio de separar totalmente ambos aspectos. Parece clara la existencia de interdependencias entre ambas funciones y que, por ejemplo, la práctica de separar los costos y productos de la investigación para poder valorar mejor la enseñanza conduce a una situación ficticia (ESCUDE-RO, 1986).

Dentro del sistema meritocrático tradicional que sustentan los grupos académicos, podría pensarse que indicadores del rendimiento de la investigación como los que se han citado en un apartado anterior, son más precisos y válidos que los utilizados para la valoración de la docencia. Sin embargo, muchos de tales índices no tienen una relación segura con el concepto de calidad; se suele reforzar en gran medida el criterio de validez interna de la investigación, pero no así la relevancia. LINDSAY (1982: 182) por ejemplo, al referirse a los indicadores del rendimiento investigador más utilizados, dice que: "...ninguno de estos refinamientos tiene una relación clara y directa con la calidad. Sin embargo,... estas medidas son las mejores y son las que se usan entre la comunidad académica. Al igual que en las medidas del rendimiento docente, la distancia entre las medidas deseables conceptualmente y las disponibles en la actualidad, no evita su uso en el proceso de toma de decisiones". El conocimiento de las deficiencias de nuestros actuales indicadores de calidad, nos debe llevar a un uso prudente —sin extrapolaciones indebidas— de los mismos, pero no a su eliminación del proceso evaluador, al menos mientras no dispongamos de indicadores superiores.

## 7. EL JUICIO DE LOS ALUMNOS SOBRE LA DOCENCIA

Los alumnos son la principal fuente de información para evaluar el rendimiento educativo de las universidades, ya que aportan información en un doble sentido: a) como elemento de referencia del análisis de los rendimientos educativos y b) como emisores de juicios desde la perspectiva de observadores y de participantes directos en la interacción didáctica. Este hecho conduce a múltiples indicadores de calidad relacionados con los alumnos, sin embargo este apartado lo queremos centrar en el procedimiento (indicador) de valoración de la docencia más habitual y más proble-

mático por las tensiones internas que puede producir. Nos referimos, lógicamente, a la valoración de la docencia con el juicio de los alumnos recogido a través de cuestionarios anónimos.

Es importante resaltar que intentamos referirnos a la evaluación de la docencia y no a la evaluación del profesorado, aunque somos conscientes del amplio rango de concomitancias y del uso indiscriminado que de ambos términos se hace, con más frecuencia de la debida, en la literatura especializada. El término de “evaluación de la docencia” creemos que resalta y se adapta mejor, que el de “evaluación de profesores” a los principios de la evaluación institucional defendidos a lo largo de estas reflexiones.

La evaluación de la docencia a través del juicio de los alumnos no solamente se ha generalizado en muchos países en las últimas décadas (MILLER, 1988), sino que se ha convertido en una línea pujante de investigación entre los especialistas en educación superior. Ciñéndonos solamente a trabajos de síntesis, podemos citar como referencia de esta realidad, el clásico libro de CENTRA (1979) sobre eficacia del profesorado, la revisión de ROTEM y GLASMAN (1979) sobre información evaluativa de los estudiantes universitarios, los meta-análisis de COHEN (1980 y 1981) sobre la eficacia de la información de los estudiantes y sobre su relación con el rendimiento académico, la revisión crítica de LEVINSON-ROSE y MENGES (1981) sobre estudios de intervención para la mejora de la enseñanza universitaria, la completa discusión de MARSH (1984) sobre la relevancia y diversos aspectos métricos de la evaluación de los estudiantes y la revisión de DUNKIN y BARNES (1986) acerca de la investigación sobre didáctica en la educación superior. En líneas generales, estos estudios de síntesis reconocen que el juicio de los estudiantes es, probablemente, el procedimiento más utilizado para valorar la docencia universitaria y el que cuenta detrás con un mayor cuerpo de doctrina y de resultados empíricos en el que fundamentarse y justificarse. Esto no evita que de cuando en cuando aparezcan trabajos en los que se defienden posturas críticas a este procedimiento. Dos recientes ejemplos vienen de la pluma de CRUSE (1987), quien destaca las malas características de la evaluación del profesorado a través de la opinión de los estudiantes y de POPHAM (1987), que hace una encendida defensa del juicio de profesionales en la evaluación del profesorado, en vez de la actual corriente (a través de los cuestionarios a estudiantes) que denomina con el término de “champagne”, porque se miden muchas cosas (burbujas), pero se evita, según este autor, el juicio global del profesor.

En conjunto, los estudios realizados nos inducen a pensar que se puede tomar cualquier decisión sobre cualquier aspecto didáctico, a partir del juicio de los alumnos como único criterio valorativo. Este criterio parece ser elemento fundamental, pero no es del todo fiable, ni válido, en todas las facetas, aunque parece ser globalmente, el más fiable y válido de los criterios posibles. Los problemas surgidos de la diversidad de resultados en la investigación con el juicio de los estudiantes, conducen mucho más a dudas sobre su validez que sobre su fiabilidad, sin embargo, esta situación es más grave con cualquier otro procedimiento evaluador.

DOWELL y NEAL (1982: 61), a pesar de ser críticos con la validez del juicio de los alumnos, se preguntan "...¿dónde están los estudios sobre la validez del proceso de evaluación del profesorado universitario por medio de colegas?", para terminar certificando la validez del juicio de los estudiantes, como forma de control del receptor de la docencia. Esta validez es más evidente en unos aspectos que en otros, por lo que el procedimiento debe enfatizarse para algunas cosas y no para otras y, en todo caso, complementarse con otro tipo de informaciones. En esta línea diferencial deben entenderse las consideraciones de COHEN (1981: 305) cuando dice que "podemos asegurar que los juicios de los alumnos sobre la instrucción son un índice válido de su efectividad. Los estudiantes distinguen muy bien a los profesores dependiendo de lo mucho o poco que han aprendido con ellos". En el mismo sentido se debe interpretar la afirmación de DUNKIN y BARNES (1986: 774) de que "hay ahora evidencia de la credibilidad de la evaluación de la enseñanza superior por los alumnos, en razón de su sensibilidad para detectar variaciones en los procesos de enseñanza". En general, los alumnos parecen ser los mejores evaluadores de la interacción didáctica propiamente dicha, y ahí es donde debe enfatizarse su utilización como elemento de evaluación institucional.

Donde parece existir un acuerdo general es en la consideración de que la docencia es multidimensional y que cualquier evaluación global que aspire a ser rigurosa debe distinguir entre las diversas dimensiones y tiene que apoyarse en diversos criterios y procedimientos de valoración que se complementen entre sí. En otras palabras, en el momento actual de desarrollo, la mejor evaluación posible de la docencia universitaria es necesariamente multidimensional en objetivos, indicadores y procedimientos.

### 7.1. La desmitificación de las encuestas a los estudiantes

A pesar de que el juicio de los estudiantes sigue ganando consideración como procedimiento sistemático para evaluar la docencia universitaria, algunos profesores siguen discutiendo el valor de esta práctica apoyándose, sobre todo, en lo que ALEAMONI (1987) denomina "mitos" que frecuentemente están lejos de la evidencia empírica.

Ya hemos apuntado que la evaluación institucional debe ganar credibilidad entre todo el profesorado para realizarse debidamente y, por lo tanto, nos parece fundamental deshacer algunos de estos mitos que no tienen fundamento y que resultan tan perjudiciales para la evaluación. Algunos ejemplos de este problema son los siguientes:

- a) Se dice que los estudiantes no pueden enjuiciar consistentemente a los profesores y a la docencia por su inmadurez y falta de experiencia, mientras que los estudios empíricos demuestran la estabilidad y consistencia de juicios.
- b) Se dice que los esquemas de evaluación de los estudiantes se apoyan fundamentalmente en la popularidad, amabilidad, amistad, sentido del humor, etc.,

del profesor, pero está demostrado que los estudiantes distinguen claramente entre las diversas dimensiones de la instrucción. Al profesor amistoso pero desorganizado, por ejemplo, los estudiantes le critican abiertamente esta característica.

- c) Se dice que los estudiantes no están capacitados para emitir juicios precisos sobre la docencia hasta que ha pasado cierto tiempo, o incluso hasta que han terminado los estudios universitarios, sin embargo, los estudios demuestran una notable similitud entre los juicios de alumnos y ex-alumnos.
- d) Se dice que las encuestas a los alumnos no son fiables ni válidas, pero es posible construir instrumentos con buena fiabilidad y razonable validez, si se usan formatos apropiados y se centran las cuestiones en las debidas dimensiones. En todo caso, esta crítica es más válida para cualquier otro procedimiento alternativo.
- e) Se dice que los juicios de los estudiantes están condicionados por factores diversos como tamaño de la clase, sexo de alumnos y profesores, horarios, rango del profesor, etc., pero no existen evidencias sistemáticas para la mayoría de estas opiniones.
- f) Se dice que los estudiantes juzgan mejor a sus profesores y a sus cursos según sean mejor o peor sus calificaciones o sus expectativas de calificación, pero los resultados de la investigación no son suficientemente concluyentes en dicho sentido. Muchos trabajos tienden a confirmar la relación, pero en la mayoría de los casos parece ser relativamente débil. Además, esta relación entre calificaciones y juicios podría ser lógica por muchos más factores que por el simple efecto de la expectativa de calificación. Desde otra perspectiva, se puede asegurar que no es muy sostenible la crítica de que la evaluación de los alumnos conduce a una elevación incorrecta de las calificaciones. De acuerdo con las evidencias empíricas, esto podría producirse en un principio en algunos casos, pero no tendría efecto una vez estabilizado el procedimiento. Además, nunca influiría en el juicio sobre otras dimensiones evaluadas no directamente relacionadas con la rigurosidad de las calificaciones. La investigación parece dejar claro que los estudiantes no dan buenas calificaciones por el simple hecho de tratarse de cursos fáciles donde no tienen mucho trabajo (WATKINS et al., 1987).

## 7.2. Dimensiones en la evaluación de la docencia por los estudiantes

Ya hemos señalado que la multiplicidad de dimensiones en la evaluación de la docencia es algo comúnmente defendido por los especialistas. Por ello, no es de extrañar que la búsqueda de dimensiones o facetas de la docencia para evaluar por los estudiantes, sea una constante preocupación de los investigadores, apoyándose, generalmente, en el análisis estadístico multivariado (análisis factorial, “cluster”, etc.).

Las dimensiones encontradas por distintos investigadores están condicionadas por el enfoque de partida, la muestra, el nivel educativo, etc., e incluso por la metodología de análisis empleada. En conjunto, se puede hablar de gran diversidad en los resultados, pero también de la aparición sistemática —casi permanente— de algunas dimensiones que parecen ser claves en todo proceso de evaluación de la docencia, especialmente cuando se utiliza el juicio de alumnos. Una selección de las dimensiones y subdimensiones de la docencia que aparecen de manera más frecuente en los estudios especializados, podría ser la siguiente: a) organización y planificación de la enseñanza, b) interacción alumnos/profesor, c) ambiente del grupo de clase, d) preparación profesional del profesor (académica y didáctica), e) metodología y recursos didácticos, f) claridad en la exposición didáctica, g) motivación y participación del alumnado, h) calidad, amplitud, dificultad, etc., del programa, i) evaluación de los alumnos, j) rendimiento global de los alumnos, k) profesionalidad del profesor (cumplimiento, preparación de las clases, etc.).

Estas facetas o dimensiones citadas —con esta u otra denominación— suelen ser recogidas de manera aislada o integradas con otras, en la gran mayoría de los instrumentos de evaluación de la docencia por los estudiantes, que se emplean por las universidades y que aparecen en los trabajos de investigación (pueden tomarse como ejemplo los trabajos de: HILDEBRAND et al., 1971; WARRINGTON, 1973; FREY et al., 1975; HASLETT, 1976; COHEN, 1981; MARSH, 1982; DE NEVE y JANSSEN, 1982; MEREDITH, 1983; BANZ y RODGERS, 1985; POONYAKANOK, et al., 1986; STREIFER e IWANICKY, 1987 y TEJEDOR, 1987). Parece lógico, por tanto, pensar en estas dimensiones a la hora de seleccionar o elaborar cuestionarios que se puedan utilizar en nuestras universidades. Además de lo anterior, también suelen recoger las encuestas las valoraciones globales del curso y del profesor, que son dos facetas que no tienen fácil interpretación. Otras veces, estas valoraciones globales se obtienen como promedios de facetas parciales, pero esto tampoco tiene un claro significado.

El análisis comparado demuestra que los esquemas de evaluación de la docencia por medio del juicio de los alumnos son parecidos en distintos países y diversos contextos (MARSH et al., 1985; WATKINS et al., 1987), aunque se pueden producir distintas ordenaciones y énfasis en las dimensiones según el contexto y el modelo universitario (DAWOUD, 1985). En consecuencia, siempre con la prudencia requerida, parece científicamente apropiado el adaptar instrumentos y modelos estudiados y utilizados con éxito en otras universidades y otros sistemas universitarios. Este planteamiento puede ser ventajoso y razonable en las primeras experiencias de evaluación de este tipo, antes de que cada universidad vaya adaptando el procedimiento de manera más precisa y elaborando sus propios cuestionarios.

## 8. SUGERENCIAS FINALES

Una escueta síntesis de lo que consideramos son algunas de las propuestas

básicas con orientación práctica, que se desprenden de los planteamientos defendidos de este trabajo, es la siguiente:

- a) Las universidades deben tomar la iniciativa en el proceso de evaluación institucional, con enfoques de auto-evaluación con orientación formativa. Ésta es, probablemente, la única forma de controlar la presión externa por la evaluación que se plantea, a menudo, de forma peligrosa, y, además, es el enfoque con más potencialidad para la mejora.
- b) Parece necesario evitar la evaluación que no sea o puede ser útil para la optimización de la institución universitaria. Evaluar solamente para comprobar que hay problemas es, generalmente, contraproducente, porque crea tensiones sin producir beneficios.
- c) Es fundamental la aceptación del modelo evaluador por todos los sectores afectados, sobre todo los profesores. En otras palabras, la evaluación debe tener credibilidad y crear expectativas positivas en los afectados. Sin duda, planteamientos como los defendidos en los dos puntos anteriores beneficiarían esta necesaria aceptación.
- d) Hay que tener presente que la comprobación de la eficiencia institucional debe ser un objetivo fundamental de la evaluación, aunque resulte una meta inabordable en muchos casos. Esto obligará a superar el simple nivel de descripción de rendimientos y resultados y a llegar mucho más que en la actualidad hasta la dimensión de la efectividad.
- e) A pesar de los problemas definicionales y técnicos existentes, la investigación sobre indicadores de calidad y de rendimiento es una fuente fundamental para los evaluadores. Aunque existe cierta mezcolanza de conceptos diversos en los indicadores más habituales, en el momento presente parece razonable seguir utilizando indicadores de todo tipo, abarcando el máximo número de aspectos funcionales y de rendimiento universitario. La literatura especializada debe marcar las orientaciones iniciales en este sentido, con el consiguiente proceso de contextualización.
- f) Se hace necesaria la informatización de los ingresos de nuevos alumnos (con seguimiento nominal), si se quiere tener datos fiables sobre abandonos, ritmos de estudio, etc. En este proceso hay que ser prudente, puesto que la consecución de algunos datos verdaderamente fiables puede ser demasiado costosa en relación con su utilidad.
- g) La evaluación de la docencia debe servir para potenciar su estatus relativo dentro de las funciones universitarias. Si dicha evaluación se aborda con el enfoque que se plantea en este trabajo, pensamos que habrá cierta incidencia en una mayor valoración de la función docente.
- h) La aplicación de cuestionarios de evaluación de la docencia (mejor que de evaluación del profesorado) a los alumnos es una práctica cada vez más generalizada en la mayoría de los sistemas universitarios desarrollados. Entendemos que no es tácticamente razonable que el profesorado se oponga a su utilización, porque crecerá la presión externa. Lo más sensato y útil parece



ser el control de tal procedimiento dentro de su verdadera potencialidad y sentido. En concreto, hay que evitar el uso indebido de los resultados y que los cuestionarios y el proceso de aplicación se ajusten al debido rigor científico.

- i) Es necesario ampliar el marco de la evaluación institucional a todo el rango de funciones universitarias (internas y externas) y evitar que siga demasiado ceñida a la evaluación de la docencia y el profesorado. Por otra parte, entendemos que si la evaluación de la docencia se lleva a cabo adecuadamente (aceptada y auspiciada por toda la institución), esto obligará necesariamente a la evaluación de otras facetas y aspectos.

## REFERENCIAS

- ALEAMONI, L. M.; 1987: Student rating myths versus research facts, *Journal of Personnel Evaluation in Education*, 1: 111-119.
- ANTIA, J. M.; 1976: Critical success factors in polytechnic performance, *Educational Administration*, 5: 14-32.
- ARUBAYI, E. A.; 1987: Improvement of instruction and teacher effectiveness: are student ratings reliable and valid?, *Higher Education*, 16: 267-278.
- BALL, R. y HALWACHI, J.; 1987: Performance indicators in higher education, *Higher Education*, 16: 393-405.
- BANZ, M. L. y RODGERS, J. L.; 1985: Dimensions underlying student ratings of instruction: A multidimensional scaling analysis, *American Educational Research Journal*, 22: 267-272.
- BECHER, A. y KOGAN, M.; 1980: *Process and structure in Higher Education*, Heinemann, London.
- BORICH, G. D.; 1985: Needs assessment and self-evaluating organization, *Studies in Educational Evaluation*, 11: 205-215.
- CAMERÓN, K.; 1978: Measuring organizational effectiveness in institutions of higher education, *Administrative Science Quarterly*, 23: 603-632.
- 1984: Assessing institutional ineffectiveness: a strategy for improvement, en R. A. Scott (Ed.), *Determining the effectiveness of campus services*, Jossey-Bass, San Francisco, 67-83.
- CENTRA, J. A.; 1979: *Determining faculty effectiveness*, Jossey-Bass, San Francisco.
- COHEN, P. A.; 1980: Effectiveness of student-rating feedback for improving college instruction: A meta-analysis of findings. *Research in Higher Education*, 13: 321-341.
- 1981: Student ratings of instruction and student achievement: A meta-analysis of multisection validity studies, *Review of Educational Research*, 51: 281-309.

- COHEN, M. D. y MARCH, J. G.; 1974: *Leadership and Ambiguity. The American College President*, McGraw Hill, New York.
- CROMBAG, H. F. M.; 1978: On defining quality of education, *Higher Education*, 7: 389-403.
- CRUSE, D. B.; 1987: Student evaluation and the university professor, *Higher Education*, 16: 723-737.
- DAWOUD, M.; 1985: La compétence du professeur d'université. Étude comparative, *Revue de Psychologie Appliquée*, 31: 1-17.
- DE NEVE, J. M. J. y JANSSEN, P. J.; 1982: Validity of student evaluation of instruction, *Higher Education*, 11: 543-552.
- DIPPELHOFER-STIEM, B.; 1986: How to measure university environment? Methodological implications and some empirical findings, *Higher Education*, 15: 475-495.
- DOWELL, D. A. y NEAL, J. A.; 1982: A selective review of the validity of student ratings of teaching, *Journal of Higher Education*, 53: 51-62.
- DRENTH, P. J. D., OS, W. VAN Y BERNAERT, F. G., 1986: Improvement of quality of education through internal evaluation, *International Journal of Institutional Management in Higher Education*, 10: 275-281.
- DUNKIN, M. J. y BARNES, J.; 1986: Research on teaching in higher education, en M. C. Wittrock, *Handbook of Research on Teaching-Third Edition*, Mac Millan Publ. Co.: New York, 754-777.
- EL PAÍS EDUCACIÓN (29 de diciembre de 1987), artículos sobre evaluación y control de la docencia universitaria.
- ESCUADERO, T.; 1980: ¿Se pueden evaluar los centros educativos y sus profesores?, *Educación Abierta* 10, ICE de la Universidad de Zaragoza.
- 1985: Aproximaciones al estudio del fracaso escolar en la universidad española, en *Fracaso Escolar*, Depto. de Sanidad, Bienestar Social y Trabajo, Diputación General de Aragón, Zaragoza: 416-436.
- 1986: Algunos criterios y evidencias del rendimiento universitario, en M. Latiesa, (Comp.), *Demanda de Educación Superior y Rendimiento Académico en la Universidad*, CIDE, Madrid: 187-204.
- 1987: Experiencias evaluativas en la Universidad de Zaragoza, ponencia presentada en las *I Jornadas Nacionales sobre evaluación y mejora de la docencia universitaria*, Valencia, 13-15 mayo.
- FERNÁNDEZ, J.; LILLO, B. y MARTÍNEZ ARIAS, M. R.; 1987: Proyecto de evaluación de la educación superior española, ponencia presentada en las *I Jornadas Nacionales sobre evaluación y mejora de la docencia universitaria*, Valencia, 13-15 mayo.
- FREY, P. W.; LEONARD, D. W. y BEATTY, W. W.; 1975: Student ratings of instruction: validation research, *American Educational Research Journal*, 12: 327-336.
- GOLDSCHMIND, M. L.; 1978: The evaluation and improvement of teaching in higher education, *Higher Education*, 7: 221-245.

- GREEN PAPER; 1985: The development of higher education into the 1990's, HMSO, Londres.
- HASLETT, B. J.; 1976: Dimensions of teaching effectiveness: A student perspective, *The Journal of Experimental Education*, 44: 4-10.
- HASTINGS, J. T.; 1973: Evaluation of instruction, en E. R. House (Ed.), *School Evaluation. The Politics and Process*, McCutchan Publishing Co. Berkeley: 140-145.
- HILDEBRAND, M.; WILSON, R. C. Y DIENST, E. R.; 1971: *Evaluating university teaching*, Center for Research and Development in Higher Education, University of California, Berkeley.
- HIND, R. R.; DORNBUSCH, S. M. y SCOTT, W. R.; 1974: A theory of evaluation applied to a university faculty, *Sociology of Education*, 47: 114-128.
- HOUSE, E. R. (ED.); 1973: *School Evaluation. The Politics and Process*, McCutchan Publishing Corporation, Berkeley.
- JARRAT, A. et al.; 1985: *Report of the Steering Committee for Efficiency Studies in Universities*, Committee of Vice Chancellors and Principals, Londres.
- JORNET, J. M.; VILLANUEVA, P.; SUÁREZ, J. M. y ALFARO, I.; 1987: Proyecto y propuestas para la implantación de un sistema de evaluación del profesorado en la Universitat de Valencia, ponencia presentada en las *I Jornadas Nacionales sobre Evaluación y mejora de la docencia universitaria*, Valencia, 13-15 mayo.
- LE COMPTE, M. D. y GOEBEL, S. D.; 1987: Can bad data produce good program planning? An analysis of record-keeping of school dropouts, *Education and Urban Society*, 19: 250-268.
- LEVINSON-ROSE, J. y MENGES, R. J.; 1981: Improving College Teaching: a critical review of research, *Review of Educational Research*, 51: 403-434.
- LEY ORGÁNICA 11/1983, de 25 de agosto, de Reforma Universitaria. (B.O.E. de 1 de septiembre de 1983).
- LINDSAY, A.; 1981: Assessing institutional performance in higher education: a managerial perspective, *Higher Education*, 10: 687-706.
- 1982: Institutional performance in higher education: the efficiency dimension, *Review of Educational Research*, 52: 175-199.
- MARSH, H. W.; 1982, SEEQ: a reliable, valid, and useful instrument for collecting students' evaluations of university teaching, *British Journal of Educational Research*, 52: 175-199.
- 1984: Students' evaluations of university teaching: dimensionality, reliability, validity, potential biases, and utility, *Journal of Education Psychology*, 76: 707-754.
- MARSH, H. W.; TOURON, J. y WHEELER, B.; 1985: Students' evaluations of university instructors: the applicability of american instruments in a spanish setting, *Teaching & Teacher Education*, 1: 123-138.
- MEREDITH, G. M.; 1983: Diagnostic value of composite student-based ratings of instruction, *Psychological Reports*, 52: 549-550.

- MILLER, A. H.; 1988: Student assessment of teaching in higher education, *Higher Education*, 17: 3-15.
- MOSES, I.; 1986: Using organisational theory in the promotion of evaluation, a case study, *Higher Education*, 15: 619-639.
- ORDEN A. DE LA; 1987: Formación, selección y evaluación del profesorado universitario, *Bordón*, XXXIX: 5-29.
- OS, W. VAN, DRENTH, P. J. D. y BERNAERT, G. F.; 1987: "AMOS: an evaluation model for institutions of higher education, *Higher Education*, 16: 243-256.
- POPHAM, W. J.; 1987: The shortcomings of champagne teacher evaluations, *Journal of Personnel Evaluation in Education*, 1: 25-28.
- POONYAKANOK, P., THISAYAKORN, N. y WINGFIELD DIGBY, P.; 1986: Student evaluation of teacher performance: some initial research findings from Thailand, *Teaching and Teacher Education*, 2: 145-154.
- RODRÍGUEZ ESPINAR, S.; 1987: La evaluación del profesorado universitario. Análisis de la opinión del profesorado de la Universidad de Barcelona, ponencia presentada en las *I Jornadas Nacionales sobre evaluación y mejora de la docencia universitaria*, Valencia, 13-15 mayo.
- ROTEM, A., y GLASMAN, N. S.; 1977: Evaluation of university instructors in the United States: the context, *Higher Education*, 6: 75-92.
- 1979: On the effectiveness of students' evaluative feedback to university instructors, *Review of Educational Research*, 49: 497-511.
- STREIFER, P. A. y IWANICKI, E. F.; 1987: The validation of beginning teacher competencies in Connecticut, *Journal of Personnel Evaluation in Education*, 1: 33-55.
- TEJEDOR, J.; 1987: La evaluación del profesorado en la Universidad de Santiago de Compostela, ponencia presentada en las *I Jornadas Nacionales sobre evaluación y mejora de la docencia universitaria*, Valencia, 13-15 mayo.
- WARRINGTON, W. G.; 1973: Student evaluation of instruction at Michigan State University, en A. L. Sockloft (Ed.), *Proceedings: The First Invitational Conference on Faculty Effectiveness as Evaluated by Students*, Measurement and Research Center, Temple University, Philadelphia: 164-182.
- WATKINGS, D., MARSH, H. W. y YOUNG, D.; 1987: Evaluating tertiary teaching: a New Zealand perspective, *Teaching & Teacher Education*, 3: 41-53.
- YORKE, D. M.; 1987: Indicators of institutional achievement: some theoretical and empirical considerations, *Higher Education*, 16: 3-20.