



# Reforma y evaluación de la universidad

Guillermo Quintás, ed.

*Guillermo Quintás, ed.*

REFORMA Y EVALUACIÓN  
DE LA UNIVERSIDAD

UNIVERSITAT DE VALÈNCIA  
1996

---

# *Evaluación de la Actividad Universitaria*

---

*J. M. Jornet, J. M. Suárez,  
J. González y A. Pérez<sup>1</sup>*  
Universitat de València

La Evaluación de la Actividad Universitaria ha constituido uno de los focos de interés en nuestro país desde hace poco más de una década. Estimulados desde la aparición de la LRU y los Estatutos de las Universidades, se han desarrollado múltiples planes de evaluación sobre la Universidad Española, desde las actuaciones de sus docentes y su actividad investigadora hasta valoraciones globales de la Universidad como institución.

La Universidad, como servicio público, ha sido la primera en asumir este tipo de procesos de control y/o revisión de su actuación, inserta en un Estado de Derecho que, como mínimo, pretende analizar el aprovechamiento de los recursos aportados por los ciudadanos y mejorar la gestión.

Como actividad, la Evaluación de la Universidad es un tema controvertido a todos los niveles (desde su planteamiento teórico y metodológico hasta sus implicaciones sociales y políticas). Fruto de esta controversia es la multiplicidad de enfoques, puntos de vista y soluciones que se han ido desarrollando en nuestro país.

Esta presentación la estructuramos sobre tres aspectos que ayudan a globalizar el análisis del problema:

- a) Enfoques de la Evaluación de la Actividad Universitaria
- b) Contextualización de la Evaluación de la Actividad Universitaria
- c) Utilización de la Evaluación de la Actividad Universitaria.

En cada uno de los apartados mencionados intentamos aportar algunas reflexiones sobre los problemas y soluciones que se han ido produciendo para abordar la Evaluación de la Actividad Universitaria. Se analizan los componentes

1. Profesores del Dpt. de Mètodes d'Investigació i Diagnòstic en Educació de la Universitat de València.

teórico-metodológicos y aplicados, que entendemos más relevantes en este problema.

La Evaluación puede constituir un instrumento que oriente a la innovación y mejora, o bien puede convertirse en un elemento disruptor en el desarrollo normal de cualquier proceso educativo. La introducción de planes de evaluación sobre diversos aspectos de la vida universitaria de nuestro país está produciendo efectos muy variados. Así, intentamos revisar las consecuencias de la puesta en marcha de planes de evaluación del profesorado (Docencia e Investigación), de las Instituciones Universitarias, etc...

Estos temas, de una u otra forma, han producido cierto grado de convulsión en la vida universitaria española. En la base de los efectos que producen están sin duda las opciones metodológicas que se han elegido para su desarrollo; las cuales devienen de las orientaciones (o desorientaciones) políticas que los han impulsado.

## **1. Enfoques de la Evaluación de la Actividad Universitaria**

La reforma de la Educación Superior en España abierta con la promulgación de la LRU, desarrollada en los diversos Estatutos de nuestras Universidades y que culminará con la reforma completa del catálogo de titulaciones que se oferta desde cada una de ellas ha conllevado una preocupación creciente en todos los estamentos universitarios acerca de la calidad de la enseñanza.

Junto a otras acciones, la evaluación de la actividad de las universidades se ha señalado como un instrumento dinamizador y orientador necesario en este proceso de renovación. Así, en diversas universidades del Estado se han puesto en marcha acciones evaluativas que atienden aspectos concretos de su acción (actividad docente y/o investigadora del profesor, calidad de programas, etc...). En este marco se han sumado los debates que se han ido desarrollando tanto en foros europeos como nacionales acerca de la evaluación de la calidad educativa de las universidades y que incluye diversos puntos de vista, desde los relativos a la calidad de la docencia impartida por el profesorado, hasta los referidos a la calidad de titulaciones, servicios o la universidad globalmente considerada. Asimismo, en nuestro país se han venido llevando a cabo desde principios de la década de los ochenta diferentes procesos evaluativos en las universidades, de forma que se ha configurando un panorama de usos de planes de evaluación con diversos grados de institucionalización.

Ahora bien, ¿qué enfoques se dan acerca de la evaluación de la calidad universitaria?

Una de las opciones que pueden utilizarse para estructurar los enfoques de Evaluación de la Docencia Universitaria es la que se deriva de la confluencia de las dimensiones que definen un proceso de evaluación (objeto, propósito, audiencia, énfasis, criterios, etc...). Por ello hemos seleccionado tres dimensiones desde las que situar en un solo espacio los principales enfoques de Evaluación Universitaria (JORNET, 1991):

a) Control de la Evaluación. Hace referencia a quien promueve y/o desarrolla y/o utiliza la Evaluación. Los polos de referencia son externo/interno. Por ejemplo, el Ministerio de Educación promueve la evaluación del profesorado (en este caso, se entiende que el que orienta la evaluación es externo), o bien, un profesor con objeto de mejorar su calidad docente promueve un proceso de evaluación de su actividad (en este caso, se entiende que quien orienta la evaluación es interno).

b) Objeto de la Evaluación. Se refiere a qué se evalúa, teniendo en cuenta graduaciones según énfasis, nivel de responsabilidad y/o unidad de análisis seleccionada. Los polos de referencia son macroanalíticos/microanalíticos, molar/molecular. Por ejemplo, la evaluación de una Universidad corresponde a un planteamiento molar, macroanalítico; mientras que la evaluación acerca de cómo imparte las clases un determinado profesor, corresponde a un planteamiento molecular, microanalítico.

c) Finalidad de la Evaluación. Está referido al propósito, utilización y objetivo de la misma. Los polos de referencia son, pues, formativa/sumativa, orientada al desarrollo-mejora profesional o innovación educativa, o bien si está orientada a la rendición de cuentas, la distribución de recursos, etc. Por ejemplo, si en la evaluación del profesorado el objetivo es aportarle información para que éste mejore su actuación, el uso es formativo; mientras que si ésta se orienta a premiar o castigar al profesor (vg., con una subida salarial), el uso es sumativo.

En cualquier caso, las polaridades aquí señaladas no tienen carácter exclusivo y únicamente pueden ser bien definidas respecto a un plan concreto de evaluación, por lo que en este caso tienen el valor de orientar el marco en el que situar genéricamente los procesos de evaluación.

Basándonos únicamente en el cruce de las dimensiones indicadas identificamos los siguientes grandes enfoques: a) Institucionales, b) De Programa y c) De Profesorado. Dentro de los cuales se contienen, a su vez, dos finalidades (Formativa-Mejora y Sumativa-Rendición de Cuentas). En cuanto al control de la evaluación, generalmente –aunque no siempre– tiende a coincidir con la finalidad, asociándose el control externo a la rendición de cuentas y el interno a la mejora.

Obviamente, estos tres grandes enfoques no constituyen opciones mutuamente excluyentes para el desarrollo de programas de evaluación universitaria. Es más, la experiencia demuestra la coexistencia de diversos planes de evaluación para diferentes propósitos (FURUMARK, 1981; BAUER, 1988; BARNETT, 1989; VROEIJENSTIJN y ACHERMAN, 1990), así como es común encontrar programas que por el cruce de éstas –y otras– dimensiones podrían ser contemplados en dos o más categorías (BANTA et al., 1986). Téngase en cuenta que en la definición de cualquier plan de evaluación influyen múltiples dimensiones –objeto, fin, promotor, demandas sociales, ideológicas de todos los implicados, posibilidades de implementación, etc.– que, combinados en su adaptación a la diversidad organizacional de cada universidad ofrecen un panorama bastante diferenciado de actuaciones evaluativas concretas.

## 1.1 Enfoques Institucionales

Este tipo de enfoques se orienta hacia el concepto “Calidad Educativa” y tienen una gran incidencia en la evaluación de la actividad universitaria en el panorama internacional. En los planes de evaluación institucional se pretende medir la acción global de la Universidad y, dependiendo del tipo de los objetivos y finalidad de la evaluación, los servicios y/o los programas y/o los profesores, así como la actuación de los órganos de gobierno y gestión. Dependiendo de la finalidad con que se desarrollen, podemos clasificarlos en formativos (de Auto-revisión), sumativos (de rendición de cuentas) y mixtos.

Los *Enfoques de Autorrevisión Institucional* se basan en modelos orientados al cambio y pretenden identificar los elementos organizacionales que orienten hacia innovaciones efectivas. Consideran la institución a nivel organizativo, identificando planos de responsabilidad, áreas de gestión, unidades de actuación, etc., así como pretenden considerar la evaluación de la institución en el marco social en que ésta debe desarrollar su labor.<sup>2</sup>

Aunque de carácter holista, se orientan hacia la determinación de propiedades y la planificación del cambio en acciones a corto, medio y largo plazo, seleccionando áreas prioritarias o de necesidades (FURUMARK, 1981; DOCHY et al. 1990). Asumen, pues, que la innovación y la mejora institucional –objetivo explícito de la evaluación en estos enfoques– para que sea profunda y estable deberá enraizarse en el cambio institucional, respetando el ritmo e integrando los elementos –estructurales y funcionales– que toda institución tiene posibilidad de dinamizar.

Elementos fundamentales para el desarrollo de este tipo de planes son:

- a) La sensibilización hacia la calidad debe realizarse desde los órganos de gobierno de la propia institución, como medio de dinamización de la participación y de integración de los individuos y los sectores en los planes de evaluación. Téngase en cuenta que los procesos de autorrevisión institucional se han de concebir como planes voluntarios de autoevaluación, por lo que es fundamental conseguir la participación de todos los estamentos de la institución.
- b) La autovaloración en los estamentos implicados –personales u organizacionales– como instrumento básico de evaluación. El esquema de desarrollo de este tipo de planes de evaluación contempla la creación de comités de compañeros, junto a expertos en evaluación y los sujetos implicados, que recogen, analizan e interpretan la información. Un programa completo de autorrevisión recorre todos los estamentos, desde el rector, profesorado, personal de administración y servicios, hasta el funcionamiento de los órganos colegiados, servicios y programas.

2. Generalmente los modelos basados en este enfoque declaran que la evaluación se debe insertar adecuadamente en el marco social. No obstante, este principio parece que no está satisfactoriamente operacionalizado, lo que comentaremos en el apartado de contextualización.

Las fuentes de información para la evaluación en este tipo de planes se basan predominantemente en el análisis conjunto realizado entre la unidad evaluada y un equipo técnico de expertos o colegas de todos aquellos datos<sup>3</sup> que reflejen la acción institucional: organización, aprovechamiento de recursos –procesos y productos–, eficacia y eficiencia institucional, etc. Las técnicas pueden ser variadas y los indicadores diversos: cuantitativos, de carácter económico o estructural –objetivables–, y cualitativos, que permitan contextualizar adecuadamente la realidad.

En definitiva, estos acercamientos, más que constituir un único programa, pueden entenderse como un marco de trabajo desde el que analizar necesidades, establecer prioridades y planificar la evaluación e innovación. Por ello pueden tender a integrar diversos enfoques, análisis de programas académicos, profesores, etc., aunque enfatizará siempre el análisis de los procesos organizativos con un carácter globalizador de la institución.

Por su parte, la finalidad de los *Enfoques Institucionales de Rendición de Cuentas* puede ser muy variada y depende, en gran medida, de los promotores de la evaluación y del tipo de universidades sobre las que se desarrolle (públicas o privadas). Los fines ligados a este tipo de planes de evaluación suelen estar vinculados a decisiones acerca de la planificación, mantenimiento/desarrollo de instituciones, distribución de recursos, acreditación, etc. Los esquemas de trabajo, asociados generalmente a la valoración de indicadores de resultados –ponderados o no, de acuerdo a otros de input, de proceso o de contexto– persiguen analizar el grado de consecución de determinados objetivos institucionales o el valor/utilidad global del producto institucional. Las fuentes de información son también múltiples, según el tipo de variables que atienda: bancos de datos de la propia universidad o nacionales, informes de expertos, etc.

Uno de los modelos de referencia obligada en este ámbito es el desarrollado en la determinación de indicadores de rendimiento institucional en el Reino Unido (CAVE et al., 1988; JOHNES y TAYLOR, 1990; CAVE et al., 1991). Una síntesis de estos indicadores se adjunta en el Cuadro 1. La aportación supone el establecimiento de un sistema de amplio espectro para valorar el funcionamiento de instituciones universitarias y llevar a cabo una planificación y distribución de los recursos, de acuerdo con las consecuencias que se derivan de su utilización. Su implantación tiene dos efectos principales: el impacto derivado sobre otros sistemas educativos europeos y el incremento desbordante del interés por el estudio y desarrollo de indicadores de rendimiento institucional. En cualquier caso, este enfoque se sustenta y se orienta hacia la idea de “rentabilidad” del producto universitario, y las consecuencias de su aplicación han provocado múltiples reacciones.

3. Los datos a tomar en consideración se pretende que sean de una gran variedad como presupuestos, costes, número de estudiantes, niveles de rendimiento de estudiantes, informes de comisiones internas, opiniones de estudiantes, opiniones de profesores, responsables de gestión, otras instituciones, etc.

Por último, denominamos *Enfoques Mixtos* a aquellos que, junto a otras informaciones, incluyen algún tipo de informe de autorrevisión realizado por la propia institución –y aceptado por la Comisión Evaluadora mediante procesos metaevaluativos y/o algún otro sistema de control adicional– y orientan su análisis con algún propósito sumativo. La evaluación se realiza por medio de un comité de expertos tomando como base las fuentes mencionadas.

Entre los tres enfoques revisados, el Mixto es el que mayor interés concita en la actualidad y se perfila como el acercamiento que mayor extensión va a conseguir, probablemente, en un futuro inmediato. Esto se debe, sin duda, a la gran flexibilidad que proporciona como marco de trabajo, permitiendo un ajuste y una negociación más eficaz entre las diversas instancias y audiencias implicadas en el proceso.

En *nuestro contexto*, la Evaluación Institucional comienza a desarrollarse a finales de la década de los 80 y principios de los 90. Presente en debates sobre la Evaluación de la calidad de Docencia Universitaria, supuso un planteamiento más global que el que se estaba imponiendo de la evaluación centrada en la actuación del profesor. En 1992, en noviembre, se da luz verde al primer programa de Evaluación Institucional en nuestro país (GARCÍA et al., 1995), orientado desde el Consejo de Universidades, al que se adscriben de forma voluntaria diversas instituciones. El carácter que inicialmente se le imprime es el de facilitar la Autorrevisión Institucional. No obstante, considerando que quien promueve la evaluación es un organismo superior a las universidades, y teniendo en cuenta el contexto de progresiva implantación de planes de evaluación para el control de calidad de las mismas, el programa se desarrolla en un clima en el que el debate interno acerca de la finalidad de este tipo de evaluaciones cree prever más su uso para la Rendición de Cuentas (concretamente, la distribución de Recursos), que para la mera innovación interna. Ciertamente, concluida su primera fase, no conocemos que se haya utilizado, hasta el momento, con otro propósito diferente al declarado en su planteamiento inicial.

En cualquier caso, la complejidad de este tipo de Evaluaciones se acrecienta por los siguientes factores:

- a) La situación autonómica del propio Estado y la autonomía universitaria. No está clara la finalidad de la evaluación institucional en un proceso de profundización autonómica, en el que buena parte de las transferencias en Educación se han realizado a muchas Comunidades Autónomas. En este contexto, más si se tiene en cuenta la autonomía universitaria que define la LRU, es muy difícil que puedan desarrollarse planes de evaluación con consecuencias efectivas orientados desde un organismo con carácter estatal. Estimamos que, lógicamente, si se desean realizar evaluaciones de las universidades desde el Consejo de Universidades no hay otro camino más que sensibilizar hacia los procesos de autorrevisión. En cualquier caso, creemos que sólo un organismo más cercano, de carácter autonómico, podría abordar la implantación de un sistema de evaluación institucional que tuviera algún otro uso (como el



### *Evaluación de la Actividad Universitaria*

- de distribución de recursos), aunque, claro está, dentro de un proceso negociado con las universidades.<sup>4</sup>
- b) No obstante, la organización interna de cada Universidad tiene sus peculiaridades, de forma que si hay diferencias en estructura, también las hay en sistemas de funcionamiento. Además, no cuentan con estructuras administrativas y de gestión de la información que puedan considerarse esencialmente equiparables entre sí. Ello dificulta establecer Indicadores Estructurales comunes; por lo que, previamente a realizar un proceso de esta índole, se debería analizar la disponibilidad de las informaciones requeridas, a no ser que se haga en el seno de un modelo exclusivamente de autorrevisión.
- c) La literatura especializada en nuestro país está repleta, en los últimos años, de recomendaciones y reflexiones acerca de la dificultad que existe de establecer una adecuada contextualización de la evaluación de la docencia que imparte el profesor y las consecuencias que ello tiene para la fiabilidad y validez de estas evaluaciones. Muchas de estas recomendaciones son ciertas, aunque también hay que señalar que existen soluciones técnicas (sin olvidar las decisiones político-administrativas) que pueden, conveniente aplicadas, ofrecer evaluaciones bastante ajustadas. Llama la atención la escasez de reflexiones que en este mismo sentido pueden hacerse acerca de las Evaluaciones de corte Institucional. No hay que olvidar que la complejidad del fenómeno que se estudia está en relación directa con las dificultades de aprehenderlo adecuadamente. Así, la difícil definición y contextualización del valor/utilidad de los diferentes "productos" universitarios, conlleva una mayor dificultad que en las evaluaciones de profesorado para establecer criterios genéricos de calidad. Asimismo, las consecuencias de aplicaciones de evaluaciones institucionales en otros países alertan hacia el hecho de que sus resultados –por la carga de interpretación subjetiva a que muchos de ellos conllevan– son especialmente susceptibles de usos políticos, de apoyo a decisiones, que en ocasiones han tenido una gran trascendencia (como el cierre de centros, disminución de presupuestos, etc.). En este sentido, entendemos que hay que aprovechar el camino iniciado y profundizarlo, procurando evitar tendencias e inercias que se pueden con toda probabilidad producir. Así, es preciso desterrar la tendencia hacia las decisiones premiosas por imperativo político, procurando crear un sistema abierto de debate, aceptando la necesidad de

4. De hecho, muy recientemente se ha hecho exposición pública, más detallada que en ocasiones precedentes, sobre los planes de establecimiento de una entidad institucionalizada en Cataluña que será responsable en exclusiva de las acciones evaluativas en este ámbito. En este mismo sentido, a finales de 1995 se ha publicado un decreto acerca de la evaluación institucional de la universidades, que por parte del Gobierno Central puede entenderse como un medio de estímulo hacia el desarrollo de estos procesos.

diferentes proyectos e iniciativas –desde diversas ópticas técnicas y políticas, centradas en diversos niveles del sistema universitario, etc.– y estableciendo los recursos para su adecuado desarrollo. Además, es fundamental posibilitar un debate en el que se impliquen realmente todas las audiencias implicadas, al contrario que otras decisiones institucionales adoptadas, para que el sistema que se derive alcance el adecuado enraizamiento dentro y fuera del ámbito universitario.

- d) La propia complejidad del problema, la evaluación de las instituciones universitarias, hace tanto más necesario el énfasis en un periodo previo, suficientemente amplio, en el que se lleven a cabo acciones orientadas a la profundización en diferentes aspectos metodológicos y procedimentales. Así, la experiencia de los desarrollos realizados en universidades españolas respecto a otros ámbitos de menor complejidad como la evaluación de la docencia –con múltiples estudios respecto al desarrollo de instrumentos, factores que afectan a la fiabilidad y validez de las evaluaciones, metodologías para el desarrollo de estándares, investigaciones sobre las relaciones de la evaluación docente con otras variables, etc.– deben servir de referente para entender el camino que queda por recorrer junto con las indudables necesidades que presenta.
- e) Los procesos de autorrevisión institucional son los únicos que se encaminan realmente en la dirección que acabamos de exponer en los puntos anteriores, al posibilitar el establecimiento de unas condiciones claves para una verdadera transformación institucional. Así, aspectos como el debate de los colectivos implicados, el desarrollo de un ajuste a las realidades concretas, la adecuada contextualización y, en definitiva, el aprovechamiento formativo inmediato proporcionan un marco mucho más productivo tanto para la institución como para la sociedad.

En cualquier caso, el camino no ha hecho más que iniciarse y hay que recordar que el producto de una institución universitaria recorre un amplio espectro de beneficios para la sociedad, desde algunos tangibles (como las patentes que provee) hasta otros que no lo son tanto (como la apertura y configuración cultural, con posiciones avanzadas en el cambio de actitudes universales o el hecho de que la Universidad como institución ha jugado históricamente un papel de vanguardia en la anticipación de soluciones a los problemas científicos, humanos y sociales...), pero que nadie puede negar que están en el sustrato de lo que conocemos como civilización. Ciertamente, los modelos de evaluación institucional no han logrado todavía integrar adecuadamente un tratamiento justo del conjunto de la Universidad (como institución). Los indicadores y técnicas desarrolladas priman excesivamente la vertiente científico-tecnológica y profesionalizadora de las Universidades, dejando en peor lugar los aspectos humanísticos y culturales –véanse los Cuadros de Indicadores anteriormente citados–. ¿Qué es una universidad de calidad?, ¿qué tipo de universidad queremos? son cuestiones que hay que responder previamente en el debate para definir qué, cómo y para qué se va a evaluar la institución universitaria.

## 1.2 Enfoques de Programas

La Evaluación de Programas se refiere a un acercamiento evaluativo de aplicación general a diversos ámbitos de evaluación. No obstante, en este contexto nos referimos a programas como unidades de actuación docente para un determinado propósito académico (materia, titulación, curso o programa de doctorado, postgrado, etc.). Sintetizando las características de este enfoque que expone PÉREZ JUSTE (1994), se pueden enumerar las siguientes:

- a) proceso sistemático de recogida y análisis de la información,
- b) diseñado intencional y técnicamente,
- c) cuya finalidad es la valoración de la calidad y los logros de un programa,
- d) entre sus posibles objetivos se pueden señalar:
  - determinar la eficacia de un programa,
  - analizar las características personales y situaciones en las que un programa puede resultar más eficaz,
  - identificar elementos de mejora en los procesos de desarrollo del programa, y
  - valorar el programa como apoyo a la toma de decisiones.

Las fuentes de información que se consideran suelen ser las propias de cualquier evaluación de la docencia, integradas y contextualizadas de forma que pueda aportarse una imagen global del mismo. Así, se sustentan, para la selección de indicadores, normalmente en el modelo CIPP –que se ha generalizado para este propósito–(DE MIGUEL, 1991), si bien son también comunes otro tipo de estrategias. De este modo, al considerar la evaluación de un programa determinado se suelen incluir indicadores de Entrada, de Contexto, de Proceso y de Producto del mismo. Un ejemplo de selección de variables a partir del modelo CIPP es el que aporta DE MIGUEL (1991) y que se puede resumir en los siguientes epígrafes:

Contexto	Entradas	Proceso	Salida
infraestructura y recursos	programas	desarrollo del programa	resultados del estudiante (a nivel medio e inmediato)
profesores	demandas de plazas	planificación del profesor	indicadores de eficiencia
estudiantes	calidad del usuario	evaluación del aprendizaje otras actividades relacionadas con los estudiantes	cambios institucionales nuevas actividades

ESTADÍSTICAS DE DIRECCIÓN E INDICADORES DE ACTUACIÓN DE LA UNIVERSIDAD. ADAPTADO DE: CAVE, M.; HANNAY, S. Y KOGAN, M. (1991);  
 THE USE OF PERFORMANCE INDICATORS IN HIGHER EDUCATION. LONDRES: JESSICA KINGSLEY PUBLISHERS.

Cuadro 1

<p><i>Gastos en Departamento Académico (Corte por Centro)</i></p> <p>FTE Staff académicos</p> <p>E1 Gastos por estudiante</p> <p>E2 Gastos por staff académico</p> <p>E3 Gastos de mantenimiento por staff académico</p> <p>E4 Gastos en equipamiento por staff académico</p> <p>E5 Ingresos por investigación por staff académico</p>	
<p><b>Estudiantes y staff (Corte por Centro)</b></p> <p>Estudiantes</p> <p>E6 Investigación de posgraduados como un porcentaje de estudiantes</p> <p>E7 Enseñanza, todos a posgraduados como porcentaje de estudiantes</p> <p>E8 Todos los posgraduados como porcentaje de estudiantes</p> <p>E9 Ratio de estudiantes y personal docente del staff</p>	
<p><b>Gastos de la administración central</b></p> <p>E10 Gastos de administración central como porcentaje de los gastos totales</p> <p>E11 Pagos efectuados como porcentaje de los gastos de la administración central</p> <p>E12 Gastos de administración central por estudiante</p> <p>E13 Gastos de administración central por staff académico</p>	
<p><b>Gastos en Bibliotecas</b></p> <p>E14 Gastos en biblioteca como porcentaje del gasto global</p> <p>E15 Gastos en publicaciones como porcentaje del gasto en biblioteca</p> <p>E16 Gastos en pagos de biblioteca como porcentaje del gasto en biblioteca</p> <p>E17 Gastos en biblioteca por estudiante</p> <p>E18 Gastos en biblioteca por staff académico</p> <p>E19 Gastos en libros por estudiante</p> <p>E20 Gastos en revistas por estudiante</p>	
<p><b>Gastos en servicios informáticos</b></p> <p>E21 Gastos en servicios informáticos como porcentaje del gasto global</p> <p>E22 Pagos en servicios informáticos como porcentaje de los gastos en servicios informáticos globales</p> <p>E23 Gastos en servicios informáticos por estudiante</p> <p>E24 Gastos en servicios informáticos por staff académico</p>	
<p><b>Gastos en locales</b></p> <p>E25 Gastos-totales en locales, como porcentaje del gasto global</p> <p>E26 Pagos en locales como porcentaje del gasto en locales</p>	
<p><b>Gastos en Servicios de facultad y organizaciones de estudiantes</b></p> <p>E27 Gastos en calefacción, agua y electricidad como porcentaje del gasto global</p> <p>E28 Gastos en servicios de limpieza y seguridad como porcentaje del gasto global</p> <p>E29 Gastos en reparaciones y mantenimiento como porcentaje del gasto global</p> <p>E30 Gastos en teléfono como porcentaje del gasto global</p> <p>E31 Gastos en locales por estudiante</p> <p>E32 Pagos en locales por estudiante</p> <p>E33 Gastos de calefacción, agua y electricidad por estudiante</p> <p>E34 Gastos de servicios de limpieza y seguridad por estudiante</p> <p>E35 Gastos de reparaciones y mantenimiento por estudiante</p> <p>E36 Gastos de teléfono por estudiante</p>	
<p><b>Primeros desechos de los graduados por Universidad</b></p> <p>E39a Desechos a 31 de diciembre después de la graduación, totales de UK por individuo académico</p>	
<p><b>Primeros desechos de los graduados por la Universidad</b></p> <p>E40 Graduados totales con desechos conocidos</p> <p>E41 Graduados con desechos "desempleado o corto tiempo"</p> <p>E42 Valor predicho del índice D41</p> <p>E43 Diferencia entre los índices D42 y D41</p> <p>E44 Diferencia por cien estudiantes</p>	
<p><b>Primeros empleos: Proporción nacional de "Desempleado a corto término" por sujetos</b></p> <p>Ocurricencias de no graduados (por Grupo académico de sujetos)</p> <p>E45 Número de abandonos</p> <p>E46 Ocurricencias como porcentaje de los que acaban sus estudios</p> <p>E47 Proporciones de tres y cuatro cursos</p> <p>E48 Términos de atención por ocurrencia</p> <p>E49 E48 relativo al valor esperado</p>	
<p><b>Cualificaciones de Principiantes no graduados a tiempo completo, 1987 a 1989 (Por grupos académicos)</b></p> <p>E50 Participantes con 3 o más A, número</p> <p>E51 Participantes con 3 o más A, puntuaciones</p> <p>E52 Participantes con 5 o más Scottish Highers, número</p> <p>E53 Participantes con 5 o más Scottish Highers, puntuaciones</p> <p>E54 Participantes con otras cualificaciones</p>	

Con propósitos de mejora e innovación, pueden ser promovidos por las propias universidades y/o departamentos responsables de la docencia e integrados en procesos más genéricos de autorrevisión institucional. Como en el enfoque institucional, el objetivo será analizar los elementos que pueden conducir a una mejora, incidiendo sobre cualquiera de los componentes del proceso. Un ejemplo de interés es el aportado por VROEIJENSTIJN y ACHERMAN (1990), que recogemos en el Cuadro 2.

Ahora bien, con propósitos sumativos, este tipo de acercamiento puede estar vinculado a decisiones de asignación/distribución de recursos y acreditación. Estarán entonces sujetos a la comprobación del grado de consecución de los objetivos marcados para el programa o a su valor/utilidad global.

Asimismo, pueden aparecer subsumidos en enfoques institucionales como unidad de análisis global de valoración de la docencia, en cuyo caso el énfasis de la función docente se evalúa de forma macroanalítica, sin referencias individuales.

Sin embargo, cuando se desarrollan enmarcados en una finalidad de carácter sumativo, se mantiene<sup>5</sup> la referencia individual en la función docente.

En nuestro contexto se han realizado diversas experiencias de evaluación de programas de carácter universitario, desde evaluaciones de la implantación de los Nuevos Planes de Estudio (SUÁREZ et al., 1995) hasta evaluaciones de programas de Doctorado y/o de Postgrado (PÉREZ CARBONELL et al., 1993, 1995; JORNET et al., 1995). En cualquier caso, las características de los programas desarrollados corresponden fundamentalmente a enfoques de autorrevisión, orientados por las universidades para mejorar los programas en cuestión. No obstante, aunque las experiencias desarrolladas en este sentido no son muy amplias, ciertamente se están manifestando como muy positivas en tanto en cuanto aportan elementos directos de información para la mejora e innovación de la oferta de los programas evaluados, pudiéndose integrar adecuadamente en la realidad de la gestión de nuestras universidades.

### 1.3 Enfoques centrados en el profesor

Son aquellos que focalizan la acción evaluativa en la función docente realizada por el profesor. Descritos con profusión en la literatura especializada tanto en nuestro contexto (VILLAR, 1987; TEJEDOR et al., 1987; MATEO, 1987; ESCUDERO, 1987, 1991; ZABALZA, 1990) como en el panorama internacional (MILLMAN, 1981; MILLER, 1987), pueden asimismo enfocarse desde las dos perspectivas —orientadas al desarrollo profesional o a la rendición de cuentas—.

Las propuestas de *enfoques y modelos formativos* son muy variadas. Se evalúa fundamentalmente la actuación docente del profesor. Las soluciones más

5. El nivel de mantenimiento de la referencia individual depende del objetivo subyacente a la rendición de cuentas. Así, si lo que se pretende es determinar la viabilidad de los programas de las instituciones evaluadas, individualmente o por sectores, se mantendrán informaciones que permitan adoptar decisiones respecto a ese nivel en concreto.

LISTADO PARA EL AUTOESTUDIO ADAPTADO DE VROEDENSTIJN, T.I. Y ACHERMAN, H. (1990): CONTROL ORIENTED VERSUS IMPROVEMENT ORIENTED QUALITY ASSESSMENT. EN C.J. GOEBEGGER, URE, P.A.M. MAASSEN Y D.F. WESTERHEIJDEN (EDS). PIER REVIEW AND PERFORMANCE INDICATORS. UTRECH. LEMMA.

Cuadro 2

<p>1. El lugar en la organización</p> <ul style="list-style-type: none"> <li>* Punto de partida del enfoque - describir la estructura organizacional específica y la posición de los miembros o de la disciplina dentro de la universidad; así como la identificación de los Departamentos implicados.</li> <li>* ¿En qué contextos está implicada la disciplina? ¿Cuáles son sus cursos y cómo pueden realizarse?</li> </ul>	<p>5. Organización y dirección del programa.</p> <ul style="list-style-type: none"> <li>-Política educativa             <ul style="list-style-type: none"> <li>* ¿Cómo está formulada la política educativa? ¿Qué contextos están implicados?</li> <li>* ¿Esta la política relacionada con la disciplina enseñada con la política institucional? ¿Contextos con los planes de mejora?</li> </ul> </li> <li>-Política de personal             <ul style="list-style-type: none"> <li>* ¿Cuál es la política respecto al ratio docente/investigación? ¿Cuál es la carga docente?</li> <li>* ¿Cuál es la política respecto al empleo en el curso propedéutico y los sujetos?</li> </ul> </li> <li>-Dirección cualitativa interna             <ul style="list-style-type: none"> <li>* ¿Cuál es la atención dada a la dirección cualitativa interna? ¿Hay una evaluación sistemática? ¿Qué sistema de evaluación se utiliza?</li> <li>* ¿Qué atención se presta a la profesionalización del staff (cursos de dirección, entrenamiento educativo o didáctico)?</li> </ul> </li> <li>-Progreso, supervisión y asesoramiento del estudio             <ul style="list-style-type: none"> <li>* ¿Han sido medidas durante los recientes años las horas de estudio inscritas en el curso propedéutico, fase de graduación, o ambos? Si es así, cuáles fueron los resultados?</li> <li>* ¿Cómo está organizada la orientación del estudio y cuál es su efecto?</li> <li>* ¿Cómo está organizada la supervisión del estudio y cuál es su efecto?</li> </ul> </li> </ul>
<p>2. Los estudiantes</p> <ul style="list-style-type: none"> <li>* El número de estudiantes.</li> <li>* Características del grupo de estudiantes (sexo, edad, origen geográfico, formación previa o experiencia profesional)</li> </ul>	<p>6. Los graduados.</p> <ul style="list-style-type: none"> <li>* ¿Puede ser la futura profesión claramente definida? ¿Han cambiado las definiciones durante los últimos años?</li> <li>* ¿Dónde van los graduados cuando acaban?</li> <li>* ¿Cuál es la tasa de desempleo entre graduados?</li> <li>* ¿Existen contactos con los alumnos?</li> </ul>
<p>3. Breve descripción de los programas</p> <ul style="list-style-type: none"> <li>-General             <ul style="list-style-type: none"> <li>* ¿Formula objetivos explícitamente?</li> <li>* ¿Dirigen los objetivos de acuerdo con los programas?</li> <li>* ¿Qué extensión tienen los objetivos logrados?</li> </ul> </li> <li>-El curso propedéutico             <ul style="list-style-type: none"> <li>* ¿Cuál es la estructura general del nivel propedéutico? ¿A qué sujetos afecta? ¿Cuáles son las proporciones relativas de los sujetos?</li> <li>* ¿Existen requisitos en términos de relaciones mutuas entre los sujetos?</li> </ul> </li> <li>-La fase de grado             <ul style="list-style-type: none"> <li>* ¿Cómo está organizada la fase de grado? ¿Cuál es la relación entre los cursos realizados y los sujetos?</li> <li>* ¿Cuál es la extensión del programa obligatorio? ¿Cuál es la variación entre los elementos optativos?</li> </ul> </li> <li>-Otros cursos ofertados             <ul style="list-style-type: none"> <li>* Cursos de posgrado (entrenamiento del profesor y vocacional)</li> <li>* Cursos de serción y contrato</li> <li>* ¿Cuántos estudiantes asisten al estudio de posgraduados?</li> </ul> </li> <li>Programas a tiempo parcial             <ul style="list-style-type: none"> <li>* ¿Se ofertan en el programa cursos a tiempo parcial?</li> <li>* Si es así, ¿en qué differs el programa a tiempo parcial en términos de material y estructura?</li> </ul> </li> </ul>	<p>Interacción entre los niveles central y de facultad.</p> <ul style="list-style-type: none"> <li>* ¿Cómo está organizada la interacción entre la facultad o disciplina y el equipo específico o de la universidad respecto a la educación y, en particular, de la calidad de promoción?</li> <li>* ¿Cómo está relacionada la política educativa de la facultad y la de la institución?</li> </ul>
<p>4. Proceso educativo</p> <ul style="list-style-type: none"> <li>* ¿Qué formas de docencia son utilizadas en la adquisición y transferencia de conocimiento (prácticas de lecturas, organización de seminarios, tutorías, trabajo práctico, etc.)?</li> <li>* ¿Cómo son los ordenadores utilizados en los cursos?</li> <li>* ¿Cómo se asesora a los estudiantes? a) en qué forma (grupo, cuestiones directas) b) cuándo (sesiones, finales)</li> </ul>	

efectivas se basan en estrategias de autorrevisión en las que, junto al profesor evaluado, participa un asesor pedagógico (que generalmente coincide en la persona del evaluador o del formador) que coadyuva en su proceso de mejora. Las fuentes utilizadas como elemento para la mejora pueden ser las mismas, integrando, además, técnicas de apoyo basadas en diferentes orientaciones –por ejemplo, procesos sustentados en planteamientos de microenseñanza dentro de un esquema clínico de apoyo a la autorrevisión–. En este tipo de planes, al requerir la participación activa y voluntaria del profesor, el proceso de implementación en la institución se orienta a la sensibilización de profesorado. Para este propósito se pueden utilizar estrategias de sensibilización basadas en evaluaciones de estudiantes que sólo se empleen como retroalimentación para el profesor,<sup>6</sup> o la aportación de recursos de autoevaluación –como sistemas informáticos que apoyen su realización (SALVADOR y GARCÍA 1989)–. Asimismo, la oferta de sistemas de apoyo individualizados y colectivos debe ser amplia y coherente con las posibles demandas –para ello, estudios de evaluación de necesidades formativas pueden constituir un elemento inicial (TEJEDOR, 1990)–. Los problemas de implementación de este tipo de planes residen en que es difícil encontrar instrumentos de dinamización del interés por la propia renovación y cualquier uso incentivador o imperativo nos traslada de inmediato a la rendición de cuentas. Así, la coherencia interna que exige un plan de este tipo limita su posible impacto y generalización.

Por su parte, los *enfoques de Rendición de Cuentas*, en general se orientan desde la propia Universidad o desde la Administración. Por sus características, normalmente requieren un proceso de institucionalización complejo, dado que se trata de llegar a evaluar a todo el profesorado con una determinada periodicidad. En estos casos, la evaluación suele orientarse hacia la toma de decisiones acerca de la promoción/mantenimiento del profesor, retribuciones y, en definitiva, aspectos académico-administrativos.

No vamos a detallar aquí los problemas y/o soluciones que presentan este tipo de evaluaciones, pues han sido ampliamente expuestos en la literatura especializada, tanto nacional (DE LA ORDEN, 1990; TEJEDOR, 1987, 1991; GONZÁLEZ SUCH et al, 1993, 1995; FERRÁNDEZ et al, 1995a y b; MARSH, H.W., TOURON, J. y WHEELER, B., 1985) JORNET et al., 1995; como internacional (MILLMAN, 1981; MILLER, 1987; CENTRA, 1980; BRASKAMP, BRANDERBURG y ORY, 1984; MARSH, 1987; MARSH y HOCEVAR, 1991). Haremos, eso sí, un breve resumen de aquellos elementos que nos sirvan para establecer el marco de referencia de estos enfoques.

Los tres componentes básicos que suelen evaluarse en la actividad del profesor universitario son su actuación como docente, la investigación y su partici-

6. Aunque este medio, a corto plazo, eliminaría –o bajaría ostensiblemente– la participación estudiantil ante la falta de otro tipo de consecuencias en la evaluación, que suelen ser reclamadas como uno de los objetivos prioritarios por esta audiencia. No obstante, dentro de este enfoque se pueden plantear esquemas negociados que impliquen consecuencias tangibles para los estudiantes –algo que no se emplea con la frecuencia y profundidad necesarias–, lo que solucionaría, en buena parte, el problema señalado.

pación en la gestión universitaria. De los tres elementos, los desarrollos mas amplios se centran en la Evaluación de la Docencia y, en segundo lugar, en la Investigación. Aunque se encuentran investigaciones sobre alguna de estas cuestiones relativas al ámbito de la gestión en el panorama internacional (MOSES y ROE, 1990; SELDIN, 1988; AL-KARNI, 1995), sin embargo en nuestro contexto se han generalizado evaluaciones de los dos primeros aspectos, pero no conocemos experiencias acerca de la evaluación de la gestión que hayan trascendido.<sup>7</sup>

Como señalan diversos autores cuando se evalúa la actividad docente del profesorado con carácter sumativo la fuente de información mas fiable es la opinión de los estudiantes, aunque estén limitados para valorar algunas variables de interés –ver Cuadro 3–. Asimismo, es la fuente mejor aceptada por el profesor, frente a compañeros o agentes evaluadores externos (JORNET et al., 1989). En este tipo de planes, aunque se recojan informaciones adicionales (como autovaloraciones, informes de Departamentos, etc...), la información central suele ser la que provee el estudiante.

Cuadro 3 VARIABLES INCLUIDAS EN LOS CUESTIONARIOS DE OPINIÓN DE ESTUDIANTES. DIMENSIONES DOCENTES PROPUESTAS POR (ABRAMI Y D'APOLLONIA, 1991)	
1 Estimulación del interés	2 Entusiasmo
3 Conocimiento de la materia	4 Expansividad intelectual
5 Preparación intelectual	6 Claridad y entendimiento
7 Habilidades de elocución	8 Nivel de progreso de la clase
9 Claridad de los objetivos del curso	10 Relevancia y valor del material del curso
11 Relevancia y utilidad de los materiales suplementarios	12 Trabajo
13 Rendimiento percibido	14 Justicia de la evaluación
15 Organización y dirección de las clases	16 Características de la personalidad
17 Feedback	18 Fomento de discusiones y diversidad de opiniones
19 Dudas intelectuales y fomento del pensamiento independiente	20 Preocupación y respeto por los estudiantes
21 Disponibilidad y ayuda	22 Curso en general
23 Docencia en conjunto	24 Miscelánea de Items

7. Aunque se le denomina Personal Docente e Investigador (PDI) ciertamente a muchos profesores se les podría identificar también con la "G" (de Gestión). A esta tercera función a la que, aunque de manera democrática, se llega más por obligación que por vocación, muchos profesores destinan buena parte de su tiempo (profesional y personal) en perjuicio, también en muchas ocasiones, de la D y de la I. En este sentido, desde estas líneas, tratamos de llamar la atención hacia este hecho que, por ende, no suele reconocerse con el suficiente peso en los currícula del profesorado.



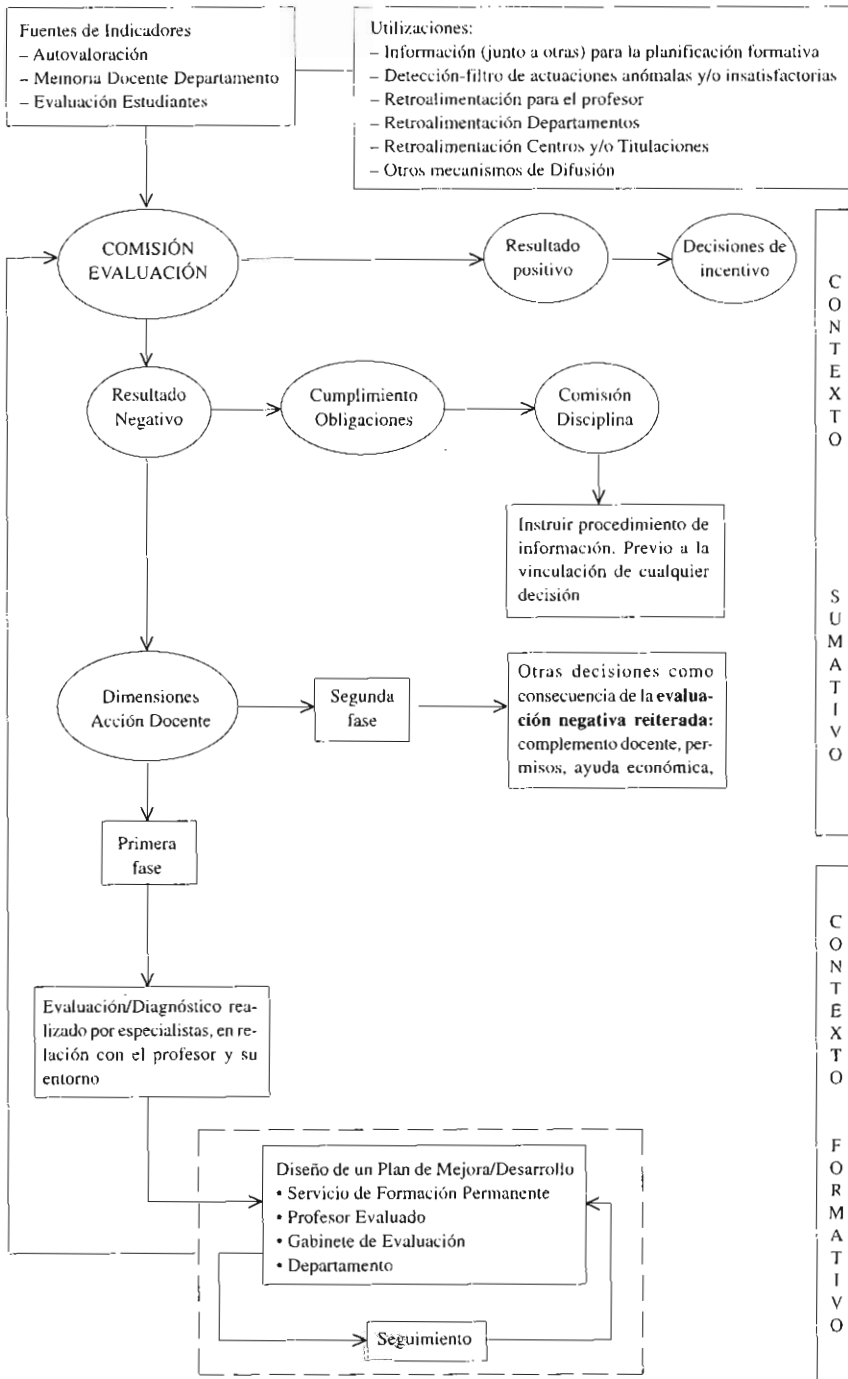


Figura 1. Esquema de senderos de decisiones vinculadas a la Evaluación de la Docencia del Profesorado. Adaptado de JORNET y SUÁREZ (1991)

Por último, es preciso señalar que ambos enfoques (Formativo/Sumativo) no son en absoluto excluyentes, al contrario, su combinación flexible puede proporcionar una solución mucho más susceptible de satisfacer las necesidades y expectativas de los diversos colectivos implicados en el proceso. En este sentido, son perfectamente posibles múltiples *esquemas mixtos* de actuación en los que, por ejemplo, se combine un plan de trabajo formativo para responder mejor a las necesidades de innovación de los individuos o de colectivos junto con unas exigencias externas –de la Universidad o del Estado– sobre determinadas cuestiones de su docencia, lo que puede entramarse en diversos grados. Así, se puede pensar en una complementariedad, junto a la compatibilidad apuntada, al referirse por sus objetivos a diferentes aspectos de la función docente –con diferente óptica, diversos instrumentos, fuentes variadas, etc.–, aun teniendo presente la existencia de un menor o mayor solapamiento entre ambos procesos. Un ejemplo de combinación entre enfoques sumativo y formativo es el que se sintetiza en la Figura 1. En él, como puede apreciarse, se parte de un planteamiento de carácter sumativo correspondiente a una evaluación administrativa del profesor en el que se consideran tres fuentes (la valoración de estudiantes, una evaluación realizada desde el departamento y la autovaloración). Esta evaluación sumativa sirve para la detección de casos problemáticos. El enfoque sumativo actúa, pues, de filtro y el árbol de decisión conduce hacia consecuencias ligadas a la rendición de cuentas y/o otras de mejora docente, dependiendo del tipo de problemática identificada.

En *Universidades Españolas* –o respecto a ellas– se han desarrollado en los últimos diez/doce años diversos planes de evaluación que, básicamente, recorren los enfoques anteriormente citados. No obstante, su implementación ha sido paulatina y en la actualidad el grado de institucionalización de estos procesos es también muy variado. En las líneas que siguen en este apartado, más que realizar una aproximación histórica de los diferentes planes que se han venido poniendo en marcha, queremos aportar unas notas de reflexión partiendo de la perspectiva de su evolución.

TEJEDOR (1991) realiza una revisión del desarrollo de Planes de Evaluación en *Universidades Españolas*. De ella se desprende que, desde las primeras realizaciones hasta el momento de la citada revisión, el foco más generalizado de las evaluaciones en nuestras universidades ha sido la actuación docente del profesor, tomando como método los cuestionarios de opinión de estudiantes.

No obstante, pueden distinguirse tres momentos: a) El desarrollo de evaluaciones orientadas desde las universidades mismas (prácticamente desde 1981 hasta 1989), b) Las evaluaciones orientadas a partir del Decreto de Retribuciones (desde agosto de 1989, hasta la actualidad) y, c) Las líneas evaluativas que se plantean desde enfoques institucionales y que se concretan con la puesta en marcha del Programa Experimental de Evaluación del Sistema Universitario por parte del Consejo de Universidades (desde Noviembre de 1992, hasta la actualidad).

El primer momento deviene de la aprobación de la LRU y la elaboración de los Estatutos de las Universidades, donde se contempla la obligatoriedad de la evaluación del profesorado teniendo en cuenta las opiniones de estudiantes. En

## *Evaluación de la Actividad Universitaria*

la mayor parte de universidades que promueven procesos evaluativos se orientan con carácter fundamentalmente formativo en el comienzo. Estos sistemas pusieron su énfasis en la concertación entre las diferentes audiencias implicadas (profesores, estudiantes, comisiones de gobierno de la universidad, responsables institucionales...), y adaptaron su ritmo de implantación a las características de cada universidad. Asimismo, en esta etapa, la mayor parte de evaluaciones se desarrolla con carácter experimental, priorizando el estudio de instrumentos a utilizar (cuestionarios, registros, medios y procesos de realización...) y el análisis técnico de las consecuencias, usos..., de las evaluaciones.

Ya en las I Jornadas Nacionales de Evaluación y Mejora de Docencia Universitaria organizadas por el Servei de Formació Permanent de la Universitat de València, en mayo de 1987 (VARIOS, 1988), se revisaron, a través de las experiencias evaluativas nacionales e internacionales, los diversos modelos de actuación para abordar este tipo de evaluaciones. Posteriormente, múltiples reuniones y debates organizados por las universidades han ido profundizando en la cuestión (Cantabria, 1990; Las Palmas, 1991; Cádiz, 1991; Castellón, 1993, por citar sólo algunos), así como han contribuido a dar solidez y mejorar las experiencias que se han ido realizando.

Los desarrollos que se han ido implementando en las diversas universidades del Estado encuentran como primer referente los trabajos de Aparicio, Tejedor y Sanmartín (1982) sobre elaboración de un Cuestionario de Evaluación de la Docencia a partir de opiniones de estudiantes para la Universidad Autónoma de Madrid. Los restantes desarrollos en buena medida corresponden a adaptaciones realizadas en las diversas universidades teniendo en cuenta las características sociales y la cultura organizacional de cada una de ellas.

Con la aparición del Decreto de Retribuciones (agosto de 1989) se identifica la evaluación de la docencia y la investigación como un elemento de mejora del sistema universitario. No obstante, los procesos evaluativos que se estaban llevando a cabo en diversas universidades fueron desvirtuados, al identificar una finalidad para la evaluación hasta el momento no prevista –de carácter sumativo (el complemento salarial)–. Un elemento positivo es que en ese mismo Decreto se amplían las fuentes de información requeridas<sup>8</sup> respecto a las que habitualmente se estaban utilizando para valorar al profesorado; necesidad, no obstante, ya sentida y reclamada tanto por evaluadores como por los colectivos implicados en el proceso. En esta nueva situación, la responsabilidad de la evaluación de la docencia se hace recaer en las universidades mientras que el gobierno central se reserva la evaluación de la investigación.<sup>9</sup>

8. Además de la opinión de estudiantes, se requiere una autovaloración así como otra información emanada de algún órgano de gestión de la Universidad en la que trabaje el profesor. En algunos casos, como en la Universitat de València, ya se estaba trabajando, a nivel técnico, en instrumentos orientados a recabar dichas informaciones (autovaloración, aprovechamiento de la Memoria Departamental anual...), pero a otro ritmo y con una finalidad diferente.

9. Probablemente la disociación de responsabilidades en la evaluación de estos aspectos pone de manifiesto la diferente valoración de la importancia que da el MEC a la docencia y la investigación. Se mantiene un control más cercano de lo que se considera verdaderamente sustantivo en la labor del profesor universitario.

Sopesando ventajas e inconvenientes, lo que pudo suponer un avance en la gestión democrática de una Institución Pública tan característica como es la Universidad, se tornó en un elemento negativo. Hay que recordar que la evaluación de la Docencia y la Investigación se propone en un contexto en el que están aún presentes las reivindicaciones salariales del profesorado universitario, que reclaman su equiparación a niveles similares del resto de la Administración Pública. La publicación del Decreto evitó una respuesta reivindicativa masiva al desarticular con un proceso individualizado los movimientos colectivos que hasta un tiempo antes habían estado actuando.

Paralelamente, los procesos evaluativos en marcha, cuyo carácter estaba orientado esencialmente a la mejora, quedaron “aparcados”, de forma que se auspició una institucionalización forzada de la evaluación administrativa del profesor. En las mesas de muchos responsables de gestión universitaria quedaron las propuestas de desarrollo y utilización formativa de la evaluación docente y fueron sustituidas por la preocupación de cómo lograr “dulcificar” el uso sumativo de la evaluación.<sup>10</sup>

Las universidades que habían puesto en marcha procesos evaluativos realizaron su adaptación a la nueva situación y las restantes se fueron adecuando a la normativa aprovechando experiencias, procesos e instrumentos de las universidades anteriores. No obstante, buena parte de las universidades continúa funcionando teniendo como única fuente real de información los cuestionarios de estudiantes,<sup>11</sup> y atemperando sus decisiones con criterios muy laxos. En nuestra opinión, ciertamente, no pueden hacer otra cosa. Se deberían destinar mayores recursos a la evaluación, para hacerla con mejor calidad, si se desea aplicar el sistema con mayor profundidad. Muchos de estos procesos han caído en el uso rutinario, como elementos de demostración de la responsabilidad institucional. En general demuestran una escasa capacidad de cambio, de innovación, derivada de su ambigüedad: no son procesos netamente de rendición de cuentas —a partir de los cuales se diferencie entre los profesores—, ni tampoco enlazan con estrategias orientadas a la mejora.

Respecto a la Evaluación de la Investigación del profesorado (o personal de investigación) los problemas son otros. El sistema empleado contraviene la mayor parte de criterios de bondad de un sistema de evaluación, lo que lo con-

10. Entiendase que este planteamiento es genérico y, por ello, no necesariamente representativo de todas las situaciones. En bastantes universidades del Estado el nivel de institucionalización no se corresponde ni tan siquiera con los requerimientos establecidos en el citado decreto. En otras, concretamente en la Universitat de València, el equipo de gobierno actual ha manifestado reiteradamente su preocupación por integrar de forma positiva la evaluación en la organización, de manera que se extraigan consecuencias de mejora a partir de la información evaluativa. No obstante, la institucionalización de la evaluación permanece en su carácter de mera respuesta a la demanda administrativa sin que a nivel técnico se hayan desarrollado procesos orientados a la mejora.

11. Prueba de que ello se utilizó con la finalidad de resolver una reivindicación salarial es que la evaluación docente se acabó convirtiendo en su uso en un “café para todos”, al no revisar el Ministerio si los procesos por los que se realiza la misma se adaptan o no al Decreto, mientras que la evaluación de la Investigación es la que produce diferencias reales —que no justas—.

## Evaluación de la Actividad Universitaria

vierte en un sistema poco justo y que no consigue los objetivos declarados para los que se propone formalmente (mejora de la investigación tanto en calidad como en productividad). Prueba de ello es la gran cantidad de recursos interpuestos en los Tribunales de Justicia y que han sido fallados a favor de los demandantes (profesores y/o investigadores).

El problema de la Evaluación de la Investigación radica en que técnicamente no es un sistema que consiga ser creíble para todos los implicados. Los elementos que mayores críticas han concitado se pueden sintetizar en los siguientes puntos:

- a) *¿Quién evalúa?* La evaluación la realizan unos Comités designados por la Administración para grandes áreas Académicas. Aunque dichos comités funcionan con el asesoramiento de investigadores también designados por la Administración, en ellos no existe en su composición representantes de todas las Áreas de Conocimiento (Estructuras de organización del conocimiento científico por las que se organiza la Universidad Española), ni tampoco de las Universidades. Ello hace que muchos miembros de los colectivos evaluados identifiquen más claramente a estos Comités por su carácter político más que académico. Asimismo, los miembros de estos Comités en las primeras evaluaciones no se hacían públicos. Sólo muy recientemente se han hecho públicos debido a las presiones sociales.
- b) *¿Qué se evalúa?* El sistema pretende acercarse a una evaluación de la calidad de la Investigación. No obstante, los Comités disponen para su valoración de cinco títulos (de trabajos) propuestos por el sujeto evaluado para el periodo que se evalúa. No se pretende evaluar la productividad, pues, aunque se solicita un currículum completo del periodo, en la Normativa no se especifica su uso. El hecho de partir de los títulos de los trabajos es un elemento que rebaja la credibilidad ya que parecen no utilizarse indicadores –Índices de impacto de revistas, grupos, personas, etc.– que objetiven (y, en su caso, automaticen) la evaluación. No obstante, en 26 de octubre de 1995 se publica el Real Decreto, en el que aparecen “Criterios de Evaluación” para las Comisiones de las diversas áreas Académicas. Sin embargo, éstos son más formales que reales y aluden al uso de catálogos de investigación como referencia para la evaluación. En cualquier caso, estos catálogos, de carácter internacional, tienen indizada poca información española y, no son pues, el mejor referente. Por ello es de suponer que dichos Comités utilizarán otros medios adicionales para valorar el impacto de la investigación. Asimismo, es bien conocido que la calidad de estos catálogos difiere entre las disciplinas de que se trate, de forma que es reconocida mayor calidad a los referentes en las áreas de Ciencias Básicas que en las de CC. Sociales (por ejemplo), y en el caso de las CC. Humanas son extremadamente problemáticos.

En cualquier caso, entendemos que el Decreto ofrece un paso sustan-

cial en la dirección que se demandaba, aunque corre el riesgo de quedarse en aspectos formales al no haberse planteado las cuestiones sustanciales que acabamos de mencionar. No obstante, quedan serios interrogantes en relación a cómo se han desarrollado las evaluaciones anteriores a la aparición del Decreto.

Previamente a la publicación del Decreto que acabamos de comentar, se publicó en 1994 –Real Decreto de 5 de diciembre de 1994– una normativa unificada de la Evaluación de la Investigación. En ella tampoco se incluyeron criterios (que como hemos señalado se publicaron con posterioridad). Un aspecto a destacar en este Decreto es que se demanda a los sujetos evaluados que, junto a los trabajos seleccionados para su evaluación, aporten lo que denominan “Indicios de Calidad”; éstos, son, esencialmente, una relación de citas que ha tenido el trabajo seleccionado, referencias publicadas de otros autores al mismo, etc... Respecto a este punto, creemos que no es precisamente serio el procedimiento en que se demande a los evaluados que hagan el trabajo que debería hacer la entidad evaluadora. Tómese como anécdota que a muchos profesores universitarios se les ha planteado la necesidad de abrir una nueva línea en su investigación: rastrear quién les cita.<sup>12</sup>

- c) Entonces, *¿cómo evalúan?*, *¿con qué criterios?* Teniendo en cuenta lo anterior, es razonable pensar que los criterios o son fundamentalmente subjetivos o de carácter intersubjetivo –consensuados dentro del Comité–; en cualquier caso, no son conocidos (dado que lo publicado no puede entenderse como criterios, sino en todo caso como líneas orientativas acerca del tipo de indicadores). Ello es la base de que exista en buena parte del colectivo evaluado la sensación de que éstos son, desiguales: dependientes del área académica (o de conocimiento incluso) de que se trate, del periodo evaluado, etc. Pero la opacidad del sistema no permite señalar cuáles son. Máxime cuando la única información que recibe el investigador acerca de su evaluación es una calificación numérica sin ninguna otra explicación de la misma. Lo cierto es que en los recursos fallados a favor de los demandantes, los Tribunales de Justicia exigen que se vuelva a evaluar al sujeto y se le explique la calificación.

*En síntesis*, el hecho de que no existan unos criterios públicos que traten de objetivar el proceso, así como que se pretenda evaluar la calidad de una investigación que los Comités pueden desconocer, conduce a que la credibilidad de todo el proceso radique sobre el funcionamiento interno de los Comités. Como se ha señalado en otros trabajos, cuando el enfoque es tan subjetivo, resulta crucial la composición y forma de designación de los Comités, así como otros aspectos relativos a su funcionamiento.

12. Dentro de unos años puede ser interesante estudiar los “Colegios Invisibles” que auspicia desarrollar este sistema. Algunos ya están funcionando de hecho.

*¿Qué componentes deberían haberse tenido en cuenta en un sistema de Evaluación de la Investigación?* La respuesta no es fácil, pero existen algunos componentes que están presentes en los sistemas de otros países que se podrían haber considerado:

- a) Si se pretende algún tipo de juicio cualitativo, es muy necesario que los Comités estén formados por representantes académicos, con presencia de todas las Áreas de Conocimiento y representantes de las Universidades. Éstos podrían ser designados entre los miembros del Cuerpo de Catedráticos con algún tipo de aleatorización pura o corregida. En el futuro, cuando ya exista una base de discriminación en función de los tramos alcanzados, este parámetro se debería de tomar en consideración para la constitución de los Comités, en lugar de la de nivel académico (por ejemplo, aquellos que tengan más de tres tramos reconocidos).
- b) Los criterios de juicio deben ser públicos, previamente conocidos por aquel que se somete a evaluación y todo el proceso de adopción de decisiones debe ser transparente. Es decir, los criterios, con el grado de especificación que sea posible, deben ser públicos, a modo de cualquier baremo. No vamos a entrar en la dificultad que tiene la evaluación de la investigación. Tanto en nuestro contexto,<sup>13</sup> como en otros países, se ha ido abordando el tema y existe una amplia metodología que permite un acercamiento más objetivo que el que se está utilizando. De hecho se propone como más efectiva una combinación de aspectos objetivables junto con los elementos de juicio más cualitativos. No obstante, el problema no está resuelto en todos sus términos con este enfoque. Así, existen importantes diferencias entre los campos científicos –por ejemplo, entre las ciencias y las humanidades y/o ciencias sociales– que debe procurarse atender al construir indicadores o criterios con los que juzgar. Otra polaridad de tensión la constituye la contraposición entre funciones básicas y aplicadas del actuar científico. Así, es fácil caer en la preeminencia de una rama sobre otra, lo que distorsionaría gravemente la evaluación. Particularmente, en nuestro contexto tiende a haber una mayor tradición de valorar las actividades de ciencia básica que las aplicadas –tanto las patentes, como informes, proyectos, asesoramientos, etc. constituyen procesos y documentos no bien reconocidos en cuanto a su valor–. Esto supone una contradicción con el planteamiento que formalmente se asume de incrementar la inserción de la universidad en el seno de la actividad social.
- c) Finalmente, entendemos que, en cualquier caso, un problema grave es el de la falta de contextualización. Así, entendemos que la Evaluación

13. En la Universitat de València, los trabajos del prof. López Piñero y su grupo de investigación han sido pioneros, a nivel nacional e internacional, del estudio de métodos bibliométricos, análisis de revistas, su impacto, etc... Métodos que pueden ayudar a hacer más objetiva la evaluación de la investigación, tanto la evaluación del producto individual como del colectivo.

del profesorado universitario debe realizarse de forma conjunta, integrando la información de las tres dimensiones que componen su actividad (Docencia, Investigación y participación en la Gestión).

Por último, los planes de evaluación que se han ido poniendo en marcha suponen avances en la mejora del sistema universitario. No obstante, la evaluación, como instrumento, puede estar al servicio de funcionamientos democráticos o autocráticos, de la innovación o de la mejora real o de la aparente y superficial. Estimamos que es necesario desarrollar una cultura de la evaluación, de base democrática, que contemple la participación efectiva de todas las audiencias implicadas en el proceso y que técnicamente busque alcanzar las mayores cotas de fiabilidad y validez. Sólo desde esa óptica tiene verdadero sentido un desarrollo de planes de evaluación.

## **2. Contextualización de la Evaluación de la Universidad**

Un problema de gran importancia que suele apreciarse en el desarrollo de Planes de Evaluación y que, por desgracia, no es un aspecto generalmente bien atendido en los mismos, es el de la contextualización de la Evaluación. Como señalan diversos autores (DOCHY, SEGERS y WIJNEN, 1990; ESCUDERO, 1991; DE MIGUEL, 1991) la contextualización constituye una condición básica para abordar la evaluación de cualquier unidad de análisis. Si tenemos presente que el objetivo de todo proceso evaluativo es recabar información, de manera estructurada, para posibilitar la consiguiente toma de decisiones,<sup>14</sup> apreciaremos inmediatamente la relevancia del problema. La contextualización de la información que se reúna será clave para otorgar significatividad a la misma y, por tanto, afectará decisivamente al conocimiento profundo del objeto evaluado. Ciertamente, el fenómeno afecta a los criterios de bondad de cualquier evaluación, fundamentalmente a su validez, y tiene implicaciones importantes para su fiabilidad. La consecuencia inmediata de una falta de contextualización de la Evaluación es que se puede producir un atentado contra uno de los pilares fundamentales del proceso evaluativo, su equidad –justicia– (JORNET, 1991).

No obstante, este hecho que, a menudo se utiliza como argumento para elegir opciones evaluadoras alternativas, generalmente no se ha abordado de manera sistemática. Sin entrar en valorar qué tipo de Planes de Evaluación atienden mejor estos requerimientos –que sería sin duda muy discutible–, parece claro que la falta de contextualización será un problema tanto más grave dependiendo del uso que se vaya a realizar de la evaluación.

En Planes de Evaluación Formativa, los elementos contextuales que afectan a la unidad que se esté evaluando tienden a estar integrados en el proceso. Es

14. Aunque no siempre la evaluación tiene como objetivo final apoyar la toma de decisiones, estas reflexiones sirven esencialmente para cualquiera de las facetas alternativas que se puedan plantear –por ejemplo, recabar información y sintetizar para describir una situación–.



más, el tipo de metodología que guía estos planes, generalmente con mayor incidencia en los aspectos cualitativos, presenta de forma natural la suficiente flexibilidad para atender variables contextuales y situacionales y, en cualquier caso, la finalidad de la evaluación restringe notablemente la gravedad de la incidencia del problema. Los defectos de contextualización repercutirán en una peor adecuación del enfoque, lo que, como máximo, desembocará en un aprovechamiento más pobre de los procesos formativos en marcha. Este hecho, sin ser menor, resulta menos grave que los problemas derivados en las evaluaciones orientadas por otros objetivos.

En los Planes de Evaluación que tengan alguna consecuencia sumativa es donde se pueden identificar los problemas más graves derivados de las deficiencias de contextualización. El hecho de tener que establecer una vinculación entre la evaluación (sus resultados) y las decisiones (sus consecuencias) conlleva que la contextualización sea un factor imprescindible en la evaluación. Es decir, situar los hechos que se reúnen en el proceso evaluativo para la toma de decisiones dentro del marco de referencia —contexto— es lo que les da un verdadero significado y permite que sean correctamente interpretados. Por ejemplo, valorar la adecuación en el funcionamiento de una institución universitaria respecto a un indicador como la capacidad de captar recursos exteriores de apoyo o financiación para la investigación, conducirá a conclusiones muy diversas si se tiene presente el contexto en el que está inmersa esta institución. Este contexto puede entenderse como el entorno geográfico, sociocultural o socioeconómico en un momento dado —si se trata de una universidad de la Comunidad Valenciana o de una universidad holandesa— y el análisis de las tendencias de evolución que muestran la institución y la sociedad. De este modo, no es lo mismo una universidad española en los años 70 que en la actualidad, pero toda visión de cualquier institución deberá entenderse como condicionada por su historia.

Contextualizar un plan o programa de Evaluación deber ser atendido a través de todas las dimensiones de su implementación (WOLF, 1990), pudiéndose identificar los factores de contextualización al menos a tres niveles: teórico-metodológico, métrico y procesual. Aunque no es nuestro objetivo en estas líneas desarrollar todos los problemas implicados en cada uno de ellos, revisaremos algunos elementos que estimamos de especial interés en el tema que nos ocupa, la Evaluación Universitaria.

## **2.1 Estrategias teórico-metodológicas**

El nivel inicial de análisis que debe considerarse para asegurar una contextualización adecuada del proceso de Evaluación es la especificación de una teoría acerca del fenómeno a evaluar. No vamos a discutir respecto al valor intrínseco de la Teoría como orientador de cualquier proceso de investigación o evaluación —cabén múltiples enfoques epistemológicos y metodológicos—. No obstante, muchos de los problemas que se dan en el desarrollo de evaluaciones universitarias podrían —deberían— ser anticipados desde un planteamiento teórico inicial. La existencia de una teoría origen acerca del objeto a evaluar determi-

nará la operacionalización de todo el proceso, desde el planteamiento de hipótesis, la selección de unidades de análisis e identificación de variables, el diseño, la instrumentación, el análisis de datos, las estrategias de utilización de la información, etc. De este modo, tomar como punto de partida un modelo de referencia, definido desde cualquier orientación, permite establecer estrategias de evaluación mejor contextualizadas, en definitiva, más válidas.

En la Evaluación Universitaria –como señalamos al comienzo de este comentario– caben diversas posiciones teóricas de identificación del objeto (institución, programa, profesor, etc.) y, en cualquier caso un análisis profundo de cada uno de ellos pone de manifiesto la complejidad a la que hay que atender en su evaluación; así como podrá proveer una imagen global del fenómeno que permita graduar –para los fines a los que se dirija la evaluación– las posibilidades de logro de evidencias de validez, dependiendo de la conjunción del objeto, fines y propósitos, medios, etc., de la Evaluación. No obstante, ni la complejidad de la situación a aprehender, ni las técnicas analíticas disponibles justifican omitir pasos en el proceso teórico-metodológico de construcción de cualquier modelo evaluativo.

Podemos, pues, referirnos al concepto campbelliano de Validez Teórica o a la Credibilidad de la Explicación de KRATHWOHL (1987) para comprender que la contextualización es una dimensión central en las garantías esenciales o criterios de bondad que debe asumir el proceso evaluador desde su inicio.

En cualquier caso, dependiendo del Enfoque evaluativo de que se trate, dentro del rol que aquí se reclama para la teoría, se integra la idea o concepto de “calidad”. Ciertamente, este concepto es ambiguo y escurridizo, de carácter multidimensional, y dependiente del medio sociocultural y del momento histórico en que se inserta. Una propuesta que intenta sintetizar los sectores que interactúan para la calidad universitaria se recoge en la Figura 2, donde se apre-

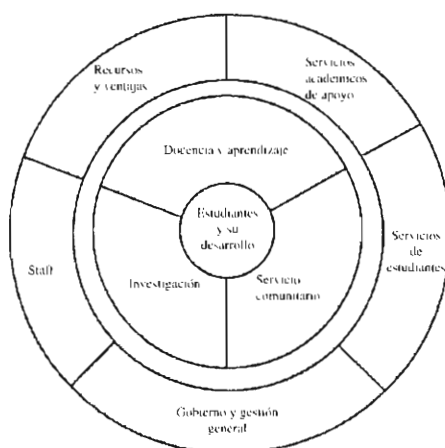


Figura 2. Sectores interactivos de la gestión de calidad de la Universidad (Meade, 1995, p. 76)

cia esta complejidad. Por ello, generalmente, se suele abordar por parte de los evaluadores a partir de estrategias que “operacionalizan” el concepto a través de las variables que se seleccionan para la evaluación. Así, por ejemplo, en evaluaciones institucionales lo más usual es abordar el concepto –y la selección de variables– desde un análisis de sistemas en los que se relacionan los conceptos de eficiencia y eficacia institucional –ver ESCUDERO, 1991–. Por su parte, en evaluaciones de la docencia del profesor, normalmente, tras definir unos conceptos mínimos a considerar (como, por ejemplo, cumplimiento de obligaciones, adecuación de las clases –claridad expositiva, estructuración...– sistemas de programación, evaluación, etc.) la expresión más clara del concepto de calidad deviene de su plasmación en los propios instrumentos, como los cuestionarios utilizados para recabar las opiniones de estudiantes, o en los indicadores utilizados para ese propósito.

No obstante, la operacionalización del “concepto de calidad” en múltiples ocasiones hace las veces de Teoría en la orientación de los procesos de evaluación. Por consiguiente es especialmente relevante en cualquier Plan de Evaluación establecer los órganos y los mecanismos necesarios para que se produzca una reflexión sobre el concepto de calidad asumido y sus implicaciones para el conjunto del proceso de evaluación.

Un segundo nivel de contextualización es el que puede obtenerse del mero análisis del objetivo de la evaluación, las demandas de los promotores, la interacción con las audiencias implicadas y la definición de la finalidad. Este segundo nivel hace referencia a un análisis de orden más pragmático, guiado por las dimensiones básicas de la evaluación que sintetiza ESCUDERO (1991) y orientado por algún modelo de evaluación. Este autor, tras revisar diversos planteamientos al respecto, propone un sistema de trece dimensiones o elementos definitorios en el proceso de construcción de un modelo evaluativo:

- objeto de la evaluación;
- propósito, objetivos; audiencia/clientelas;
- énfasis/aspectos prioritarios o preferentes;
- criterios de mérito o valor;
- información a recoger; métodos de recogida de información;
- métodos de análisis;
- agentes de la evaluación;
- secuenciación del proceso;
- informe de la evaluación/utilización de resultados;
- límites de la evaluación; y
- control y evaluación de la propia evaluación/meta-evaluación.

La propia complejidad de las situaciones que ordinariamente se abordan en los programas de Evaluación de carácter sumativo, junto con la presión de las diversas audiencias implicadas sobre los procesos de puesta en marcha, conlleven que no se establezcan de forma satisfactoria las condiciones de contextualización. Paradójicamente son precisamente estas carencias las que se constituyen en fuentes generadoras de tensión y rechazo respecto al plan evaluativo. Es

preciso, pues, destacar la importancia de establecer mecanismos— que permitan alcanzar una adecuada contextualización de los mismos. Ello debe realizarse desde la puesta en marcha de los programas de Evaluación Universitaria, especialmente si conllevan decisiones sumativas.

La adopción de períodos amplios de revisión del Plan de Evaluación que se está desarrollando, permite analizar los elementos que tienen un impacto crítico sobre la contextualización de los Planes de Evaluación. Como, por ejemplo, la puesta en marcha, la generación de bancos de información e instrumentos, la construcción progresiva del modelo y su aval con información empírica, el establecimiento de procesos de implicación de las audiencias con el proceso, etc. En definitiva, se trata de establecer un marco de trabajo en el que se avance progresivamente de manera armónica —implicando tanto a los responsables institucionales, las audiencias, el equipo técnico de evaluación, etc.—, especificando los elementos del modelo de evaluación en cada caso, reuniendo las evidencias de validez necesarias y tomando las decisiones que sean coherentes con la solidez del modelo en cada momento.<sup>15</sup>

En consecuencia entendemos que la mejor forma de dar solidez a los Planes de Evaluación es enmarcarlos, tanto en su desarrollo como en su utilización, en procesos de investigación evaluativa, en los cuales se mantenga un análisis del propio sistema de evaluación que aporte informaciones contextuales y evidencias de validación. La contextualización, dimensión esencial de la Validez Teórica del proceso evaluativo, debe estar presente en la planificación, desarrollo y conclusión de todo proceso evaluativo. De hecho, modelización teórica y desarrollo teórico-empírico, no son excluyentes y suelen darse concomitante o secuencialmente. Es más, pueden ocurrir problemas de adaptación entre el modelo teórico y las demandas de evaluación. Un ejemplo de esta tensión —entre el planteamiento teórico de calidad docente y su evaluación desarrollada en las universidades— es la que se puede apreciar con la evaluación de la docencia centrada en el profesor. En este caso, aunque existan concepciones teóricas sobre la calidad docente operacionalizables, sería bastante absurdo pretender evaluar a un profesor universitario tomando como referencia, además de otras dimensiones, habilidades específicas en técnicas docentes. En la actualidad este enfoque estaría totalmente descontextualizado. Los profesores universitarios, en general, no han recibido formación como docentes —lo que es común a otros países de nuestro entorno (VAN OS, 1990)—, el colectivo evaluado necesariamente reaccionaría por lo injusto del tratamiento —al menos—, y ni tan siquiera los requerimientos actuales de las instituciones universitarias se lo plantean más allá de unos mínimos aceptables en la realización de la función docente.

Si bien la contextualización es un aspecto a exigir en cualquier proceso evaluativo, se debe tener en cuenta que los elementos y/o variables que ayuden a

15. De ahí que en la implementación de sistemas de evaluación de profesorado, en nuestras universidades, los equipos técnicos encargados reclamaron insistentemente que no se pretendiera vincular los resultados de la evaluación a decisiones sumativas durante el proceso de desarrollo del sistema y que se utilizara prioritariamente con una orientación formativa, acompañándolo de sistemas de apoyo para la mejora docente.

la misma deberán ser minuciosamente concretados en función del objeto a evaluar. Por ejemplo, para la evaluación de instituciones universitarias se atenderá a cuestiones normalmente muy diferentes de las que se deben tener en cuenta para la contextualización de la evaluación de la actividad investigadora del profesorado. De esta forma, la evaluación de una institución universitaria atenderá a cuestiones como índices comparativos de asignación económica de otras instituciones similares –en tamaño, tipo de estudios, etc.– o índices anuales –durante la última década– respecto a las ratio profesor/estudiante que ha tenido esta institución. Si nos referimos a la evaluación de la producción investigadora de un profesor se pueden atender a aspectos como la productividad del ámbito o área en la que trabaja o la evolución de la productividad modulada en función de determinadas infraestructuras disponibles –por ejemplo, la existencia de servicios electrónicos de información–.

En definitiva se trata de establecer, en primer lugar, en el modelo teórico las condiciones de contextualización y, en segundo lugar, desarrollar en el plano teórico–metodológico las circunstancias concretas a que se atenderá para garantizar la significatividad del contexto para la función evaluativa que se pretende.

## 2.2 Estrategias métricas

Separar el plano métrico del metodológico, aunque es una estrategia de muy frecuente utilización por la importancia decisiva del hecho de medir en el ámbito de las ciencias humanas y sociales, es un artificio que no debe distorsionar la unidad de conjunto respecto al problema que pretendemos abordar. De ser así, si acabamos percibiendo ambos discursos como ámbitos estanco separados, esta estrategia obraría justo al contrario del objetivo que pretende cubrir: ahondar más en un punto especialmente crítico del mismo proceso cual es la medición. Es preciso asumirlo como una faceta o nivel del mismo proceso global de contextualización como lo es de la acción evaluativa en general.

Existen al menos dos temas nucleares a abordar en el proceso de contextualización desde el plano métrico. En primer lugar, partiendo de la determinación previa de las variables, conseguir el adecuado desarrollo de sus indicadores concretos que se van a utilizar desde esta perspectiva. En segundo lugar, el establecimiento de un proceso para la determinación del estándar de calidad frente al cual se adoptarán, consecuentemente, las decisiones. Vamos a revisar someramente algunas consideraciones en torno a estos dos polos de acción.

La *determinación* en el modelo evaluativo de las *dimensiones que afectan al problema*, sus roles y relaciones es una parcela fundamental que tiene su continuidad en el plano métrico en cuanto al desarrollo de los instrumentos precisos para obtener esta información. Como ya hemos indicado previamente, el tema de la medida es crucial en el ámbito de las ciencias humanas y sociales y, por consiguiente, en los procesos evaluativos. En este sentido, el desarrollo de indicadores se constituye en un punto fundamental que ya hemos visto concita una buena parte del interés de trabajo en éste ámbito.

Las aportaciones en el campo de la formalización y medición de variables

contextuales son más bien escasas y en general se aprecia una visión restringida del contexto. Además, como el conjunto de propuestas y estrategias evaluativas se están viendo sometidas a presiones considerables que reclaman insistentemente una reducción de los sistemas de evaluación esgrimiendo razones de coste y/o utilidad. Estos argumentos, sin duda relevantes, suelen ejercer una influencia en el sentido de sobresimplificar los sistemas que se emplean y acortar los tiempos y posibilidades de desarrollo. Entre otros aspectos, el análisis del contexto y su traducción en indicadores que permitan considerarlo en relación al objeto evaluado es uno de los máximos candidatos a cualquier reducción ya que no se trata de indicadores directos de rendimiento/funcionamiento de la unidad –institución, programa, profesor, etc.– a evaluar.

Es evidente que los problemas más graves se vuelven a producir, de nuevo, cuando se pretenden extraer consecuencias sumativas en el proceso evaluador, a cualquier nivel. Así, en los estratos que hemos establecido respecto a los enfoques evaluativos –institucional, de programas o de profesores– se han producido reflexiones acerca de la necesidad de utilizar y medir variables que sitúen contextualmente la decisión. YORKE (1991, pág. 246), desde una perspectiva de evaluación institucional, demanda la necesidad de aplicar de forma consistente y habitual “juicios enraizados contextualmente a los datos de rendimiento, ‘triangulándolos’ donde sea posible con pruebas cruzadas”.

Con respecto a la contextualización existen escasas propuestas estructuradas respecto a qué y cómo medir para aportar un adecuado contexto a la decisión. Una de las más rigurosas la aportan DOCHY, SEGERS y WIJNEN (1990) dentro de un enfoque de gestión enmarcado en la corriente de evaluación institucional y se concreta en un listado de variables e indicadores orientados a este propósito, diferenciados por su carácter cuantitativo o cualitativo –ver Cuadro 4–. Como puede apreciarse,<sup>16</sup> incluso dentro de este enfoque resulta un planteamiento muy restrictivo que sería poco eficaz para el objetivo que formalmente se declara. En este sentido, es preferible acogerse a un planteamiento de evaluación mixta y fiar la contextualización en las informaciones derivadas de las autoevaluaciones (NEDWEK y NEAL, 1993). No obstante, esto no elimina la problemática por completo, ya que sería igualmente decisivo establecer el sistema generalizado o integrador para interpretar a nivel global los datos así contextualizados.

Otro nivel de contextualización que puede y debe tenerse en cuenta es el que se aborda desde la *determinación de criterios de evaluación*. La existencia de unos criterios de evaluación y unos procedimientos para su aplicación, implí-

16. Como hemos señalado anteriormente, la definición completa del problema resulta crucial y afecta a muchos de los criterios a emplear en la determinación de las dimensiones de contextualización, sus medidas y sus roles. En el ejemplo del Cuadro 4 se incluyen algunos indicadores que otros expertos en evaluación catalogarían como de Entrada (siguiendo el modelo CIPP). Este es el caso del conocimiento previo que tienen los estudiantes. En otras propuestas, en cambio, la construcción de este indicador se sustentaría sobre pruebas estandarizadas, con lo que sería una medida de tipo cuantitativo. Debe entenderse, pues, como una propuesta que intenta ejemplificar una aproximación bastante sistemática, respecto a lo que resulta habitual en su ámbito, de determinación de indicadores de contextualización.

Cuadro 4  
 MEDIDAS, DIFERENCIADAS POR TIPOS, PARA LA  
 CONTEXTUALIZACIÓN EN LA EVALUACIÓN INSTITUCIONAL  
 (Adaptado de Dochy, Segers y Wijnen, 1990, Pág. 149)

Tipos	Indicadores	Variables
Medidas Cuantitativas	Estudiantes que entran	<ul style="list-style-type: none"> <li>• Número total de matrículas por institución, dedicación completa (ft), dedicación parcial (pt)</li> <li>• Ídem por facultad, carrera (ft, pt)</li> <li>• estructura de la población de estudiantes por edad</li> <li>• distribución porcentual de estudiantes varones/mujeres</li> <li>• calificación final en secundaria/selectividad</li> </ul>
	Personal  Temas financieros	<ul style="list-style-type: none"> <li>• ratio estudiantes/profesor</li> <li>• número de horas de clase por tutor</li> <li>• estadística del número de horas lectivas perdidas por el profesorado/profesorado soporte por enfermedad</li> <li>• costes de enseñanza comparados con otros costes</li> <li>• porcentaje de coste por equipamiento</li> <li>• porcentaje de coste por biblioteca, porcentaje de coste por centro audio-visual</li> <li>• porcentaje de gasto en proyectos de innovación</li> <li>• ratio de gasto total en relación al presupuesto para la innovación</li> <li>• porcentaje de presupuesto gastado en la formación continua</li> <li>• análisis del gasto por objetivos educativos, revisión de los presupuestos por tipos de costes</li> <li>• presupuestos frente a gastos reales</li> <li>• proporción de metros cuadrados para la docencia por profesor, por institución</li> </ul>
Medidas Cualitativas	Infraestructura y posibilidades materiales	<ul style="list-style-type: none"> <li>• procedimientos para la admisión en educación superior</li> </ul>
	Estudiantes que entran	<ul style="list-style-type: none"> <li>• calidad de los estudiantes: conocimiento preuniversitario</li> <li>• origen geográfico de los estudiantes</li> <li>• conocimientos previos de los estudiantes</li> </ul>
	Personal Temas financieros  Infraestructura y medios materiales	<ul style="list-style-type: none"> <li>• Posibilidades (del profesorado) para el soporte educativo</li> <li>• gestión financiera</li> <li>• gasto presupuestado respecto al real</li> <li>• centros audiovisuales, centros de medios de comunicación, etc., como medios de soporte para la instrucción</li> </ul>

citos o explícitos, es una característica común esencial a cualquier acercamiento evaluativo. Si el resultado de la evaluación de un programa, en alguna de sus facetas, pretende vincularse a algún tipo de decisión o propósito, ésta deberá estar referida a algún criterio. Existen diversos medios de determinar un criterio con el que comparar una ejecución individual y, bien es cierto, que en el panorama de utilidades suelen encontrarse suficientes usos indebidos como para rechazar las evaluaciones. Un ejemplo habitual es el que pone ESCUDERO (1991) relativo a la comparación de niveles promedio de profesores de departamentos diferentes como una utilización claramente indeseable y, como él señala, muestra de una evidente descontextualización.

No obstante, hay que tener presente que este tipo de ejemplos, aunque en ocasiones se da, no refleja el marco métrico de utilización correcta y posible de criterios evaluativos. Para desarrollar adecuadamente un proceso de evaluación de cualquier índole, se deberá acompañar de un sistema de determinación de estándares y niveles mínimos de aceptación. En definitiva, esto exige el establecimiento de algún criterio que soporte el juicio de valor que tendrá al menos una expresión dicotómica (evaluación positiva y negativa, bien/mal evaluado). En este sentido, el problema que se plantea no es muy diferente del que se da en cualquier otra área de evaluación educativa: se trata de crear un vínculo entre las mediciones realizadas y la decisión final, estableciendo patrones estables y congruentes para cada una de las modalidades posibles de decisión.

Sin embargo, aunque los procesos de determinación de estándares pueden ser similares para los planes evaluativos institucionales, de programas o de profesorado, hace falta realizar algunas consideraciones previas:

- a) Un estándar puede constituir un resumen más adecuado para la evaluación cuanto más se aproxima su utilización a unidades de análisis simples. Por ejemplo, la opinión de los estudiantes sobre la docencia impartida por un profesor. Actualmente, no existen indicios que nos permitan precisar mejor cuáles son los límites en los que resultan eficaces los modelos matemáticos que apoyan la asociación de un estándar, un punto de corte y los criterios de bondad de la decisión. En contextos en los que por la complejidad de la situación que se aborda –evaluación de instituciones, servicios y programas amplios,...– no se tenga la suficiente certidumbre sobre el procedimiento, se puede utilizar como estrategia la aplicación de este procedimiento a cada dimensión o conjunto de dimensiones que se concreten dentro del modelo. Esto conduciría al establecimiento de un perfil combinado que finalmente se podría sintetizar en un único criterio de decisión si así lo exige el modelo.
- b) Siempre que se desarrolle un proceso de evaluación orientado a apoyar decisiones, es imprescindible determinar los estándares que reflejen los criterios de calidad.
- c) Ello es ineludible en planes de carácter sumativo, mientras que tiene características particulares en planes formativos, en los que se integra en los procesos de análisis en la interacción unidad evaluada/evaluador.
- d) Los mayores problemas de contextualización se producen indefectiblemente al abordar el establecimiento de estándares en procesos orientados a la Rendición de Cuentas. Sobre ellos centraremos nuestro análisis posterior.

No vamos a extendernos en revisar las soluciones al respecto pues son ampliamente conocidas. Únicamente queremos presentar una reflexión acerca de aplicaciones en este contexto. De esta manera, entre los métodos de determinación de estándares podemos distinguir tres grandes agrupaciones de Modelos: a) Normativos, b) Opináticos, sin referencia empírica y c) Criteriales. Pasamos a revisarlos brevemente:



Los *estándares normativos* son aquellos orientados a determinar la posición relativa de un individuo respecto a su población de referencia. El hecho de que se basen completamente en una posición relativa provoca que puedan ser poco deseables en multitud de situaciones educativas. De hecho existen situaciones muy diversas relacionadas con el hecho educativo en las que se desea alcanzar determinados niveles fijados y la valoración depende esencialmente de si se consigue o no, dejando a un lado lo que puedan conseguir otros referentes de actuación. Por ejemplo, si una de las condiciones para valorar positivamente el programa se basa en la satisfacción de los usuarios y se utiliza una escala valorativa para averiguarlo (Muy satisfactorio, Satisfactorio, Indiferente, Insatisfactorio, Muy Insatisfactorio) todo nivel que se sitúa por debajo de la satisfacción declarada explícitamente debe conducir a una valoración no positiva respecto a esta dimensión. Parece poco adecuado en este contexto proporcionar una valoración positiva a un programa cuyo promedio se sitúa en la categoría de Insatisfactorio por el hecho de que otros programas *semejantes* han obtenidos este nivel u otro inferior.

No obstante, un trabajo minucioso de referencia normativa es de gran valor como retroalimentación para el proceso evaluativo. En este sentido, aportar a un departamento universitario información acerca de su posición relativa respecto a otros de su entorno, puede ser una gran ayuda para enriquecer su interpretación, contextualizarlo, y ayudarle a construir unas estrategias de evolución mucho mejor ancladas en su realidad inmediata.

Considerando que el tipo de criterio para la evaluación debe ser, pues, de carácter absoluto –entendido como un nivel mínimo, consensuado como expresión de calidad– referido a la calidad, el marco de trabajo posible se sitúa en los dos restantes. El planteamiento será pues comparar la ejecución individual con un nivel criterio establecido por los responsables del Plan de Evaluación. Para este cometido se debe elegir un sistema que garantice la concreción de este criterio en función de un acuerdo intersubjetivo alcanzado por las diferentes audiencias implicadas en cada caso. Así, para la valoración de un programa universitario, deberán atenderse las posiciones del personal de la universidad (profesores, personal de administración y servicios, departamentos, gestores,...), de los estudiantes, de los profesionales del entorno que se pretende abordar, de las empresas y/o entidades para cuyo contexto se pretende preparar, de los responsables de la política educativa... A tal efecto se abordará la construcción de un proceso que permita la adecuada intervención de estas partes en la definición de lo que se considera como calidad en la entrada, proceso y producto de la actividad de formación que constituye el programa.

La determinación *opinática* supone establecer un nivel mínimo sin considerar referencias empíricas acerca de las consecuencias de aplicación del criterio. Por ejemplo, para evaluar si la opinión de los estudiantes acerca de un profesor es positiva, partir de que las medias de respuestas deban ser iguales o superiores al punto medio de la escala.<sup>17</sup> O bien, cuando se combinan diversas

17. En una Escala del tipo I: 1: Muy en desacuerdo, 2: En desacuerdo, 3: Indiferente, 4: De acuerdo, 5: Muy de acuerdo, considerando ítems que formulan descripciones positivas de la acción del

fuentes de información, asignar, a priori, un valor o ponderación a cada fuente. Este tipo de opciones, se configuran como procesos ciegos de determinación de estándares—criterios de evaluación— y de consecuencias imprevisibles en su aplicación, además de suponer una acción evaluativa absolutamente descontextualizadora, por ejemplo, de la realidad en que desarrolla la docencia cada profesor. Por todo ello, constituyen también aproximaciones no deseables para la evaluación docente. De hecho, supongamos que el comportamiento tipo de profesores de su entorno en condiciones homologables a las que tiene este profesor—experiencia, infraestructura, etc.— alcanza un nivel de satisfacción en las valoraciones de estudiantes y compañeros claramente superior al de un profesor—que, no obstante, se sitúa en el punto medio de valoración de la escala antes mencionada—. En este caso, podríamos convenir en que la valoración, una vez contextualizada con este tipo de informaciones, puede no resultar tan satisfactoria.

Finalmente, en esta pequeña revisión, nos restan los *Estándares Criteriales*. Por estándar criterial en este contexto entendemos aquellos niveles absolutos de calidad que, establecidos a priori por medio de un consenso intersubjetivo<sup>18</sup> han sido adoptados como nivel mínimo de aceptación, teniendo en cuenta las consecuencias institucionales de su aplicación, así como condiciones de fiabilidad y validez.

De entre las alternativas de que se dispone para el establecimiento de estándares criteriosales (JORNET y SUÁREZ, 1989), los métodos de Compromiso de HOFSTEE (1980, 1983), BEUCK (1984) y DE GRUIJTER (1980, 1985), aunque suponen aportaciones estadísticas diferentes, podrían constituir soluciones razonables al problema que nos ocupa. Todos los métodos parten de un trabajo con jueces (Comisión de Determinación de Estándares), quienes determinan cuál es la puntuación mínima aceptable (nivel criterio/estándar), así como la Razón de Pase o proporción de unidades o individuos que obtendrían una evaluación positiva. Posteriormente, se trata de analizar cuál es la mejor aproximación que presente ambos criterios sobre la distribución empírica. Es decir, se pretende establecer un compromiso entre los requerimientos planteados por la Comisión de Determinación de Estándares y la distribución empírica. Estos procedimientos podrían ser convenientemente adaptados para este propósito, asumiéndolos como parte de un proceso iterativo (elicitación de juicio, análisis de consecuencias de la aplicación de estándares) basado en la retroalimentación de información a la Comisión de Determinación de Estándares.

JAEGER (1978, 1982, 1990) propone un método de ajuste progresivo del estándar a partir de un proceso iterativo que se inicia con una estrategia de determinación subjetiva por parte de un grupo de jueces (Comisión de Determinación de Estándares), análisis de las consecuencias de aplicación del estándar sobre el colectivo evaluado, retroalimentación a la Comisión y nueva propuesta de estándar.

---

profesor, se podría asumir un criterio absoluto en tres o más puntos de promedio de valoración en la dimensión.

18. Y en el que han participado todas las audiencias implicadas.

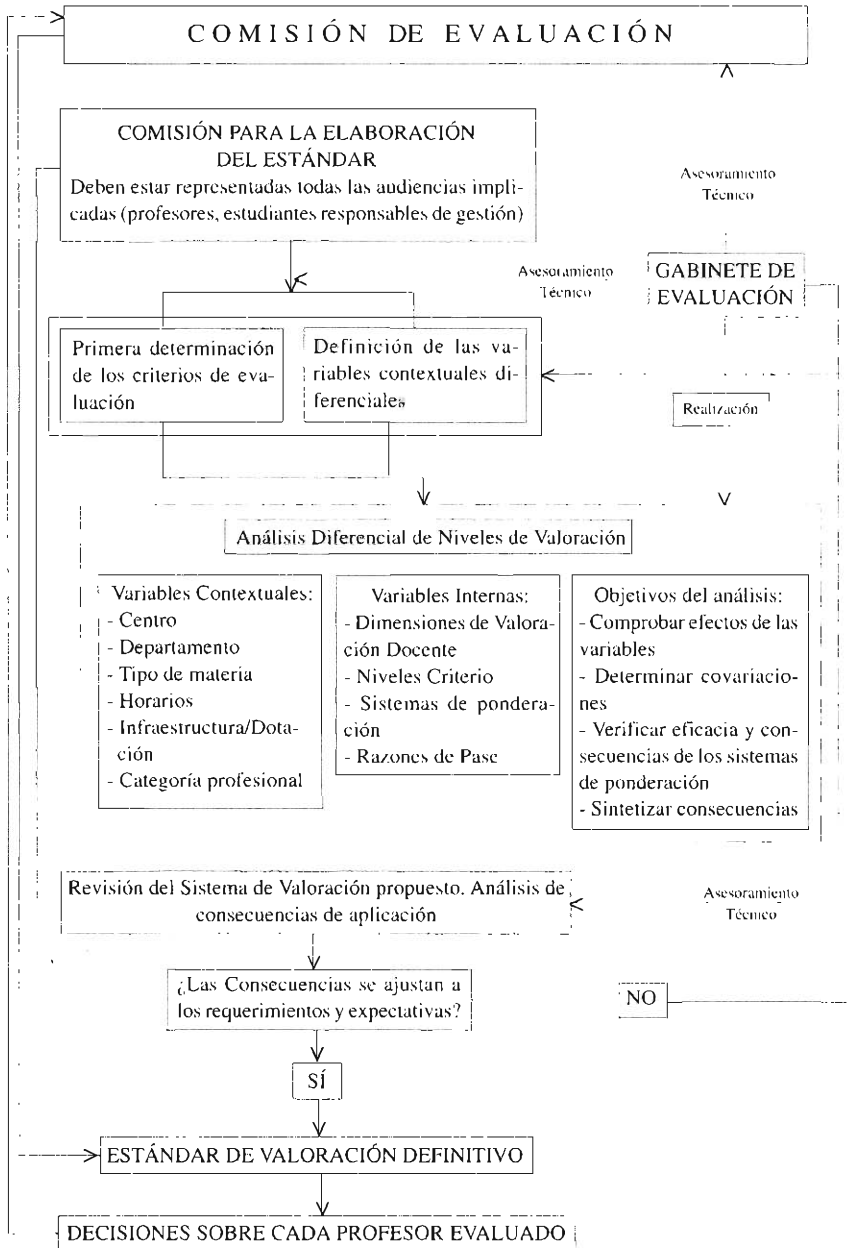


Figura 3. Esquematización de un proceso de determinación de estándares y niveles mínimos (Puntos de corte). Adaptado de JORNET y SUÁREZ (1991)

Este tipo de procedimientos pueden llegar a ofrecer soluciones razonablemente contextualizadas para el establecimiento de criterios de evaluación, al conjugar en el mismo la realidad normativa, los efectos y covariaciones que explican –al menos en parte– las valoraciones y los requerimientos y expectativas de la institución. Asimismo tienen la ventaja de poder ser automatizados –una vez determinados– y flexibilizados respecto a criterios métricos de fiabilidad (error de medida). Una esquematización del proceso propuesto (JORNET y SUÁREZ, 1991) para la Universitat de València se reseña en la Figura 3.

Este tipo de estrategias, además de proveer una razonable contextualización de la evaluación, suponen un compromiso entre los niveles mínimos requeridos por la Institución con los senderos de decisiones que, como consecuencia para la evaluación, se hayan previsto.<sup>19</sup>

### 2.3 La contextualización en el proceso de evaluación

Junto a las estrategias anteriormente citadas, es preciso tener en cuenta que la contextualización debe ser atendida a lo largo de todo el proceso evaluativo. En estos procesos pueden distinguirse dos etapas: a) desarrollo del sistema de evaluación, y b) aplicación del programa. En la primera fase se trata de determinar los componentes del Plan de Evaluación, sus fases, instrumentos, medios de implementación, utilidades previstas para cada una de ellas y compromisos entre las audiencias implicadas (WOLF, 1990). Elementos contextualizadores de esta etapa pueden considerarse (JORNET, 1991):

#### *1. Elementos técnicos:*

- a) Evaluación de las necesidades y análisis del estado de la cuestión.
- b) Desarrollo y depuración de instrumentos.
- c) Análisis progresivo y determinación de criterios, retroalimentado por las consecuencias de su aplicación.

19. En esta línea, una utilización adecuada de este sistema debe contemplar diversos factores en su implantación.

Desde el sujeto evaluado:

- Información de sus propios resultados como retroalimentación durante un periodo de tiempo suficiente para que pueda adaptarse a los requerimientos institucionales.
- Información al sujeto evaluado y a las instituciones implicadas acerca del proceso de determinación del estándar.
- Información acerca de las decisiones vinculadas al resultado de la evaluación, y posibilidades de actuación individual ante ellas.

De la Institución.

- Acciones de modificación de condiciones que perfeccionen la acción evaluativa.
- Compromiso de usos y consecuencias consensuados con el colectivo evaluado y las audiencias implicadas.
- Desarrollo de normativas que aseguren los compromisos establecidos.
- Comprometerse con la revisión periódica de estándares y procedimientos.

- d) Preparación y desarrollo de infraestructura.
- e) Análisis y seguimiento de los procesos de recogida, análisis e interpretación de la información.
- f) Adecuación de patrones de informes diferenciados según audiencias y usos.

*II. Elementos sociales y político-administrativos:*

- g) Conocimiento por parte de la unidad evaluada (institución, programa, sujeto) del proceso, procedimientos, técnicas e instrumentos de la Evaluación, atención a sugerencias y propuestas.
- h) Utilización de resultados únicamente como retroalimentación para la unidad evaluada
- i) Desarrollo de sistemas de apoyo a la mejora de la calidad del funcionamiento de la unidad evaluada, dotados de suficientes medios.
- j) Implicación del colectivo evaluado y audiencias en la determinación de estándares.
- k) Desarrollo de normativa que asegure los componentes establecidos.
- l) Desarrollo de normativa que especifique claramente los senderos de decisión vinculados a los resultados, en qué condiciones –y cuándo– podrán asociarse.
- m) Desarrollo de normativa acerca de la información que se archiva, cómo, dónde, quién puede tener acceso, bajo qué circunstancias y con qué límite.

Entre los aspectos mencionados se atiende, aunque no de forma exhaustiva, a cuestiones que son de interés diferencial según el enfoque, sumativo o formativo, que oriente el proceso de evaluación. No obstante, hay que tener presente que aun dentro de un proceso con orientación esencialmente sumativa, deben tomarse en cuenta elementos formativos para cubrir una serie de necesidades que presenta el propio proceso –como son el aprovechamiento de la información y recursos disponibles, la implicación de las audiencias en el proceso, etc–.

Determinar adecuadamente estos aspectos, y otros que sin duda hemos obviado, conlleva necesariamente un periodo importante de tiempo. Ello debe ser entendido por las partes implicadas en el proceso. Acelerar este tipo de desarrollo más allá de lo que suponga la optimización de factores que realmente puedan optimizarse, puede ser causa de una grave descontextualización que conlleve la desaparición del propio sistema.

Finalmente, durante la aplicación del Plan de Evaluación es necesario ser muy estricto en cuanto a los requerimientos que se hayan considerado para la recogida de información, análisis, tiempos de entrega de informes, etc. Paralelamente, mantener la máxima flexibilidad para atender e integrar adecuadamente los elementos no previstos en el sistema, constituye otra clave fundamental de contextualización y de mejora de la Validez.

Podemos considerar, a modo de síntesis, un conjunto de recomendaciones extractadas por ORMALA (1989):<sup>20</sup>

- a) El método, la organización y los criterios para la evaluación se deben elegir y ajustar para la situación concreta que se evalúa.
- b) Los diferentes niveles de evaluación requieren diferentes métodos de evaluación.
- c) Los objetivos del programa y de la evaluación son consideraciones importantes para llevar a cabo un estudio de evaluación.
- d) Se deben comunicar claramente a todas las partes implicadas tanto el motivo básico que subyace a la evaluación como las relaciones entre la evaluación y la toma de decisiones.
- e) Se deben formular explícitamente los propósitos de la evaluación.
- f) Siempre se debe establecer cuidadosamente la credibilidad de la evaluación.
- g) Se deben tomar en consideración en el diseño de evaluación los requisitos para el uso efectivo de los resultados de la evaluación.

### **3. Utilización de la Evaluación**

Dentro del curso de cualquier proceso de evaluación la última fase es siempre la utilización de los resultados y constituye uno de los temas de mayor polémica en la implantación y/o desarrollo de cualquier plan evaluativo. Es un aspecto clave para la consecución de los objetivos de la Evaluación. Considerando que los planes de evaluación se insertan —o deben insertarse— en la realidad a la que pretenden acercarse, la adecuación del uso de los resultados de la evaluación a las posibilidades que marcan las características del objeto o fenómeno evaluado, será pues un elemento a anticipar en la evaluación: el uso de resultados supone necesariamente la operativización práctica de la finalidad de la evaluación, plasmada en las consecuencias que se suponen para la misma y el tratamiento que se da de la información evaluativa con respecto a los colectivos —audiencias— implicados.

El tratamiento de este tema puede ser susceptible de múltiples enfoques (pedagógicos, sociológicos, administrativos, técnico-metodológicos, etc.) sin embargo hemos optado por un acercamiento basado en los tópicos que suelen suscitar mayor controversia en la implantación o desarrollo de planes de evaluación en la Universidad. La estructura de presentación de estas reflexiones se basa en la dualidad tradicional señalada por Scriven respecto de la finalidad de la evaluación (formativa/sumativa) y, dentro de cada enfoque, intentaremos presentar algunas de las soluciones que se han planteado como respuestas a las cues-

20. Aunque el autor realiza su propuesta para un contexto más reducido, como es la evaluación del impacto de la investigación basada en revisión por colegas, es perfectamente generalizable para este propósito.

tiones que enmarcan el problema: ¿para qué se evalúa?, ¿quién puede o debe tener acceso a los resultados?, ¿cómo se puede realizar su difusión?, ¿qué consecuencias y trascendencia tiene la evaluación?

Asimismo, aportamos algunas reflexiones específicas sobre tres problemáticas que nos parecen de especial interés: proporcionar información a la unidad evaluada (retroalimentación), la difusión de la información y el impacto de la evaluación.

### **3.1 Los dos grandes usos**

La dualidad formativa-sumativa –o más específicamente entre las orientaciones hacia el desarrollo/mejora o hacia la rendición de cuentas– constituyen los dos grandes polos de referencia para la evaluación. Estas dos finalidades requieren y/o conllevan utilizaciones marcadamente diferenciadas y que se explicitan en la determinación de las consecuencias de la evaluación, cómo se sintetiza y archiva la información evaluativa y, por último, a quién y cómo se trasmite dicha información.

En cualquier caso estos aspectos presentan matizaciones dependientes del objeto de evaluación. Por ello, y aun a riesgo de caer en reduccionismos siempre indeseables, haremos referencia a los grandes ámbitos de aplicación de Programas de Evaluación en la Universidad con el fin de clarificar las apreciaciones de utilización de los resultados de la evaluación.

Como referencias a Planes de Evaluación tomamos las acotadas en el apartado en que se revisan los enfoques de la evaluación: a) Evaluación Institucional como un ámbito en el que el énfasis se sitúa en los aspectos estructurales y funcionales de la Universidad como organización y en el que las referencias personales son menores y, en todo caso, tienen valor en cuanto colectivo, no tanto como individuos, b) Evaluación de Programas como un ámbito en el que el análisis enfatiza la comprobación del valor/utilidad de un determinado programa (Título, Curso, etc.), como expresión de unidades estructurales y funcionales (contexto, condiciones, procesos y productos) y su interacción en el que las referencias personales se consideran por su valor individual en la organización, y c) Evaluación del Profesorado como aquella en la que el énfasis se sitúa en la acción individual del profesor y, en todo caso, utiliza referencias institucionales o de programa como elementos de ponderación de dicha actuación. Obviamente entre estos enfoques, que nos van a servir como referencia para revisar algunos problemas implicados en la utilización de los resultados de la evaluación, recuérdese que pueden situarse diversos tipos de planes de evaluación intermedios que, por ende, suelen ser los más habituales.

Evidentemente, las consecuencias derivadas de utilizar uno u otro enfoque serán necesariamente diferentes.

### **3.2 Utilización de resultados en Planes de Evaluación Formativa**

En primer lugar, respecto a Planes de Evaluación formativa, tanto a nivel institucional como de programas y de profesorado, hay que señalar que requieren una gran delicadeza en el planteamiento y desarrollo –para asegurar su objetivo– por lo que generalmente la información evaluativa queda entre el sujeto –o institución– evaluado y el evaluador – y/o los agentes de evaluación que hayan participado en la misma. Es necesario por tanto que el sujeto acepte y asuma su papel en el proceso evaluador, una especie de “contrato inicial de evaluación”. Es necesario ante todo que el evaluado crea en los resultados de esa evaluación (GARAWAY, 1995). La utilización de resultados se centrará fundamentalmente en aquellos aspectos que ambas partes hayan establecido como susceptibles de mejora y se aportarán únicamente al sujeto evaluado, no siendo frecuente que los resultados trasciendan a otros ámbitos.

Ahora bien, ¿qué consecuencias e impacto pueden tener estos acercamientos? Este tipo de planes de evaluación, si se desarrollan de forma adecuada, constituyen la única opción evaluativa que puede conllevar cambios realmente profundos y duraderos, en cualquiera de los ámbitos mencionados. Respecto a evaluaciones institucionales y de programas porque pueden incidir precisamente en la modificación de las condiciones y unidades procesuales que orienten hacia la mejora (FURUMARK, 1981) y, en relación a los profesores, porque se orientan, en un proceso recurrente de autorrevisión, hacia la detección de los problemas del sujeto en su desarrollo profesional, aportando un marco de intervención e indicando medios de mejora (VILLAR, 1987).

En cualquier caso la utilización formativa de la evaluación –de índole institucional, de programas o de profesorado– es especialmente útil en las fases iniciales de implantación de un sistema de evaluación estable. En estas fases iniciales los procesos de innovación son a la par necesarios y útiles y constituyen un momento imprescindible, previo, a la implantación de cualquier sistema de evaluación. Los usos fundamentales de resultados en estas fases son entre otras: a) toma de contacto inicial con la realidad a evaluar, b) sensibilización de las audiencias implicadas, c) retroalimentación –individual o colectiva– (en condiciones de confidencialidad y ausencia de consecuencias–, d) conocimiento por parte de los sujetos implicados de los procedimientos, criterios y usos previstos del futuro plan de evaluación, e) experimentación, análisis de criterios de bondad (fiabilidad y validez) y depuración de los instrumentos a utilizar, f) análisis y puesta a punto de infraestructura, medios, etc. De desarrollarse con éxito esta fase, las consecuencias posibles derivan generalmente en un aumento de confianza en el futuro sistema de evaluación, con la consecuente disminución de resistencias.

Sin embargo, el impacto o viabilidad de este tipo de acercamientos creemos que es muy diferente, según se oriente hacia una evaluación institucional, de programas o de profesorado y tendrán una incidencia a niveles y en grados muy diversos.

En cuanto a los usos orientados a la mejora en Evaluaciones Institucionales, se trata de identificar aquellos elementos en la organización que tienen posibili-



dades de innovación. Puede orientarse a aspectos como la identificación de necesidades de infraestructura y/o su mejor aprovechamiento, adecuación de plantillas de profesorado y de personal de servicios, logros de objetivos y adecuación de los órganos de gobierno (cargos unipersonales y órganos colegiados), funcionamiento de servicios (bibliotecas, documentación, información y orientación a estudiantes, unidades de administración...), planificación de la oferta educativa, etc. (MACE, 1993; TEICHLER, 1994; ELTON y CRYER, 1994). En definitiva se trata de identificar los elementos de optimización de la eficiencia y la eficacia institucional.

Por su parte, los Planes de Evaluación de programas con un enfoque orientado a la mejora centran su análisis en el funcionamiento global del programa con el fin de facilitar información al Coordinador o responsable del programa que le ayude a mejorarlo. La unidad a evaluar en este caso hace referencia a Titulaciones (SUÁREZ et al., 1995), Programas de Doctorado (BAID, 1990; ISAAC et al., 1992), Postgrados (PÉREZ CARBONELL, 1993), etc. Al ser la unidad de análisis más concreta que una institución las informaciones se centran en aspectos procesuales, como cuestiones acerca de la organización, logro de objetivos específicos en unidades concretas curriculares y en el global, adecuación de las actuaciones docentes, pertinencia social del programa, etc. Asimismo se tienen en cuenta usos en relación a variables de resultado, entre las que se puede encontrar el aprendizaje como producto (basado en pruebas estandarizadas o valoraciones cualitativas de trabajos prácticos) (JORNET et al., 1993a), o el grado de penetración en el mercado laboral de los egresados. Este tipo de indicadores se suele orientar a mejorar la adecuación de los objetivos del programa y su diseño curricular.

En cualquier caso, los usos orientados a la mejora de programas, no suelen darse independientemente de algunas decisiones sumativas, que deben ejercer los responsables del programa.

Por su parte, cuando referimos las aplicaciones únicamente formativas para la evaluación del profesorado, hay que ser conscientes de que su alcance puede estar muy limitado por diversos factores (MURRAY, 1984). Ciertamente los cambios más profundos y duraderos pueden producirse en aquellos profesores involucrados en un proceso de evaluación formativa; la participación voluntaria y activa del profesor es fundamental, y ha de ser la propia institución la que motive al profesorado con ofertas de interés, ya que resultaría una ingenuidad pensar que el concepto de mejora o innovación tiene la misma relevancia para todo el profesorado. Por este motivo, en algunas universidades se han propuesto procesos asimilables a planes sumativos como elementos que operativicen una evaluación y enlacen, como consecuencia, con procesos formativos de carácter selectivo.

En este sentido, resulta conveniente que la institución motive al profesor mediante los mecanismos adecuados para facilitar su formación en aspectos docentes, como planes de formación para profesores noveles, con ofertas de actuación desde la universidad, planes específicos de innovación en unidades docentes específicas o Departamentos, facilite los medios necesarios como instalaciones, recursos, especialistas, etc. Evidentemente, esta motivación no puede deter-

minarse únicamente en términos de promoción, ya que en ese caso entraríamos en una dinámica de tipo sumativo.

Asimismo hay que tener presente que este tipo de planes requieren una gran cantidad de tiempo y medios, así como la armonización de intereses y disponibilidades. De esta forma no es fácil que un profesor mediatizado por requerimientos de su promoción dedique más tiempo del preciso a su mejora docente, a no ser que ello tuviera también un valor asociado a la promoción. Lo cual constituiría ya un uso sumativo. Por otra parte, es difícil que la Universidad en la situación actual disponga de suficiente personal técnico especializado que pueda actuar como asesor de estos procesos. Evidentemente, estas aproximaciones útiles para la Universidad deberán ser enmarcadas en cada contexto e integrarse en situaciones de compromiso que configuren un marco organizacional posible para la generalización de la mejora e innovación.

En cualquier caso, aunque un uso fundamental en cualquier plan de evaluación formativa es la retroalimentación de información, dado que es una actividad que se ha generalizado por diferentes motivos a ambas perspectivas, lo desarrollaremos independientemente.

### **3.3 Utilización de resultados en planes de evaluación sumativa**

Cualquier Plan de Evaluación que nazca amparado en algún propósito de carácter sumativo se encuentra con diversos problemas en su implementación que en gran parte se derivan de la utilización de resultados que produzcan. Por este motivo intentaremos centrar esta pequeña revisión en algunos problemas y tópicos relacionados con la utilización de resultados evaluativos que se enmarcan en propósitos sumativos.

Una evaluación de orientación institucional (Universidades, Centros, Departamentos, etc.) o de programas (titulaciones, cursos de postgrado, etc.) y carácter sumativo necesariamente pretende utilizar sus resultados sobre factores de gestión y organización. La revisión de la consecución de los objetivos del gobierno de la propia institución, infraestructura y dotaciones –necesidades y aprovechamiento– oferta de la institución e impacto y demanda social, etc... pueden ser factores involucrados en este tipo de acercamientos. Las utilizaciones de resultados se orientan entonces hacia la asignación de recursos, el control de funcionamiento, revisión de la oferta de titulaciones, reestructuración de plantillas, demostración de responsabilidad institucional, etc.

Los usos sumativos en Evaluación Institucional tienden a estructurar información que apoye decisiones tales como la distribución de recursos entre universidades, el ajuste de plantillas, el mantenimiento de servicios, etc.

En otros contextos se ha llegado a la supresión de centros basándose en Indicadores de Rendimiento Institucional. No obstante, este tipo de usos de la evaluación (que pueden calificarse como “duros”) sirven fundamentalmente en la planificación política.

Una crítica común a este tipo de usos es que basan su análisis principal-

mente sobre indicadores de carácter económico que reflejan muy pobremente la calidad educativa —o el valor social de la educación—.

En cuanto a la evaluación de programas con propósitos sumativos (titulaciones, cursos de doctorado y cursos de postgrado) existen también diversos acercamientos (JOHNSTON, 1991), aunque fundamentalmente se estructuran en torno a la utilización de los resultados para poder establecer unos indicadores de eficacia que posibiliten optimizar la oferta de cursos incluyendo los mejores o los más rentables... Un uso también frecuente es el de la Evaluación para la Acreditación de programas.

Por su parte una evaluación de profesorado de carácter sumativo se orienta hacia consecuencias de tipo laboral. Será pues requerida, en su caso, para apoyo de decisiones acerca de la contratación —o renovación de contratos—, estabilidad y promoción, asignación de sueldo —o complementos—, y control de profesorado. Además, aunque paradójicamente es menos frecuente su uso, puede constituir un elemento de apoyo sobre factores de organización docente (asignación de profesores a materias, información a estudiantes, etc.) o de desarrollo profesional, como asignación de años sabáticos, dotaciones específicas —individuales o colectivas— para la docencia y formación, etc. (MAGNUSEN, 1987).

Un caso especial es el de la Evaluación de la Investigación. Sus usos se centran fundamentalmente en la vertiente sumativa. La identificación de los indicadores de rendimiento que van a ser evaluados puede ser el primer uso que se puede hacer de la evaluación: la identificación de indicadores de investigación. En general, se identifican unos objetivos en un programa que la institución debe cumplir. Estos objetivos tienen dos funciones. Primero, como guías de acción por la administración central; segundo, como objetivos o prioridades para las divisiones académicas (departamentos, etc.) (LEVIN, 1991). Se utilizan diferentes tipos de recompensas, normalmente económicas en cuanto a salarios. En este sentido, encontramos diferentes enfoques, si bien generalmente aplican parámetros económicos en la evaluación de la investigación (BRADBURY y MANN, 1993; ALTSCHULD y ZHENG, 1995).

En todos los casos, en la utilización de resultados de planes sumativos se asocian cuestiones que tienen que ver tanto con la ética de la evaluación como con su impacto, junto a cuestiones sociales e institucionales, normalmente interactivas, que condicionan en gran medida la utilidad de la evaluación.

### **3.4 Retroalimentación**

La retroalimentación de información al sujeto evaluado —individuo, colectivo o institución— parece claro que es una de las condiciones básicas de utilización de resultados de evaluación, dado que es un derecho básico de éste, así como constituye el elemento clave en cualquier proceso evaluativo que persiga una mejora.

La posición más generalizada es, que la retroalimentación de información, por sí misma, siempre produce un efecto de mejora. Sin embargo, estudios reali-

zados al respecto (ALEAMONI, 1985; COHEN, 1980; STEVENS y ALEAMONI, 1985), tanto con retroalimentación a corto como a largo plazo, no parecen concluyentes y, como señala MARSH (1987), la utilidad de la misma aumenta sólo si existe algún tipo de ayuda al profesor para su desarrollo. Así, en sistemas de formación, en donde la evaluación actúa como elemento dinamizador del proceso, la propia participación del sujeto evaluado en el análisis e interpretación de resultados en relación con el asesor que coadyuva en la mejora, imprime a la retroalimentación, por su carácter de inmediatez y de guía para la mejora, un máximo valor de utilidad (VILLAR, 1987).

Asimismo, una retroalimentación colectiva –con resultados globales, no individuales– también muestra efectos positivos. En el estudio de implantación del sistema de evaluación de la docencia en la Universitat de València se contrastaron los niveles medios obtenidos en las dos primeras aplicaciones del cuestionario de evaluación a partir de opiniones de estudiantes. Encontramos en la mayoría de centros un aumento estadísticamente significativo en la mayor parte de las dimensiones. Entre ambas aplicaciones la retroalimentación se dirigió en un informe para cada uno de los centros en el que junto a la síntesis de sus resultados globales –no individuales– se aportaba una valoración comparativa de sus niveles en cada una de las dimensiones respecto a los niveles medios del conjunto de la Universitat de València, además de unas recomendaciones acerca de áreas deficitarias y sugerencias de cambio. Este informe fue discutido en cada centro por el equipo técnico con una Comisión nombrada por las Juntas de Centro. Descartadas las variaciones significativas que pudieran tener una explicación basada en aspectos muestrales –variaciones en el tamaño de la muestra, efectos de regresión hacia la media, etc.– quizá la relevancia del resultado reside en que los aumentos cuantitativamente más significativos se produjeron precisamente en aquellas dimensiones en las que cada centro expresamente había recibido recomendaciones en su informe. Obviamente los resultados no podemos entenderlos como concluyentes, pero consideramos que son un indicador del efecto de la retroalimentación realizada, si bien para valorar su alcance hace falta una mayor cantidad de estudios en nuestro contexto universitario (GONZÁLEZ SUCH et al., 1990a, b y c).

Un factor de interés es el tiempo transcurrido entre la recogida de información y la devolución de resultados. La recomendación más usual es que cuanto antes se produzca la retroalimentación mayor será su efectividad (HOFMAN y KREMER, 1983; WHORTON et al., 1981). La necesidad de inmediatez en el aprovechamiento de resultados se podrá cumplir probablemente en estudios de investigación evaluativa institucional de carácter esporádico y en evaluaciones formativas de profesorado. En sistemas de evaluación de profesorado universitario con aplicaciones periódicas, la limitación más importante reside en el derecho de los estudiantes a que sus opiniones no puedan revertirse en actitudes hostiles por parte del profesor; de este modo, los resultados no pueden aportarse con anterioridad a que hayan transcurrido los exámenes oficiales de los alumnos implicados en la evaluación, aunque ésta se realice en condiciones de anonimato. Se ha sugerido que una revisión de los resultados de evaluación antes de las notas finales puede

afectar a la objetividad del profesor (SCHWIER, 1982). En cualquier caso, en este contexto parece claro que de producir alguna mejora en el profesor, no serán nunca los estudiantes del curso en el que ha sido evaluado quienes disfruten las consecuencias.

Por último, un problema quizá poco tratado en el uso de resultados es el relativo a cómo se informa de los resultados al propio sujeto evaluado. La claridad de la información aportada y su contextualización son condiciones básicas para que el feedback sea efectivo. De este modo es habitual aportar la información evaluativa en medidas estadísticas que no siempre son igualmente comprensibles para el conjunto de la población evaluada. Sin embargo, alejarse de estos usos puede conllevar una pérdida en la precisión de la información que se aporta. En este sentido, algunos autores recomiendan intentar configurar informes evaluativos en los que se hayan “traducido” los indicadores estadísticos a etiquetas resumen de actuación y, si ello no es posible, apoyar la presentación de resultados en aproximaciones gráficas que ayuden al sujeto a entender mejor la información (BRINKERHOFF et al., 1983).

La información del sujeto evaluado será tanto más útil si se le aporta contextualizada, tanto respecto a comportamientos normativos de sus colectivos de referencia (por ejemplo, comparar los niveles de un profesor con los de su universidad, su centro, departamento, tipo de asignatura que imparte –trunca/optativa, teórica/práctica, etc.– tipo de grupo impartido –ratio de profesor/alumno, horario de mañana/tarde, etc.–, categoría del profesor, años de docencia, etc.), como en relación a los criterios marcados en la evaluación como niveles mínimos de aceptación. Asimismo, informaciones técnicas tales como la fiabilidad de la evaluación (establecida por algún indicador de consistencia) y representatividad (de acuerdo con el número de sujetos evaluadores, etc.) y la tendencia-evolución de la evaluación (si ésta es de aplicación periódica), son indicadores que aumentan la utilidad a la información aportada. Sin embargo, estos últimos aspectos por desgracia suelen tener bastantes descuidos en las evaluaciones –bien por problemas de tratamiento de información, bien porque no se les ha considerado relevantes– tanto en universidades extranjeras como españolas que realizan evaluaciones periódicas.

### **3.5 Difusión de la información**

La publicidad de resultados es probablemente el tema más conflictivo con que se enfrenta cualquier Plan de Evaluación. En él se produce el conflicto inevitable entre el derecho del sujeto o unidad evaluada a su intimidad y prestigio con los derechos e intereses de las audiencias implicadas, junto con otras cuestiones que enmarcan el problema (objetivos del proceso, demandas de la institución, etc.). Definir qué información se difunde, cómo y a quién, son las tres preguntas sobre las que se articula el problema.

Respecto a qué información se difunde, las posiciones oscilan entre la confidencialidad de la información evaluativa y la máxima difusión. Nuestra

experiencia nos indica que ambas posiciones llevadas a un extremo causan problemas. al menos en el contexto en que sé desenvuelve actualmente la actividad evaluadora en la Universidad española. En el polo de la información confidencial, un fallo en la transmisión de la información puede desencadenar pérdida de credibilidad del sistema. Este extremo de confidencialidad puede ser más razonable en algunos enfoques (como en el caso formativo) o en determinadas etapas de una institución y del propio proceso de evaluación (desarrollo, aplicación piloto, etc.).

Otro fenómeno que suele observarse, si se mantiene la confidencialidad de los resultados más allá del mínimo tiempo necesario, es el de la evaluación espontánea (ESCUDERO, 1991) y que se manifiesta como la respuesta de una de las audiencias que espera y no recibe información. Uno de los efectos importantes y deseables de la puesta en marcha de un sistema de evaluación debe ser la reducción al mínimo razonable de este tipo de fenómenos. Así, en la medida en que la institución madure y el sistema de evaluación se consolide y evolucione puede integrar perfectamente –y dar soporte técnico y supervisión general que es lo más importante– a las demandas de valoración que emanen de las diversas audiencias con derecho a la información y/o partes de la propia institución universitaria.

En el otro extremo –aportar toda la información disponible– se encuentran sin duda graves problemas éticos, dado que puede vulnerar el derecho a la propia imagen de un individuo o una institución. Asimismo debe tenerse en cuenta que la información evaluativa puede, por sí misma, configurar una imagen y condicionar las actuaciones futuras del sujeto o entidad evaluada.

En nuestra opinión las soluciones son en extremo delicadas por lo que siempre deben ser de compromiso y tendrán que estar recogidas por acuerdo entre las partes, con anterioridad a la puesta en marcha de cualquier sistema de evaluación (SIMONS, 1984). Es decir, previamente al desarrollo de la evaluación habrá que identificar las audiencias implicadas en la misma –colectivos con derecho a la información– y seleccionar y pactar qué, cómo y para qué información se les aportará como resultado de la evaluación. Con ello abordamos la segunda cuestión ¿quién debe ser informado?

Parece claro que tanto el sujeto evaluado como la institución –sus órganos de gobierno– que encarga o promueve la evaluación inicialmente, tienen derecho no sólo a toda la información disponible, sino también a conocer en profundidad los procedimientos y criterios que se han seguido en la evaluación. Más allá de ellos con toda seguridad podrán identificarse audiencias implicadas; en el caso de la evaluación de la actividad universitaria –sea de institución, de un título o programa, o del profesorado– las audiencias inmediatas son, como mínimo, los estudiantes, instituciones públicas y privadas y los ciudadanos. El tipo de interés de cada una de las audiencias es necesariamente diferente y, por ello, las informaciones que se pueden requerir desde cada instancia también. Por ello es una labor previa a la implementación de un programa de evaluación, determinar –y a ser posible acordar entre las partes implicadas– la información que se aportará como resultado (WOLF, 1990).

Una solución de compromiso puede consistir en seleccionar la información para cada una de las audiencias de acuerdo con las expectativas e intereses primordiales de cada una de ellas, al valor/utilidad que les pueda reportar la información, atendiendo a todos estos factores en su interacción y respecto al sujeto o entidad evaluada/audiencias. Así, por ejemplo, la necesidad de información que manifiestan como prioritaria los estudiantes en evaluaciones de profesorado es como guía para sus elecciones –tanto de asignaturas como de profesores– y, en las restantes audiencias anteriormente mencionadas –entidades públicas o privadas y los ciudadanos que puedan subvencionar total o parcialmente algunas actividades de la Universidad–, el interés se suele centrar en conocer si en definitiva la Universidad responde a la demanda y apoyo social prestado. En el primer caso, la información será sin duda más detallada y específica y puede aportarse la síntesis de las valoraciones realizadas por los propios estudiantes, es decir, una información acerca de la opinión estructurada del propio colectivo que evalúa y, en el segundo caso, la información necesariamente tenderá a ser más globalizadora.

En ambos casos, la forma en que se aporte la información (¿cómo se difunde?), condicionará en gran medida que se cumplan los objetivos previstos de la evaluación. De esta manera, los aspectos mencionados anteriormente respecto a la retroalimentación, son aplicables aquí, así como otros relacionados con el medio de difusión elegido.

Los medios de difusión de la información evaluativa –en países donde exista una cierta tradición– son variados, desde la realización de informes técnicos no publicados, pero que quedan a disposición de las audiencias implicadas, pasando por informes técnicos de agencias o institutos de evaluación, elaboraciones de rankings, la inclusión de síntesis de valoraciones en guías de estudiantes, hasta la publicación de síntesis de resultados en periódicos y medios de comunicación de masas.

En cualquier caso se han realizado diversas recomendaciones respecto a los requisitos que deben guardar los informes de evaluación (SMITH y SMITH, 1980; WOLF, 1990). No vamos a extendernos aquí en ellas. Sin embargo, debe tenerse en cuenta que un requisito imprescindible al presentar la información es que, junto a la síntesis de resultados, se aporte información precisa acerca de las características del Plan de Evaluación: promotores, objeto y finalidad, descripción de los procedimientos de recogida, análisis y síntesis de información, criterios de bondad de los instrumentos utilizados, etc., criterios de valoración y consecuencias de la evaluación, etc.

Un primer aspecto está en determinar la adecuación de la forma de la distribución utilizada para las audiencias que se habían considerado con derecho a la información, así como la pertinencia o utilidad de canales alternativos. Teniendo en cuenta las audiencias anteriormente mencionadas, y la posición que hemos adoptado respecto a la información que puede aportarse, se han planteado también algunos usos diferenciados. En el caso de información a estudiantes, son muy utilizadas las guías de estudiante\$, que incluyen información de cada asignatura (programa, actividades, sistema de evaluación, etc.), opiniones sobre la

asignatura de algunos estudiantes así como en algunos casos la síntesis de valoraciones sobre el profesor.

Los rankings de instituciones o profesores son otra práctica común: su valor es relativo ya que se presentan algunos factores que pueden conducir a sesgos importantes como son el ordenamiento por un criterio único<sup>21</sup> (en general no ponderado respecto a otras variables), que en el caso de las universidades públicas no tiene mucho sentido al no existir distrito único; o si se trata de instituciones públicas y privadas la absoluta independencia de la agencia evaluadora. Los rankings de profesores, que son utilizados en otros contextos universitarios para la selección –contratación– de profesorado entre diversas universidades, en el contexto de la Universidad española –con sistemas diferentes de reclutamiento de profesorado– no parece que puedan llegar a tener interés. Adicionalmente, este tipo de síntesis puede ser demandado por colectivos de estudiantes. No obstante, en este caso, tampoco entendemos que ofrezca beneficio a esta audiencia y, sin embargo, es una práctica que puede causar graves perjuicios al profesorado. Por todo ello entendemos que en este caso tampoco pueden haber razones para recomendar su uso.

En otro orden de cosas, si existen razones (en función del objeto de la evaluación, interés de la institución, derechos de información de audiencias, expectativas creadas, etc.) que indiquen la conveniencia de aportar información con la máxima difusión, puede ser de interés considerar en el Plan de Evaluación la emisión de un informe destinado a los medios de comunicación. Ello presenta diversas ventajas que tienen que ver con las consecuencias e impacto de la evaluación. Si se ha analizado a qué información tiene derecho –y puede interesar con valor/utilidad y sin menoscabo de otros derechos– la audiencia a que se va a atender desde tales medios, la selección de dicha información, debidamente sintetizada y contextualizada en un informe, puede facilitar que no se produzcan utilidades no deseadas de resultados y, con ello, evitar consecuencias imprevistas en la evaluación.

Con la puesta en marcha de sistemas de evaluación de aplicación periódica en las universidades españolas se asocia un problema crucial de la sociedad de la información: el control y archivo de los datos evaluativos. Las necesidades de operativización de los planes evaluativos conllevan el uso y creación de bancos de datos informatizados donde se archiva la información que proviene de las evaluaciones. Ello, junto a los problemas meramente técnicos que indudablemente tiene, plantea otros de orden ético que, desgraciadamente, no suelen atenderse por los promotores políticos en la génesis de los sistemas de evaluación: cuál es la información que se archiva, cómo, para qué, quién es el garante de los datos archivados, quién puede tener acceso a los mismos y bajo qué condiciones de uso; en definitiva, qué límites de utilización se marcan para el futuro, son cuestiones que deberían responderse anteriormente a la puesta en marcha de un sistema de evaluación. Aunque este tema ya lo planteamos en trabajos ante-

21. Aunque hay intentos de establecimiento de rankings de carácter multidimensional, este problema no puede considerarse superado.



riores (JORNET et al., 1987; JORNET, 1991), y aunque la situación social e institucional ha variado desde entonces, algunos de los problemas que entonces se apuntaban y que se dejaban entre paréntesis precisamente se han agravado con las soluciones provisionales que se les ha dado en nuestra sociedad —claramente desfasadas respecto a las necesidades existentes y no integradas en el contexto europeo—.

Este tipo de aspectos y previsiones deberían formar parte del Plan de Evaluación y, de manera institucional, deberían garantizarse por el promotor de la evaluación al colectivo evaluado. En el contexto de la Universidad española el desarrollo de una normativa al efecto es prioritario. Pese a ser reciente la realización de evaluaciones se llevan ya suficientes como para considerar urgente el desarrollo de normas y la creación de algún tipo de organismo que garantice que estos resultados no son utilizados fuera de los objetivos marcados en los planes de evaluación. En este sentido, hasta la aparición de normativas de carácter más general, entendemos como responsabilidad prioritaria del equipo de evaluación en cualquier universidad promover en la institución la discusión que conduzca a una normativa interna eficaz para atender esta problemática que constituya un compromiso público entre las diversas partes implicadas.

### **3.6 Consecuencias e impacto de la evaluación**

Un concepto fundamental para el desarrollo de evaluaciones es que éstas no constituyen nunca un fin en sí mismas, sino que son elementos instrumentales para algún propósito (ofrecer información, toma de decisiones, etc.). Desde esta perspectiva, no tiene sentido llevar a cabo una evaluación que no implique consecuencias —formativas o sumativas, en sentido lato—. Por ello es preciso que los órganos de gobierno que actúen como promotores de evaluaciones prevean, antes de hacer efectivo un plan evaluativo, qué objetivos se pretenden cubrir con el resultado de la evaluación y qué tipo de acciones pueden o deben abordarse según el mismo. De igual modo esta previsión debe ser conocida por todos los participantes en la evaluación (evaluados, audiencias implicadas, etc.). Este aspecto creemos que es importante porque tanto los sujetos o entidades evaluadas como aquellos a los que se les pide información y los que tienen derecho a ella, deben participar consciente y libremente a lo largo del proceso, no alentados por falsas expectativas ni confiados en la falta de consecuencias. Asimismo debe asegurarse que, realizada la evaluación, las consecuencias previstas se van a llevar a cabo.

¿Qué tipo de consecuencias suelen ir ligadas a planes de evaluación sumativa? Generalmente, los usos señalados al comienzo de este apartado, dependen en su operativización, como consecuencia tanto de la decisión política de los promotores de la evaluación como de las posibilidades reales de determinación, de consecuencias asociadas al propio objeto de la evaluación. Sin embargo, la tendencia más generalizada es respaldar estas evaluaciones más sobre la idea de incentivo que sobre la de sanción, considerando que la primera aporta mayores beneficios para el sistema u organización.

Otra cuestión de interés es si los planes de evaluación sumativa conllevan algún tipo de mejora. Obviamente para responder a ella habría que precisar para quién (sistema o institución, profesor, estudiantes, otras audiencias) y en qué consiste el beneficio. No vamos a entrar en consideraciones específicas. El tema, complejo y susceptible de múltiples enfoques, debe ser abordado y definido también al promover la evaluación. Únicamente y en relación a la evaluación del profesorado, parece que, en todo caso, los promotores de planes sumativos se interesan en mayor grado por mejoras globales del sistema que por mejoras de tipo individual (si ello se consigue o no, o el grado de interacción de ambas —sistema o individuo— sería otro aspecto a debatir). Ahora bien, en nuestro país, los planes de evaluación en la Universidad si bien se han desarrollado impulsados desde posiciones sumativas, han encontrado un estado de opinión que exige, responsablemente, que como consecuencia prioritaria, la evaluación produzca alguna mejora institucional, de programas e individual o, desde otra perspectiva, no es posible una mejora institucional sin la mejora individual, siendo además un derecho del sujeto evaluado el que la institución armonice cauces para su mejora. Ello se ha manifestado por diversos autores (CARRERA, 1987; MATEO, 1987; RODRÍGUEZ ESPINAR, 1987; TEJEDOR, 1987) en múltiples foros y reuniones y, eventualmente, se ha traducido en diferentes soluciones organizativas en nuestras universidades —unas en proyecto, otras en desarrollo—, que persiguen extender la mejora a otros individuos. En este sentido, aunque la utilización de resultados de la evaluación no la decide el evaluador —ni puede, ni debe hacerlo—, sino los que tienen responsabilidad política en la gestión, es responsabilidad de éste avisar del grado de fiabilidad y validez de la evaluación y, por ello, señalar el valor/utilidad de los resultados evaluativos para los usos y decisiones que se pretendan llevar a cabo.

En cualquier caso, en la Universidad española el problema está planteado y ha devenido en más complejo, como señalamos con anterioridad, por la interferencia que algunas decisiones ministeriales han producido en el desarrollo de los sistemas de evaluación propios de las universidades. Así como señalamos anteriormente, si bien la evaluación estaba planteada desde la LRU desde la utilización de resultados para la promoción y se había desarrollado en los Estatutos de las Universidades, es a partir del Decreto del MEC 1086/1989 de 28/8, donde se plantea el complemento retributivo quinquenal como consecuencia para la evaluación docente e investigadora.

En la Evaluación de la Investigación, basada en una Comisión Estatal, las deficiencias que han presidido el proceso, han conllevado la respuesta de los colectivos evaluados, con la consiguiente pérdida de credibilidad del sistema. En la Evaluación de la Docencia, la falta de coordinación con las Universidades se ha traducido en que las consecuencias se han aplicado sin utilizar los sistemas en desarrollo ni ningún otro alternativo, lo cual ha conllevado el replanteamiento desde los colectivos implicados de la utilidad que puedan tener realmente los planes que están desarrollando.

La determinación de una consecuencia en la evaluación, sin haber planteado los medios que den coherencia al sistema ni aportado a las universidades apo-

vos para la realización de los cometidos previstos, reduce considerablemente el impacto que pudieran tener los planes de evaluación. En este contexto, quizá la posición más razonable es la adoptada por diversas universidades, que han asumido que el ritmo y temporalización marcado para el desarrollo del propio plan de evaluación debe ser independiente de las consecuencias hasta ahora señaladas y, por ello, su valor/utilidad se demostrará cuando sea posible. Con todo, es difícil mantener tan siquiera el desarrollo de planes de evaluación alentados por propósitos sumativos y carentes de consecuencias; pueden darse diversos efectos negativos producidos por esta situación.

Desde esta perspectiva, una posición cautelosa para el buen desarrollo y credibilidad de los sistemas de evaluación, así como para el mantenimiento de la confianza en los órganos promotores de las evaluaciones, es la de plantear las consecuencias que se puedan llevar a cabo, omitiendo –aunque pudieran ser muy deseables– todas aquellas que la propia institución no tenga seguridad de poder cumplir.

No obstante, el impacto de la evaluación puede y debe estar más allá que en su mera utilización vinculada a decisiones (BRASKAMP et al., 1980; PATTON, 1978; PATTON, 1988), como elemento dinamizador, revelador de situaciones deficitarias, como apoyo/orientación para reducir la incertidumbre de quienes toman las decisiones y como un elemento de refuerzo político.

#### **4. Notas finales**

Probablemente, el reto más importante que tiene ante sí en la actualidad la Universidad española es apostar, desde su autonomía, por orientar esfuerzos hacia la calidad. Un elemento instrumental para este propósito es, sin duda, el desarrollo de programas internos de evaluación que puedan conseguir credibilidad en todas las audiencias implicadas, asumiendo niveles de rendición de cuentas internas y externas, integrados con acciones formativas sólidas.

En la actualidad los Planes de Evaluación presentan diversos grados de desarrollo en nuestras universidades. Únicamente cuando alcancen un grado adecuado de institucionalización podrán demostrar su verdadero papel.

Asegurar su utilidad es la primera responsabilidad de los promotores que orientan los Planes de Evaluación. Deben garantizar que los mismos se alejen de los usos de mera demostración de responsabilidad institucional, alentando sólo aquellos procesos que tengan una finalidad clara.

Por su parte, los técnicos es preciso que asuman su rol, desde un planteamiento deontológico de la Evaluación, evitando suplantar la decisión política y encarando su verdadera responsabilidad: garantizar que los Planes de Evaluación son adecuados a nivel metodológico y técnico.

Todo ello enmarcado en el desarrollo de una cultura evaluativa, de carácter democrático, que base su acción sobre los siguientes principios: a) la participación de todas las audiencias implicadas, b) la orientación prioritaria hacia la mejora y la innovación, c) la transparencia de los procesos, y d) la calidad técnica de los mismos.

Desde un punto de vista fundamentalmente metodológico creemos que, teniendo en cuenta las características de nuestro sistema universitario, cualquier enfoque (institucional, de programas y centrado en el profesor) maximizará su utilidad sobre la base de modelos mixtos (formativos/sumativos). Deben ser las mismas universidades, desde su autonomía, las que definan finalidades y objetivos a la evaluación y determinen los modelos que les resultan más adecuados.

La integración en la gestión universitaria de diversos Planes de Evaluación para diferentes propósitos (por ejemplo, autoestudios dirigidos a la mejora de la gestión y optimización de procesos, evaluaciones de programas –titulaciones, cursos de doctorado, postgrado...– para mejorar la oferta y evaluaciones de profesorado –como control y con consecuencias orientadas a la mejora...–) puede constituir un elemento clave para la innovación universitaria, especialmente en este momento en el que la Universidad está inmersa de pleno en un amplio proceso de renovación.

En cualquier caso no sólo con desarrollos de planes de evaluación se va a trabajar en pro de la mejora de la calidad universitaria. Es necesario un compromiso, por parte de todos, con la investigación básica y metodológica de la Educación Superior, como referente inicial para abordar la mejora de los procesos educativos en la Universidad.

## Referencias bibliográficas

- ABRAMI, P.C. Y D'APOLLONIA, S. (1991): Multidimensional Students' Evaluations of Teaching Effectiveness: Generalizability of "N = 1" Research: Comment on Marsh (1991). *Journal of Educational Psychology*, 83 (3), 411-15.
- ALEAMONI, L.M. (1985): Peer evaluation of instructors and instruction. *Instructional evaluation*, 8.
- ALKIN, M.C. y ELLET, F.S. (1990): Development of Evaluation Models, en WALBERG, H.J. y HAERTEL, G.D. (eds)(1990): *The International Encyclopedia of Educational Evaluation*. Oxford: Pergamon.
- ALTSCHULD, J.W. Y ZHENG, H.Y. (1995): Assessing the effectiveness of research organizations. *Evaluation Review*, 19 (2), 197-216.
- APARICIO, J.J.; TEJEDOR, F.J. y SANMARTÍN, R. (1982): *La enseñanza universitaria vista por los alumnos: un estudio para la evaluación de los profesores en la enseñanza superior*. Cuadernos de Capacitación Docente, Madrid: O.E.I.
- BAIRD, L.L. (1990): Disciplines and doctorates: the relationships between program characteristics and the duration of doctoral study. *Research in Higher Education*, 31 (4), 369-385.
- BANTA, T.W.; FISHER, H.S. y MINKEL, L.W. (1986): Assessment of institutional effectiveness at the University of Tennessee, Knoxville. *International Journal of Institutional Management in Higher Education*, 10: 262-271.
- BARNETT, R. (1989): *Towards an educational audit*. Informative Series Discussion Paper 3. Londres: Council for National Academic Awards.

- BAVER, M. (1988): *Evaluation in a Higher Education System undergoing decentralitation. Outlines and attempts*. Paper presented at the 10th European Forum of the Association for Institutional Research, Bergen: Norway, August.
- BEUCK, C-H. (1984): A method for reaching a compromise between absolute and relative standars in examinations. *Journal of Educational Measurement*, 21: 147-152.
- BEBOU, L. et al (1975): User studies in the humanities. *Research Quarterly*, 15, 40-44.
- BRADBURD, R.M. Y MANN, D.P. (1993): Wealth in Higher Education Institutions. *Journal of Higher Education*, 64 (4), 472-493.
- BRASKAMP, L.A.; BRANDERBURG, D.C. y ORY, J.C. (1984) *Evaluating teaching effectiveness*. Beverly Hills, Cal.: Sage.
- BRASKAMP, L.A.; BROWN, R.D. Y NEWMAN, D.L. (1980): *Studying evaluation utilization through simulations*. Illinois: University of Illinois and University of Nebraska-Lincoln.
- BRINKERHOFF, R.O.; BRETHOWER, D.M.; HLUCHYJ, T. Y NOWAKOWSKI, J.R. (1983): *Program evaluation. A practioner's guide for trainers and educators*. Boston: Kluwer-Nijhoff Publishing.
- CARRERA, M.J. (1987): Presentación. En Varios, *Consideraciones metodológicas sobre Evaluación y Mejora de la Docencia Universitaria*. Valencia: Servei de Formació Permanent. Universitat de València.
- CAVE, M; HANNEY, S. y KOGAN, M. (1991): *The Use of performance Indicators in Higher Education. A critical Analysis of developing practice*. Londres: Jessica Kingsley Publishers. (2ª edición revisada)
- CAVE, M; HANNEY, S.; KOGAN, M. y TREVETT, G. (1988): *The Use of Performance Indicators in Higher Education: A critical Analysis of developing practice*. Londres: Jessica Kingsley Publishers.
- CENTRA, J.A. (1980) *Determining faculty performance*. San Francisco: Jossey-Bass.
- 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.
- CHUBIN, D.E. y HACKETT, E.J. (1990): *Peerless science: Peer review and U.S. science policy*. Albany: State University of New York Press.
- DE GRUIJTER, D.N.M. (1980): A two-stage testing procedure. en L.J.T. VAN DER KAMP et al. (eds): *Psychometrics for Educational Debates*. Nueva York: John Wiley.
- (1985): Compromise models for establishing examination standards. *Journal of Educational Measurement*, 22: 263-269.
- DE LA ORDEN, A. (1990): Evaluación, selección y promoción del profesor universitario. *Revista Complutense de Educación*, 1 (11), 11-29.
- DE MIGUEL DÍAZ, M. (1991): *Indicadores de calidad de la Docencia Universitaria*. Ponencia en el I Congreso Internacional sobre Calidad de la Enseñanza Universitaria. 6-8 Marzo. Puerto de Santa María (Cádiz)

- DOCHY, F.J.R.C.; SEGERS, M.S.R. y WIJNEN, V.H.F.W. (eds) (1990): *Management Information and Performance Indicators in Higher Education: An International Issue*. Assen: Van Gorcum.
- ELTON, L. Y CRYER, P. (1994): Quality and change in higher Education. *Innovate Higher Education*, 18 (3), 205-220.
- ESCUDERO, T. (1987): Experiencias evaluativas en la Universidad de Zaragoza, en Varios: *Consideraciones Metodológicas sobre Evaluación y Mejora de la Docencia Universitaria*. Servei de Formació Permanent. Informes de Investigació Evaluativa nº 1. Universitat de València.
- (1987): Experiencias evaluativas en la Universidad de Zaragoza. En Varios (Eds.), *Consideraciones metodológicas sobre Evaluación y Mejora de la Docencia Universitaria*. Valencia: Servei de Formació Permanent. Universitat de València.
- (1991): Enfoques modélicos en la evaluación de la enseñanza universitaria. Ponencia en *III Jornadas Nacionales de Didáctica Universitaria "Evaluación y Desarrollo Profesional"*, Las Palmas de Gran Canaria, 23-26 de Septiembre.
- (1991): Enfoques modélicos en la Evaluación de la Enseñanza Universitaria. Documento presentado en *III Jornadas Nacionales de Didáctica Universitaria "Evaluación y Desarrollo Profesional"*. Las Palmas de Gran Canaria, 23-26 de septiembre de 1991.
- FERRÁNDEZ, M.R.; JORNET, J.M.; PÉREZ CARBONELL, A. y GONZÁLEZ SUCH, J. (1995): Estudio del Cuestionario de Evaluación de la Docencia a partir de opiniones de estudiantes en la Universitat Jaume I (I) El Cuestionario y planteamiento metodológico, en AIDIPE (Comp.), *Estudios de investigación educativa en intervención psicopedagógica*. Valencia: AIDIPE.
- FERRÁNDEZ, M.R.; JORNET, J.M.; PÉREZ CARBONELL, A. y GONZÁLEZ SUCH, J. (1995): Estudio del Cuestionario de Evaluación de la Docencia a partir de opiniones de estudiantes en la Universitat Jaume I (II) Estudio factorial, en AIDIPE (Comp.), *Estudios de investigación educativa en intervención psicopedagógica*. Valencia: AIDIPE.
- FURUMARK, A.M. (1981): Institutional Self-Evaluation in Sweden. *International Journal of Institutional Management in Higher Education*, v 5 nº 3.
- GARAWAY, G.B. (1995): Participatory Evaluation. *Studies in Educational Evaluation*, 21, 85-102.
- GARCÍA, P.; MORA, J.G.; RODRÍGUEZ, S. y PÉREZ, J.J. (1995): Experimenting institutional evaluation in Spain. *Higher Education Management*, 7, (1), 101-118.
- GONZÁLEZ SUCH, J.; JORNET, J.M.; SUÁREZ, J.M. y PÉREZ CARBONELL, A. (1990a): Análisis de la estabilidad de niveles y dimensiones observados en el cuestionario de evaluación de la docencia a partir de opiniones de estudiantes en dos aplicaciones anuales sucesivas en centros del área de CC. Humanas de la Universitat de València. *Revista de Investigación Educativa*, 8, (16), 681-688.

- SUÁREZ, J.M.; SÁEZ, A.; ALIAGA, F.; ORELLANA, N.; BELLOCH, C. y SALAVERT, L. (1995): Evaluación de la implantación de la licenciatura de Pedagogía en la Universitat de València en AIDIPE (Comp.), *Estudios de investigación educativa en intervención psicopedagógica*. Valencia: AIDIPE.
- TEICHLER, U. (1994): Quantity and Quality of Staff in Higher Education. *Higher Education Policy*, 7 (2), 19-23.
- TEJEDOR, F.J. (1987): *Evaluación del profesorado de la Universidad de Santiago por los alumnos (curso 86-87)*. Universidad de Santiago de Compostela.
- (1990): Perspectiva Metodológica del Diagnóstico y Evaluación de Necesidades en el ámbito Educativo. *Revista de Investigación Educativa*, 8: 15-37.
- (1991): *Experiencias españolas de evaluación de la enseñanza universitaria y nuevas perspectivas*. Ponencia presentada en las III Jornadas Nacionales de Didáctica Universitaria. Las Palmas de Gran Canaria, 23-26 de Septiembre.
- TEJEDOR, F.J.; JATO, E. y MÍNGUEZ, C. (1987): Evaluación del Profesorado Universitario por los alumnos en la Universidad de Santiago, en VARIOS: *Consideraciones Metodológicas sobre Evaluación y Mejora de la Docencia Universitaria*. Servei de Formació Permanent. Informes de Investigación Evaluativa nº 1. Universitat de València.
- VAN OS, M. (1990): Strategies for Quality Assessment: The Human Factor. *Revista Española de Pedagogía*, nº 186: 281-293.
- VARIOS (1988) *Informes de Investigación Evaluativa nº 1: Consideraciones metodológicas sobre la evaluación y mejora de la docencia universitaria*. Vicerrectorado de Estudios, Universidad de Valencia.
- VILLAR, L.M. (1987): Evaluación de la conducta docente y del clima psicosocial de clase, en VARIOS: *Consideraciones Metodológicas sobre Evaluación y Mejora de la Docencia Universitaria*. Servei de Formació Permanent. Informes de Investigación Evaluativa nº 1. Universitat de València.
- (1987): Evaluación de la conducta docente y del clima psicosocial de clase. En Varios (Eds.), *Consideraciones metodológicas sobre evaluación y mejora de la docencia universitaria*. Valencia: Servei de Formació Permanent. Universitat de València.
- VROEIJENSTIJN, T.I. y ACHERMAN, H. (1990): Control oriented versus improvement oriented quality assessment, en C.J. GOEDEGEBUURE, P.A.M. MAASSEN y D.F. WESTERHEIJDEN (eds): *Peer Review and Performance Indicators*. Utrech: Lemma.
- WHORTON, J.E.; SIDERS, J.A. Y MCCALLUM, R.S. (1981): Repeated student ratings: Course improvement now, not later. *College Student Journal*, 15, 343-346.
- WOLF, R.M. (1990): *Evaluation in Education. Foundations of Competency Assessment and Program Review*. New York: Praeger.
- (1990): *Evaluation in Education*. Nueva York: Praeger.
- YORKE, M. (1991) Performance indicators: towards a sinoptic framework. *Higher Education*, 21 (2), 235-248.
- ZABALZA BERAZA, M.A. (1990): Evaluación orientada al perfeccionamiento. *Revista Española de Pedagogía*, XLVIII, 186: 295-317.

- GONZÁLEZ SUCH, J.; JORNET, J.M.; SUÁREZ, J.M. y BELLOCH, C. (1990b): Análisis de la estabilidad de niveles y dimensiones observados en el cuestionario de evaluación de la docencia a partir de opiniones de estudiantes en dos aplicaciones anuales sucesivas en centros del área de CC. Psicopedagógicas de la Universitat de València. *Revista de Investigación Educativa*, 8, (16), 689-696.
- GONZÁLEZ SUCH, J.; SUÁREZ, J.M.; JORNET, J.M. y FERRÁNDEZ, M.R. (1990c): Análisis de la estabilidad de niveles y dimensiones observados en el cuestionario de evaluación de la docencia a partir de opiniones de estudiantes en dos aplicaciones anuales sucesivas en centros del área de CC. de la Salud de la Universitat de València. *Revista de Investigación Educativa*, 8, (16), 643-650.
- GONZÁLEZ SUCH, J.; JORNET, J.M.; PÉREZ CARBONELL, A. Y FERRÁNDEZ, (1994): Factores intervinientes en la valoración del profesor por parte del estudiante. *Revista de Investigación Educativa*, 8 (16), 681-688.
- GONZÁLEZ SUCH, J.; JORNET, J.M. y PÉREZ CARBONELL, A. (1995): Análisis de un cuestionario de evaluación de clases prácticas experimentales. en AIDIPE (Comp.). *Estudios de investigación educativa en intervención psicopedagógica*. Valencia: AIDIPE.
- HOFMAN, J.E. y KREMER, L. (1983): Course evaluation and attitudes toward college teaching. *Higher Education*, 12, 681-690.
- HOFSTE, W.K.B. (1980): *Policies of education selection and grading: The case for compromise models*. Paper presented at the Fourth International Symposium on Educational Testing, Antwerp, Belgium
- HOFSTE, W.K.B. (1983): The case for compromise in educational selection and grading, en S.B. ANDERSON y J.S. HELMICJ (eds) *On educational Testing*. San Francisco: Jossey-Bass, 109-127.
- HURT, C.D. (1985): Methodological citation differences in science, technology, and social sciences literatures. *Library and Information Science Research* 7: 345-355.
- ISAAC, P.D.; QUINLAN, S.V. Y WALKER, M.M. (1992): Faculty perceptions of the doctoral dissertation. *Journal of Higher Education*, 63 (3), 241-268.
- JAEGER, R.M. (1978): *A proposal for setting a standar on the North California Hingh School Competency Test*. Paper presented at the Spring meeting of the North California Association for Research in Education, Chapel Hill.
- (1982): A iterative structured judgement process for establishing standards on competency test: Theory and application. *Educational Evaluation and Policy Analysis*, 4(4): 461-475.
- (1990): Setting standards on teacher certification tests, en J. MILLMAN y L. DARLING-HAMMOND (eds) *The New Handbook of Teacher Evaluation*. Newbury Park, CA: Sage.
- JOHNES, J. y TAYLOR, J. (1990): *Performance Indicators in Higher Education*. Bristol: SRHE & Open University.
- JORNET, J.M. (1991): *Enfoques de evaluación universitaria*. Ponencia presentada en las III Jornadas Nacionales de Didáctica Universitaria. Las Palmas de Gran Canaria, 23-26 de Septiembre.



- JORNET, J.M. y SUÁREZ, J.M. (1989): Revisión de Modelos y Métodos en la determinación de estándares y en el establecimiento de un Punto de corte en Evaluación Referida al Criterio (ERC). *Bordón*. Vol. 41, nº2, pp. 277-301. 1989.
- (1991): Propuesta de desarrollo del Sistema de Evaluación de la Universitat de València. Documento interno. Departamento MIDE. Universitat de València.
- JORNET, J.M.; GONZÁLEZ SUCH, J. y PÉREZ CARBONELL, A. (1995): Análisis de la consistencia de cuestionarios de opinión de estudiantes para evaluaciones docentes, en AIDIPE (Comp.), *Estudios de investigación educativa en intervención psicopedagógica*. Valencia: AIDIPE.
- JORNET, J.M.; SUÁREZ, J.M.; GONZÁLEZ SUCH, J.; PÉREZ CARBONELL, A. Y FERRÁNDEZ, M.R. (1993): *Evaluation Report of the Project: Communication and presentation skills for technological transfer agents*. Euro-Innovations Manager Project. COMETT.
- JORNET, J.M.; VILLANUEVA, P.; SUÁREZ, J.M. Y ALFARO, I.J. (1987): Proyecto de implantación de un sistema de evaluación del profesorado en la Universitat de València. En Varios (Eds.), *Consideraciones metodológicas sobre Evaluación y Mejora de la Docencia Universitaria*. Valencia: Servei de Formació Permanent. Universitat de València.
- KOSTOFF R.N. (1994): Assessing research impact: semiquantitative methods. *Evaluation Review*, 18 (1), 11-19.
- KRATHWOHL, D.R. (1985): *Social and behavioral science research*. San Francisco: Jossey-Bass Publishers.
- LEVIN, H.M. (1991): Raising productivity in Higher Education. *Journal of Higher Education*, 62 (3), 241-262.
- MACE, J. (1993): University funding changes and university efficiency. *Higher Education Review*, 25 (2), 7-22.
- MAGNUSEN, K.O. (1987): Faculty Evaluation, Performance, and Pay: Application and Issues. *Journal-of-Higher-Education*, 58 (5), 516-29.
- MARSH, H.W. (1987) Students' evaluations of university teaching: research findings, methodological issues, and directions for future research. *International Journal of Educational Research*. 11, 3 253-388.
- MARSH, H.W. y HOCEVAR, D. (1991) The multidimensionality of students' evaluations of teaching effectiveness: The generality of factor structures across academic discipline, instructor level, and course level. *Teaching and Teacher Education*, 7, 1, 9-18.
- MARSH, H.W.; TOURON, J. y WHEELER, B. (1985) Students' evaluations of university instructors: The applicability of american instruments in a spanish setting. *Teaching and Teacher Education*, 1,2, 123-138.
- MATEO, J. (1987): La Evaluación del Profesorado Universitario. Algunas reflexiones en torno al estado de la cuestión, en VARIOS: *Consideraciones Metodológicas sobre Evaluación y Mejora de la Docencia Universitaria*. Servei de Formació Permanent. Informes de Investigació Evaluativa nº 1. Universitat de València.
- MEADE, P. (1995): Managing quality by devolution. *Higher Education Management*, 7, (1), 63-80.

- MILLER, R.I. (1987): *Evaluating Faculty for Promotion and Tenure*. San Francisco: Jossey-Bass.
- MILLMAN, J. (ed) (1981): *Handbook of Teacher Evaluation*. Beverly Hills: Sage
- MOSES, I. y ROE, E. (1990) *Heads and chairs. Managing academic departments*. St. Lucia: Univ. of Queensland Press.
- MURRAY, H.G. (1984): The Impact of Formative and Summative Evaluation of Teaching in North American Universities. *Assessment and Evaluation in Higher Education*, 9 (2), 117-32.
- NEDWEK, B.P. y NEAL, J.E. (1993): *Performance indicators and rational management tools: A comparative assessment of projects in North America and Europe*. Documento presentado en el Annual Forum of the Association for Institutional Research. Chicago, Mayo.
- ORMALA, E. (1994): Impact assessment. European experience of qualitative methods and practices. *Evaluation Review*, 18 (1), 41-51.
- PATTON, M.Q. (1978): *Utilization-focused evaluation*. Beverly-Hills: Sage.
- (1988): The evaluator's responsibility for utilization. *Evaluation Practice*, 9 (2),
- PÉREZ CARBONELL, A.; JORNET, J.M. y SUÁREZ, J.M. (1993): *Diseño de un sistema de evaluación de los cursos de Postgrado en la Universitat de València*. Vicerrectorado de Estudios. Doc. interno. Universitat de València.
- (1995): Evaluación de cursos de postgrado e indicadores estructurales, en AIDIPE (Comp.), *Estudios de investigación educativa en intervención psicopedagógica*. Valencia: AIDIPE.
- PÉREZ JUSTE, R. (1994): Investigación evaluativa en V. García Hoz, *Tratado de educación personalizada. Problemas y métodos de investigación en educación personalizada*. Madrid: Rialp.
- RODRÍGUEZ ESPINAR, S. (1987): La evaluación del profesorado universitario. Análisis de la opinión del profesorado de la Universidad de Barcelona, en VARIOS: *Consideraciones Metodológicas sobre Evaluación y Mejora de la Docencia Universitaria*. Servei de Formació Permanent. Informes de Investigación Evaluativa nº 1. Universitat de València.
- SALVADOR, L. y GARCÍA, A. (1989): *Evaluación de la Docencia*. ICE Universidad de Cantabria.
- SCHWIER, R.A. (1982): Design and use of student evaluation instruments in instructional development. *Journal of Instructional Development*, 5, 28-34.
- SELDIN, P. (1988) *Evaluating and development administrative performance*. San Francisco: Jossey-Bass.
- SIMONS, H. (1984): Negotiating conditions for independent evaluations. En C. Adelman (Eds.), *The politics and ethics of evaluation*. Croom Helm.
- SMITH, D.M. Y SMITH, N.L. (1980): *Writing effective evaluation reports*. Portlan, OR: Northwest Regional Educational Laboratory.
- STEVENS, J.J. Y ALEAMONI, L.M. (1985): The use of evaluative feedback for instructional improvement: A longitudinal perspective. *Instructional Science*, 13, 285-304.