



***CALIDAD
Y RENDIMIENTO
EN LAS
INSTITUCIONES
UNIVERSITARIAS***

José-Ginés Mora Ruiz



CONSEJO DE UNIVERSIDADES
Secretaría General

**CALIDAD Y RENDIMIENTO EN LAS INSTITUCIONES
UNIVERSITARIAS**

JOSE-GINES MORA RUIZ

CONSEJO DE UNIVERSIDADES
Secretaría General
1991

© Consejo de Universidades. Secretaría General
NIPO: 176-91-022-7
I.S.B.N.: 84-369-1976-9
Depósito Legal: M-17799-1991
Imprime: Pedro Cid

INDICE

	<u>Páginas</u>
PROLOGO	9
CAPITULO I: INTRODUCCION	
I.1. Visión general sobre la evolución de la educación superior	11
I.2. Aspectos peculiares de la evolución de la educación superior en España	14
I.3. Algunos problemas actuales de la educación superior	16
I.4. Programa del presente estudio	17
CAPITULO II: LA CALIDAD EN LA EDUCACION SUPERIOR	
II.1. Introducción	21
II.2. Objetivos y beneficios de la educación superior	23
II.2.1. Distintas perspectivas	23
II.2.2. Los objetivos de la educación superior	25
II.2.3. Los beneficios de la educación superior	27
II.3. La calidad en la educación superior	32
II.3.1. Distintos enfoques	33
II.3.2. Factores que afectan a la calidad	37
II.3.3. La evaluación de la calidad	40
CAPITULO III: EVALUACION INSTITUCIONAL	
III.1. Introducción	43
III.2. El concepto de evaluación institucional	44
III.3. Factores a considerar en la evaluación	46

	<u>Páginas</u>
III.3.1. Aspectos a evaluar	47
III.3.2. Elementos a evaluar	49
III.4. Distintos tipos de evaluación	52
III.5. La puesta en práctica de la evaluación	57
III.5.1. Criterios básicos de la evaluación	57
III.5.2. Las líneas maestras para evaluar	58
III.5.3. Algunos problemas de la evaluación	60

CAPITULO IV: RENDIMIENTO INSTITUCIONAL E INDICADORES DE RENDIMIENTO

IV.1. Rendimiento institucional	65
IV.1.1. Introducción	65
IV.1.2. Efectividad, eficiencia y rendimiento	67
IV.1.3. La evaluación del rendimiento institucional y sus problemas	69
IV.1.4. La evaluación del rendimiento institucional	72
IV.1.5. Objetivos de la evaluación del rendimiento institucional.	74
IV.1.6. El proceso de la evaluación institucional	76
IV.2. Indicadores de rendimiento	78
IV.2.1. El concepto de indicador de rendimiento	78
IV.2.2. Tipos de indicadores de rendimiento	80
IV.2.3. La utilización de los indicadores de rendimiento	84
IV.2.4. Discusión sobre el uso de los indicadores	88

CAPITULO V: EVALUACIONES AUTORREGULADAS

V.1. Introducción	95
V.2. La acreditación en USA	97
V.2.1. El sistema de acreditación	97
V.2.2. Los criterios para la acreditación	99
V.2.3. Propósitos del sistema de acreditación	101
V.2.4. Ventajas e inconvenientes de la acreditación	102
V.3. Autoestudios	103
V.3.1. El concepto de autoestudio	104
V.3.2. Los objetivos del autoestudio	106
V.3.3. Características del autoestudio	106
V.3.4. El proceso del autoestudio	108

	<u>Páginas</u>
V.3.5. Ventajas e inconvenientes del autoestudio	112
V.4. Revisión de programas	115
V.4.1. El concepto de revisión de programas	115
V.4.2. Objetivos de la revisión de programas	119
V.4.3. Distintos enfoques a la revisión de programas	120
V.4.4. Criterios a utilizar en la revisión	122
V.4.5. Algunos problemas de la revisión de programas	125

CAPITULO VI: VALORACION DE LA CALIDAD EN LAS UNIVERSIDADES ESTADOUNIDENSES

VI.1. Introducción	131
VI.2. Estudios basados en encuestas	132
VI.2.1. Evaluación de programas de postgrado	134
VI.2.2. Evaluación de programas de pregrado	142
VI.2.3. Evaluación de programas profesionales	145
VI.3. Estudios basados en medidas objetivas	146
VI.3.1. Estudios que usan indicadores relativos al profesorado.	147
VI.3.2. Estudios que usan indicadores relativos al alumnado ...	152
VI.3.3. Estudios que utilizan una combinación de indicadores.	154
VI.4. Estudios que relacionan el enfoque reputacional con las medidas objetivas	156
VI.5. Discusión sobre los distintos enfoques	160
VI.5.1. Crítica a los métodos reputacionales	161
VI.5.2. Crítica a los métodos objetivos	165
VI.5.3. Requerimientos para una evaluación ideal	167

CAPITULO VII: LA EVALUACION DE LA EDUCACION SUPERIOR EN DIVERSOS PAISES

VII.1. Introducción	171
VII.2. Estados Unidos, Canadá, Australia	172
VII.2.1. Estados Unidos	172
VII.2.2. Canadá	179
VII.2.3. Australia	182
VII.3. Países de Europa Occidental	184
VII.3.1. Visión general de los países europeos	184
VII.3.2. Holanda	187

	<u>Páginas</u>
VII.3.3. Reino Unido	196
VII.3.4. Finlandia	200
VII.3.5. Alemania (R.F.)	201
VII.3.6. Suecia	203
VII.3.7. Francia	204

CAPITULO VIII: CONCLUSIONES Y PROPUESTAS PARA EL SISTEMA UNIVERSITARIO ESPAÑOL

VIII.1. Introducción	207
VIII.2. Algunas conclusiones generales sobre la evaluación institucional	209
VIII.2.1. El problema de la calidad de la educación superior.	209
VIII.2.2. La evaluación institucional	210
VIII.2.3. Clasificación de las universidades por criterios de calidad	212
VIII.2.4. Evaluaciones autorreguladas y evaluaciones de control	213
VIII.2.5. Indicadores de rendimiento y revisiones por colegas (peer review)	214
VIII.3. Particularidades del sistema universitario español	215
VIII.3.1. Algunas características del sistema en relación con la evaluación	216
VIII.3.2. Algunas experiencias relacionadas con la evaluación institucional	218
VIII.4. Una propuesta de evaluación institucional para las universidades españolas	222
REFERENCIAS BIBLIOGRAFICAS	231

PROLOGO

El presente trabajo es un estudio sobre temas relacionados con el funcionamiento de las instituciones de educación superior. Concretamente se investigan aspectos de las instituciones universitarias tales como calidad eficiencia, efectividad, evaluación institucional, autoestudios, etc., haciendo especial hincapié en los mecanismos utilizados en diversos países para la valoración de tales conceptos.

Los métodos de evaluación de las instituciones universitarias están adquiriendo una notable importancia en los últimos años en la mayoría de los países desarrollados en los que, tras un período de importante expansión de la educación superior, se ha llegado a una situación de relativa estabilidad financiera y humana que permite dedicar la atención debida a aspectos del sistema educativo superior que, de algún modo, se habían descuidado en los momentos de fuerte crecimiento de la demanda que se dieron en las últimas décadas.

El objetivo de este estudio, y el de la mayoría de los que se referenciaran en las siguientes páginas, es el de proporcionar herramientas que posibilitan la mejora de la calidad y de la eficiencia de las universidades, de modo que éstas puedan afrontar los problemas inherentes al desarrollo científico y tecnológico, y a la nueva exigencia de la sociedad, con la mayor efectividad posible.

Este estudio ha sido realizado por encargo del Consejo de Universidades. La elaboración del trabajo ha sido facilitada en gran medida por la utilización de las bibliotecas y servicios de la Stanford University, en cuyo Center for Educational Research el autor ha sido Visiting Scholar durante el curso 1989-90. La estancia en este centro de investigación ha sido posible gracias a la beca concedida por al Consellería de Cultura, Educació y Ciencia de la Generalitat Valenciana, y al permiso de estudios otorgado por la Universitat de Valencia. El trabajo ha sido completado en una visita al Center for Educational Research and Innovation (CERI) de la OCDE en París, y con la participación del autor en el fórum anual de la European Association for Institutional Research (EAIR) celebrado en Lyon en septiembre de 1990. El Center for Higher Education

Policy Studies (CHEPS) de la Universidad de Twente en los Países Bajos, nos suministró una gran parte de la información sobre los sistemas de evaluación holandés y británico.

A todas estas instituciones, y especialmente a las personas que en cada una de ellas han colaborado o facilitado mi trabajo, les estoy profundamente agradecido. En correspondencia, espero que el presente estudio puede ser de utilidad a las instituciones que lo han posibilitado.

CAPITULO I. INTRODUCCION

I.1 VISION GENERAL SOBRE LA EVALUACION DE LA EDUCACION SUPERIOR

Es un hecho evidente que la educación superior ha sufrido en todo el mundo transformaciones radicales en las últimas décadas. Desde la óptica del presente estudio, se pueden establecer tres períodos de tiempo que diferencian claramente la evolución de la educación superior en el mundo:

1. El primer período se inicia con la propia existencia de las universidades, y su final se podría situar en torno a 1960; fecha aproximada que varía según los distintos países y los diferentes niveles de desarrollo económico y demográfico. Puede afirmarse que, salvo contadísimas excepciones, la educación superior fue hasta ese momento un privilegio de unos pocos: concretamente en España, en un año tan cercano como 1962 sólo un 4,3 % de los jóvenes en el grupo de edad correspondiente a la del ingreso en la universidad se matriculaban en los centros universitarios (Mora, 1989). Como consecuencia de su reducido alcance, los costes que esta educación suponía para la sociedad en su conjunto eran mínimos, lo que no suponía ningún grave problema social ni económico. Las universidades, algunas de ellas centenarias, llevan durante todo este período un vida relativamente apacible, sin que se produzcan cambios sustanciales, ni en su funcionamiento interno, ni en su relación con el entorno social.

2. Con ligeras diferencias de tiempo según los países, la universidad inicia un período de cambios radicales a finales de los años cincuenta y principios de los sesenta. Dos hechos se alían para generar este fenómeno:

2.1. El fuerte incremento del número de nacimientos que se da a partir del final de la Segunda Guerra Mundial en los países más desarrollados, genera un aumento considerable en el tamaño de las cohortes de edad que empiezan a ingresar en la universidad hacia los años sesenta.

2.2. Un desarrollo económico generalizado, que exige, por una parte, la

participación de mano de obra cada vez más cualificada en el mundo laboral, mientras que por otra facilita la adquisición de educación mediante la elevación del nivel de vida medio de la sociedad. Dicho en otros términos: la inversión en educación se convierte en una necesidad económica, mientras que el consumo de educación es estimulado fuertemente por los mayores niveles de renta de los individuos.

La consecuencia de ambos efectos es doble: no sólo cada vez hay más jóvenes en edad universitaria, sino que éstos, cada vez en mayor proporción, demandan educación superior.

En esta situación todas las fuerzas ideológicas, políticas y económicas se pusieron implícitamente de acuerdo en promover el desarrollo de la educación superior, aunque por motivaciones no siempre coincidentes: para unos se trataría básicamente de mejorar el rendimiento de las fuerzas productivas, mientras que otros veían la situación propicia para el progreso del nivel cultural de la sociedad, profundizando en la igualdad social y en el desarrollo de la democracia. Como consecuencia del interés creciente de los individuos y de las fuerzas sociales en la educación superior, los gobiernos de todos los países se aprestaron a satisfacer la fuerte demanda social mediante la creación de nuevos centros universitarios, la ampliación de los existentes y la asignación de cada vez mayores fondos a la educación superior. El desarrollo de la educación universitaria, durante las dos décadas aproximadas en las que se puede cifrar este período expansivo es impresionante, tanto si se mira el incremento en el número de estudiantes, como si se atiende a los gastos totales del sistema educativo superior.

3. En la segunda mitad de la década de los setenta se conjuntan una serie de hechos que provocan un cambio sustancial en la tendencia y en los puntos de vista que sustentaron la expansión de los años precedentes.

Sintéticamente son estos:

3.1. La expansión de la educación superior que tuvo lugar en el período anterior produjo dos efectos simultáneos: el incremento del gasto público en las universidades y la mayor democratización en el acceso a los niveles educativos superiores. Ambos efectos se combinan con la crisis económica de los sesenta y la llegada al poder de gobiernos conservadores en varios países, pero especialmente en los dos con mayor capacidad de crear opinión pública (USA y el Reino Unido, a través de sus potentes medios de comunicación científica y de masas), para que empiece a extenderse una corriente de opinión generalizada que, cuestionando las ventajas de la educación superior, propug-

na la reducción del gasto público en ella. En algunos países como el Reino Unido se producen drásticas reducciones del presupuesto de las universidades, mientras que en otros aparece una importante tendencia a estabilizar los presupuestos, frenándose el fuerte ritmo de crecimiento que se mantenía en los años anteriores. En casi todos los países desarrollados se puede apreciar como la participación de la educación superior en el PNB, se va reduciendo durante estos años (Eicher, 1988).

3.2. Una disminución en el tamaño de las cohortes de jóvenes en edad de ingresar a la universidad, lo que podría suponer un descenso en el número de estudiantes ingresando en las universidades. Sin embargo, a pesar de que numerosas previsiones en todo el mundo hablaban, y siguen hablando, de esta disminución de estudiantes, lo cierto es que tal hecho no ha tenido lugar todavía en ningún país desarrollado (UNESCO, 1989): sirva como ejemplo que en Estados Unidos había en 1987 un millón de universitarios más que en 1977 (Carnegie Foundation, 1987). La explicación hay que encontrarla en el hecho, ampliamente demostrado para numerosos países (Mora, 1989), de que la demanda de educación superior crece con el nivel de desarrollo económico, de modo que el aumento de la demanda entre los individuos en edad universitaria y la asistencia cada vez más numerosa de adultos a las aulas, compensa la disminución de la población juvenil. Sin embargo, aunque la temida (o tal vez deseada) disminución del número de estudiantes no se ha producido en ningún país desarrollado, y nada hace sospechar que se vaya a producir en un futuro inmediato, se está ya observando un estancamiento en las matriculaciones en las universidades europeas, lo que ha producido un cambio importante en la situación de expansión continua que había estado viviendo la educación superior durante los últimos treinta años de un modo ininterrumpido.

En este tercer período, las universidades se enfrentan a dos situaciones nuevas: por un lado, una estabilización del número de alumnos; por otro, un estancamiento, cuando no una reducción, de las aportaciones estatales al presupuesto universitario. Estos dos hechos suponen un nuevo planteamiento para la organización de unas instituciones que han vivido y crecido durante muchos años en un ambiente de continuo desarrollo expansivo.

Otros autores, Coutler y Moore (1987), en un estudio referido expresamente a la educación superior en USA, distinguen también tres períodos diferenciados respecto a su evolución, haciendo notar el cambio en los objetivos y las necesidades experimentados en aquel país. Son éstos:

— En los años sesenta se desarrolla el sistema educativo para proporcionar oportunidad de acceso a todos los ciudadanos. Las metas son accesibili-

dad y equidad en la oferta de educación superior. Lógicamente el sistema de financiación fue dirigido por el número de matriculaciones. Las valoraciones de las universidades se realizaban básicamente en términos cuantitativos, para lo que se desarrolló un sistema de recogida de datos sobre todo el proceso educativo superior.

— En la siguiente década, aunque la expansión continuó, los estados empezaron a preocuparse por el continuo incremento del gasto en educación superior, de modo que la contención del gasto y el freno al crecimiento se convirtieron en dos preocupaciones básicas. Instaurar un sistema de contabilidad de todo el proceso educativo, empezó a ser la meta básica de los responsables del sistema universitario, juntamente con la preocupación por la medida de la productividad y de la eficiencia en las instituciones de educación superior.

— En los años ochenta los problemas se han centrado en la necesidad de adaptar el sistema de educación superior a un cambio demográfico y a la preocupación generalizada por la calidad de la educación. La utilización de sistemas de asignación de recursos basados en índices de efectividad se convierte en una práctica habitual en las instituciones de educación superior. Otra característica de este período es la renovada importancia que se le otorga a la educación superior como fuente de crecimiento y de desarrollo tecnológico.

1.2. ALGUNAS PARTICULARIDADES DE LA EVOLUCION DE LA EDUCACION SUPERIOR EN ESPAÑA

Para evitar extrapolaciones ilegítimas de otras realidades, es necesario puntualizar que la situación española presenta algunas peculiaridades respecto a la de otros países de su entorno en cuanto a la evolución del sistema universitario en las últimas décadas. Dos son las diferencias más destacables:

1. En primer lugar, la gran expansión demográfica de la postguerra mundial se produjo en España con un retraso que se podría cifrar en unos diez años. Igualmente, el inicio del descenso de los índices de natalidad se dio con cierto retraso respecto al de otros países, de tal manera que mientras que en ellos llevan ya algunos años con cohortes en edad universitaria menos numerosas, en España hasta la segunda mitad de la década de los noventa no llegará a cumplir dieciocho años la primera cohorte con un tamaño menor que la precedente.

2. En segundo término, como consecuencia de la peculiar situación política vivida en España, junto con su menor nivel de desarrollo económico respecto a otros países europeos y norteamericanos, la oferta de educación superior fue considerablemente inferior, sobre todo en recursos económicos, a la de aquellos otros países. Un dato ilustra perfectamente esta situación: en el año 1980 España dedicaba a la educación superior un gasto público del 0,32 % del PNB, mientras que otros países europeos invertían proporciones del PNB que doblaban y triplicaban el porcentaje español, llegándose en el caso canadiense al 1,95 % del PNB (Eicher, 1988).

Aunque es evidente que en España se ha producido también un extraordinario desarrollo de la educación superior, la expansión se inició con algunos años de retraso respecto a otros países desarrollados, y nunca alcanzó el esplendor de medios económicos que conocieron otros sistemas universitarios. Esta diferencia de *tempo* tendrá una importante repercusión sobre la actual situación española:

1. Mientras que en otros países se puede hablar actualmente de un estancamiento de la población universitaria, en España este fenómeno no se presentará hasta dentro de algunos años, de modo que todavía cabe esperar en los próximos ocho años un apreciable incremento en el número universitarios. Por otra parte, no hay que olvidar que los niveles de demanda reales de educación superior en España son todavía inferiores a los valores medios de los países desarrollados, con la notable excepción del Reino Unido, por lo que puede esperarse razonablemente un incremento en la demanda que compense, e incluso supere, la disminución demográfica. La prolongación de la educación obligatoria hasta sólo dos cursos previos al ingreso en la universidad ayudará, sin lugar a dudas, a este incremento en la demanda. Unidos ambos efectos, se puede esperar que la estabilización del número de universitarios no se producirá en España antes del año 2000.

2. Por otra parte, la crisis económica de los setenta y la tendencia a la reducción de los gastos en educación superior llegó a España cuando aquí no se habían alcanzado todavía los niveles de gasto necesarios para disponer de un sistema universitario dotado de los suficientes medios para que se pudiera considerar de una calidad aceptable. En consecuencia, mientras que en otros países desarrollados se puede hablar con cierta lógica del alto gasto que supone la educación superior y de la necesidad de estabilizarlo, en España no se pueden hacer las mismas afirmaciones. Es evidente que se puede controlar mucho mejor la eficacia y la calidad del sistema universitario español, y sin duda debe hacerse, pero a diferencia de otros países, no parece posible que la calidad deseable pueda conseguirse con los mismos niveles de gasto actua-

les, que a pesar de los más recientes esfuerzos, todavía esta lejos de eliminar la diferencia que nos separa de otros países (Pedró, 1988).

1.3. **ALGUNOS PROBLEMAS ACTUALES DE LA EDUCACION SUPERIOR**

Sin duda se puede afirmar que, con algunas diferencias temporales según los países, estamos ante una nueva era para la educación superior caracterizada por la estabilidad en su crecimiento numérico y por la amplitud de la población que se va a beneficiar de ella: sería de desear que se caracterizara también por la madurez y por la calidad, tanto en los aspectos docentes como investigadores. Para que así sea será necesario afrontar los nuevos problemas que está generando una situación realmente distinta a la anterior.

Posiblemente, el principal problema que tendrán que afrontar las universidades en esta nueva época es el de adaptarse continuamente a un mundo con una enorme capacidad para los cambios científicos, sociales y tecnológicos. Este hecho exige dos tipos de repuestas: unas están relacionadas con los aspectos que se podrían considerar curriculares y que, aunque de indudable importancia, no van a tratarse en este estudio; el segundo tipo de respuestas se relaciona con la necesidad de convertir las universidades en organizaciones efectivas para los propósitos que la sociedad demanda, eficaces en su gestión y con el nivel de calidad necesario para formar individuos con una educación flexible que les capacite a adaptarse con facilidad a un mundo cambiante.

Otro tipo de problema al que se enfrenta actualmente la universidad es el de su poca eficiencia. Esta es una cuestión teóricamente muy debatida para la que existen dos puntos de vista generalizados (Levin, 1989): el que considera que la ineficiencia se debe al incremento continuo de los costes, dada la incapacidad de las universidades a aumentar la productividad en un sistema intensivo en trabajo y con elevados costes de personal (Baumol, 1968); y el que considera que las universidades, dados los intereses personales de profesores y administradores que son *de facto* sus propietarios, son incapaces de transformarse para aumentar la eficiencia. A pesar de esta problemática situación, Levin (1989) cree que es posible que aumenten simultáneamente tanto la productividad como la calidad de la educación superior. En cualquier caso dar los pasos para que el rendimiento de las universidades mejore es una obligación ineludible de los gestores de las universidades y de los responsables de los sistemas públicos de educación superior.

Con razón o sin ella, es evidente que existe una cierta idea generalizada, aunque difícilmente contrastable, de que, en la época de rápida expansión del

sistema universitario, éste se adaptó a la creciente demanda social con una cierta pérdida de calidad (en el segundo capítulo de este estudio se tratará de este concepto tan evanescente, y se discutirá si realmente se produce o no disminución de «calidad» cuando la educación superior se generaliza). Aunque tal afirmación sea dudosa, si parece lógico que, en cualquier caso, se aproveche el panorama actual de estabilización financiera y humana del sistema educativo superior para desarrollar y aplicar, como se está haciendo en los países más avanzados, técnicas de control de calidad que permitan mejorar la calidad, aumentar la efectividad y hacer más eficiente la totalidad del sistema de educación superior, de modo que el rendimiento social y económico de las universidades se optimice, sin aumentar necesariamente los gastos del sistema.

Otra cuestión que aumenta el interés actual por la calidad de la educación superior es la tendencia generalizada por parte de los empleadores a exigir algo más que la mera titulación. Al menos en algunos países con sistemas universitarios más diversificados, el exigir una mayor acreditación de los niveles de calidad de la educación recibida por los que buscan empleo esta empezando a ser algo común por parte de las empresas. Ante este cambio de las expectativas sociales respecto de la formación recibida en las universidades, éstas deberán responder con una mayor preocupación por la calidad de la educación que imparten.

El interés por la calidad de la educación superior en los últimos tiempos es algo generalizado en todos los países: «nunca antes en nuestra historia se había concentrado tanto interés en la calidad de la educación» (Miller, 1986: 242). Por otra parte, una gran cantidad de esfuerzos investigadores se están realizando en este mismo sentido por parte de organismos y asociaciones científicas; como ejemplo se podría citar que las dos organizaciones europeas con mayor dedicación a los aspectos institucionales de la educación superior, el *International Management in Higher Education Program* y la *European Association of Institutional Research*, dedicaron sus últimas reuniones anuales en 1990 exclusivamente al tema de la evaluación de la calidad de la educación superior.

1.4. PROGRAMA DEL PRESENTE ESTUDIO

En los capítulos siguientes se van a estudiar con detalle los esfuerzos que en los países occidentales se han venido realizando en la última década para la estimación de la calidad de las universidades y para la estimación de su rendimiento, efectividad y eficiencia. Los métodos que han sido propuestos por

investigadores e instituciones en diferentes países han creado una abundante literatura científica que será revisada con exhaustividad en las siguientes páginas.

Este estudio no se ha limitado de antemano a los países desarrollados, pero la realidad es que sólo en estos países, y no en todos, se han realizado investigaciones sobre calidad y rendimiento de las universidades; aunque tampoco el estudio se ha limitado a la última década, el hecho es que prácticamente sólo durante ella, se han venido realizando estudios sistemáticos para la evaluación de las instituciones universitarias. Sin embargo, en un país, los Estados Unidos, algunos de estos estudios se remontan a bastante antes. Dado su peculiar sistema de educación superior, fuertemente competitivo, hace ya casi un siglo que se realizan cierto tipo de estimaciones de la calidad de sus instituciones universitarias más señeras. Aún así, puede afirmarse que los métodos más rigurosos y de mayor interés no superan casi nunca los diez años de edad.

El programa con el que se presentan los resultados de este estudio es el siguiente:

1. En el capítulo II se pasa revista a las ideas que subyacen bajo el concepto de calidad, encuadrándolas dentro de los objetivos que persigue la educación superior.
2. En el capítulo III se trata del concepto de evaluación institucional, de los diferentes enfoques y de los principales métodos de evaluación.
3. El capítulo IV está dedicado al rendimiento institucional y a los indicadores de rendimiento que han sido propuestos en diversos países como medidas objetivas, aunque discutidas y contestadas, del rendimiento con la finalidad de asignar recursos racionalmente o de adoptar cualquier otro criterio de política universitaria.
4. En el capítulo V se trata la evaluación de las universidades desde otra perspectiva: los métodos que se han llamado autorregulados, es decir, promovidos por las propias instituciones con el fin de mejorar, a través del autoestudio, su efectividad y eficiencia.
5. En el capítulo VI se describen y analizan los métodos que se han desarrollado en los Estados Unidos para la estimación de la calidad de las instituciones de educación superior, generalmente estableciendo comparaciones y ordenaciones jerárquicas entre ellas, ya sea mediante escuestas reputacionales o utilizando medidas objetivas de estimación de la calidad.

6. El capítulo VII está dedicado al análisis de los métodos aplicados en diversos países para evaluar las instituciones universitarias y para medir su rendimiento. Dado que estos métodos de evaluación son objeto actual de trabajo en muchos países, se dedicará una parte importante del capítulo a revisar, país por país, las diferentes puestas en práctica.

7. Por último en el capítulo VIII, tras un sumario de las principales ideas que se desprenden de todo el trabajo, se formulan algunas ideas sobre el diseño de un método para la evaluación de la educación superior en España.

CAPITULO II. LA CALIDAD EN LA EDUCACION SUPERIOR

Quality... you know what it is, yet you don't know what it is. But that's selfcontradictory. But some things are better than others, that is, they have more quality. But when you try to say what the quality is, apart from the things that have it, it all goes proof! There's nothing to talk about. But if you can't say what Quality is, how do you know what it is, or how do you know that it even exist? If no one knows what it is, then for all practical purposes it really does exist. What else are the grades based on? Why else would people pay fortunes for some things are better than others... but what's the «betterness»? ...So round and round you go, spinning mental wheels and nowhere finding anyplace to get traction. What the hell is Quality? What is it?

Pirsing, 1974: 184

II.1. INTRODUCCION

Aunque el tema de la calidad de las instituciones de educación superior estará presente a lo largo de todo el estudio, en este capítulo se va a hablar fundamentalmente de qué significa este concepto, de cómo se puede evaluar en los distintos niveles institucionales (universidades, facultades, departamentos), y de qué se puede hacer para mejorar la calidad de esas instituciones. Aunque el estudio tiene como objetivo final colaborar en la mejora del sistema universitario español, se hará continua referencia a otros sistemas educativos, básicamente de los países desarrollados occidentales, en los que estos temas se han tratado con cierta intensidad en los últimos años. La experiencia de otros países, aunque tampoco demasiado extensa, puede servirnos de ayuda inestimable.

No toda la información que se recoge en este estudio puede ser aplicada a España; los diferentes sistemas universitarios carecen de la homogeneidad suficiente como para pensar en trasladar mecánicamente ideas y técnicas de unas realidades a otras. Así, mientras que en algunos países europeos existen sistemas universitarios con similitudes al español, en otros como el Reino

Unido, se dan diferencias estructurales e históricas importantes. Justamente el país del que más información se dispone, los Estados Unidos, tiene un sistema educativo radicalmente distinto al nuestro, con más de 3.000 instituciones universitarias de los más diversos tipos: un sistema diversificado, con diferentes niveles de calidad, y, sobre todo con una gran autonomía e independencia de cualquier tipo de poder. Por todo esto, las ideas y los hechos que se van a tratar aquí no siempre serán aplicables en el sistema universitario español.

Es necesario hacer otra salvedad previa: cuando se trata de la calidad de la educación superior, la mayor parte de la literatura científica se refiere a la calidad global de las instituciones, y no a la calidad de la formación que reciben los estudiantes. El punto de vista estrictamente de los estudiantes ha sido mucho menos tratado, posiblemente porque tiene menos interés para los responsables de la gestión universitaria, mientras que, por otro lado, la dificultad técnica de tener en cuenta el resultado final del proceso educativo es notoriamente elevada. La calidad de la formación de los estudiantes sólo ha intervenido como variable, junto con otras muchas, en la construcción de algunos índices de medida de calidad institucional (Kuh, 1981). Evidentemente ésta es una deficiencia de las valoraciones usuales de la calidad: el efecto global de la institución sobre sus usuarios debería ser una pieza esencial para evaluar la calidad institucional, indudablemente mucho más adecuado que el valorar, como suele hacerse habitualmente, la calidad de cada una de sus componentes por separado.

De un modo general puede decirse que existen dos enfoques diferentes para tratar el problema de la calidad y del rendimiento en las instituciones de educación superior:

1. El primero de ellos, que podríamos llamar *enfoque competitivo*, tiene su expresión más notoria en Estados Unidos, en donde tradicionalmente se ha dado una situación competitiva entre sus universidades, posiblemente como reflejo de la propia manera de ser de la sociedad: «Es interesante especular sobre por qué la educación superior norteamericana ha tomado su aspecto particular. De alguna manera, el sistema refleja la sociedad. Muchas instituciones no académicas en los Estados Unidos son competitivas y jerárquicas. Nosotros mismos parecemos preocupados de ser los mejores, tanto en deporte como fabricando automóviles, en armamento militar, en espectáculos de televisión o en moralidad» (Astin, 1985: 13). La consecuencia de esta actitud, en referencia a la calidad de las instituciones universitarias, se traduce en los numerosos *rankings* de universidades, de departamentos o de escuelas profesionales que se realizan constantemente en USA, intentando valorar o comparar la calidad de las distintas instituciones. Como el propio Astin (1985)

afirma, esta visión de la calidad ha conducido en muchos casos a una mera competencia por la adquisición de más medios materiales o por la búsqueda de los rasgos más brillantes, aunque no siempre mejores. A la consideración y análisis de los métodos de evaluación que ha generado en USA este enfoque, se va a dedicar expresamente el capítulo VI.

2. El otro enfoque sobre la calidad de la educación superior centra su atención en la mejora de las instituciones, sin establecer comparaciones entre unas y otras. Este enfoque es el que se ha utilizado en los países europeos, con sistemas públicos y homogéneos de educación superior, y en ciertos sistemas estatales dentro de los Estados Unidos. En estos casos la preocupación por la calidad de la educación superior ha venido de la mano de la necesidad de satisfacer una demanda creciente, conteniendo al mismo tiempo los gastos. Bajo esta perspectiva se ha hecho especial hincapié en cuestiones tales como la efectividad y la eficiencia del sistema universitario. Hasta no hace mucho, el uso de los indicadores de rendimiento constituía la herramienta esencial de este enfoque. Recientemente están desarrollándose otros métodos alternativos, como, por ejemplo, las autoevaluaciones de las instituciones. Aunque ya se volverá sobre este tema con más detalle en el capítulo VIII, este enfoque es sin lugar a dudas el más adecuado para el sistema universitario español, aunque no hay que descartar *a priori* la posibilidad de recurrir en el futuro a ciertas dosis de competitividad entre las instituciones universitarias españolas. A los métodos de estimación que se basan en este enfoque se les dedicará expresamente los capítulos IV y V; por su parte, en el capítulo VII se describen ejemplos de la puesta en práctica de esos métodos en diversos países.

Antes de seguir tratando conceptos como los mencionados conviene establecer el marco en el que se van a utilizar y definirlos de acuerdo con los criterios que son habituales en el campo de la investigación institucional. Por esta razón, en este capítulo se van a discutir algunas ideas que ayudan a justificar la necesidad de que la educación superior tenga los adecuados niveles de calidad; concretamente se hablará de los objetivos básicos del sistema universitario y de los beneficios que esta educación reporta a la sociedad en general y a los individuos que la adquieren, en particular.

II.2. OBJETIVOS Y BENEFICIOS DE LA EDUCACIÓN SUPERIOR

II.2.1. Distintas perspectivas

De un modo general se puede afirmar, y en esto casi todo el mundo estaría de acuerdo, que las misiones esenciales de la educación superior son enseñar,

investigar y servir a la sociedad mediante la trasmisión de los conocimientos. Sin embargo, si se trata de profundizar en los objetivos más inmediatos que son necesarios para alcanzar esas metas, se encuentran distintas posturas que, siguiendo a Astin (1985), pueden ser agrupadas bajo dos modelos antitéticos: el modelo de producción industrial y el modelo del desarrollo de la capacidad de los individuos. Veamos en qué consisten ambos:

1. *El modelo de producción industrial*

Para una parte importante de los individuos relacionados con la educación superior, sobre todo en los últimos treinta años, ésta tiene como objetivos básicos la producción de mano de obra altamente cualificada que pueda satisfacer las necesidades del desarrollo tecnológico. Bajo este punto de vista, una institución que tiene por objetivo la producción de las cualificaciones técnicas necesarias para el mercado laboral, debería ser gestionada como una industria. Una consecuencia de este enfoque es la buena cantidad de técnicas de gestión de instituciones universitarias que se han creado en los últimos tiempos con el objeto de mejorar su rendimiento (Hopkins y Massy, 1981; Mora, 1987). Lógicamente, cabe preguntarse lo apropiado que es este enfoque para la universidad. Parece obvio que las diferencias entre la producción industrial y las universidades son lo suficientemente importantes como para desechar el modelo de producción industrial como una concepción global de la educación superior: ni los estudiantes son como las materias primas que se pueden elaborar en un proceso perfectamente evaluable, ni el objeto de la educación es algo tan objetivamente medible como el beneficio económico en una empresa industrial.

2. *El modelo del desarrollo de las capacidades*

Bajo este punto de vista el objetivo básico de la educación superior es el de desarrollar el talento de los estudiantes y profesores hasta su mayor potencial. Utilizando una terminología económica: el objetivo de la educación superior es el de aumentar los recursos humanos globales de la sociedad. Bajo esta perspectiva, valorar la calidad de la universidad deberá traducirse en evaluar la capacidad de las instituciones para mejorar la capacidad de los individuos que la integran, de modo que la mejor institución sería aquella que mayor impacto tenga (mayor *valor añadido*, como diría un economista) sobre los conocimientos y el desarrollo personal de los usuarios (Astin, 1989). Este es posiblemente el punto de vista que mantienen con mayor frecuencia aquellas personas más involucradas e interesadas en el proceso educativo. Así, por ejemplo, Mark Blaug afirma que: «La enseñanza en la universidad sólo puede

ser valorada en términos del *valor añadido* a los estudiantes entre el momento de la entrada y el de la graduación» (Blaug, 1966: 401). Igualmente, Miller (1986) afirma que el rápido interés por el *valor añadido* como punto de vista sobre la educación superior seguirá siendo dominante en los próximos años.

Si consideramos que el desarrollo de la capacidad de los individuos es uno de los objetivos básicos de la educación superior, será necesario complementarlo con otro de no menos importancia, como es el de proporcionar la igualdad de oportunidades a todos los miembros de la sociedad para que tengan acceso a la educación superior y poder así desarrollar sus capacidades intelectuales. Aceptar el punto de vista del *valor añadido* tiene implicaciones importantes, puesto que habrá que aceptar como consecuencia que un sistema educativo es tanto mejor cuando maximiza conjuntamente número y calidad de los individuos que se benefician de él.

II.2.2. Los objetivos de la educación superior

Evaluar la educación superior exige, como paso previo, el determinar cuáles son los objetivos perseguidos por ésta. Evidentemente sus objetivos son múltiples, pero además son diferentes las valoraciones que pueden hacer de ellos. Como una muestra del amplio abanico de criterios, veamos algunas propuestas formuladas por diferentes autores que han tratado este tema:

El *Robbins Report* (1963) es una de las primeras visiones generales que se formularon sobre la educación superior. En este informe se consideró el sistema universitario británico globalmente y se formuló una política coherente para todo el conjunto. Respecto a los objetivos de este sistema educativo se establecieron las siguientes metas:

- Capacitar a los individuos para desempeñar las funciones que exige la división del trabajo.
- Enseñar a los alumnos de modo que se promuevan las potencialidades de la mente.
- Avanzar en el conocimiento.
- Transmitir la cultura y los valores de la sociedad.
- Proporcionar educación superior a todos los que puedan beneficiarse de ella y deseen hacerlo.

La *Carnegie Comission* (1973), una institución norteamericana dedicada a los temas educativos, establecía los siguientes objetivos para la educación superior estadounidense:

- Proporcionar oportunidades para el desarrollo de las habilidades personales, para el desarrollo intelectual, estético y ético de los estudiantes individuales, y proporcionar un ambiente que pueda ayudar positivamente a un desarrollo general del individuo.
- Avanzar en la capacidad de la sociedad considerada como un todo.
- Estimular la equidad educativa para todos los jóvenes en edad universitaria.
- Transmitir y avanzar en el conocimiento y en la sabiduría.
- Valorar críticamente la sociedad y actuar persuasivamente para promover la autorrenovación de la propia sociedad.

También en referencia a la educación superior en los Estados Unidos, Leslie y Brinkman (1988) enumeran los siguientes objetivos básicos:

- Promover las oportunidades educativas para todos los individuos.
- Promover el crecimiento económico y la productividad del sistema económico.
- Proporcionar mano de obra capacitada.
- Conseguir objetivos sociales específicos.
- Desarrollar una ciudadanía educada.
- Crear conocimiento y estimular el aprendizaje.

Pero la educación superior no sólo tiene unas metas internas, sino que también realiza una serie de servicios a la comunidad, que también pueden considerarse como objetivos del sistema universitario. En un informe de la OCDE sobre los efectos que la universidad ofrece a la comunidad (Ceri, 1982) se enumeran los siguientes:

- Poner al servicio de la comunidad sus medios y recursos humanos, y disseminar en ella nuevos conocimientos y mano de obra preparada.

- Contribuir al análisis de las necesidades de la comunidad.
- Participar en el análisis de los problemas y proponer posibles soluciones.
- Participar en la implementación de las soluciones cuando esto es compatible con el estatus de la universidad.

En un estudio empírico sobre los objetivos de las universidades, Romney (1978) realizó una encuesta entre universitarios (gestores y profesores) sobre cuáles debían ser las metas de las instituciones de educación superior. Aunque las respuestas varían según el tipo de institución y el grupo de encuestados, a continuación se citan las 10 que fueron consideradas más importantes por el profesorado, según orden de valoración:

- Desarrollo académico del alumno.
- Preparación para servir a la comunidad.
- Orientación intelectual.
- Formación para la libertad.
- Ambiente intelectual y estético del campus.

Desde el punto de vista de los tres niveles a los que se dirigen los objetivos de la educación superior (individual, institucional y social), Weert (1990) considera dos tipos de metas: internas y externas. En el cuadro II.1 se esquematizan estas ideas.

III.2.3. **Los beneficios de la educación superior**

La valoración de la calidad de la educación superior deberá hacerse en función de su capacidad para alcanzar adecuadamente los beneficios que cabe esperar de ella. Por ello conviene señalar brevemente cuáles son algunos de estos beneficios. Numerosos autores los han señalado y analizado con detalle, como por ejemplo, Astin (1977), Bowen (1977), Haveman y Wolfe (1984).

Los beneficios que proporciona la educación superior pueden agruparse en dos categorías básicas: los beneficios de carácter individual y los de carácter social.

CUADRO II.1

Metas de la educación superior

NIVEL SOCIAL

Internas

- Total de educación proporcionada a la sociedad.
- Variedad y diversidad del sistema de educación superior.

Externas

- Relación adecuada entre oferta y demanda de graduados.

NIVEL INSTITUCIONAL

Internas

- Disponibilidad de recursos.
- Política de admisiones
- Eficiencia interna.

Externas

- Mejora de las cualificaciones educativas y profesionales de los alumnos.
- Desarrollo de la independencia de mente y de la autonomía personal.
- Respuestas institucionales al medio exterior.

NIVEL INDIVIDUAL

Internas

- Contribución del programa educativo al valor añadido.
- Aprendizaje de estrategias y procesos.
- Disponibilidad de oferta de cursos, asesoramiento y apoyo.

Externas

- Programas de acuerdo con las necesidades del mercado laboral.
 - Adquisición de capacidades extracurriculares.
 - Clima de innovación.
 - Desarrollo personal del individuo.
 - Gobierno democrático de la institución.
 - Preparación profesional.
 - Formación humana y altruista.
-

A. *Beneficios sobre los individuos*

Son aquéllos que más directamente afectan a las personas que adquieren educación superior, pero que evidentemente tienen un efecto social, ya que, de algún modo, se transmiten al resto de la sociedad. Estos beneficios se pueden clasificar en tres categorías:

1. Beneficios educativos

Son los cambios beneficiosos que se generan en las capacidades, conocimientos, valores, actitudes, intereses, hábitos, etc., de los estudiantes como consecuencia de su paso por las instituciones universitarias. Es evidente que estos beneficios sobre los individuos, en una educación superior generalizada, tienen una clara repercusión sobre toda la sociedad. Aunque no es fácil valorar cuantitativamente todos estos efectos que los estudios superiores tienen sobre los individuos, las estimaciones realizadas por diversos autores demuestran su enorme importancia. Por otra parte, no se puede olvidar el efecto de las instituciones universitarias sobre el desarrollo de la capacidad del profesorado, tanto sobre sus cualidades docentes, como en su capacidad para el trabajo académico e investigador.

2. Beneficios marginales

Son los beneficios adquiridos por los estudiantes por el hecho de recibir unos credenciales que les otorguen ventajas sociales y económicas que, en principio, no tienen que ver con su formación específica. Este es un hecho notorio en sistemas universitarios tan variados y jerarquizados como el estadounidense, en donde los títulos recibidos de universidades situadas en la cúpula jerárquica proporcionan muchos más beneficios de este tipo que los títulos de universidades menos prestigiosas. Esta clase de beneficios se extiende también al profesorado, de una manera semejante a la citada para los alumnos. La importancia de estos beneficios también ha sido evaluada por diversos autores.

3. Beneficios existenciales

Son los que produce la experiencia universitaria sobre los individuos que la viven, con independencia de los cambios en las capacidades intelectuales o en las ventajas sociales. Es indudable que la vida universitaria representa para los individuos que la disfrutan una experiencia profunda motivada por las relaciones humanas entre compañeros, por las actividades extracurriculares y por el ambiente cultural que rodea la vida universitaria. Podría decirse que los beneficios existenciales vienen dados por el conjunto de experiencias subjetivas vividas por el estudiante mientras permanece en la universidad y que configurarán de modo importante su personalidad. Este tipo de beneficios son esenciales, especialmente en el caso de determinadas instituciones universitarias de indudable calidad. Por otra parte, es obvio que los beneficios existenciales son básicos para explicar la situación y actitud de una gran parte

del profesorado universitario respecto a sus propias instituciones. Confirmando la idea de los beneficios existenciales de la educación, se puede recurrir a uno de los escritos más antiguos que se pueden citar: el faraón Amenemmes I, de la XII Dinastía, que subió al trono en el 1990 a. de C., alentaba la formación de escribas que fueran capaces de dirigir la Administración del Estado. Al crear una escuela para su formación (lo que venía a ser una especie de universidad de la época), el faraón, en unos escritos que se conservan, afirmaba que los escribas «sea cual sea su empleo en la administración, no es en ella desdichado» y demuestra posteriormente que el oficio de escriba es muy superior a cualquier otro en cuanto a sus condiciones de vida (Cassin y cols., 1983: 280).

B. *Beneficios para la sociedad*

Siguiendo la clasificación que hace Shattock (1990) sobre las contribuciones de las universidades al progreso de la sociedad, podemos hablar de tres clases de beneficios sociales:

1. Contribución al progreso social

Aunque es discutida la contribución de la educación a ciertos aspectos del progreso social, pueden citarse algunas aportaciones de indudable importancia:

— Igualdad social en el acceso a la educación superior. Aún con muchas limitaciones por causas externas a las propias instituciones universitarias, éstas mantienen una tendencia creciente a recibir a estudiantes de todas las clases sociales, de modo que el elitismo en las universidades está disminuyendo paulatinamente.

— Igualdad en el acceso para las mujeres y para las minorías étnicas. Las universidades han colaborado a través de programas de acción a la igualdad en el acceso a ellas de ambos grupos sociales.

— Educación para adultos. En muchos países las universidades ejercen una importante misión de justicia intergeneracional, proporcionando cursos especiales para aquéllos que no tuvieron posibilidades de acceso en su juventud, o que necesitan un reciclaje profesional.

— Influencia sobre la enseñanza previa a la universidad. La universidad ejerce una influencia directa en la enseñanza primaria y secundaria, a través de los requisitos académicos de entrada en ella, e indirecta, por medio de la

formación que suministra a los que serán profesores del sistema educativo no universitario.

— Impacto de la investigación. La investigación universitaria sobre temas sociales, su influencia sobre los cambios sociales y sobre las actitudes, aunque de difícil evaluación, es muy significativa.

— Impacto sobre le entorno local y regional de las instituciones. Las universidades ejercen una indudable influencia ambiental, especialmente las situadas en áreas más aisladas. Shattock (1990), como caso extremo de la importancia que puede revestir esta influencia sobre el entorno, recuerda el caso del asesinato del rector y cinco profesores de la Universidad Centroamericana de San Salvador, ocurrido en 1989, a causa del peso que sus opiniones en defensa de los derechos humanos ejercían sobre la sociedad salvadoreña.

2. Contribución al crecimiento económico

Existe una extensa literatura sobre la influencia de la educación superior en el crecimiento económico [véase, por ejemplo, Leslie y Brinkman (1988)]. Muy sintéticamente se pueden recordar los siguientes efectos:

— Impacto económico de las instituciones universitarias sobre el entorno local y regional. El caso de Stanford sobre el llamado *Silicon Valley* es paradigmático, pero no único. No hay que olvidar tampoco el impacto del gasto universitario sobre la economía local o regional.

— Efecto de la explotación de la investigación científica y tecnología. Es evidente que, actualmente más que nunca, el desarrollo económico depende del desarrollo científico, que está preferentemente unido a la actividad de las universidades.

— La contribución de la universidad a la formación de mano de obra preparada es sin duda una de las causas más influyentes en el crecimiento económico de los pueblos.

3. Contribución al desarrollo cultural

La influencia de las universidades al desarrollo cultural es imposible de estimar, pero es obvia en todos los aspectos del pensamiento humano, de la actividad política, de la transmisión de las ideas e incluso de las manifestaciones artísticas. Shattock (1990) recoge un amplio muestrario de estos efectos de la educación superior.

II.3. LA CALIDAD EN LA EDUCACIÓN SUPERIOR

La frase de Pirsing que precede este capítulo es sin duda todo un éxito de citación entre los investigadores que se han acercado al problema de la calidad en la educación superior. Al menos Lawrence y Green (1980), Soimon (1981), Chambers (1984), Conrard y Blackburn (1985a), Elton (1986), Billing (1986), Tan (1986), Weert (1990) utilizan las palabras de Pirsing como presentación de su discusión sobre la calidad. Las dudas del filósofo respecto a qué es la calidad son asumidas por la mayoría de los que intentan abordar el problema. Algún investigador llega a afirmar: «La calidad como la belleza esta en el ojo de quien la observa» (Tan, 1986). Para otros, responder a la pregunta de qué es la calidad exigiría «una piedra de Rosetta, que ni los administradores ni los profesores han desenterrado todavía» (Chambers, 1984). Hay incluso quien opina (Os y cols. 1987) que la utilización de la palabra calidad sirve para esconder incertidumbres y diferencias de opinión acerca de a lo que se están refiriendo en realidad, una especie de refugio para aquéllos que evitan una discusión abierta sobre temas más concretos, como formulación y elección de objetivos, análisis de efectividad, evaluaciones, etc. Algunos simplifican la cuestión y, desde un punto de vista eminentemente práctico, afirman que calidad es simplemente la adecuación de las instituciones a un propósito educativo predeterminado (Ball, 1985), con lo que su evaluación se reduce a la medida del grado de aproximación entre la realidad y esos objetivos de las instituciones.

Por otra parte, dado que la calidad está relacionada en muchos casos con la obtención y reparto de fondos, su definición no sólo tiene problemas conceptuales como los expuestos, sino también políticos (Blackburn y Lingelfelter, 1973) por lo que su definición, o el concepto que se tenga de ella, tiene consecuencias importantes para el funcionamiento de los sistemas universitarios.

Sin una idea demasiado clara de qué se entiende por calidad de una institución universitaria, cientos de investigadores en todo el mundo han tratado de evaluarla. Evidentemente, los problemas que produce la valoración de la calidad en la educación superior generan gran cantidad de contradicciones, no siempre resueltas satisfactoriamente. Así, por ejemplo, cuando se utiliza una visión funcional de la calidad, se buscará lógicamente el adecuar el funcionamiento de una universidad a unos ciertos objetivos explícitos que deberán incluir la consideración de todos los aspectos valiosos del trabajo universitario; sin embargo, el problema de marcar unos objetivos explícitos es que éstos no necesariamente resaltarán los aspectos más meritorios, que no siempre son los más útiles, de la labor universitaria (Kuh, 1981). Por otra parte, intentar adecuar

la universidad a los resultados que dan simples indicadores de rendimiento, podría equivaler a convertir la literatura en la mera producción de *best-sellers* (Elton, 1986). Esta es una consideración importante que habrá que tener siempre presente, incluso cuando por puras razones prácticas, la evaluación de la calidad se reduzca a la utilización de estos indicadores.

Como señala George (1982), la mayor fuente de confusión a la hora de plantearse la mejora de la calidad de las instituciones universitarias nace de la incapacidad para definir previamente que se entiende por calidad. Por esta razón muchos autores son partidarios de las evaluaciones subjetivas de la calidad realizadas por expertos; quienes así opinan afirman que, en un proceso tan complejo como el de la educación superior, no existe un acuerdo sobre qué es la calidad, pero aunque lo hubiera estaría tan lleno de matices que la evaluación a través de unas medidas objetivas sería prácticamente imposible. Del mismo modo que la calidad de una orquesta sinfónica es juzgada por expertos y no por unas medidas objetivas, se debe medir la calidad de una universidad (Williams, 1986).

II.3.1. **Distintos enfoques sobre la calidad**

Existen diversas maneras básicas de encarar el problema de la calidad en las instituciones de educación superior según sean las características institucionales que se creen más destacables. Los diversos enfoques que han sido utilizados pueden encuadrarse en las categorías que proponen George (1982) y Astin (1985):

1. *La calidad como reputación*

La idea básica de este enfoque es la de que la calidad de una institución de educación superior puede ser evaluada mediante encuestas realizadas a un buen número de profesores de otras instituciones, a los que se les supone la capacidad de juzgar sobre la calidad de cada institución desde su punto de vista subjetivo. Este es un enfoque típicamente norteamericano, ya que en su sistema universitario es común confundir calidad con reputación. Son tan numerosos los estudios sobre reputación de las universidades norteamericanas, que han pasado a formar parte de la visión que se tiene de ellas en todo el mundo. Para muchos este punto de vista debe ser incluido dentro del folklore típico de un país en el que son algo común las encuestas de opinión sobre las más variopintas materias, la mayoría de las veces con el objetivo de establecer clasificaciones jerárquicas sobre que es lo mayor o lo mejor. Las valoraciones de la calidad realizadas por Carter (1966), Roose y Andersen (1970), Jones y

cols. (1982) son las piezas básicas de la versión más seria de este enfoque que es ya centenario, y que encuentra expresiones menos científicas en multitud de clasificaciones jerárquicas que aparecen en revistas de actualidad y en libros de dudosa seriedad. Valoraciones de carácter reputacional sobre las instituciones también son realizados en el Reino Unido como parte del sistema de valoración de la calidad de los departamentos con el objetivo de convalidar las titulaciones que conceden (Church, 1988). A los estudios de carácter reputacional en USA, a sus críticas y a los modelos derivados de este enfoque se dedicará íntegramente el capítulo VI.

2. *La calidad como disponibilidad de recursos*

El punto de vista reputacional de la calidad es frecuentemente rechazado por subjetivo. Frente a él se consideran otros métodos de evaluación de la calidad basados en medidas objetivas, como puede ser el caso de los recursos de que dispone la institución para la realización de sus funciones docente e investigadora. Cuatro tipos básicos de recursos son habitualmente considerados: personal, medios físicos, recursos financieros, y estudiantes. El profesorado, la parte más importante del personal, suele ser evaluado mediante magnitudes como la tasa de publicación científica, la proporción de doctores, u otras medidas como premios o distinciones académicas conseguidas. Los estudiantes suelen ser evaluados a través de las notas medias de acceso a la institución. Los medios físicos se evalúan típicamente en términos de número de aulas, su calidad y tamaño, disponibilidad y tamaño de las bibliotecas, laboratorios, residencias, etc. Otras medidas usuales de carácter mixto son el número medio de alumnos por clase, el coeficiente profesor/alumno, el gasto por alumno y los salarios de los profesores.

Parece obvio que, como señala Williams (1986), en una institución con pocos recursos se darán necesariamente algunas deficiencias relacionadas con la calidad de los estudiantes que eligen ese centro, la calidad del equipamiento, y la calidad y cantidad de los empleados.

Bajo este punto de vista se valora la calidad de una institución por la calidad académica de los alumnos que admite y por los medios de que dispone, aunque no por la adecuada e intensiva utilización de esos medios. Aunque exista una cierta correlación entre medios disponibles y calidad de una institución, es evidente que más importante todavía es el uso que se hace de los medios disponibles. Es más, parece bastante evidente que la mayoría de las instituciones de educación superior podrían mejorar su rendimiento y su calidad simplemente haciendo mejor uso de sus disponibilidades actuales (Williams, 1986).

Si se sigue este enfoque se tiende a valorar mejor las instituciones grandes, que son habitualmente las que disponen de mayores medios, lo que no tiene por qué coincidir siempre con la calidad.

3. *La calidad a través de los resultados*

Otro modo de enfocar la valoración de la calidad de las universidades es a través de la evaluación de su producto, es decir, del resultado global de sus graduados cuando se incorporan al mundo laboral una vez han acabado su vida universitaria. Este enfoque valora magnitudes tales como la facilidad en encontrar empleo de los graduados, los salarios que éstos consiguen, los doctorados que obtienen, o incluso en USA se llega a contabilizar la cantidad de ex alumnos de una institución que aparecen en publicaciones como *Who's who*. Es evidente que este enfoque sería el más adecuado si realmente se valorase lo que las instituciones universitarias aportan a los individuos que pasan por ellas. Sin embargo, existe un problema: los resultados del sistema no son independientes de la calidad y formación de los individuos que ingresan. La realidad es que las instituciones con mayor reputación atraen los mejores y, en consecuencia, producen los graduados con mejores posibilidades de empleo y de conseguir buenos salarios. El mismo caso se presenta si consideramos la cantidad y calidad de la producción científica del profesorado en los sistemas universitarios jerarquizados y con grandes diferencias de medios y de salarios. Un último problema de este tipo de enfoque es que las valoraciones de los resultados se hacen con demasiado retraso respecto al momento en el que el estudiante pasaba por las aulas, por lo que el efecto de retroalimentación sobre el sistema educativo es muy bajo, no consiguiéndose uno de los principales objetivos de la evaluación de la calidad: su mejora.

En un sistema universitario como el español, con una notable similitud en los *inputs* de las distintas universidades, medir la calidad por los resultados podría ser un método bastante aproximado de valorar la calidad de las instituciones, pero el problema del retraso entre el proceso de formación del alumno y el momento real de la evaluación lo hace poco interesante, salvo en algún caso concreto como puede ser el de los MIR. Este tipo de enfoque ha sido utilizado en el Reino Unido para evaluar sus universidades, pero sin resultados prácticos de interés (Johnes y Taylor, 1987; Johnes y cols., 1987; Taylor y Johnes, 1989).

4. *La calidad por el contenido*

Otro punto de vista de abordar el problema de la calidad de las instituciones universitarias es a través de la consideración de lo que se enseña en ellas, de

cuál es su nivel docente, de cuál es su currículum, de cuál es su sistema pedagógico o de cuál es el clima que se respira en el campus. Este enfoque se enfrenta a la grave dificultad de evaluar los conceptos implicados en el proceso de aprendizaje. Lo que es evidente es que existe una tendencia a la búsqueda de la calidad por parte de los estudiantes que se inclinan por los *currícula* más prestigiosos, no siempre porque sean los profesionalmente más interesantes. Así, por ejemplo, en USA se está dando un curioso fenómeno en este aspecto (Astín, 1985): en los últimos tiempos está teniendo lugar una pérdida de interés de los estudiantes por temáticas demasiado específicas, mientras que se están desarrollando, en cambio, los *currícula* basados en ciencias básicas y en humanidades (las llamadas *liberal arts*), juntamente con la ingeniería, por estar cada vez consideradas más prestigiosas que otros tipos de estudios que eran hasta hace poco tradicionales, como es el caso de los estudios de primer ciclo de escuelas de profesorado y de negocios *business*, que, por el contrario, han sufrido drásticas reducciones en prestigio.

5. *La calidad por el «valor añadido»*

Este punto de vista dirige su atención a lo que la institución ha contribuido a la formación global del estudiante, es decir, a lo que el estudiante ha aprendido mientras ha estado en la universidad. Bajo esta perspectiva, la calidad de una institución deberá ser evaluada por la diferencia entre el nivel de formación de los estudiantes al ingreso y a la salida de las instituciones universitarias. La dificultad de esta visión está en las dificultades de medida de este proceso, dado el gran problema que supone el intentar aislar los efectos de la institución de otros efectos ajenos a ella. Aunque este punto de vista parece sin duda el más correcto, sus implicaciones prácticas son escasas por las dificultades de evaluación.

Habría que preguntarse ahora cuál es el enfoque de entre los citados que es más consistente con los objetivos fundamentales que persiguen las instituciones y el sistema universitario como un todo. Evidentemente el punto de vista de la reputación sólo es válido en sistemas muy diferenciados y competitivos; pero incluso en ellos es dudoso el que se esté valorando la calidad; básicamente lo que se hace es establecer órdenes jerárquicas basadas en la reputación, que no siempre está unida a la calidad. El punto de vista de los recursos parece más aceptable, ya que resulta evidente que un cierto nivel de ellos son imprescindibles para alcanzar un nivel mínimo de calidad; sin embargo, no está demostrado que exista una correlación demasiado estrecha entre el gasto en educación y los buenos resultados, una vez superados esos mínimos imprescindibles. Como Bowen (1981: 27): afirma «...nadie conoce con certeza la relación existente entre el dinero gastado y los auténticos resultados educati-

vos». Por otra parte, aunque el tamaño de las instituciones beneficia las disponibilidades de recursos para los alumnos, éstos posiblemente sean menos utilizados por los alumnos de las grandes universidades que las pequeñas, mientras que las relaciones dentro del campus son más efectivas en las pequeñas instituciones (Astin, 1985). Por lo que respecta al contenido de los programas existe poca evidencia de que unos sean mejores que otros, sobre todo si se miran a largo plazo.

El problema del enfoque óptimo de la calidad sigue sin resolverse. En su estudio sobre esta cuestión, Astin (1985: 53) acaba concluyendo que: «...las instituciones que son excelentes en términos de reputación, recursos, resultados o contenido curricular no son necesariamente efectivas en desarrollar el talento de los estudiantes y del profesorado. De hecho, la única magnitud consistentemente relacionada con los beneficios educativos parece ser el tamaño institucional, y resulta que tiene un amplio efecto negativo sobre el desarrollo del estudiante». Podría afirmarse, para resumir, que la búsqueda de la calidad a través de medidas como las indicadas tiene poco que ver con la mejora del sistema como un todo y con la calidad del programa educativo. Sin embargo, todas ellas juntas, debidamente ponderadas, sí deben dar una cierta idea de la calidad de una institución; cada una de ellas contiene alguna información válida y puede contribuir a entender el problema de la calidad de la educación superior.

II.3.2. Factores que afectan a la calidad

Independientemente de los enfoques sobre la calidad que se han citado en el apartado anterior, existen toda una serie de factores que, con mayor o menor peso, influyen sobre la calidad de las instituciones universitarias. Kuh (1981), analizando el sistema universitario estadounidense, clasifica los factores que influyen sobre la calidad de las instituciones universitarias en cuatro tipologías. Son las siguientes:

A. *Índices sobre características de los alumnos*

- Capacidad intelectual.
- Rasgos demográficos.
- Actitudes y preferencias personales.

Aunque las medidas sobre capacidad de los alumnos son frecuentemente usadas como medida de calidad, es evidente que ellas por sí solas no miden la calidad del proceso educativo, aunque si se puede inferir alguna relación

entre la calidad del centro y la de los alumnos que demandan entrar en él. Las características demográficas como sexo, proporción de extranjeros o variedad cultural de origen pueden tener alguna influencia sobre la calidad, especialmente en los países multiculturales. Las actitudes, gustos y preferencias de los alumnos respecto a actividades académicas y no académicas de las instituciones, o a sus características ambientales, pueden ser una buena manifestación del correcto funcionamiento del proceso educativo de las instituciones.

B. *Indices referentes al funcionamiento*

- Relación entre compañeros.
- Métodos de aprendizaje.
- Relación informal profesor-alumno.
- Esfuerzo de los estudiantes.

Es obvio que las características relacionadas con el funcionamiento deberán tener la máxima influencia en la calidad de la institución. La relación de los alumnos entre ellos, así como la relación entre profesores y alumnos, están fuertemente relacionadas con la calidad, dada la importancia del aprendizaje fuera de las aulas. La calidad del proceso de trasmisión de los conocimientos es otra característica esencial en la evaluación de la calidad.

C. *Indices referentes a las instituciones*

- Tamaño de la institución.
- Propósitos explícitos de la institución.
- Organización formal e informal.
- Recursos financieros.
- Servicios residenciales para los estudiantes.

Varias manifestaciones de calidad han sido asociadas con el tamaño de la institución; por un lado, en las mayores universidades existe una mayor disponibilidad de recursos y de alternativas, aunque, por otra, en las menores la participación de los estudiantes en las actividades extracurriculares es mayor, así como la participación en las decisiones institucionales. Algo semejante se puede decir de los objetivos de las instituciones, normalmente mucho mejor definidos en las pequeñas instituciones. Las condiciones ambientales de los estudiantes en el campus también están relacionadas con la calidad del proceso educativo, dada la gran cantidad de tiempo que los estudiantes pasan en él. Por último, el nivel y la eficacia de la organización son también manifestaciones claras de la calidad del proceso.

D. *Indices sobre resultados*

- Persistencia.
- Logros.
- Desarrollo intelectual y social.
- Situación social y laboral de los ex alumnos.

El nivel de retención de alumnos en las instituciones puede ser usado como un índice de la calidad, ya que frecuentemente el rechazo de los estudiantes puede estar relacionado con el poco atractivo de la institución. Aunque es evidente que una manifestación de la calidad de una institución son los logros académicos de sus estudiantes, las medidas de éstos son difíciles de realizar. Basándose en la hipótesis de que el estudiante con una mejor formación es más abierto, sofisticado y libre en su forma de pensar, se han diseñado algunas valoraciones del desarrollo intelectual y social del individuo en relación con el tipo de institución. Gran número de investigadores han estudiado los beneficios económicos de los ex alumnos de diferentes centros como indicador de la calidad; aunque parece existir una relación positiva, es difícil asignar los diferenciales a la calidad de la institución y no a otra característica.

Resumiendo estas ideas de Kuh (1981) sobre los factores que influyen sobre la calidad, se podría decir que la calidad de las instituciones universitarias esta básicamente relacionada con:

- El tamaño de la institución.
- El grado de asociación de los estudiantes con la actividad universitaria extra-académica.
- El ambiente general de aprendizaje que rodea al alumno.
- La calidad del esfuerzo de los profesores y alumnos.
- La coherencia y consistencia de los objetivos del proceso educativo.

Sin embargo, ninguno de esos factores es por sí solo suficiente para explicar la calidad de las universidades. Realmente la calidad debería ser medida por el total de «valor añadido» que cada institución genera sobre los individuos que pasan por ella (Solmon, 1981), lo que supondría algo tan complejo como controlar todo el proceso educativo.

A pesar de todas estas ideas generales sobre la calidad, en una reciente encuesta realizada a directores de departamentos de algunas de las más prestigiosas universidades estadounidenses (Heveron, 1987), éstos daban sus opiniones sobre el mejor método de mejorar la calidad de sus departamentos. Las cuatro estrategias que los directores proponían eran, por orden de más o menos valoradas, las siguientes:

- Contratar profesores prestigiosos.
- Invertir en la formación de jóvenes.
- Hacer ambas cosas.
- Disponer de un fuerte liderazgo.

Además, los directores de departamento también valoraban altamente estas otras estrategias:

- Mejorar la reputación de la institución.
- Planificar y revisar el proceso.
- Organizar el departamento.
- Tener cierto grado de buena suerte.

Como se puede apreciar, los inmediatos responsables de los departamentos, al menos de las consideradas universidades prestigiosas, difieren en sus opiniones sobre el camino hacia la calidad que tienen los expertos en la tema: la visión de los directores de departamento es mucho más pragmática y basada claramente en un enfoque reputacional de una indudable estrechez de miras.

II.3.3. **La evaluación de la calidad**

Aún sin una idea demasiado clara del concepto de calidad en la educación superior, una gran cantidad de esfuerzo investigador ha sido dedicado a su evaluación. De todos modos, como afirma George (1982), lo más importante en este asunto es que aquellas personas involucradas en la evaluación de la calidad tengan una idea razonablemente clara de lo que se proponen antes de desarrollar las técnicas y procedimientos de la evaluación, si no se quiere que este proceso sea un cúmulo de frustraciones, confusión y contiendas. A pesar de todas las limitaciones, si se toman las debidas precauciones, es posible de algún modo aproximarse al problema de la evaluación de la calidad de las universidades.

Evaluar la calidad exige dos pasos previos: el primero es establecer las metas y objetivos que se propone el sistema universitario, para en segundo lugar definir qué magnitudes o aspectos del sistema pueden valorar el grado de aproximación de la realidad a esas normas o criterios previamente establecidos. Un tercer paso consistirá en decidir la ponderación que se va a asignar a cada una de las magnitudes propuestas para la evaluación. Establecer esta serie de criterios es un proceso delicado en el que hay que tener en cuenta no sólo todos los factores que intervienen en el proceso, sino muy esencialmente el trasfondo de intereses sobre la educación superior que cada enfoque o magnitud propuesta conlleva.

La mejora de la calidad de las universidades exige, según Levin (1989), dos tipos de información: información para el conocimiento e información sobre las alternativas. La primera necesita un sistema de medidas, resultado de evaluaciones sobre las instituciones, que digan a los responsables de éstas como está funcionando la institución globalmente y cada una de sus unidades en particular; la segunda es necesaria para poder conocer el efecto que sobre la mejora de la calidad y del rendimiento pueden tener las medidas alternativas que se pueden adoptar.

En general los enfoques metodológicos a la evaluación de la calidad se pueden clasificar en cuantitativos y cualitativos. Ambos métodos son válidos y proporcionan una visión complementaria de la institución que se quiere evaluar. Una visión ecléctica de la evaluación de la calidad que reconozca la validez y las limitaciones de los distintos métodos sería la más recomendable (Kuh, 1981). Sin embargo, de entre las múltiples facetas de la calidad, sólo unas pocas de ellas han sido objeto de estudio empírico, y preferentemente lo han sido las relacionadas con las medidas cuantitativas.

Existe, por tanto, una cierta preponderancia de estas medidas, a veces demasiado simplificadoras, subestimándose la gran variedad de factores no cuantificables que influyen sobre la calidad de la educación superior. El principal problema de una visión simple sobre la calidad es la de no dedicar la debida consideración a ideas o sugerencias que mejoren el proceso educativo, que, a pesar de su importancia esencial, no suelen estar incluidas en las valoraciones de carácter cuantitativo más usuales. Es necesario también prestar atención a las evaluaciones cualitativas sobre las instituciones universitarias, que, aunque sean indudablemente más complejas de analizar, proporcionan una visión más rica, que es complementaria de la de los estudios cuantitativos.

CAPITULO III. EVALUACION INSTITUCIONAL

The most important point is that those persons involved must be reasonably clear about what is meant before assessment techniques and procedures are developed. To do otherwise is almost to guarantee confusion, frustration and contention.

George, 1982: 49.

III.1. INTRODUCCION

En la mayoría de los países desarrollados existe en la actualidad un notable interés hacia la evaluación de las instituciones de educación superior. Dada la importancia decisiva que la investigación y la educación superior están adquiriendo para el desarrollo de los países, y dado también el gran volumen de gasto que esta educación representa, todos están de acuerdo en que la calidad y la eficiencia deben de ser garantizadas y controladas (Bauer, 1988). En esta situación la evaluación se ha convertido en una cuestión clave que parece aplicarse a todo tipo de situaciones y de sistemas, casi una especie de bálsamo de Fierabrás que puede curar todos los males del sistema universitario.

A pesar de que los intentos por evaluar las instituciones de educación superior no son nuevos, las restricciones económicas al sector y el cambio en la tendencia expansiva de años anteriores, han llevado a los investigadores a buscar técnicas para evaluar el rendimiento de las instituciones con el objetivo básico de usar los resultados para la mejora de la calidad y eficiencia del sistema (Lindsay, 1981). Aunque también es cierto que el creciente interés por la calidad, y, por tanto, por su evaluación, es una consecuencia de la necesidad de una educación superior más flexible, accesible y preparada para emprender nuevos caminos (Sizer, 1982). Por unos y otros motivos, la actual preocupación por la calidad en la educación superior ha ido generando en los últimos años, preferentemente en los países occidentales, toda una serie de sistemas de control de calidad de las instituciones universitarias, conocidas bajo la denominación general de evaluaciones institucionales. Estas evaluaciones se refieren, tanto a las unidades individuales, como a los sistemas universitarios

globales, y se han llevado a cabo, tanto mediante valoraciones individuales de cada institución, como estableciendo comparaciones entre ellas.

La evaluación institucional, como mecanismo de control de la calidad de uso generalizado, sólo tiene algo más de una década de antigüedad, y aunque nació en una época de restricciones económicas, el interés por ella sigue creciendo aunque haya cambiado el panorama económico general. Por otra parte, aunque la evaluación institucional ha provocado todo tipo de respuestas en el mundo académico, desde las más airadas hasta las más optimistas (Romney *et al.*, 1979), hoy ha empezado a formar parte del bagaje habitual del funcionamiento de las universidades en los países más preocupados por la calidad de su educación superior.

En este capítulo se revisarán algunas de las ideas básicas sobre la evaluación institucional, que servirán de preámbulo a las metodologías específicas de evaluación que se tratarán en los siguientes capítulos, y a la concretización de estas evaluaciones en las experiencias de diversos países que se verán en el capítulo VII.

III.2. EL CONCEPTO DE EVALUACION INSTITUCIONAL

La evaluación institucional se define típicamente como un intento de medir cómo se están cumpliendo las metas preestablecidas de una institución, en nuestro caso de una institución universitaria. Evidentemente, y como se desprende de la propia definición, la evaluación exige haber definido previamente y con claridad cuáles son las metas que persigue la institución. Esta definición de objetivos es en sí misma un grave problema, dado que existe menos acuerdo del aparente sobre las metas de las instituciones de educación superior, por lo que es frecuente encontrar que en una institución no se han definido con precisión cuáles son las metas buscadas (Billing, 1986).

Los esfuerzos evaluadores han surgido tanto desde el Estado, como desde la sociedad o de las propias instituciones. La consecuencia de esta múltiple actividad evaluadora es que se está creando en todos los países desarrollados una situación que ha sido denominada del *estado evaluativo* que «... intenta insertar una forma particular de *ética competitiva*, definida externamente como la primera fuerza directriz del desarrollo de la educación superior» (Neave, 1988). En cierto modo, esta predisposición hacia la evaluación es consecuencia de un cambio en la perspectiva bajo la que se considera la educación superior, habiéndose pasado de una visión burocrática del sistema a una visión más autónoma y orientada a métodos de mercado (Hüfner y Rau, 1987), lo que

implica no sólo una mayor y mejor evaluación de las instituciones, sino también una perspectiva diferente que suponga, más que una mera verificación del funcionamiento del sistema, un método para mejorarlo: «Libertad implica responsabilidad», afirma Bauer (1988), refiriéndose al incremento de independencia de las universidades suecas, pero que es aplicable a una situación generalizada en los países occidentales que disponían de sistemas universitarios fuertemente centralizados, pero que ahora están empezando a cambiar en el sentido de una mayor autonomía: Suecia, Holanda, España, Alemania, son algunos ejemplos de esta situación. (No conviene olvidar que una clara excepción a esta situación generalizada en Europa es la del Reino Unido, que justamente se mueve en una dirección opuesta: de una situación de gran autonomía de sus universidades se está pasando a otra de mayor control estatal.)

Se pueden encontrar dos enfoques básicos para las evaluaciones institucionales: aquéllos que centran su atención en las magnitudes contables y de control, y las que enfatizan los aspectos organizativos del proceso de aprendizaje (Premfors, 1986). Para algunos ha existido una tendencia general a desarrollar los primeros, en conexión con valoraciones de la eficiencia y de los factores económicos, en detrimento de otros aspectos como la efectividad y el desarrollo profesional y educativo (Pollit, 1987). Mientras que los partidarios del primer enfoque acusan a los partidarios del segundo de no satisfacer las necesidades informativas de una adecuada gestión (Lindsay, 1982), los segundos afirman que una visión contable de la educación superior deja de lado aspectos esenciales del proceso de las instituciones. Aunque estos dos enfoques básicos pueden darse combinados, existe obviamente una cierta tensión entre ellos.

Mientras que para algunos la finalidad de la evaluación es básicamente la de mejorar la calidad de las instituciones, para otros la evaluación tiene como fin esencial ayudar a facilitar la gestión de esas mismas instituciones. Así, por ejemplo in't Veld (1990) señala como objetivos de la evaluación los siguientes:

- Ser una herramienta para la gestión de la institución.
- Ser un apoyo para la toma de decisiones.
- Ser un instrumento para la gestión del personal de las universidades.

Respecto a las distintas situaciones en las que se pueden desarrollar las evaluaciones, Frackmann y Maassen (1987) señalan la existencia de tres posibilidades. Para estos autores, según sea la posición de dependencia de la institución, la evaluación institucional puede ser:

- *Regulada por el Estado*. El Estado es responsable de todas las normas

financieras, educativas, de personal y las referentes a los estudiantes. Apenas existe diferencia entre que se trate de un gobierno central o de uno regional.

— *Regulada a través de un organismo «amortiguador».* Ciertas normas importantes son adoptadas por un organismo intermedio entre el Gobierno y las instituciones, en el que participan ambas partes. Estos cuerpos pueden verse como el brazo del Gobierno o como un modo voluntario de coordinación por parte de las universidades. Un claro ejemplo de este tipo de organismos es el Consejo de Universidades español o la Asociación de Universidades de los Países Bajos. Pero, a diferencia del Consejo de Universidades español, fuertemente relacionado con el Ministerio de Educación, el organismo holandés es absolutamente independiente de las autoridades gubernamentales.

— *Autorreguladas.* Es el caso que se da para instituciones autónomas, en las que la mayor parte de las decisiones son tomadas por las propias universidades. Así suele suceder en la mayor parte de las universidades norteamericanas, tanto públicas como privadas.

Como se verá posteriormente con más detalle, la evaluación institucional puede ser de dos tipos básicos según sea el agente que la realiza: evaluación interna y evaluación externa. La primera es realizada por las propias instituciones como parte de sus responsabilidades por mejorar la calidad de la educación, mientras que la segunda es llevada a cabo por agentes externos a las universidades, habitualmente por mandato de organismos estatales o públicos, o fruto del encargo a expertos en evaluación por la propia universidad que desea ser evaluada (Weert, 1990). En ciertos casos se utiliza una combinación de ambos criterios como método para realizar una más completa evaluación institucional. Existe un importante debate que discute los pros y contras de cada uno de los dos enfoques (Premfors, 1986; Maassen, 1987; Weert, 1990), que en el fondo representa la lucha entre los partidarios del control externo (básicamente estatal) y los del control interno ejercido por la propia institución como organismo autónomo.

III.3. FACTORES A CONSIDERAR EN LA EVALUACION

Realizar una evaluación institucional supone fijar la atención del evaluador sobre aquellas cuestiones que se consideren básicas para la consecución de los objetivos de la institución. Algunos de los aspectos básicos que toda evaluación debe considerar se exponen a continuación, juntamente con algunos de los puntos concretos que deben ser objeto de valoración en la evaluación.

III.3.1. Aspectos a evaluar

En una evaluación institucional se deberán elegir previamente aquellos aspectos de las instituciones que hay que tomar en consideración por su supuesta relación con la calidad. Típicamente son éstos (Miller, 1980):

A. Resultados de los estudiantes

La postura más ampliamente aceptada respecto a los *outputs* del sistema universitario es la del *valor añadido*. Este punto de vista considera todos los efectos que la institución universitaria ejerce sobre el alumno no sólo sobre su formación profesional, sino también sobre otros aspectos culturales. Bowen (1979) enumera siete principios básicos que deben tenerse en cuenta para evaluar los resultados del proceso educativo en las universidades. Son éstos:

— Debe evitarse la habitual confusión entre *inputs* y *outputs*. La única medida correcta es la que valora cómo cambian las personas a través de la experiencia universitaria.

— La evaluación debe hacerse en relación con los objetivos básicos de la educación, y no reducirse a aquellos aspectos del desarrollo humano que pueden medirse más fácilmente o que están relacionados con el éxito económico.

— Al tener en cuenta los resultados del proceso educativo debe evaluarse el desarrollo integral de las personas como un todo, y no sólo las características más estrictamente académicas.

— La valoración debe hacerse basándose tanto en los estudiantes actuales de la institución como en los antiguos alumnos.

— La valoración debe basarse estrictamente en el incremento en formación de los estudiantes, y no en el nivel absoluto del rendimiento del alumnado.

— La evaluación debe ser razonablemente factible, tanto en cuanto al tiempo necesario para llevarla a cabo como en los costes involucrados en su realización. Igualmente, la evaluación debe concentrarse en los objetivos más importantes.

— La evaluación debe partir de la propia base de las universidades, y no ser impuesta desde arriba por ningún organismo extraacadémico.

B. *Gerencia*

La administración y gerencia de la educación superior es una pieza básica para el funcionamiento de las instituciones universitarias. Sin embargo, las universidades, como otras instituciones que no buscan el beneficio económico, tienen una tendencia a ser poco efectivas y eficientes. La evaluación administrativa debe formar parte de la evaluación institucional bajo los mismos supuestos que la evaluación del profesorado: valorar la competencia y mejorar los rendimientos. Miller (1980) recomienda las siguientes líneas maestras para la evaluación de los aspectos administrativos de las instituciones:

- Basarse en las características de cada institución.
- Establecer un propósito general de carácter positivo, que persiga como objetivo básico la mejora del sistema.
- Evaluar el rendimiento de las personas en cada empleo según las características y expectativas del puesto, que son razonablemente específicas y diferenciables.
- Emplear en la evaluación tanto medidas objetivas como subjetivas.
- Recoger información de quienes están capacitados en cada caso, debido a su posición de responsabilidad en el sistema.
- Evaluar en colaboración con los individuos que están siendo evaluados.
- Mantener la confidencialidad a lo largo de todo el proceso evaluador.

C. *Percepciones de los interesados en el sistema universitario*

La evaluación institucional debe recurrir a las opiniones de aquellas personas interesadas en el proceso educativo: padres, antiguos alumnos, organizaciones públicas y representantes del mundo empresarial e industrial. De este modo se conseguirá una visión mucho más amplia sobre la institución objeto de evaluación.

D. *Impacto en la comunidad*

La influencia de una universidad sobre la comunidad en la que está instalada, debe ser también objeto de consideración en una evaluación institucional. A pesar de la dificultad de evaluación, varios mecanismos de

medida han sido diseñados para este fin, especialmente en los Estados Unidos, en donde el impacto de la institución sobre el entorno ha sido considerado desde siempre como muy importante.

E. *Indicadores financieros*

Aunque evidentemente estos indicadores son necesarios e importantes para evaluar una institución, frecuentemente son supervalorados, y se tiende a considerar los recursos económicos de una institución como sinónimo de su nivel de calidad. Establecer un paralelismo entre recursos y calidad es algo que está lejos de estar demostrado; aunque es evidente que un buen nivel de calidad exige un cierto nivel de recursos económicos, no es cierto que mayores niveles de gasto produzcan necesaria y linealmente mejoras de la calidad.

III.3.2. **Elementos a evaluar**

Desde un punto de vista más concreto, Miller (1979) establece 45 elementos de las instituciones que deben ser objeto de revisión en toda evaluación institucional. Estos elementos concretizan los aspectos institucionales citados en el apartado anterior. Para Miller es necesario comprobar que en las instituciones evaluadas se cumplen los siguientes requisitos:

A. *Metas y objetivos*

1. Las metas establecidas sirven como una guía efectiva para el presente y el futuro de la institución.
2. Los objetivos sirven como refuerzo de las metas.
3. La política y los procedimientos de admisión de alumnos son consistentes con las metas y objetivos de la institución.
4. La institución tiene una adecuada capacidad de planificación.
5. Las metas y objetivos de la institución tienen una razonable identidad con las de otras instituciones del sistema universitario.

B. *Aprendizaje de los estudiantes*

6. Los estudiantes valoran positivamente el sistema de asesoría y tutoría de que dispone la institución.

7. Las tasas de retención de alumnos son razonables.

8. Existe disponibilidad de un conjunto de recursos compensatorios de aprendizaje individualizado para los alumnos que lo necesitan.

9. La administración para asuntos de los estudiantes es efectiva.

10. Existe un progreso satisfactorio hacia las metas educativas.

C. *Rendimiento del profesorado*

11. Existe una política y unos procedimientos satisfactorios para evaluar el profesorado.

12. Hay programas en marcha para la mejora del profesorado.

13. La política de personal respecto al profesorado es adecuada.

14. Los salarios y los beneficios marginales de los profesores son competitivos.

15. La calidad general del rendimiento del profesorado es óptima.

D. *Programas académicos*

16. La institución tiene políticas efectivas para desarrollar nuevos programas.

17. La institución tiene métodos de revisión efectivos para evaluar los programas existentes.

18. La educación general de carácter formativo que se proporciona a los estudiantes es estimulante, y forma parte esencial del currículum.

19. La calidad y el tamaño del programa de graduados es consistente con las metas y objetivos institucionales.

20. La biblioteca y otros recursos de enseñanza sirven eficazmente a la comunidad académica.

E. *Servicios*

21. Los edificios y los servicios son adecuados para el número de estudiantes y para la naturaleza de los programas.

22. La institución tiene un plan a largo plazo para el desarrollo y mantenimiento de los recursos físicos.

23. Los salarios y otros beneficios del personal de administración y servicios son suficientes para atraer y retener a personas competentes.

24. Se realizan sistemáticamente evaluaciones del personal de administración y servicios.

F. *Gerencia*

25. La gerencia dedica la atención adecuada a la planificación.

26. Los gerentes de mayor rango tienen relaciones de trabajo efectivas con los administradores de otras instituciones.

27. La política general de la institución permite una gerencia efectiva.

28. Existen unos mecanismos satisfactorios de evaluación administrativa y de desarrollo profesional.

29. La institución tiene un programa de acción positiva efectivo.

G. *Aspectos financieros*

30. Las tasas de matrícula son adecuadas a las necesidades de la institución y a las posibilidades de los alumnos.

31. La institución tiene un sistema eficiente de contabilidad.

32. Los costes y los gastos son compatibles con las posibilidades de la institución.

33. Las inversiones están bien administradas.

34. La institución tiene un sistema efectivo para mostrar su contabilidad.

H. *Consejo de gobierno*

35. La política para dirigir el consejo de gobierno es satisfactoria.

36. Los regentes de la institución comprenden las diferencias entre la formulación de la política institucional y su implementación, y aplican este conocimiento.

37. El consejo de gobierno trabaja con efectividad con otros organismos en el entorno de la institución.

38. El consejo de gobierno contribuye efectivamente a mejorar la situación de la institución.

I. *Relaciones externas*

39. Las actividades de la institución ayudan a mejorar la calidad de vida de la población circundante.

40. La institución tiene una relación efectiva con las autoridades estatales (regionales, en nuestro caso) de educación superior.

41. La institución tiene una relación efectiva con las autoridades nacionales de educación superior.

42. La institución es capaz de conseguir una cantidad aceptable de fondos de empresas privadas.

J. *Preocupación por la mejora*

43. La institución busca el mejorar por medio de la innovación y la experimentación.

44. Los distintos grupos que configuran la institución tienen una actitud positiva hacia las mejoras.

45. La institución tiene establecidos mecanismos para evaluar su efectividad.

III.4. **DISTINTOS TIPOS DE EVALUACION**

Clasificar es algo complejo, básicamente por lo que puede tener de simplificador. Sin embargo, se va a intentar aquí, en orden a dar una idea general de los diversos métodos de evaluar las instituciones de educación superior.

En principio se pueden clasificar los métodos para la evaluación de la educación superior en dos categorías, según sea el objeto básico de la evaluación: la que considera programas individuales (habitualmente denomi-

nada revisión de programas) y la que considera las instituciones como un todo (la evaluación institucional propiamente dicha).

Por otra parte, las evaluaciones pueden ser clasificadas bajo dos tipologías según quien las realice: evaluaciones internas y externas. La diferencia esencial entre ellas radica en quién es el agente que lleva a cabo la evaluación: los propios miembros de la institución, en el primer caso; personas de una agencia especializada o de una comisión exterior, en el segundo. Actualmente son frecuentes las evaluaciones mixtas en las que se unen ambas perspectivas, de modo que la evaluación de una institución es el resultado combinado de un autoestudio y de una evaluación externa.

Desde otro punto de vista se podrían clasificar las evaluaciones institucionales según el fin perseguido por la evaluación: evaluaciones para el control de la institución por parte de instituciones gubernamentales; evaluaciones para conocer la situación real de la institución con el objetivo de mejorar la calidad, promovidas por la propia institución o por otras instancias, gubernamentales incluidas.

Puede hacerse también otra distinción importante entre las evaluaciones según sea el grado de obligación con el que la institución debe responder a los resultados de la evaluación. Las evaluaciones pueden ser imperativas cuando la institución tiene que seguir fielmente los resultados de la evaluación; esto es algo que suele suceder con una gran parte de las evaluaciones gubernamentales. En otros casos las evaluaciones tienen como objetivo recomendar unas normas, y no el imponer unos criterios: a éstas se les denomina evaluaciones normativas. Por último, están aquellas evaluaciones que, habitualmente solicitadas por la propia institución o incluso realizadas por ella misma, tienen como objetivo básico el conocer la propia situación y el encontrar soluciones que puedan ayudar a mejorarla: a estas evaluaciones se les denomina consultivas (Kogan, 1986; Frackmann y Maassen, 1987).

Aún se pueden encontrar otros tipos de evaluación. Por ejemplo, unas evaluaciones están enfocadas en la mejora de la efectividad institucional (grado de aproximación a los objetivos), mientras que otras se preocupan principalmente por la eficiencia (lo bien que se están aprovechando los recursos disponibles). Desde otros puntos de vista se puede clasificar una evaluación según sea cuantitativa o cualitativa, formativa (concentrada en el proceso) frente a sumarias (en los resultados), interactiva o no, a nivel estatal, local o institucional, etc. (Premfors, 1986).

Utilizando criterios estructurales, Escotet (1990) clasifica los métodos de evaluación institucional en cuatro modelos según los elementos básicos que se

tomen en consideración para la evaluación: modelo de entradas-operaciones; de operaciones-productos; de entradas-productos; y entradas-operaciones-productos.

En el cuadro III.1 se esquematizan las más notables clasificaciones de la evaluación institucional que se han citado aquí.

CUADRO III.1

Tipos de evaluación institucional

Según el agente	Evaluación externa Evaluación interna Evaluación mixta
Según el objetivo	Control burocrático Mejora de la calidad
Según la obligatoriedad	Evaluación imperativa Evaluación normativa Evaluación consultiva
Según el enfoque	Evaluaciones de eficiencia Evaluaciones de efectividad
Según el objeto	Evaluaciones formativas Evaluaciones sumarias

Aunque en la práctica evaluadora de las instituciones se dan diversos grados de combinaciones entre todos los tipos de evaluaciones citadas, hay cuatro variantes de evaluación que son muy utilizadas:

Evaluaciones externas

El mecanismo habitual para este tipo de evaluación es el de la utilización de una agencia u organismo especializado en técnicas de evaluación. El caso más elemental sería la típica auditoría usada en las empresas y que ha sido utilizada en ciertas universidades. Sin embargo, las más frecuentes y de interés son las evaluaciones específicas para las instituciones universitarias, consistentes en la utilización de una comisión de expertos de otras universidades nombrada *ex profeso* para realizar la tarea evaluadora.

El uso de revisiones externas tiene la ventaja de la imparcialidad y de la utilización de expertos que son competentes en técnicas de evaluación. Sin

embargo, el análisis de un organismo tan complejo como es una universidad, es difícilmente realizable desde fuera, de modo que, como dice Miller (1980), los expertos pueden dejar al acabar la evaluación más problemas que soluciones, debido a su falta de comprensión del funcionamiento de la institución como un todo y a las posibles suspicacias levantadas entre el personal. Por otra parte, una indudable ventaja de este tipo de evaluaciones es que pueden ser iniciadas y realizadas más rápidamente, ya que normalmente se utiliza personal familiarizado con las técnicas de evaluación (Premfors, 1986). Igualmente la flexibilidad y coherencia de las evaluaciones externas son encomiadas por Williams (1986) como el medio menos arbitrario de evaluar y de asignar fondos a las instituciones; para este autor, las revisiones realizadas por expertos son la mejor manera de evaluar la calidad de las instituciones, porque un sistema complejo es evaluado más correctamente por conocedores del sistema, que por una serie de medidas objetivas que no recogen toda la complejidad de una institución universitaria.

Evaluaciones internas

La evaluación interna es aquélla que es promovida y realizada por las propias instituciones con el objetivo de mejorar la calidad de la institución. Hay que señalar que este tipo de evaluaciones puede incluir, y de hecho lo hace habitualmente, la utilización de evaluadores ajenos a la propia universidad, junto con otros de la institución que está siendo evaluada. Bajo la forma de autoestudios ha sido ampliamente utilizada, especialmente en los Estados Unidos, ya que forma parte del proceso de acreditación de las instituciones de educación superior de aquel país. (El método de la acreditación en los Estados Unidos, junto con los autoestudios, será tratado con detalle en el capítulo V.) Independientemente del proceso de acreditación, también se están realizando autoestudios en numerosas universidades de todo el mundo, de modo que este mecanismo se ha convertido en los últimos tiempos en el método más generalizado de evaluar las instituciones bajo la perspectiva de la mejora de su calidad. Debido a su uso generalizado, se han desarrollado también una serie de métodos para realizar los autoestudios, promovidos por diversas universidades y otros organismos relacionados con la educación superior.

Evaluaciones gubernamentales

Otro tipo de evaluación externa es la que promueven organizaciones gubernamentales y que ejercen sobre las instituciones de carácter público o, en algunos países, también sobre las de carácter privado, con el objeto de otorgarles los permisos docentes necesarios, para asignar recursos, o para ejercer una función controladora sobre el rendimiento de las instituciones.

Uno de los problemas de este tipo de evaluación es que existe una acusada tendencia a trasladar métodos evaluativos desarrollados para industrias o para otros servicios públicos, pero que tienen un valor limitado en la evaluación de instituciones de educación superior (Miller, 1980; Kogan, 1986).

Es evidente que la sociedad tiene el derecho de conocer cómo está funcionando la educación que ella apoya y financia, y en un Estado democrático, la evaluación necesaria para ello puede ser asumida legítimamente por las autoridades centrales o regionales. En la mayoría de los países desarrollados se han incrementado en los últimos veinte años este tipo de evaluaciones como consecuencia del considerable aumento del gasto público en la educación superior. En consecuencia, existe una cierta tendencia a desarrollar evaluaciones externas de control. Este hecho se ha manifestado más claramente en los países anglosajones, en los que tradicionalmente había existido un gran nivel de libertad y autonomía en las universidades públicas. Incluso en los Estados Unidos ha tenido también lugar un importante incremento en los estudios evaluativos promovidos por los estados que poseen sistemas estatales (Boyer *et al.*, 1987). La situación en los Estados Unidos durante los últimos tiempos es descrita claramente por Johnson (1988: 107): «Durante los años sesenta muchas universidades habían experimentado crecimiento en tamaño, recursos y autonomía. Durante la siguiente década el crecimiento se estabilizó, los costes continuaron subiendo y la autonomía del profesorado fue atacada. El arma de ese ataque por parte de los ministerios fue el de desarrollar fórmulas para sustentar las universidades, revisar los programas, aprobar los grados, controlar los puestos de profesorado, establecer políticas uniformes en campus diferentes, separar presupuestos de investigación de los de docencia (y éstos, a su vez, separados en presupuestos para formación de graduados y de pregraduados), unificar sistemas de datos, e iniciar estudios de coste-efectividad. Hoy las autoridades educativas y las universidades están buscando modos de incrementar los resultados de la educación superior sin devolver la autonomía al profesorado, mientras que éstos desconfían de cualquier tipo de revisión de su trabajo por miedo a perder más autonomía». Curiosamente en otros países, que han gozado de poca o ninguna autonomía en el pasado, se está dando la tendencia en sentido inverso, pasándose de un fuerte control estatal a mayores niveles de autonomía, en los que es mucho más propicia la autoevaluación por las propias instituciones autónomas (Kogan, 1986).

Evaluaciones controladoras

En un sistema integrado de control de calidad de la institución universitaria, parece razonable utilizar una combinación de evaluaciones internas y externas. Sin embargo, los sistemas de evaluación utilizados por los gobiernos, especial-

mente algunos europeos, son frecuentemente criticados porque le dan primacía a una evaluación externa, con frecuencia de carácter simplificador, llevada a cabo utilizando métodos autoritarios (Weert, 1990). En estas evaluaciones los objetivos son más el control de las instituciones universitarias, que la mejora de la calidad de su funcionamiento y de sus resultados.

Cuando las evaluaciones son simplificadoras, como suele ser el caso en las evaluaciones de control, se tiende a utilizar exclusivamente medidas o indicadores fácilmente obtenibles, con lo que se distorsionan las metas últimas del proceso educativo, que difícilmente están bien representadas por esos indicadores. Hasta en sistemas poco regulados, como son los norteamericanos, se da esta situación: centrar la medida de la calidad en los aspectos más medibles y no en los que realmente son importantes en el proceso educativo (Boyer, 1987). Por otro lado, estos indicadores suelen ser exclusivamente datos cuantitativos, que no suelen ser excesivamente fiables (Weert, 1990), son de interpretación ambivalente e incluso pueden ser confusos en cuestiones tan básicas como la diferencia entre medidas de *inputs* y de *outputs* (Ball y Halwachi, 1987).

III.5. LA PUESTA EN PRACTICA DE LA EVALUACION

La experiencia de la puesta en práctica de la evaluación en instituciones universitarias por todo el mundo permite conocer algunos de los criterios, de las recomendaciones y de los problemas que los expertos en la evaluación institucional han encontrado al aplicar estas técnicas. Algunas de sus conclusiones se exponen en esta sección.

III.5.1. Criterios básicos de la evaluación

La idea de evaluación lleva implícita la de una preselección de criterios para ser valorados (Holdaway, 1986). Por tanto, decidida la realización de una evaluación institucional, hay una serie de preguntas previas que hay que contestar, de tal modo que se puedan elegir los mecanismos más adecuados de acuerdo con las respuestas que se hayan dado a esos interrogantes. Kogan (1986) recomienda plantearse estas 11 cuestiones previamente al inicio de una evaluación institucional:

1. ¿Quién establece los criterios para la evaluación?
2. ¿Cuáles son los criterios de la evaluación?
3. ¿Qué se pretende mejorar con la evaluación?

4. ¿Cuál es el enfoque: entradas, proceso, salidas, impacto, etcétera?
5. ¿Cuáles son las etapas de la evaluación?
6. ¿Existen controles formales?
7. ¿Qué fuentes de información y de datos van a ser usados?
8. ¿Se usan medidas cuantitativas para evaluar los cambios?
9. ¿Existen hipótesis predefinidas?
10. ¿Se busca que los resultados sean objeto de discusión?
11. ¿Cuál es el tipo de evaluación deseada?

Contestadas estas preguntas, el evaluador puede elegir entre diferentes técnicas evaluativas y entre los diferentes factores que va a considerar, según sean los objetivos básicos de la evaluación que se propone llevar a cabo.

Aunque las técnicas de la evaluación pueden ser construidas *ad hoc*, y sobre ello se tratará más adelante, existen algunos modelos ya experimentados. Brigder (1989) cita los siguientes que se están aplicando en las instituciones universitarias norteamericanas:

— El *ACT Evaluation/Survey Service* que proporciona la *American College Testing Program (ACT)*.

— El *Student Outcomes Information Service (SOIS)* realizado por el *National Center for Higher Education Management System (NCHEMS)*.

— El *Educational Testing Service* también ofrece diversos paquetes para distintos tipos de unidades a ser evaluadas.

Existen también otras propuestas adecuadas para las universidades europeas, como es el caso del sistema de evaluación programado por el *Center for Higher Education Policy Studies (CHEPS, 1990)*.

III.5.2. **Las líneas maestras para evaluar**

Una evaluación institucional debe hacerse siguiendo algunas premisas que garanticen el éxito de la evaluación o, al menos, unos resultados aceptables. Siguiendo a Miller (1986), se pueden enumerar algunas de las líneas maestras que es necesario tener en cuenta antes de emprender una evaluación institucional. Son éstas:

— Los planes teóricos y operativos para la evaluación de la calidad de una institución universitaria deben de ser compatibles con la idiosincrasia de cada

institución. Es decir, debe usarse la máxima «*adaptar, no adoptar*». Aunque se pierda eficacia al reformar el plan, la experiencia demuestra (donde han sido puestos en práctica tales métodos) que es mejor la planificación *a medida*.

— Es necesario un liderazgo institucional enérgico para que sea posible mejorar la calidad de la institución. Este liderazgo debe dar el ímpetu inicial, dirigir los esfuerzos y trasladar las recomendaciones a la práctica. Una buena dirección exige además respeto y conocimiento de los dos constituyentes básicos de la institución: profesorado y estudiantes.

— Un plan general para mejorar la calidad de la institución universitaria debe ser participativo en todos sus aspectos. Es decir, los miembros de la comunidad académica deben tener amplia oportunidad de conocer los esfuerzos evaluadores, y algunos de ellos deben participar en tareas concretas; asimismo, deben darse explicaciones públicas del proceso que se está siguiendo. Todas las personas afectadas por el proceso deben ser conocedoras de los pasos que se dan, y deben poder estar involucradas en las distintas fases del proceso.

— Los planes de evaluación institucional deben estar relacionados con la asignación o con la reasignación de recursos.

— Las evaluaciones institucionales deben usar datos objetivos cuando están disponibles, pero no se deben descartar los subjetivos cuando se crea que son convenientes.

— Los modelos para evaluar la calidad deben contemplar la posibilidad verosímil de cometer errores.

— Los estudios evaluadores deben estar orientados hacia la acción. Deben incluir toda una serie de recomendaciones para tomar medidas que mejoren la calidad de la institución.

— La evaluación debe incluir un plan para evaluar la propia evaluación.

La evaluación institucional puede tener en algunos casos la función de asignar fondos a las distintas universidades de un sistema estatal. En estos casos algunas precauciones suplementarias deben tenerse en cuenta. Frackmann (1987) señala éstas:

— Dejar que la asignación de recursos la realicen los miembros de la propia comunidad académica por medio de algún tipo de institución *amortiguadora*.

— Usar sólo valores normalizados, y no valores ideales que podrían permitir que las instituciones se salieran del proceso evaluador.

— Confiar lo más posible la asignación de recursos a la utilización de criterios normalizados.

— Dar a las instituciones tiempo para remediar sus puntos débiles antes de proceder a variar las asignaciones de recursos.

— No poner nunca en relación las evaluaciones de instituciones que tienen un funcionamiento y estructura reguladas obligatoriamente por autoridades externas, con la asignación de recursos en función de esas evaluaciones.

III.5.3. **Algunos problemas de la evaluación**

En la puesta en práctica de una evaluación surgen lógicamente multitud de problemas. Algunos de los que con más asiduidad aparecen son los señalados por Miller (1986):

— *Las metas y objetivos mal definidos* pueden obstaculizar seriamente los esfuerzos evaluadores. Algunos de los fallos más usuales a este respecto son éstos:

— Los objetivos de la evaluación no están relacionados con los objetivos generales de la institución.

— Los objetivos no están demasiado relacionados con las actividades que pueden realizarse para conseguirlos.

— Los objetivos son muy poco ambiciosos o demasiado utópicos.

— Se da más importancia a las actividades que a los resultados.

— Tienen demasiada palabrería o «lenguaje académico», o se centran en cuestiones poco importantes.

— Están escritas más para impresionar que para ser implementadas.

— *Las diferencias de actitud y mentalidad* entre administradores y académicos puede ser una causa de dificultades a la hora de poner en práctica los resultados de la evaluación.

— *Sobrevalorar una orientación o un aspecto de la evaluación* como una panacea para todos los problemas puede ser una fuente de desencanto y de fracaso de la evaluación.

— *Hacer o planear demasiados cambios, demasiado deprisa*, puede ser contraproducente. La institución universitaria se mueve lentamente y no hay que ser impaciente.

— *No prever los problemas* puede ser causa de euforia a corto plazo y de desinterés a largo plazo. Varios de los problemas que es necesario prever, para evitar el fracaso de los intentos evaluadores, son los siguientes:

— Los planes pueden estar mal hechos.

— Los planes no están suficientemente apoyados desde la dirección de la institución.

— Los planes son vistos como un asunto privado de los planificadores.

— Los planes para desarrollar la calidad no tienen en cuenta todas las necesidades administrativas que serán necesarias para desarrollar, mantener, evaluar y modificar los planes.

— Los planes realmente están siendo usados como cortinas de humo para esconder los verdaderos problemas.

— Un motivo de fracaso es el diluir las responsabilidades, de modo que no exista una persona encargada de impulsar cada una de las partes del plan.

— La falta de atención sostenida hacia las metas finales puede ser otra causa de fracaso. Es necesario pensar estratégicamente para llegar a buenos resultados.

Por otra parte, Romney *et al.* (1979), agrupan las precauciones y limitaciones que hay que tener en cuenta antes de realizar una evaluación de las instituciones, en estos cuatro apartados:

— *Precauciones políticas*

Hay que considerar los posibles efectos negativos que la evaluación puede tener sobre la marcha general de la institución, sobre las relaciones de poder, los compromisos y el consenso, además de sobre las relaciones de la

institución con los organismos estatales. Más específicamente, se pueden citar los posibles efectos negativos siguientes:

— La obsesión por la evaluación cuando las cosas funcionan bien, puede originar disonancias entre los grupos y deshacer el adecuado equilibrio entre productividad, satisfacción personal y los esfuerzos positivos dentro de la institución.

— Se puede destruir la estructura de incentivos que se ha generado con el tiempo, y que puede ser positiva.

— Puede aumentar innecesariamente la burocracia, la rigidez y la inflexibilidad del sistema.

— Se pueden destruir peculiaridades absolutamente legítimas de cada institución.

— Un escaso nivel de aceptación de los resultados por la comunidad universitaria puede tener implicaciones políticas negativas importantes.

— Se pueden provocar expectativas poco realistas después de una evaluación, ya que quedan por hacer las cosas que se recomiendan en los resultados.

— Se produce una reducción de las prerrogativas a las que están habituados los responsables de la institución, ya que se han de someter a los resultados de la evaluación, por lo que pueden tender a interferir la evaluación.

— Se pueden provocar intromisiones externas al preocupar a toda la comunidad con los problemas internos de la institución.

— *Precauciones metodológicas*

Algunas de las más importantes dificultades de las evaluaciones de carácter metodológico son las siguientes:

— Aplicaciones erróneas debidas a la utilización de criterios inapropiados, o a la trasposición de métodos aplicados en otros centros, pero que no se han adaptado debidamente.

— Limitaciones que existen para la medida de algunas de las más interesantes e importantes variables del proceso y del resultado educativo.

— Dificultades para evaluar actividades típicas de la educación superior que generan a la vez diversos beneficios importantes.

— Problemas en la elección del nivel que debe ser evaluado, ya que en la agregación se pierde capacidad de discriminación.

— Considerar, antes de iniciarla, la oportunidad de realizar la evaluación, ya que existe la posibilidad de que no se trate de un momento oportuno para los intereses de la institución.

— Considerar la posibilidad de que el personal de la institución no esté capacitado para adaptarse a los resultados de un programa de evaluación.

— *Limitaciones económicas*

Algunas limitaciones de carácter económico para la evaluación pueden ser las siguientes:

— Los beneficios económicos de la evaluación pueden ser sólo marginales, ya que el sistema es demasiado rígido y las consecuencias de la evaluación son económicamente insignificantes, por lo que la evaluación no compensa los costes.

— Los costes de la evaluación pueden ser demasiado elevados, justamente cuando los recursos son escasos y pueden existir otras necesidades más perentorias.

— *Precauciones generales*

Dos importantes consideraciones de carácter general deben hacerse ante una evaluación:

— La autonomía y la flexibilidad de las instituciones universitarias se verán de algún modo menoscabadas si se inician evaluaciones de rendimiento, que en cierto modo no tienen ya marcha atrás.

— Es necesario preguntarse antes de iniciar una evaluación institucional si los resultados van a afectar realmente a las cuestiones en las que la propia institución tiene realmente algún control, como puede ser la enseñanza u otros temas semejantes.

CAPITULO IV. RENDIMIENTO INSITUCIONAL E INDICADORES DE RENDIMIENTO

All over the country these groups of scholars, who would not make a decision about the shape of a leaf or the derivation of a word or the author of a manuscript without painstakingly assembling the evidence, make decisions about the admissions policy, size of universities, staff-student ratios, content of courses, and similar issues, based on dubious assumptions, scrappy data and mere hunch... although dedicated to the pursuit of knowledge, they have until recently resolutely declined to pursue knowledge about themselves.

Eric Ashby, 1963: 6.

IV.1. RENDIMIENTO INSTITUCIONAL

De entre las diferentes maneras de enfocar el problema de la evaluación de las instituciones universitarias, se recogen bajo la denominación de evaluación del rendimiento institucional aquellos aspectos relacionados básicamente con la medida, en la mayoría de los casos con carácter cuantitativo, de la relación existente entre los objetivos que se proponen las instituciones, los resultados reales que se obtienen y las aportaciones de todo tipo que son necesarias para el proceso educativo.

En este capítulo se pasa revista a las principales ideas subyacentes al concepto de evaluación del rendimiento institucional: en primer lugar a los conceptos generales de rendimiento, efectividad, eficiencia, etcétera, y a los modos de evaluarlos; a continuación, se habla de los indicadores de rendimiento, que constituyen una herramienta esencial de los intentos que se han realizado para evaluar el rendimiento de las universidades.

IV.1.1. Introducción

A partir de la segunda mitad de los años setenta se despertó en todos los países desarrollados una notable preocupación respecto al rendimiento de las

universidades. Las causas fueron básicamente de tipo económico: elevados costes, gran cantidad de alumnos y una situación general de crisis económica, cuya consecuencia más directa para las universidades fue que se generara un clima de restricciones económicas, con una tendencia a la limitación de sus fondos y al mayor control de los recursos. Además, en esos momentos se preveía una próxima disminución del número de alumnos como consecuencia de la disminución del tamaño de las cohortes de jóvenes alcanzando la edad de ingreso en las universidades. La adaptación de las instituciones a una nueva situación restrictiva se veía como inminente, y había que prepararse adoptando medidas de tipo gerencial que elevaran el rendimiento y disminuyeran los costes. Sizer (1982) resume así la situación que se presentaba, mucho más acuciadamente en las universidades británicas: «Dentro de las instituciones es necesario compensar la presión por la eficiencia a corto plazo, que genera el aumento de los costes y las restricciones en las admisiones de estudiantes, con las acciones que deben ser tomadas por las instituciones para ser efectivas a largo plazo» (Sizer, 1982: 34).

En la situación descrita por Sizer no sólo intervinieron factores económicos, sino también un cambio de criterio de los gobiernos de algunos países hacia la autonomía de las universidades. La restricción de la autonomía universitaria en los últimos tiempos es una situación que se ha dado especialmente en el Reino Unido; este estado de cosas es descrito así por Kogan (1986: 125): «El giro hacia la adopción de mecanismos gerenciales está fuertemente reforzado por aquellos gobiernos que no pueden aceptar la idea de que existan instituciones públicas con libertad para crear sus propias misiones y estilo, al menos a expensas del gasto público, lo que hace que los gerentes de las universidades estén ansiosos de encontrar bases racionales para justificar sus afirmaciones». Afortunadamente no es este el caso de España, en donde partiendo de una situación de absoluto control gubernamental sobre las universidades, la única dirección de movimiento posible era justamente en el sentido que se ha dado: el aumento de la autonomía de las instituciones.

Aquel estado generalizado de inquietud ha decaído en cierto modo en los últimos tiempos por dos razones esenciales: la primera es que la situación económica general cambió favorablemente en los pasados años; la segunda es que las previsiones que se hicieron sobre descenso del número de estudiantes fueron exageradas, de tal modo que en la realidad, salvo en algún país como el Reino Unido (a causa de drásticas reducciones presupuestarias en las universidades), la situación general ha sido mucho más estable de lo que se esperaba, tanto en cuanto al número de alumnos como a nivel financiero. Aunque en estos momentos los esfuerzos investigadores dentro del campo de las instituciones de educación superior se encaminan más hacia las cuestiones

de calidad institucional, los estudios sobre rendimiento forman parte de bagaje que se ha ido acumulando en los últimos años para evaluar las instituciones de educación superior, y a los que conviene dedicar la oportuna atención.

Antes de discutir más detenidamente sobre la evaluación del rendimiento de la educación superior es necesario señalar que, incluso por encima de las discusiones sobre la oportunidad y las distintas maneras de llevar a cabo la evaluación del rendimiento institucional, existe otra discusión más general todavía: el de si la educación superior, al igual que la ciencia o el arte, pueden o deben ser objeto de un tratamiento gerencial, y en el caso de que sea adecuado este tratamiento, si los gerentes de las instituciones universitarias deben o no ser personas ajenas al ámbito universitario. Unos piensan que los universitarios son unos individuos lo suficientemente motivados por la idea de calidad que no necesitan ser inducidos externamente; es decir, que los universitarios tienen la actitud responsable y la capacidad suficiente para dirigir sus actividades de modo correcto. Otros, sin embargo, opinan que las estructuras actuales en las que se plantean los problemas científicos y educativos tienen un nivel de complejidad tal, que los individuos inmersos en ellas no pueden tener una visión lo suficientemente amplia como para dirigir el proceso educativo e investigador de un modo eficaz. Aunque no entraremos en esta discusión con más detalle, parece evidente que en la actualidad, teniendo en cuenta el importante coste para la sociedad de los procesos científico y educativo, es necesario al menos algún tipo de organización dotada de mecanismos capaces de controlar gerencialmente todo el complejo proceso de las instituciones universitarias (Frackman, 1987).

IV.1.2. **Efectividad, eficiencia y rendimiento**

Eficiencia y efectividad son conceptos cuyas diferencias conviene remarcar, ya que suele darse una cierta confusión en su uso. En términos sencillos: la eficiencia es la razón entre las salidas y las entradas de un sistema, mientras que la efectividad es la razón entre los resultados reales obtenidos por el sistema y los que son posibles o ideales, es decir, la razón entre resultados y metas (Cowan, 1985). Utilizando otras palabras: «Eficiencia es hacer las cosas correctamente y efectividad es hacer las cosas correctas» (Miller, 1980: 409). Normalmente el concepto de eficiencia se valora en términos monetarios (eficiencia en precios), aunque es posible utilizar otras valoraciones alternativas como puede ser la utilización de recursos (eficiencia en uso de recursos) (Lindsay, 1981).

Algunos autores utilizan también el concepto de productividad, que se podría definir como la razón entre el valor producido y el valor de los recursos

consumidos (Romney, 1978), y que envuelve los conceptos de eficiencia y de efectividad, ya que la valoración de recursos producidos ha de hacerse manteniendo unos ciertos niveles de calidad de acuerdo con las metas que se propone la institución. A pesar de sus diferencias, el concepto de productividad se confunde frecuentemente con el de eficiencia: sin embargo, mientras que la productividad está referida a la capacidad de producir que tiene un sistema dado, incluso en diversos momentos del tiempo, la eficiencia se refiere a la mejor utilización de los recursos disponibles en cada momentos (Blaug, 1968). De un modo más amplio los conceptos de efectividad, eficiencia y productividad se pueden englobar bajo el concepto general de medidas del rendimiento, que cuando es evaluado a nivel de una institución pasa a denominarse *rendimiento institucional*.

Algunos autores consideran el rendimiento institucional constituido por cuatro elementos, tal y como se representan en el cuadro IV.1 tomado de Calvert (1979); estos elementos del rendimiento son, además de la efectividad y la eficiencia ya citados, otros dos: dotación (relación entre objetivos y recursos asignados); utilización (relación entre recursos asignados y utilizados).

CUADRO IV.1

Algunas magnitudes relacionadas con el rendimiento institucional

RECURSOS ASIGNADOS	Utilización	RECURSOS USADOS
Asignación		Eficiencia
OBJETIVOS	Efectividad	RESULTADOS OBTENIDOS

Es necesario señalar, para finalizar este apartado, que una buena parte de los educadores mantienen que la educación tiene unos objetivos demasiado abstractos como para poder ser objeto de medidas cuantitativas. El abuso de ciertas evaluaciones de rendimientos demasiado simplistas, junto con una actitud lógica, aunque no siempre justificada, por parte del mundo académico, han hecho que el uso de valoraciones del rendimiento institucional esté poco aceptado en estos medios. Sin embargo, si se evitan las simplificaciones excesivas, puede decirse que las evaluaciones del rendimiento cubren una imperiosa necesidad de contabilidad del sistema, junto con la posibilidad de usar esta información para mejorar el proceso educativo (Lindsay, 1980).

IV.1.3. **La evaluación del rendimiento institucional y sus problemas**

Conocer los objetivos que persigue una institución universitaria y evaluar el grado en el que son alcanzados en la realidad, constituye la esencia de la evaluación del rendimiento de la educación superior (Romney *et al.*, 1979). Ahora bien, dado que las instituciones de educación superior son organizaciones de servicios que no buscan el beneficio económico, medir su rendimiento, como el de todas las instituciones de este tipo, es notablemente más complicado que medir los beneficios económicos de una empresa típica. En el caso de las instituciones universitarias, es necesario en primer lugar, identificar, cuantificar y ponerse de acuerdo con exactitud en cuál es el objetivo primordial de las instituciones; desarrollar, en segundo lugar, una jerarquía de objetivos secundarios que se deriven del primario; y, como consecuencia, comparar y medir finalmente los rendimientos reales frente a todos los objetivos propuestos. Estos pasos necesarios para evaluar el rendimiento, son especialmente complicados cuando se trata de instituciones de educación superior, ya que la educación superior es una actividad basada justamente en un proceso en el que las entradas y salidas son, en última instancia, las mismas: los estudiantes, por lo que la aplicación de las habituales funciones de producción para medir el rendimiento son inadecuadas (Kogan, 1986).

Ciertamente las instituciones universitarias tienen toda una serie de características que hacen que las evaluaciones de rendimiento sean especialmente complejas: evaluar el logro institucional supone evaluar toda una serie de factores interconexados entre sí y con magnitudes internas y externas. Por otra parte, la frecuente carencia de una estructura lógica y clara en una institución, es otro factor que dificulta cualquier estudio de su rendimiento (Yorke, 1986).

Que la aplicación de los conceptos de efectividad y eficiencia a la educación superior no es sencilla lo ilustran los siguientes puntos tomados de un estudio realizado para el sistema universitario australiano (Commonwealth Tertiary Education Commission, 1986):

— La determinación de la efectividad supone que se conocen cuáles son los objetivos y que existe acuerdo sobre ellos, lo que no es siempre el caso. No sólo los objetivos pueden ser diferentes para los distintos elementos que constituyen el sistema educativo (alumnos, profesores, administradores, empleadores externos, etc.), sino que también existen metas necesariamente excluyentes que, sin embargo, se persiguen a la vez, como es el caso de intentar extender la educación universitaria, mientras que a la vez se quiere evitar el exceso de oferta de graduados; o intentar realizar una investigación

científica de calidad, junto con la necesidad de satisfacer la demanda inmediata de tecnología por parte de la industria o de la comunidad.

— Dado que los objetivos son múltiples, es difícil encontrar criterios simples que puedan medir el éxito de las instituciones en conseguirlos.

— No existe tampoco un acuerdo sobre la escala de tiempo en la que es razonable medir los efectos de la educación, que realmente son a largo plazo.

— Por último, no existen unos criterios absolutos respecto a los que medir la eficiencia y la efectividad, por lo que las instituciones de educación superior no pueden ser siempre objeto de un riguroso análisis estadístico y financiero, aunque se disponga de los datos suficientes.

Los estudios que evalúan aspectos relacionados, bien con la efectividad, bien con la eficiencia de las instituciones universitarias, han sido desarrollados en los últimos años por numerosos autores. Dada la dificultad en valorar todas las componentes implicadas, el éxito de estos intentos ha sido relativo. Sin embargo, son mucho más frecuentes los estudios relacionados con la eficiencia que los relacionados con la efectividad (Cameron, 1978), justamente por su menor dificultad relativa, pero sin duda también por razones de tipo político: los gestores, gubernamentales o no, demuestran mayor interés por cómo se utilizan los recursos, que para qué se utilizan. Incluso las medidas de calidad (cómo se está funcionando) son mucho más frecuentes que las de la efectividad (cómo puede conseguirse que las cosas funcionen mejor) (Cameron, 1984). Sin embargo, el interés por la efectividad de las instituciones universitarias está aumentando en la actualidad, y ya empieza a ser evaluada en las universidades, sobre todo en aquéllas que tienen implantado un sistema regular y propio de investigación institucional (Rogers y Gentemann, 1989). Incluso recientemente existen ya tratados donde se describen metodologías prácticas para la realización de evaluaciones de la efectividad institucional, como es el caso del trabajo de Nichols (1989).

Dada la complejidad del sistema educativo superior se han propuesto también evaluaciones del rendimiento diferenciadas según los tres tipos de actividades básicas que constituyen la actividad de las instituciones universitarias: uso de recursos, actividad docente y actividad investigadora (Calvert, 1979). Sin embargo, este punto de vista puede conducir a resultados erróneos, ya que no necesariamente las tres funciones básicas de las universidades apuntan en la misma dirección; como afirma Blaug (1968: 398): «Las universidades sirven objetivos múltiples y sus operaciones pueden ser evaluadas (...) en términos de la efectividad con la que cada objetivo es alcanzado (...) No hay

garantía, sin embargo, de que diferentes razones de coste-efectividad apunten todas en la misma dirección, en cuyo caso habrá que establecer distinciones entre los objetivos de acuerdo con su importancia».

Los problemas de criterios concretos son también para Cameron (1978) las mayores dificultades que se plantean al intentar evaluar la efectividad institucional. Estos problemas están relacionados con cuatro cuestiones:

— *Aspectos de organización*. Elección de qué aspectos de la institución universitaria van a servir para evaluar los resultados institucionales y los logros de las metas que tenía propuestas.

— *Universalidad de los criterios*. Es necesario elegir criterios de carácter suficientemente amplio como para que puedan ser válidos en cualquier otra institución universitaria.

— *Criterios descriptivos o normativos*. Elección del más adecuado, en cada caso, de ambos enfoques.

— *Naturaleza dinámica o estática de los criterios*. Elección de variables de uno u otro tipo.

Incluso superados los problemas conceptuales señalados, en el momento de realizar la evaluación del rendimiento institucional surgen algunos problemas inmediatos entre los que se pueden citar los siguientes (Bogue, 1982):

— La evaluación de la efectividad exige el uso de una metodología que no está bien desarrollada.

— Es difícil de obtener un consenso sobre los criterios adecuados para evaluar la efectividad.

— No todas las variables relacionadas con el rendimiento de los estudiantes o de los programas están bajo control interno de las instituciones.

— Existe una tendencia a creer que sólo lo medible es importante.

— La utilización de definiciones de efectividad demasiado estrechas puede conducir a resultados contraproducentes para el propio rendimiento de la institución.

— Existe el peligro de un mal uso de los datos manejados.

Una dificultad adicional para la realización de una evaluación del rendimiento institucional es que, como afirma Cowan (1985), existe una tendencia a medir el rendimiento de las instituciones amalgamando medidas que son de efectividad con otras que son de eficiencia, por lo que los resultados son en muchos casos inadecuados. Debe cuidarse el que las medidas que se utilicen para medir alguna de las dos magnitudes, las valoren separadamente con nitidez.

Romney (1978), en un estudio empírico realizado para medir los objetivos de las universidades norteamericanas, encuentra que el acuerdo entre los objetivos abstractos de las instituciones era muy alto entre todos los estamentos encuestados (profesores y gerentes). Este parece ser un hecho frecuente, sin embargo, como señala Sizer (1979b), esa facilidad para llegar a un acuerdo sobre metas generales no se concreta en la práctica en acuerdos sobre las medidas cuantificables que deben usarse para valorar esas mismas metas. Es decir, parece evidente que el acuerdo es mucho más posible en los principios generales que en la medida de cómo deben evaluarse el que esos principios se alcanzan o no.

Por último, señalar un ejemplo que demuestra la especificidad de las instituciones universitarias frente a otro tipo de instituciones. Se trata del estudio que hizo Cameron (1983) sobre el funcionamiento de un conjunto de universidades norteamericanas. En este trabajo analizó las características que diferenciaban a las universidades que estaban en un momento de auge, de las que estaban en situación de declive; en estas últimas se daba el hecho curioso de que existía una mayor preocupación por la eficiencia, es decir, que sus administradores habían hecho hincapié en las cuestiones fiscales, de presupuesto y de asignación de fondos, mientras que habían dedicado poca importancia a la relación con los medios externos, públicos o privados. Es decir, que según este interesante estudio, la preocupación primordial por la eficiencia en las universidades y su declive como institución van unidas de la mano.

IV.1.4. **La evaluación del rendimiento institucional**

Cabe preguntarse de un modo general cuál es la función de las evaluaciones del rendimiento institucional, es decir, bajo qué perspectivas principales y con qué objetivos primordiales se elaboran estas evaluaciones. Según Pollitt (1987) tres tipos de perspectivas son posibles para la evaluación del rendimiento de las instituciones universitarias:

- Puede ser una evaluación obligatoria o voluntaria.

— Puede estar enfocada sobre la organización como un todo o sobre los individuos que la constituyen.

— Puede tener como función establecer un método de incentivos y penalizaciones, o simplemente el conocer la marcha de la institución.

Ante esta misma cuestión, Gardner (1977) distingue cinco tipos posibles de evaluaciones del rendimiento, según sea la función primordial que se persigue con ella:

— Evaluación como medida.

— Evaluación como juicio profesional.

— Evaluación como medida de congruencia entre objetivos y resultados.

— Evaluación como medida de la utilidad social.

— Evaluación orientada a la toma de decisiones.

Por su parte, Feasby (1980) añade a las cinco anteriores posibilidades otras cuatro más que completan el cuadro:

— Evaluación como un modo de resolver conflictos en la institución.

— Evaluación como un medio para reducir la complacencia sobre el propio funcionamiento.

— Evaluación para impulsar el cambio.

— Evaluación como un ritual para proporcionar una imagen de racionalidad que produzca un sentimiento de seguridad.

Otra cuestión que cabe preguntarse en relación con la evaluación del rendimiento institucional es quiénes están especialmente interesados en su puesta en marcha. Frackmann (1987) da las siguientes explicaciones:

— El estado en el caso de sistemas financiados por los fondos públicos.

— Los organismos *amortiguadores* cuya misión es la asignación de fondos dentro del sistema de educación superior.

— Los clientes (alumnos y empleadores) que desean conocer el valor de lo que ofrece la universidad.

- Los gerentes de la propia universidad que desean conocer el funcionamiento de la institución como una organización.
- Los gerentes de los departamentos u otras subunidades semejantes.
- El profesorado con propósitos de autoevaluación.

Frackmann (1987) señala también que las discusiones sobre el rendimiento institucional son, más frecuentemente promovidas por la sociedad (o por el propio estado en los casos de sistemas públicos), que por la propia comunidad universitaria; es decir, es más deseada por aquéllos que proporcionan los recursos para que las instituciones funcionen, que por aquéllos que los reciben. Por el contrario, la comunidad universitaria parece sentirse de algún modo amenazada o vulnerable ante la perspectiva de que sea evaluado su rendimiento. Ante esta posibilidad, tres tipos de respuestas son posibles por parte de esta comunidad (Frackmann, 1987):

— *A la defensiva*

La organización objeto de evaluación intenta esconder información y evitar la evaluación externa. Esta estrategia sólo puede tener éxito si no existe ningún sistema de castigo (fundamentalmente financiero) para las instituciones que evitan dar información.

— *De colaboración incompleta*

Colaboración en actividades de evaluación sólo en cuanto que son pruebas, estudios piloto y *sin consecuencias*. El realizar actividades evaluadoras sin buscar resultados prácticos es típica de situaciones en las que ninguna de las dos partes ve que estos resultados puedan ayudar a la mejora de las instituciones.

— *De transformar el cambio*

La institución objeto de la evaluación colabora al considerar ventajosos los cambios para su supervivencia, sugiriendo canales y utilización de la información, siguiendo más el modelo de autonomía y autorregulación que el de control centralizado. Es la estrategia necesaria cuando el cambio no puede ser evitado.

IV.1.5. **Los objetivos de la evaluación del rendimiento institucional**

La necesidad de planificación y control es evidente en cualquier estructura con recursos escasos, pero sobre todo es importante actuar en aquellas áreas

de las instituciones en las que es posible mejorar la efectividad y la eficiencia. En el caso de las instituciones de educación superior la necesidad de evaluar el rendimiento se puede expresar de modo más concreto señalando con claridad los objetivos que se persiguen, y que son alcanzables mediante la evaluación. Así, por ejemplo, Sizer (1979a), señala como objetivos de los esfuerzos de la evaluación del rendimiento institucional los siguientes:

- Analizar históricamente el rendimiento.
- Prever el futuro económico, sociopolítico, tecnológico, industrial y educativo, de modo que se puedan identificar las necesidades a largo plazo de la sociedad.
- Desarrollar los objetivos a largo plazo que son relevantes para satisfacer dichas necesidades.
- Formular estrategias para alcanzar dichos objetivos, y aproximar la situación actual a la que se prevé según las necesidades futuras ya determinadas.
- Transformar estas estrategias en planes operativos a medio plazo.
- Motivar a todos los individuos implicados, en la necesidad de alcanzar esos planes, teniendo en cuenta qué planes y control influyen en el comportamiento positiva, pero también negativamente.
- Comparar de un modo continuo la realidad de cada momento con los planes diseñados, estableciendo los procesos de retroalimentación oportunos para modificar sobre la marcha lo que sea necesario.

Romney (1978), por otra parte, distingue dos tipos de objetivos en las instituciones de educación superior que, a su vez, dan lugar a diferentes medidas de rendimiento:

— *Objetivos relacionados con los resultados*

Dan lugar a medidas que informan sobre la efectividad y eficiencia de la institución valorando los resultados finales.

— *Objetivos relacionados con el propio proceso*

Para ellos se utilizan medidas que informan de cómo la institución está funcionando internamente.

Los objetivos concretos de la evaluación del rendimiento institucional pueden establecerse utilizando como punto de partida los diferentes objetivos de tipo político que pueden querer adoptarse frente al funcionamiento de la institución. Basándose en este criterio, Pollitt (1987) identifica los 10 siguientes objetivos de política institucional:

1. Clarificar los objetivos de la propia institución.
2. Evaluar los resultados de la actividad de la institución.
3. Encontrar áreas de posible ahorro económico mediante un estudio de eficiencia.
4. Plantear cuestiones sobre la asignación de recursos.
5. Diseñar esquemas de incentivos para los empleados.
6. Ayudar a conseguir el nivel menos costoso con el que se pueda conseguir un objetivo dado.
7. Establecer o controlar el cumplimiento de ciertas normas por parte de organismos privados.
8. Descubrir cómo contribuyen diferentes servicios a un mismo fin.
9. Proporcionar mejor información a los consumidores.
10. Proporcionar al personal de la institución mecanismos para desarrollar y mejorar sus capacidades.

Como el propio Pollitt (1987) remarca, la utilización de estos objetivos políticos es muy desigual en la práctica real de la evaluación del rendimiento institucional. Mientras que los propósitos 3 y 4 han sido muy usados, otros, como los 2, 6, 8 y 9 son auténticas rarezas. Esta asimetría en el uso de los indicadores demuestra que, como ya se dijo con anterioridad, existe una tendencia a efectuar evaluaciones del rendimiento enfocadas en los aspectos de eficiencia, con menoscabo de los dedicados a analizar la efectividad.

IV.1.6. **El proceso de la evaluación**

Fijados los objetivos de la evaluación del rendimiento, es necesario poner en marcha el proceso de la evaluación. Según Romney *et al.* (1979), este proceso consta de al menos cuatro pasos:

1. Establecer las metas de la evaluación.
2. Asignar los recursos necesarios para la obtención de esas metas.
3. Utilizar los recursos para la consecución de las metas.
4. Obtener los resultados.

En el proceso real de evaluación del rendimiento institucional, habrá que ir contrastando en cada uno de los sucesivos pasos que se vayan dando, la relación entre los logros que se han alcanzado, las metas que estaban programadas y los distintos costes que se van generando en cada paso. Es decir, se debe utilizar una sistemática de la evaluación de procesos que tiene su origen en las técnicas gerenciales habituales, que proporcionan un camino que ya ha sido experimentado en otros campos y que manejado con precaución puede ser también de interés. Aunque, como señala Lindsay (1981), la utilización de estas técnicas gerenciales para evaluar el rendimiento de la educación superior no ha estado nunca exenta de polémica.

La manera más usual de poner en práctica las evaluaciones del rendimiento institucional es a través de la utilización de los indicadores de rendimiento, a los que se dedicará la segunda parte de este capítulo. Con mucho es el método más generalizado a pesar de las controversias que provoca. Aunque otras técnicas han sido también ensayadas, es lamentable, como indica Lindsay (1981), que no hayan marchado al mismo paso las necesidades de evaluación con la puesta a punto de técnicas adecuadas para esa misión. Entre estas otras técnicas que se han aplicado a las evaluaciones cuantitativas del rendimiento de las instituciones universitarias, destacan por su interés la utilización de funciones de producción, de técnicas *input-output*, de métodos estadísticos multivariantes y otras más sofisticadas de complejo aparato matemático (Lindsay y Bailey, 1980; Lindsay, 1982). El análisis *portfolio* ha sido también propuesto para la resolución de algunos problemas (Sizer, 1982).

Por último, reseñar las recomendaciones que da Lindsay (1981) para conseguir que sea fructífera la puesta en práctica de las evaluaciones institucionales. Son las siguientes:

— Cuando la información es dispar o conflictiva es necesario sopesarla previamente y tratar de eliminar las contradicciones que aparecen.

— La evaluación debe evitar producir resultados para cada una de las componentes del sistema. Los resultados, por el contrario, deben ser globales para toda la institución.

— Los resultados deben presentarse de modo que faciliten la elección de alternativas a los que han de tomar decisiones.

- Las unidades de análisis en la evaluación deben coincidir con las unidades en las que se van a tomar las decisiones.
- Los resultados deben ser válidos, es decir, correctos e insesgados.
- Los resultados deben ser contrastables.
- La información proporcionada debe ser económica, es decir, que los beneficios superen los costes.
- La información debe ser oportuna.

IV.2. INDICADORES DE RENDIMIENTO

Una de las herramientas que más se han utilizado para la evaluación del rendimiento de las instituciones de educación superior han sido los llamados indicadores de rendimiento. Dada la amplitud de su uso, por otro lado fuertemente controvertido, conviene dedicarles una atención especial. Esta segunda parte del capítulo va a dedicarse a estos indicadores de rendimiento: su definición, su utilidad, su utilización, y una breve discusión acerca de sus ventajas e inconvenientes más destacados.

IV.2.1. El concepto de indicador de rendimiento

De una manera general se llaman indicadores de rendimiento a las medidas objetivas, usualmente cuantitativas, de los logros de una institución o de todo un sistema de educación superior (Ball y Halwachi, 1987). Para Cuenin (1987) un indicador de rendimiento es un valor numérico que proporciona una medida para evaluar el rendimiento de un sistema de educación superior.

Como ya se ha señalado con anterioridad, la utilización de indicadores de rendimiento para evaluar el rendimiento de las instituciones universitarias se desarrolló enormemente a partir de la segunda mitad de los años setenta. Las causas de este desarrollo hay que buscarlas, por una parte, en el elevado coste de los recursos de todo tipo dedicados a la educación superior, y, por otra, en el incremento del valor de esta educación como puerta a casi todos los puestos de trabajo valiosos: para Williams (1986) estas dos razones legitiman suficientemente las exigencias del uso de indicadores objetivos del rendimiento institucional.

A pesar de sus posibles deficiencias, es cierto que el uso de los indicadores de rendimiento puede ayudar considerablemente a establecer un cierto orden

y a hacer previsiones en un sistema tan complejo como es en la actualidad cualquier institución universitaria. Cuando los indicadores son suficientemente desarrollados ayudan a entender la situación de los problemas con respecto a unas variables determinadas, a ver cómo es su evolución en el tiempo y a hacer previsiones para el futuro (Jacobson, 1978).

Otro de los objetivos más interesantes que se han propuesto para justificar el uso de los indicadores de rendimiento en las instituciones universitarias es el de que pueden ayudar a detectar con rapidez las tendencias de los cambios en las necesidades de la sociedad, en las preferencias por determinados tipos de cursos o estudios, y, en consecuencia, en la reasignación de recursos a las partes del sistema universitario más necesitadas en cada momento (Sizer, 1982).

Uno de los problemas más frecuentes que se plantean al hablar de la utilización de los indicadores de rendimiento en las universidades es el de la posibilidad de que sea factible construir tales indicadores. Así, por ejemplo, Sizer (1979a, 1982) se pregunta si es posible llegar a un acuerdo sobre los objetivos a largo plazo culturales, económicos, político-sociales, tecnológicos, ecológicos y educativos de una sociedad. Y, en el caso de llegar a un acuerdo, si es posible trasladarlos a medidas cuantificables en la forma de indicadores de rendimiento, con pesos adecuados para cada objetivo, de modo que podamos medir la efectividad y la eficiencia de la institución. Esta discusión sigue abierta años después de que se iniciara, y el uso de indicadores sigue teniendo partidarios y enemigos acérrimos.

Dadas las dificultades de establecer objetivos globales para las instituciones de educación superior, se han ido estableciendo medidas para sectores del sistema, es decir, indicadores de rendimiento para algunas partes de la institución que son más fácilmente medibles y sobre las que es más fácil llegar a acuerdos. Esta es una tendencia que puede ser útil en un momento dado, pero que pudiera ser peligrosa si no se tiene en cuenta al usar estos indicadores parciales que, en un sistema complejo como el universitario, el optimizar las partes no significa necesariamente optimizar el sistema como un todo (Sizer, 1979a).

Por otra parte, los indicadores de rendimiento deben tener unas cualidades que los hagan utilizables con fiabilidad y confianza por todas las partes implicadas en sus resultados. Sizer (1979a) establece los criterios que deben satisfacer un indicador de rendimiento institucional para que sea apropiado para evaluar una situación determinada. Son éstos:

— *Relevancia*

El indicador debe ser útil para aquéllos que están interesados en los aspectos gerenciales de la actividad a la que se refiere el indicador.

— *Verificabilidad*

Individuos cualificados, usando los mismos datos y trabajando separadamente, deben llegar a resultados y valores semejantes para los indicadores.

— *Inssegados*

El cálculo de los indicadores debe carecer tanto de sesgo estadístico como de sesgo personal.

— *Cuantificabilidad*

Los indicadores deben estimarse mediante cantidades numéricas.

— *Económicamente factibles*

De modo que el beneficio de utilizar el indicador supere los costes que su cálculo o utilización reporte.

— *Aceptabilidad institucional*

Las personas afectadas por el uso de los indicadores deben haberlo aceptado, previamente a su utilización, que los indicadores son útiles y correctos.

IV.2.2. **Tipos de indicadores de rendimiento**

Existen diversos tipos de indicadores de rendimiento según el aspecto del sistema que estén tratando de analizar o según la función que tenga su uso. Diversos autores han propuesto diferentes clasificaciones para estos indicadores. Algunas de ellas se exponen a continuación.

La clasificación más amplia de los indicadores de rendimiento es la basada en estos criterios (Dochy *et al.*, 1990):

— Según las funciones de la educación superior: educación, investigación, servicios.

— Según el objeto: indicador de *input*, de proceso, de *output*.

— Según el tipo de medida: indicador cuantitativo, indicador cualitativo.

Una de las más importantes clasificaciones es la que se puede hacer según la parte del proceso universitario que están evaluando los indicadores (Jacobson, 1978). Según este criterio pueden definirse tres tipos de indicadores de rendimiento:

— *Indicadores de esfuerzo o de input*. Son los que miden las entradas de todo tipo que la institución está absorbiendo para su funcionamiento.

— *Indicadores de funcionamiento*. Miden cómo está funcionando la institución.

— *Indicadores de efecto o de resultados*. Son los que valoran los resultados que se están consiguiendo mediante el proceso que se está evaluando.

Una clasificación distinta, basada en la fuente de la información de la que se extraen los datos para la construcción de los indicadores, es la que ofrece Cuenin (1987). Esta clasificación está basada en un estudio empírico sobre la utilización real de indicadores de rendimiento en las universidades pertenecientes al programa IMHE (*International Management in Higher Education*) de la OCDE. Según este estudio, los indicadores utilizados en la práctica por estas universidades pueden ser agrupados dentro de cuatro categorías:

— *Indicadores internos*

Aquéllos que utilizan información obtenida en la propia institución educativa.

— *Indicadores de funcionamiento*

Son los que describen aspectos del funcionamiento interno del sistema.

— *Indicadores externos*

Aquéllos que usan información que se obtiene fuera de la institución.

— *Indicadores de investigación*

Están basados en los datos de la actividad científica de la institución.

Otra clasificación de carácter funcional muy interesante es la dada por Sizer (1979a). En ella se utiliza como criterio de clasificación la utilidad concreta que tienen los indicadores dentro del proceso evaluador. Según este criterio, los indicadores de rendimiento pueden ser agrupados en nueve categorías:

— *Disponibilidad*

Indicadores que informan sobre qué puede ser obtenido de la utilización de las instituciones de educación superior.

— *Publicidad*

Informan sobre el grado en el que son conocidos por los usuarios los servicios de que disponen las instituciones.

— *Accesibilidad*

Miden el nivel de accesibilidad a los servicios que ofrece la institución a los potenciales usuarios.

— *Utilización*

Comparan el nivel real de utilización y la demanda potencial de los servicios existentes.

— *Adecuación*

Informan si es adecuada la relación entre la cantidad y calidad de los servicios, por una parte, y las necesidades de las instituciones, por otra.

— *Eficiencia*

Comparan las entradas en la institución como un todo, con el producto final obtenido.

— *Efectividad*

Comparan los logros conseguidos con los objetivos propuestos.

— *Impacto externo*

Tratan de identificar y valorar el efecto y los beneficios que las instituciones de educación superior pueden ejercer sobre la sociedad que las rodea.

— *Aceptabilidad*

Miden el grado de aceptación por parte de los usuarios de los servicios ofrecidos por las instituciones.

Otra clasificación similar para los indicadores de rendimiento es la que propone Antia (1976). Para este autor los indicadores de rendimiento deben ser los adecuados para cubrir la información necesaria sobre estas nueve áreas:

— *Tono social*

Algunos indicadores deben medir la adaptación de la institución a la sociedad, así como su capacidad para la previsión de las necesidades de esa sociedad.

— *Coste-efectividad*

Medidas relacionadas con los costes de funcionamiento interno de la institución.

— *Desarrollo de programas*

Valoraciones sobre el diseño de los cursos, la evaluación de su marcha y la capacidad de innovación de la institución.

— *Reputación*

Estimaciones del nivel de reputación de la institución desde los puntos de vista académico y de servicio a la sociedad.

— *Inversión en capital humano*

Medidas de la capacidad de la institución para preparar profesionalmente a sus estudiantes.

— *Desarrollo de los servicios materiales*

Valoraciones de la amplitud y calidad de los servicios esenciales para el funcionamiento de las instituciones, como bibliotecas, laboratorios, servicios de ordenadores, etcétera.

— *Relaciones con los estudiantes*

Medidas de la extensión, variedad y calidad de los servicios complementarios que se ofrecen a los estudiantes como restaurantes, alojamiento, tutorías,

participación y representación en los organismos dirigentes, servicios deportivos y recreativos, instalaciones para el estudio, existencia de sistema de quejas institucionalizado, etcétera.

— *Relaciones con los empleados*

Valoraciones del clima y funcionamiento de las relaciones de la institución con sus empleados.

— *Responsabilidad pública*

Estimaciones del cumplimiento por parte de la institución de los deberes ciudadanos respecto a la sociedad que la circunda y que la sustenta.

IV.2.3. **La utilización de los indicadores de rendimiento**

Cabría ahora preguntarse cuál es el grado real de utilización de los indicadores de rendimiento y qué tipos de estos indicadores son utilizados preferentemente por las universidades. Para contestar a la segunda cuestión recurriremos a la información que aporta Cuenin (1987) sobre el uso de indicadores en un conjunto de universidades de países de la OCDE. Según esta encuesta, los indicadores son utilizados realmente para valorar toda una serie de cuestiones que, clasificadas por tipologías, son las siguientes:

1. *Indicadores para la asignación interna de recursos dentro de la institución:*

a) En relación con el profesorado y los departamentos:

- Carga docente del profesorado.
- Número de estudiantes matriculados.
- Resultados de los exámenes.
- Razón profesor/alumno.
- Costes unitarios.
- Tipos de asignatura.
- Fondos externos de investigación que se consiguen.
- Número de publicaciones.
- Número de empleados del departamento.
- Número de estudiantes de postgrado.
- Número de disertaciones doctorales.
- Número de proyectos de investigación.

b) En relación con la asignación de plazas de profesorado y de empleados:

- Razón profesor/alumno.
- Carga docente.
- Razón entre personal docente y no docente.

2. *Indicadores de planificación institucional:*

a) Relacionados con los estudiantes:

- Normas académicas de admisión.
- Número de solicitudes.
- Tendencias de las matriculaciones.
- Tasa de abandono.
- Número de graduados.
- Origen geográfico y étnico.

b) Relacionados con las fuentes financieras de la institución.

c) Relacionados con planes de retiro del personal.

3. *Indicadores de organización docente:*

- Carga de trabajo que soportan los estudiantes.
- Valoración realizada por los alumnos de los cursos que reciben.
- Valoración de la enseñanza recibida por los alumnos que ya están graduados.

— Duración media de los estudios.

— Solicitudes de matrícula para cada curso y asignatura.

4. *Indicadores de la relación con organismos externos:*

a) Respecto a los estudiantes:

— Número de estudiantes clasificados por tipo de estudios y nivel.

b) Respecto a los profesores y empleados:

— Profesorado clasificado por categorías.

— Otro personal, también clasificado por categorías.

c) Respecto a aspectos académicos:

— Número de graduados.

— Razón profesor/alumno.

d) Respecto a aspectos materiales:

— Area construida.

— Area no construida.

e) Respecto a aspectos financieros:

— Fondos recibidos clasificados por tipos.

En uno de los recientes informes generales sobre la situación de la educación universitaria en el Reino Unido, el *Informe Jarratt* (Jarratt, 1985), se recomienda, tras las experiencias ya conocidas en aquel país sobre el uso de indicadores de rendimiento, que se utilicen por parte de las universidades un conjunto de ellos. Clasificados en tres categorías, son los siguientes:

1. *Indicadores de rendimiento interno:*

— Cuota de mercado que mantiene la institución respecto a solicitudes de matriculación de estudiantes en todo el país.

— Tasa de graduación y tipos de grados conseguidos.

— Atracción de estudiantes de doctorado y de *masters*.

— Tasa de éxito en los grados superiores y tiempo necesitado para completarlo.

— Atracción de fondos para investigación.

— Calidad de la docencia.

2. *Indicadores de rendimiento externo:*

— Aceptabilidad de los graduados en empleos.

— Primer destino de los graduados.

- Reputación de los graduados según juicio de evaluadores externos.
- Publicaciones y citaciones del profesorado.
- Patentes, inventos y consultorías.
- Premios y distinciones.
- Trabajos en congresos.

3. *Indicadores de rendimiento operativo:*

- Costes unitarios.
- Razón profesor/alumno.
- Tamaño de las clases.
- Disponibilidad de asignaturas para elegir.
- Carga docente.
- Disponibilidad de volúmenes en la biblioteca.
- Disponibilidad de ordenadores.

Como consecuencia del *Informe Jarratt*, se formó en el Reino Unido un comité encargado de formular una lista de los indicadores de rendimiento que debían servir para evaluar las universidades británicas. El resultado del estudio de este comité fue el establecimiento de una lista de 54 indicadores con alto grado de concreción, que debían servir para evaluar las universidades. Así, por ejemplo, se consideran indicadores del tipo de «gasto telefónico como porcentaje de gasto total» o «gasto en revistas por estudiante», lógicamente junto con otros indicadores de mayor entidad (Sizer, 1990).

De entre todos los indicadores de rendimiento que se han propuesto, hay evidentemente algunos más usados que otros. Cuenin (1987) cita los cinco que están siendo más utilizados por las universidades de los países de la OCDE:

— *Primer destino de los graduados*

Obtenido por medio de encuestas realizadas a los antiguos alumnos. El objetivo es el de reestructurar los cursos, las materias, cambiar las cuotas de ingreso en las distintas especialidades y modificar la asignación de recursos de acuerdo con las mayores o menores demandas de la sociedad.

— *Razón profesor/alumno*

Este indicador es muy usado para realizar comparaciones entre instituciones. El principal problema de este indicador es su interpretación (exceso de alumnos o pocos profesores) y la carencia de unas normas generales de referencia sobre qué valores son los más adecuados.

— *Publicaciones del profesorado*

Utilizadas para la asignación de recursos a las unidades de investigación.

— *Costes unitarios*

Se usan diversos tipos: costes por estudiante, costes por profesor, costes por asignatura y algunos otros. El problema del cálculo y de la interpretación de estos indicadores de costes es importante y ha de realizarse con sumo cuidado.

— *Fondos obtenidos del exterior*

Número y cantidad de los contratos realizados por los departamentos. Estos indicadores son usados principalmente para la asignación de los recursos de la universidad entre los diversos departamentos.

IV.2.4. **Discusión sobre el uso de los indicadores**

El uso de los indicadores de rendimiento en la educación superior se puso realmente de moda en la segunda mitad de la década de los setenta y, aunque su uso sigue siendo contestado en la actualidad por muchos universitarios, son recomendados con insistencia por aquellos organismos que de alguna manera están más relacionados con los aspectos gerenciales de la educación superior. En este sentido cabe mencionar el *Informe Jarratt* (Jarratt, 1985) en el que se afirma la necesidad de utilizar indicadores de rendimiento fiables y consistentes que puedan ser usados para la planificación y para la asignación de recursos en las universidades británicas.

Los indicadores de rendimiento no presentan especiales problemas cuando existe acuerdo sobre su significado y su uso, pero es mucho más problemática su utilización en otros casos. Los indicadores de rendimiento son por su propia naturaleza ambiguos, al menos que sean usados en un contexto de diálogo adecuado entre las partes implicadas y sean utilizados para objetivos aceptables por todos (Bormans *et al.*, 1987). Como afirma In't Veld (1990: 40): «Los indicadores de rendimiento son siempre aproximaciones imperfectas de una característica final, de modo que siempre hay lugar para el debate sobre los méritos o deficiencias de los indicadores de rendimiento». Sizer (1979a) se plantea de un modo general la viabilidad del uso de los indicadores de rendimiento: «¿Es una institución de educación superior efectiva si alcanza los objetivos adecuados al entorno económico, sociopolítico, tecnológico, ecoló-

gico y educativo en que opera? ¿Deben ser esos objetivos congruentes con las necesidades a largo término de la sociedad? Muchos de aquellos individuos que están relacionados con la organización de la educación superior responderán positivamente a esta cuestión. Sin embargo, es poco verosímil que ellos fueran capaces de alcanzar un acuerdo sobre las necesidades a largo plazo de la sociedad, sobre la contribución que cada institución debe hacer para satisfacerlas y sobre qué objetivos debe perseguir su propia institución. Incluso si se pusieran de acuerdo sobre los objetivos generales para sus instituciones, ¿podrían trasladarlos de común acuerdo a metas cuantificables, y posteriormente a indicadores de rendimiento? Si ese fuera el caso, ¿cómo se podrían establecer objetivos e indicadores de rendimiento más detallados para los componentes de la institución? No sólo es poco verosímil que se alcancen acuerdos sobre los objetivos generales de la institución, sino que en los departamentos, especialmente en los multidisciplinares, es inverosímil que se puedan alcanzar acuerdos» (Sizer, 1979a: 51).

A estas dificultades habría que añadir otra de carácter más general incluso: la dificultad de planificar a largo plazo para unas instituciones necesariamente lentas como las educativas, inmersas en un mundo tremendamente cambiante como el actual. Ante esta situación es necesario seguir investigando en las tendencias de la sociedad y de las instituciones, de modo que se puedan convertir en objetivos cuantificables que se puedan expresar por medio de algún tipo de indicador (Sizer, 1981).

Uno de los problemas más importantes que acarrea la utilización de indicadores de rendimiento es el de la interpretación de sus resultados. La definición incorrecta y la interpretación estrecha de los indicadores de rendimiento puede distorsionar las actividades de las instituciones de educación superior con un efecto fuertemente negativo (Williams, 1986). Algunos autores piensan que la perspectiva básicamente contable de la educación superior que proporcionan los indicadores de rendimiento es inapropiada, por ser demasiado simplificadora de una realidad compleja (Romney, 1978).

Otro de los problemas importantes que puede ocasionar la utilización de los indicadores de rendimiento, especialmente cuando son usados por organismos externos preocupados esencialmente por las cuestiones financieras, es el énfasis que se le puede prestar a las medidas de eficiencia y de efectividad a corto plazo en contra de las metas a largo plazo que deben tener las universidades. El problema puede surgir cuando el logro de estas metas entre en contradicción con medidas a corto plazo sobre las que pueden estar centrados los indicadores (Sizer, 1982).

Algunos indicadores de rendimiento son de dudosa aplicabilidad, ya que

pueden ser utilizados en sentidos opuestos según quien los use (Yorke, 1986). Especialmente problemático es el uso de indicadores relacionados con la asignación de fondos. Su utilización ha de ser extremadamente cuidadosa, porque, como señala Williams (1986), existe una frecuente malinterpretación o círculo vicioso respecto al uso de esos indicadores: pueden estar indicando malos resultados en una institución porque los recursos son bajos, lo que puede dar lugar a que se le asignen menos recursos todavía. En el mismo sentido, Cuthbert y Birch (1979) citan el caso de un indicador que se usaba en el Reino Unido para la asignación de fondos a las universidades, y que se conseguía mejorar disminuyendo el número de horas de clase impartidas, aumentando el tamaño de las aulas o matriculando más alumnos en cualquier condición. Como señalan Birch y Calvert (1977) para el uso de indicadores relacionados con los costes, se puede inducir una «mejora» en la eficiencia de la institución aumentando simplemente el número de alumnos aunque no se incremente, e incluso disminuya, el valor añadido a cada uno de ellos.

Dadas las dificultades que rodean la definición y el diseño de los indicadores de rendimiento, la tendencia más habitual ha sido la de usar aquéllos que, o bien son más sencillos de obtener, o están relacionados con la medida de partes del sistema universitario que son más fácilmente medibles, descuidándose de este modo aspectos más esenciales pero que tienen mayor dificultad para traducirse en indicadores cuantitativos (Sizer, 1979b). En este mismo sentido, es digno de mencionarse que la mayoría de los indicadores que se utilizan están relacionados con la productividad investigadora y muy pocos están referidos a la calidad de la docencia (Weert, 1990), como consecuencia lógica de la mayor dificultad que tiene el evaluar el rendimiento de la enseñanza que el de la investigación.

Según Frackmann (1987), aunque los indicadores de rendimiento son vistos por la comunidad universitaria con criticismo y escepticismo, y a pesar de todas sus posibles deficiencias, no hay que descartar su uso, simplemente hay que ser cuidadoso con su utilización. Las críticas más frecuentes al uso de indicadores de rendimiento son sintetizadas así por el mismo Frackmann:

— Los indicadores de rendimiento intentan medir cuantitativamente los resultados de la educación superior, algo que es esencialmente cualitativo.

— Los indicadores de rendimiento comparan algo que no es comparable, ya que a diferencia de otros procesos productivos, el de la educación superior no intenta producir el mismo resultado masivamente.

— Los indicadores están simplificando un proceso tan complejo y diverso como es el de la educación superior y la producción científica.

— Existe el peligro de que los indicadores de rendimiento se conviertan en una meta por sí mismos y no simplemente en una manifestación de una realidad compleja.

Como consecuencia de estas dificultades inherentes a los indicadores de rendimiento, y debido también a su uso por organismos externos a las universidades con fines de control, la aceptación de los indicadores por los universitarios es bastante reducida. Sobre la utilidad de los indicadores afirma Cuenin en su estudio sobre las universidades de países de la OCDE que «las opiniones están divididas (...), pero prevalece la actitud de escepticismo» (Cuenin, 1987: 126). Esta misma actitud de escepticismo era encontrada por Romney (1978) en un estudio empírico realizado en USA en la que analizaba la respuesta de diferentes universidades del país. Muchas de las críticas más vehementes vienen de aquéllos que consideran que los indicadores de rendimiento son un producto y una herramienta de grupos políticos de presión que están intentando dirigir la educación superior en la dirección de la educación en cadena: el punto de vista del «*input-proceso-output*» no es la más adecuada para la calidad de la educación superior. Para algunos los indicadores de rendimiento son un síntoma de una cultura educativa «thacheriana» que ha creado una mentalidad obsesiva de autoanálisis y autojustificación (Findlay, 1990).

A pesar de la poca aceptación de los indicadores de rendimiento, algún estudio demuestra la validez de algunos de ellos. Así, por ejemplo, Taylor y Johnes (1989), en un estudio realizado en el Reino Unido sobre el primer empleo de los graduados por las universidades, llegan a la conclusión de la validez de los indicadores utilizados por las universidades, tras comprobar la relación entre esos indicadores y el éxito en las carreras profesionales posteriores de un grupo de individuos. Aunque este estudio está referido a unos indicadores muy concretos, los resultados son concluyentes.

Sea cual sea la actitud hacia el uso generalizado de los indicadores de rendimiento en las instituciones universitarias, existe un problema para el cual parece evidente que deben usarse: la asignación de recursos a las universidades en los sistemas universitarios de carácter público. Parece necesaria la utilización de indicadores de rendimiento que clarifiquen la situación de equidad distributiva entre todas las partes del sistema. Curiosamente éste es un problema frecuente que no siempre está adecuadamente solucionado en los diferentes países. Así, por ejemplo, Mertens (1979) realizó un estudio sobre la asignación de recursos entre las universidades alemanas, comparando los resultados de tres tipos de indicadores de rendimiento; encontró que existían dispersiones muy importantes entre estos indicadores y las correspondientes

asignaciones económicas, lo que indicaba la poca atención prestada hasta entonces a la utilización de criterios objetivos para la asignación de recursos y para el propio funcionamiento de las instituciones. Estudios semejantes realizados en España llegan a parecidos resultados (Moltó y Oroval, 1982).

IV.2.5. **El futuro de los indicadores de rendimiento**

En los últimos dos o tres años un nuevo factor ha empezado a intervenir en las discusiones sobre indicadores de rendimiento, se trata de la cuestión de la calidad de las instituciones universitarias, y de cómo los datos estadísticos o gerenciales pueden servir para evaluarla. La discusión está haciendo pasar el núcleo del interés sobre el uso de los indicadores desde los problemas de eficiencia a los de efectividad, y de las comparaciones entre instituciones al autoestudio de cada una de ellas para la mejora de su nivel de calidad (Findlay, 1990).

En este nuevo clima de preocupación por la calidad, y habiendo cambiado las perspectivas bajo las que se proponía su uso al final de la década de los setenta, el uso de los indicadores, fuertemente contestado con anterioridad, vuelve a recobrar auge. Los indicadores de rendimiento, adecuadamente definidos y adecuadamente interpretados pueden jugar un papel útil en la determinación de la política educativa, en la gestión de una institución y en la optimización de la calidad educativa (Dochy *et al.*, 1990).

Uno de los problemas que actualmente se ven como cruciales para la puesta en marcha de los indicadores de rendimiento, es el de la forma de su implantación. Mientras que los indicadores de rendimiento son mirados con sospecha y escepticismo cuando son impuestos por agencias externas, e incluso pueden distorsionar el proceso que están intentando controlar, cuando su uso es acordado por las partes y su necesidad es aceptada desde la base, su efectividad puede ser mucho más alta. El modo como se introduzcan los indicadores de rendimiento en el contexto de las instituciones es posiblemente más importante que el diseño de los indicadores por sí mismos: una actitud de consulta y de acuerdo parece la más adecuada antes de iniciar el uso de indicadores en una institución (Findlay, 1990).

Otro punto de vista respecto a los indicadores que también está empezando a cambiar, es el de su misma definición. Se está proponiendo que pasen de ser exclusivamente una información numérica de tipo estadístico, a algo más completo que incluya juicios subjetivos y opiniones consensuadas. El principio debería ser el de «medir lo que debe ser evaluado», en vez de «medir lo que

es fácil de medir», aunque para ello haya que recurrir a sistemas no necesariamente numéricos de evaluar.

En esta misma línea Dochy *et al.* (1990) amplían la definición de los indicadores de rendimiento: «Indicadores de rendimiento son datos empíricos de naturaleza cuantitativa o cualitativa. Estos datos se convierten en indicadores de rendimiento si expresan las metas contempladas por las partes implicadas. Eso significa que ellos tienen una importancia contextual además de temporal. Ellos pueden hacerse operacionales mediante variables» (Dochy *et al.*, 1990: 136). Según esta definición hay que diferenciar entre lo que son datos estadísticos, que proporcionan información, de los indicadores de rendimiento que describen el funcionamiento de una institución, el modo en el que esta institución se aproxima a sus metas.

En estos momentos existe una amplia experiencia e información sobre el uso de indicadores de rendimiento, sobre los problemas que plantean, sobre su validez y sobre su interpretación. Dado que su utilización es una base necesaria para las relaciones correctas entre los gobiernos y las instituciones públicas de educación superior, el principal problema que queda sin resolver es el de encontrar en cada caso los indicadores más idóneos y el de llegar a un acuerdo entre todas las partes implicadas, tanto para su uso como para su interpretación.

CAPITULO V: EVALUACIONES AUTORREGULADAS

Colleges and universities have studied everything from cabbages to kings, but rarely have they paused willingly to reflect on the effectiveness of their own educational program and services.

Johnson and Christal, 1985: 2

V.1. INTRODUCCION

En este capítulo se tratan algunos tipos de evaluaciones de las instituciones universitarias, que se pueden agrupar bajo la denominación general de evaluaciones autorreguladas. El término expresa la idea de que son evaluaciones promovidas por la propia institución, realizadas por ella misma o por agencias externas, pero, en cualquier caso, con carácter eminentemente voluntario.

Este tipo de evaluación surge históricamente en USA, donde su peculiar sistema universitario, absolutamente anárquico en apariencia, exige para su buen funcionamiento, y para información de los usuarios, algún tipo de guía o control de la calidad y efectividad de las instituciones. El mecanismo una agencia externa especializada en esta misión. Una de las piezas claves del proceso de acreditación lo constituye el autoestudio que realiza la propia institución previamente a la visita de los miembros de la agencia acreditadora. Este autoestudio, que sirve de base para la evaluación externa, se ha independizado, en cierto modo, del proceso de la acreditación, de forma que hoy en día es uno de los métodos que con más frecuencia están siendo propuestos y puestos en práctica, tanto en USA como en otros países, para la evaluación de las instituciones universitarias.

Otra manera, muy relacionada con las anteriores, de afrontar la evaluación las instituciones es la llamada revisión de programas, es decir, una valoración del funcionamiento de programas específicos o de los departamentos universitarios, realizados por las propias instituciones o, en el caso de algunos

sistemas universitarios de carácter público, por los responsables de esos sistemas.

En este capítulo se dedica una sección a cada uno de estos tres temas: la acreditación en USA, por su interés como generador de otros métodos de evaluación, los autoestudios y las revisiones de programas. Estos dos últimos por su posible utilidad, ya que pueden ser transportables, con las debidas modificaciones, a otros sistemas universitarios.

Antes de pasar a revisar cada uno de estos mecanismos de evaluación institucional, veamos el marco general en el que estos procesos autorregulados tienen sentido. Es decir, cuáles son las condiciones generales que debería cumplir todo sistema de evaluación autorregulado. Siguiendo a Kells y Vught (1988) estos serían los atributos ideales:

— El propósito de la evaluación debe ser doble: asegurar la calidad y conseguir mejoras.

— La evaluación debe estar centrada en la institución, más que ser comparativa o tener carácter gubernamental.

— La evaluación debe tener aspectos internos y externos (autoevaluación y evaluaciones por visitantes externos).

— La autoevaluación debe ser la pieza fundamental de la evaluación.

— Las medidas de la calidad que se utilicen en la evaluación deben ser consensuadas, y tanto de carácter general como referidas a metas de los programas.

— El sistema de evaluación debe hacerse regular y cíclicamente.

— La evaluación debe ser global. Todas las unidades deben ser evaluadas en cada ciclo.

— El proceso debe tener consecuencias. Los planes de la institución deben ser afectados por la evaluación.

— Los equipos externos de revisión deben estar constituidos por colegas experimentados, elegidos sin sesgos.

— Los resultados detallados de las visitas y de los estudios deben ser confidenciales.

— Una información resumida de la situación de la institución, obtenida del análisis del proceso de evaluación, deberá proporcionarse al público.

V.2. LA ACREDITACION EN USA

La acreditación es un mecanismo típicamente norteamericano que de algún modo ejerce la función de inspección sobre las instituciones universitarias que en los países europeos realizan organismos gubernamentales creados para ese fin. La acreditación es la forma más antigua de evaluación de la calidad en las universidades estadounidenses (Lawrence y Green, 1980), y sigue teniendo actualmente una gran vitalidad. En un principio consistía básicamente en la revisión, por parte de un equipo de expertos, de la capacidad de los recursos de las instituciones universitarias para realizar su cometido; posteriormente se incluyó un autoestudio realizado por la propia institución en el que no sólo se hace referencia a los recursos disponibles, sino al uso que se hace de ellos y al rendimiento de la institución (Tritschler, 1981).

La acreditación refleja el carácter diverso y relativamente independiente de la educación superior en USA, extendiéndose a todo tipo de instituciones de educación superior, públicas y privadas. En las instituciones públicas de algunos estados norteamericanos, también se realizan revisiones de los programas por medio de agencias de evaluación de carácter gubernamental, cuyas actividades se solapan con frecuencia con las de las agencias de acreditación. A pesar de que la acreditación, aunque sólo sea por razones históricas, no es transportable a otros sistemas universitarios, tiene interés su idea y funcionamiento por su influencia en el control de la calidad de las universidades estadounidenses y porque ha generado otros sistemas derivados, los autoestudios, que son de gran interés por sí ser trasladables a otros sistemas educativos.

V.2.1. El sistema de acreditación

La acreditación de las instituciones de educación superior estadounidenses es realizada por agencias especializadas para tal función. Las agencias de acreditación son organismos de carácter privado, sustentadas por las propias instituciones que han sido acreditadas. Su misión esencial es conceder el *sello de la calidad* a las instituciones que han superado la revisión realizada por un equipo de la propia agencia. La acreditación es solicitada voluntariamente por las instituciones universitarias, y aunque no son necesarias para el funcionamiento de las instituciones, existen toda una serie de beneficios que no es posible obtener si se carece de la acreditación pertinente.

Una definición del proceso de acreditación puede ser ésta: «La acreditación es un proceso voluntario, realizado por colegas, a través de agencias no gubernamentales, para conseguir al menos dos cosas: valorar hasta qué punto las instituciones alcanzan los criterios de calidad establecidos, e intentar periódicamente contabilizar los logros alcanzados por la institución o por algunos de sus programas en concreto. La principal misión de este proceso es la de promover las mejoras e identificar aquellas instituciones y programas que están alcanzando las propias metas que se han establecido» (Kells, 1988b).

La acreditación es el sistema más universal que existe en USA para evaluar las instituciones. Aunque existen diversas agencias de acreditación todas tienen en común la preocupación por el buen funcionamiento de las instituciones universitarias. Estas agencias centran su atención tanto en la estructura como en el proceso, por lo que sus evaluaciones están en cierta manera asegurando la calidad de las instituciones (Troutt, 1981).

Existen dos tipos básicos de acreditación: la institucional y la especializada en programas de carácter profesional. La primera se refiere a la acreditación de instituciones de un modo global, es decir, de sus programas docentes, de su actividad investigadora, de los servicios que se ofrecen a los estudiantes y de los servicios administrativos. El segundo tipo de acreditación está dedicado a la evaluación de programas específicos (medicina, derecho, etc.) incluidos o no dentro de instituciones más complejas; el objetivo básico, en la mayoría de los casos, es el de certificar la calidad, con fines profesionales, de la formación que se proporciona a los estudiantes en esos programas. En ambos tipos de acreditaciones el proceso es llevado a cabo por una comisión formada por miembros elegidos de la comunidad académica, pertenecientes a otras instituciones universitarias, asistidos por algún empleado de la agencia. Las dos fuentes básicas de información que utiliza la comisión para su evaluación son el informe de autoevaluación, que presenta la institución que desea ser acreditada, y una visita del equipo a la propia institución.

Existen varias agencias regionales y alguna nacional que realizan el proceso de acreditación del primer tipo, y más de 40 agencias de acreditación especializadas en programas o grupos de programas, todas ellas con carácter nacional, y muchas controladas por las propias asociaciones profesionales del ramo correspondiente.

Algunos pasos comunes a todos los procesos de acreditación son éstos:

— La institución o programa que desea ser acreditado presenta un autoestudio en el que se describe.

— Un equipo de evaluación formado por colegas elegidos de otras instituciones no competidoras visita la institución o programa, evaluándola a la luz de los documentos del autoestudio presentado por la institución y de las propias normas de la agencia. Como resultado el equipo elabora un informe que envía a la propia institución y a la agencia de acreditación.

— La institución, o el programa, responde formalmente al informe del equipo de evaluación.

— La comisión de acreditación de la agencia decide conceder, denegar o confirmar la acreditación a la luz del autoestudio, del informe del equipo y de la respuesta de la institución.

La acreditación ha de ser confirmada cada diez años, pero las propias agencias han estimulado en los últimos tiempos métodos de autoestudio continuos para que las propias instituciones estén controlando continuamente el funcionamiento de la institución y de sus programas. De este modo, las agencias de acreditación institucional confían cada vez más en los autoestudios que en las visitas personales a la institución. Por el contrario, las acreditaciones de programas profesionales están más enfocadas a comprobar el grado en el que cada programa se acerca a los criterios de calidad que los profesionales de cada rama consideran relevantes, por lo que la visita personal es considerada primordial en estas agencias.

Es necesario constatar que la acreditación nació a principios de siglo con unos fines claramente elitistas, pero que en la actualidad ha perdido ese carácter y es fundamentalmente el método más generalizado que utiliza el sistema universitario estadounidense para la mejora de la calidad (Kells, 1988b).

V.2.2. **Los criterios de la acreditación**

Las agencias de acreditación funcionan utilizando unos criterios comunes para todas las instituciones. Dada la gran diversidad de instituciones que han de evaluar estas agencias, los criterios que utilizan son necesariamente muy generales. Harris (1978) recoge en estos siete puntos, cuáles son los criterios básicos de la acreditación:

— *Metas y objetivos*

Si la institución posee unas metas y objetivos claros y explícitos, de acuerdo con los cuales va a ser evaluada.

— *Gobierno y estructura*

Las instituciones deben cumplir los requisitos de cualificación del profesorado y las estructuras adecuadas para sus propósitos

— *Validez de los títulos*

Si existe una adecuada relación entre los logros académicos de los estudiantes y la titulación que se concede.

— *Recursos adecuados*

Si existe la suficiente cantidad de recursos humanos, económicos y materiales en la institución para realizar su misión.

— *Estabilidad*

Si se trata de una institución estable, lo que se supone que representa mejor los valores académicos.

— *Estudiantes y programas*

Si las necesidades de los estudiantes, sus intereses y aspiraciones están reflejados en los programas institucionales, y si los servicios relacionados con la misión de la institución y las necesidades de los alumnos les son proporcionados debidamente.

— *Integridad*

Si la integridad institucional es reflejada en las metas de la institución; si existe una gestión transparente; si los métodos de admisión y de promoción del personal son éticamente correctos; si los códigos y las reglas se aplican consistentemente y si existe una práctica continua de autoevaluación y control.

Siendo todos estos criterios comunes para el proceso de acreditación, se podría afirmar, en consecuencia, que toda institución que ha superado la acreditación cumple al menos los siguientes requisitos (Kells, 1988b):

— Tiene establecidos unos propósitos y objetivos claros.

— En gran medida los está alcanzando.

— Estudia continuamente sus logros, sus procesos, sus problemas, e intenta usar los resultados de estos estudios para mejorar la institución.

— Parece satisfacer los criterios de calidad definidos por la agencia de acreditación.

— Parece tener los recursos humanos, financieros y materiales necesarios para continuar alcanzando esos criterios por un razonable período de tiempo.

V.2.3. **Propósitos del sistema de acreditación**

Los dos objetivos básicos del proceso de acreditación ya han sido citados: la mejora de las instituciones y de los programas, identificando aquéllas que están alcanzando mejores niveles de calidad. En Estados Unidos, «la comunidad académica y el público han mirado tradicionalmente a las asociaciones de acreditación para la evaluación de la calidad» (Millard, 1983: 33).

Sin embargo, se pueden enumerar toda una serie de objetivos complementarios que se alcanzan como consecuencia del proceso de acreditación. Siguiendo a Miller (1980) y a Kells (1988b), se pueden citar los siguientes objetivos complementarios, agrupados según que su función sea interna o externa:

A. *Objetivos internos*

— Identificar las instituciones o los programas que están alcanzando los criterios preestablecidos.

— Ayudar a las instituciones a determinar la aceptabilidad de las convalidaciones de créditos obtenidos por los estudiantes en otras instituciones universitarias.

— Proporcionar consejo y asistencia para el establecimiento de nuevas instituciones y programas o para el desarrollo de las ya existentes.

— Estimular la mejora de las instituciones y de los programas por medio de los continuos autoestudios y de la planificación, así como desarrollar el interés del profesorado y de los empleados en estas evaluaciones.

— Estimular la automejora y elevar la calidad de la educación postsecundaria a través del desarrollo de criterios de calidad y de efectividad.

— Proteger a las instituciones contra el anquilosamiento que puede poner en peligro la efectividad educativa y la libertad académica.

B. *Objetivos externos*

— Asesorar a los potenciales estudiantes a elegir la institución más adecuada para ellos.

— Ayudar a identificar las instituciones y programas más adecuados para la inversión de fondos de investigación o para la asignación de fondos federales.

— Garantizar a la comunidad educativa, al público en general, y a otras organizaciones, que una institución o programa tiene unos objetivos apropiados y claramente definidos, mantienen unas condiciones bajo las cuales estos objetivos son alcanzables, que de hecho son alcanzados sustancialmente, y se puede prever que lo seguirán haciendo en el futuro.

— Estimular la diversidad de la educación postsecundaria, permitiendo a las instituciones alcanzar sus objetivos particulares.

— Actuar como un criterio para la concesión de licencias para el ejercicio profesional.

V.2.4. **Ventajas e inconvenientes de la acreditación**

Para unos el proceso de acreditación es el mayor factor de control de calidad en USA, tanto para las instituciones de educación superior, como para los programas profesionales especializados (Harcleroad y Dickey, 1975); protege a las instituciones del control público; aúna a los miembros de la comunidad académica en la búsqueda de las mejoras institucionales; actúa como guardián de la calidad y ha evitado que el proceso de evaluación institucional haya seguido criterios rígidos y no adaptados a cada institución en particular (Millard, 1983).

Para otros, sin embargo, la acreditación es criticable por una serie de motivos (Millard, 1983, Miller, 1979, 1980; Scott, 1983; Troutt, 1981):

— Las asociaciones de acreditación están constituidas por las propias instituciones que son evaluadas, por lo que existe el peligro de autobeneficiarse.

— La acreditación no ha sido capaz de definir la calidad de la educación salvo refiriéndose a criterios específicos que no han sido demostrados.

— Al hacerse la evaluación en función de las metas declaradas por la propia institución, no se evalúa el valor intrínseco de éstas.

— El sistema de acreditación no ha dedicado atención a los resultados del proceso educativo, de modo que no se encuentra relación entre los criterios de acreditación y el éxito profesional posterior de los graduados de las instituciones acreditadas.

— El proceso de acreditación depende de una evaluación subjetiva por parte de otros miembros de la comunidad académica, que por su propia naturaleza puede ser motivo de errores.

— La acreditación utiliza criterios preestablecidos, de modo que tiende a desalentar la innovación y la experimentación.

— El proceso de acreditación es demasiado dependiente de los objetivos que declara tener la propia institución evaluada, de modo que no puede establecer comparaciones con otras instituciones o valoraciones más absolutas respecto a una norma general.

— La acreditación está cada vez más relacionada con la obtención de fondos federales, que con una valoración voluntaria del propio proceso educativo.

— El sistema de acreditación puede ser usado por grupos de interés, más para beneficios privados que para los legítimos propósitos públicos.

— La acreditación es demasiado costosa. La cantidad de esfuerzos humanos y de recursos que exige el proceso de acreditación es elevado.

— La acreditación no sirve como método de evaluación selectivo de la calidad, ya que el 90 % de las instituciones están acreditadas, y la mayoría de las que no lo están es por razones ajenas a la calidad.

De todos modos la crítica más seria que se puede hacer al proceso de acreditación, como a muchos otros mecanismos de evaluación, es que se valoran más las condiciones y recursos de la institución, que los resultados del proceso educativo, es decir, los logros intelectuales, profesionales y humanos de los estudiantes (Troutt, 1981; Scott, 1983).

V.3. AUTOESTUDIOS

El autoestudio, como método para la evaluación de las instituciones universitarias, nace en los años cincuenta en USA como uno de los requisitos

para la realización del proceso de acreditación de las instituciones. Desde entonces ha ido evolucionando y se ha convertido en una práctica habitual en muchas instituciones dentro y fuera de USA, incluso con independencia del proceso de acreditación. «Aunque la acreditación es un proceso útil que puede promover el autoanálisis, es algo periódico y motivado externamente. El autoestudio, por otra parte, puede convertirse en una parte integral y continua de la conciencia de una institución sobre sí misma» (Johnson y Christal, 1985: 2).

V.3.1. **El concepto de autoestudio**

El autoestudio es un método para estimular la autorreflexión, ofreciendo a una institución la oportunidad de preguntarse acerca de sí misma, con el objetivo básico de mejorar la calidad de sus programas y servicios. El proceso de autoestudio es una revisión completa de medios, funcionamiento y resultados de la institución. La realización de este proceso requiere que la institución tenga articuladas sus metas de la forma más concreta y comprensible que sea posible, porque sólo entendiendo las metas de una institución pueden ser adecuadamente valorados sus resultados. Un autoestudio conecta las metas de una institución con los resultados que obtiene, los planes con el producto final. El autoestudio permite encontrar los puntos débiles de una institución, facilitando de este modo la misión de los que han de tomar decisiones para mejorar la calidad de los servicios.

Para algún autor el proceso de autoestudio es tan importante, que es una prueba evidente de la calidad de una universidad: «Las instituciones que fallan en realizar un autoestudio objetivo, crítico y profesional, son, con toda verosimilitud, centros faltos de integridad, y, en poco tiempo, faltos de calidad» (Chambers, 1984: 14).

Un rasgo básico del autoestudio es que debe ser una empresa decidida y conformada por las propias instituciones, y no ordenada por organismos externos. La institución debe encontrar en el autoestudio un mecanismo que satisfaga sus propias necesidades internas de autoconocimiento para saber exactamente qué se está haciendo, con qué grado de perfección y a qué precio, de modo que se pueda establecer un sistema racional de prioridades para aumentar la efectividad de la institución (Furumark, 1981).

Cada vez es más frecuente la realización cíclica de autoestudios, normalmente cada año, como método interno de las instituciones para mantener el control sobre todas sus partes, estimulando así la calidad de su funcionamiento. A pesar del esfuerzo que eso pueda suponer, los beneficios que proporcionan a la institución compensan el esfuerzo (Chambers, 1984). De hecho, el

autoestudio ha pasado de ser una descripción de los recursos disponibles, como era en un principio, a ser una autoevaluación del rendimiento de la institución, más enfocado al estímulo de la calidad de la institución, que a la formación de una imagen externa favorable (Tristchler, 1981).

Hay que puntualizar también que el autoestudio es hoy en día una de las formas más frecuentes de evaluación institucional por parte de las universidades, entre otras causas, ya señaladas, por el temor de las instituciones a ser controladas por organismos gubernamentales, si ellas no se anticipan emprendiendo decididamente la tarea de la autoevaluación (Miller, 1979; Nichols, 1989).

El objetivo básico de un autoestudio es el de analizar globalmente una institución, lo que significa considerar cada una de sus partes esenciales. Pueden citarse las siguientes como objeto esencial del autoestudio (Chambers, 1984):

- La misión, metas y objetivos de la institución.
- Los programas y *currícula*.
- Los resultados que se consiguen y que se esperan conseguir.
- La política de admisión de alumnos.
- Los servicios que se proporcionan a los estudiantes.
- El papel del profesorado en relación con las metas de la institución y con las necesidades de los alumnos.
- La organización administrativa como vehículo que facilita las funciones docente y de aprendizaje.
- El papel del equipo dirigente como garante de que los objetivos están siendo cumplidos.
- Las políticas y prácticas de presupuestos, de planificación y de contabilidad.
- Los servicios de biblioteca y de ordenadores para uso de los estudiantes.
- Lo adecuado de sus edificios y equipamiento.

- Sus catálogos, publicaciones y relaciones públicas.
- Su apertura a la innovación, la experimentación y al crecimiento futuro.

V.3.2. **Los objetivos del autoestudio**

Los objetivos de un autoestudio, en su práctica habitual en USA, se pueden agrupar en dos categorías: una que recoge los objetivos relacionados con la buena marcha de la institución y de sus programas; y otra en la que se incluyen los relacionados con el proceso de la acreditación. En el presente estudio, como es obvio, sólo se considerarán los primeros. Entre ellos cabe destacar los siguientes (Kells, 1988b):

- Ayudar a las instituciones y a sus programas a mejorar clarificando metas; identificando problemas; revisando programas, procesos y recursos; introduciendo los cambios necesarios durante y después del autoestudio para aumentar la efectividad institucional.
- Estimular a las instituciones a que tengan de un modo continuado sistemas de análisis de su situación interna.
- Fundamentar sólidamente todos los intentos de planificación que haga la institución.
- Hacer partícipe al personal de la institución de las metas y de los logros que se alcanzan.
- Apertura de la institución a los criterios y opiniones de todos sus miembros, abriendo estructuras estables de comunicación.
- Revisar las políticas, prácticas y procedimientos del funcionamiento de las instituciones y que la realización del autoestudio puede ayudar a estimular.

V.3.3. **Características de autoestudio**

Un hecho esencial y diferenciador de los autoestudios es que mientras que las evaluaciones del rendimiento educativo hacen hincapié en los aspectos de eficiencia, los autoestudios tienden a fijarse fundamentalmente en la efectividad de las instituciones.

Existen algunas ideas generales que marcan las características esenciales del autoestudio. Siguiendo a Kells (1981) son éstas:

— La motivación primaria de un autoestudio es la que decide el que sea o no efectivo finalmente; es decir, depende de cuáles sean los motivos subyacentes a la evaluación el que ésta tenga o no éxito. Sólo en el caso de que el motivo emane de la propia institución, y esté guiado por la idea de mejorar la marcha de la institución, puede esperarse el éxito final del proceso.

— La efectividad de una institución debe ser evaluada mediante dos modos complementarios: analizando la relación entre metas y logros, y desde un punto de vista organicista, es decir, considerando a la institución como una organización viva.

— Las instituciones están formadas por personas y toda evaluación debe tenerlas en cuenta y utilizarlas en el proceso evaluador.

— Una evaluación debe conducir siempre a estrategias y recomendaciones específicas para cambios útiles.

Basándose en su larga experiencia como promotor de autoestudios, Kells (1988a, 1988b) señala como características deseables para el éxito de un autoestudio, bien de una institución, bien de un programa, los siguientes:

— El proceso debe ser motivado internamente. Si el autoestudio es meramente la respuesta a un imperativo externo, pocos de los objetivos podrán ser alcanzados dada la presumible actitud de recelo de las personas hacia una evaluación externa. Es imprescindible que el autoestudio sea visto por los implicados como una necesidad de la institución, no como una obligación que se impone desde fuera.

— Los dirigentes de las instituciones deben estar comprometidos en el proceso y manifestarlo a todos los componentes de la institución. Si no es así, la utilidad del estudio será muy pequeña. En el caso de universidades dependientes del Estado, se exigirá además para el éxito del proceso que las personas que ejercen algún tipo de autoridad sobre las instituciones estén también implicadas en el proceso de autoevaluación.

— El diseño del autoestudio debe ser apropiado a las circunstancias de cada institución y debe estar «bien hecho», lo que no es fácil de conseguir al primer intento. La experiencia positiva que se saque de una autoevaluación es de una gran ayuda para que en los nuevos intentos la colaboración sea más entusiasta.

— El proceso debe contener un verdadero propósito de clarificar las metas de la organización, y de valorarlas con el fin de mejorar la situación de la institución.

— Deben intervenir en el proceso de autoestudio personas adecuadas, útiles y representativas de todos los estamentos de la institución.

— El proceso debe estar bien dirigido por un grupo eficaz.

— La capacidad de una institución para funcionar efectivamente debe ser estudiada y estimulada.

— Algunas mejoras se deben producir ya durante el proceso de autoestudio.

— Los resultados del proceso deben presentarse en un informe que sea legible para audiencias de diversos niveles e intereses.

— El principal resultado del proceso debe ser la puesta en marcha de un sistema de investigación institucional, de autoanálisis y de automejora.

Jamieson (1988) señala cuáles son las componentes básicas de un autoestudio, es decir, qué información básica debe contener. Son éstas:

— La historia y la situación actual de la unidad evaluada.

— Los objetivos actuales de la unidad.

— Los puntos fuertes y débiles.

— Las matriculaciones previstas en cada unidad docente.

— Los métodos de reclutamiento del personal y flexibilidad de éste.

— La productividad investigadora.

— Las necesidades de recursos, incluyendo las de espacio físico.

— Las perspectivas futuras y las oportunidades de alcanzarlas.

V.3.4. **El proceso del autoestudio**

Poner en práctica un proceso de autoestudio exige seguir todo un conjunto de directrices y métodos que, en el caso de las universidades norteamericanas, están más o menos programados por medio de informes, textos y todo tipo de ayudas, como pueden ser, por ejemplo, los trabajos de Tristchler (1981) o Kells (1988b), en los que se explicitan detalles de cómo llevar adelante el proceso de un autoestudio.

Sólo algunas ideas básicas del procedimiento del autoestudio se pueden señalar aquí, más para dar una idea general que como explicación detallada del proceso.

Según Kells (1988b), un autoestudio debe contestar básicamente a estos cuatro grupos de preguntas:

— ¿Cuáles son las metas de la institución? ¿Son claras, apropiadas y útiles? ¿Existe consenso y comprensión sobre ellas?

— ¿Son los programas y los servicios consistentes con los objetivos? ¿Están designados para alcanzarlos? ¿Funcionan correctamente? ¿Cuáles son los problemas? ¿Cómo pueden resolverse?

— ¿Son los recursos humanos, financieros y materiales disponibles suficientes para llevar adelante los programas y los servicios propuestos? ¿Continuarán estando disponibles en el futuro?

— ¿Están siendo alcanzadas las metas? ¿Cómo puede usarse la evidencia de los resultados ya obtenidos para mejorar la institución?

El mismo autor señala (Kells, 1988b) los pasos que deberán darse para responder a estas preguntas. Son los siguientes:

A. *Preparación y diseño del autoestudio*

— Establecer previamente:

- Dirección.
- Motivación interna.
- Lista específica de necesidades.

— Diseñar el autoestudio:

- Seleccionando el enfoque y las metas.
- Decidiendo el área y la profundidad del estudio.
- Eligiendo la secuencia de actividades.
- Decidiendo la naturaleza de la participación.
- Relacionando el estudio con los problemas actuales de la institución y con su agenda de trabajo.

B. *Organización del proceso del autoestudio*

— Definir tareas y misiones de los participantes.

— Establecer un método como guía del estudio.

- Seleccionar a los participantes, orientarlos y entrenarlos.
- Obtener los recursos necesarios y la asistencia de personal auxiliar.
- Establecer grupos de trabajo.
- Establecer coordinación y mecanismos de comunicación entre grupos e individuos.

C. *Realización del proceso de autoestudio*

- Trabajar con metas preestablecidas.
- Estudiar las entradas al sistema, su entorno, los programas de estudio y el funcionamiento.
- Revisar los criterios que están instituidos y cuyo cumplimiento es obligatorio.
- Recoger hechos, pero también opiniones.
- Discutir los resultados y preparar un informe final.
- Usar los resultados e implementar los cambios que se recomienden.

D. *Uso de colegas de otras instituciones*

- Usar a colegas de otras instituciones como consejeros del proceso durante el autoestudio, e incluso después.
- Utilizar una visita de un equipo externo para promover los cambios que se proponen en las conclusiones.
- Trabajar en colaboración con otros organismos externos implicados en procesos semejantes.

E. *Ciclos de estudio y planificación*

- Usar el autoestudio como base para iniciar un proceso de planificación.
- Incrementar, a partir del autoestudio, la investigación institucional.

Abundando en los últimos puntos citados, Johnson y Christal (1985) proponen el establecimiento en las universidades de una *Oficina de Investigación Institucional* como ente permanente de la institución universitaria, tanto para la realización del autoestudio como para la posterior continuidad en la investigación del propio funcionamiento de la institución. Para estos autores, el paso de un esporádico autoestudio a un estudio continuo sobre el funcionamiento de la institución, con una continua recogida de datos, sólo es posible con una oficina de este tipo.

Furumark (1981), relatando la experiencia sueca en autoevaluación, señala los pasos a seguir para iniciar el proceso:

- Establecer, en una reunión general de los miembros de la institución, la idea y los propósitos de una autoevaluación, asegurándose un consenso sobre el diseño y organización del proyecto.

- Identificar los aspectos débiles y fuertes de todas las actividades de la institución y de la organización de todas las unidades que la componen. Esto se hará utilizando encuestas sobre la opinión de los miembros de la propia institución, pero también se deberá recoger la opinión de personas externas a ella, pero conocedoras de su funcionamiento.

- Formular metas factibles para los cambios que se propongan como necesarios, y recabar la opinión sobre ellas de todas las unidades implicadas en los posibles cambios.

- Proponer y adoptar un programa de acción a corto y medio plazo.

Para el caso de los países europeos, con una estructura de poder diferente al de las universidades norteamericanas, un experto en autoestudios a ambos lados del Atlántico, Kells (1988a), recomienda algunas medidas específicas:

- Con el actual desarrollo de la autonomía de las universidades en muchos países europeos, es necesario crear la capacidad y el deseo de las propias instituciones en su autoevaluación. Paralelamente, es necesario desarrollar la capacidad técnica para la realización de evaluaciones, tanto desde el punto de vista de la Administración central como desde la de las propias instituciones universitarias.

- Quizá la mejor manera de empezar la realización de autoestudios sea a nivel de facultades, preferiblemente que para toda la universidad. En las facultades los datos suelen estar más disponibles, y el temor a los controles

externos es menor. Realizados con éxito estos estudios, se puede pasar a otros más amplios que abarquen toda la universidad.

— La actitud de los gobiernos centrales debe ser la de estimular la realización de los autoestudios, más que la de exigirlos a modo de una inspección. El autoestudio debe realizarse con el único objetivo básico de la mejora de la calidad y efectividad de la institución.

— Quizá algunos profesores europeos, no necesariamente rectores o administradores centrales y locales, deberían conocer *in situ* (USA o Canadá) cómo funcionan los autoestudios.

— Sería muy provechoso establecer conexiones entre distintos países europeos para compartir experiencias en la puesta en práctica de los autoestudios.

V.3.5. **Ventajas e inconvenientes de los autoestudios**

Las ventajas de un sistema interno de evaluación como el que se realiza con un autoestudio son recogidas así por Bauer (1988):

— Las actividades que son evaluadas y los problemas que necesitan resolverse son continuos y no momentáneos, por lo que requieren un seguimiento continuo.

— Con la autoevaluación no sólo se está más cerca de las actividades que son evaluadas que con otros sistemas, sino que también es más fácil pasar a la acción, ya que existe menos resistencia a los cambios que en el caso de las evaluaciones externas.

— La autoevaluación conduce a una amplia pluralidad de valores y jerarquías que, aunque ocasionalmente pueda ser molesto, garantiza la inexistencia de abusos de poder a cualquier nivel.

— Las responsabilidades asumidas por las personas durante el proceso de autoevaluación sirven también como estimulantes de la preocupación de los individuos sobre la marcha de las actividades de la institución que han sido o están siendo evaluadas.

— La iniciativa de la autoevaluación puede partir de diferentes niveles dentro de la institución, y no necesariamente de la cabeza.

Una de las ventajas más importantes de los autoestudios frente a otras técnicas de evaluación, señalada por diversos autores que tratan este tema, está en el hecho de que la actitud del personal evaluado es mucho más positiva hacia el autoestudio que hacia cualquier otra forma de evaluación (Tristchler, 1981).

En cuanto a los problemas que se plantean en la realización de autoestudios, Furumark (1981) señala los siguientes:

— *Definición de los problemas*

Pasos previos a un proceso de autoestudio son: ponerse de acuerdo sobre cuáles son los objetivos del autoestudio, detectar los problemas importantes, llegar a un consenso sobre ellos, sacar a la luz los problemas que están escondidos.

— *Participación y apoyo*

Conseguir participación y apoyo, al menos consentimiento, de todas las instancias implicadas en el proceso del autoestudio, es otra de las dificultades básicas para su realización. Las resistencias al autoestudio deben ser vencidas para que el estudio tenga éxito. Para ello es necesario convencer de los beneficios que se van a obtener, e incluso puede ser útil establecer algún sistema de recompensas que puedan estimular el interés.

— *Uso de los resultados*

Suele ser una queja frecuente la poca repercusión que en la práctica tienen los autoestudios sobre la toma de decisiones en la institución. Es decir, la carencia de un cambio real en la dirección recomendada por los resultados del autoestudio.

— *Grado de institucionalización*

El grado de institucionalización del autoestudio es otro de los problemas que conlleva. Hay que elegir el punto idóneo entre un proceso rígidamente controlado y otro en el que tenga cabida la espontaneidad en cada uno de los pasos que se van dando.

— *Métodos de evaluación*

La autoevaluación exige que todas las personas implicadas en el proceso sean capaces de seguirlo en todas sus etapas sin necesidad de asistencia externa.

— *La credibilidad de la autoevaluación*

Otro problema importante surge del hecho de que las decisiones importantes sobre las instituciones universitarias son ejercidas en los sistemas europeos por organismos gubernamentales fuera de la propia universidad, mientras que los autoestudios son realizados por los individuos de la propia institución con una perspectiva interna. En consecuencia, no está claro que los que toman las decisiones en los organismos externos vayan a confiar en los resultados del autoestudio, y no prefieren complementarlos con estudios externos.

Un firme defensor de la visión gerencial de las universidades, Lindsay (1982), opina que los autoestudios adolecen de los siguientes defectos:

- No utilización del concepto de rendimiento.
- Pocas posibilidades de generar un método amplio aplicable a diferentes instituciones.
- Resultados en forma de un conjunto inconexo de recomendaciones.
- No satisfacen los requerimientos de información gerencial.

Entre las dificultades de los autoestudios, Sizer (1982) señala el conflicto entre la evaluación a distintos niveles institucionales (departamentos, facultades, universidad), tanto en cuanto a los recursos necesarios en cada nivel como a los objetivos y a las perspectivas de futuro, por lo que acusa a estos métodos de poco realistas. Sin embargo, este problema ya es tenido en cuenta en el planteamiento de los autoestudios que se realizan últimamente, como es el caso del modelo de autoevaluación formulado por O'Brien y Waters (1983).

Un problema inherente a los autoestudios es consecuencia del hecho de que no son transportables a cualquier institución sin antes realizar un proceso previo de adaptación. Los problemas concretos de la transposición de estos métodos, originalmente norteamericanos, a la realidad de las universidades europeas, básicamente públicas y con un profesorado funcionario, son señalados por Kells (1988a) en un trabajo expresamente dedicado a esta cuestión.

Algunos autores son muy críticos hacia los sistemas de autoevaluación, ya que utilizan para el proceso de evaluación a los propios miembros de la profesión. Así, por ejemplo, Wright (1989), refiriéndose al sistema de control de la calidad de las universidades británicas, considera que se está perpetuando los intereses particulares de los miembros de la comunidad académica.

Igualmente, Lieberman (1987) critica duramente la utilización de estos sistemas para el caso de los Estados Unidos por razones similares.

V.4. REVISION DE PROGRAMAS

Se utiliza la denominación específica de autoestudio para las evaluaciones dedicadas a analizar las instituciones universitarias como un todo, mientras que para los métodos diseñados para evaluar los distintos programas académicos que están funcionando en una institución, se reserva el término de revisión de programas, revisión departamental o autoestudio departamental.

La revisión de programas es una de las más recientes aportaciones a la evaluación de la educación superior. Su origen se encuentra también en los Estados Unidos, en donde su uso se está extendiendo a la mayor parte de las instituciones universitarias (Conrad y Wilson, 1985). Su interés e importancia radica en el hecho de que analiza unas entidades, los departamentos, en los que se toman la mayor parte de las decisiones que realmente afectan a los estudiantes, a la investigación y a la docencia (Kells, 1988a).

V.4.1. El concepto de revisión de programas

Por revisión de programas se entiende el proceso de recoger y analizar información acerca de un programa académico existente, o de una unidad no docente, para llegar a un juicio sobre la continuación, modificación, promoción o finalización del programa o de la unidad (Craven, 1980). Otra definición alternativa es esta: «Una revisión departamental es una investigación sistemática de la estructura y funciones de un departamento económico con el fin de evaluar y mejorar su rendimiento» (Roe *et al.*, 1986). La palabra sistemática en esta última definición es significativa: indica planificación y continuidad en el proceso de la revisión. Por departamento económico se quiere significar que se trata de una unidad con una cierta autonomía en los recursos.

Holdaway (1988) señala una diferencia entre las palabras evaluación y revisión: mientras que la evaluación puede ser considerada como un control de la unidad evaluada frente a unos criterios previamente establecidos, la revisión indica un estudio más global que describe y valora los antecedentes de la unidad, su rendimiento actual y los planes futuros. Como afirma Slovacek (1988): «Las revisiones de programas representa el esfuerzo institucional más adecuado, serio y continuado para evaluar la esencia de una universidad: sus programas académicos».

Las revisiones de programas, realizadas de un modo sistemático, son algo relativamente reciente. En el caso de los Estados Unidos están relacionadas con el establecimiento de organismos estatales para el control o la coordinación de la enseñanza superior pública en algunos estados, hecho que se inició a partir de los años sesenta. Pero fue incluso posteriormente cuando la mayoría de las instituciones y los sistemas estatales implantaron un modelo sistemático de revisión de programas (Conrad y Wilson, 1985). En estos momentos la revisión de programas se ha convertido en una de las actividades más comunes en las universidades de los países anglosajones (Holdaway, 1988).

Las revisiones de programas que se realizan en los Estados Unidos pueden ser agrupadas en tres categorías, según sea el ente organizador de la revisión:

Revisiones estatales

Un fenómeno reciente es la constitución de organismos estatales para la aprobación de programas en las universidades públicas de algunos estados norteamericanos, e incluso en las privadas de unos pocos. Anteriormente, la responsabilidad sobre los programas estaba exclusivamente en manos de las propias instituciones, pero actualmente existe tendencia a un cierto control público.

Revisiones de sistemas multicampus

También recientemente los gestores de las universidades *multicampus* han desarrollado sistemas de control de programas sobre las diversas unidades, territorialmente diferenciadas, que componen este tipo de universidades (Balderston, 1985).

Revisiones realizadas por las propias instituciones

La mayor parte de instituciones universitarias realizan algún tipo de revisión de programas, tanto para aprobar nuevos programas como para revisar los ya existentes.

La actual extensión y desarrollo de las revisiones de programas en los Estados Unidos se debe a la confluencia de diversos factores que han actuado en la misma dirección (Conrad y Wilson, 1985):

— El generalizado interés en mantener y mejorar la calidad de la educación superior, tanto desde dentro del mundo académico como desde fuera de él, especialmente en momentos de restricciones financieras para las universidades.

— La puesta en marcha de nuevas técnicas de organización y de gestión de las universidades.

— Algunas restricciones económicas causadas por el estancamiento en las matriculaciones, por el incremento de los costes, y por cambios en la financiación de las instituciones.

— Mayor demanda de control externo por parte de estudiantes, consumidores y legisladores.

— Mayor demanda de los gestores de las universidades para que los programas sean más eficientes y efectivos en el uso de los limitados recursos.

Aunque la revisión de programas no es un método definido con exactitud y homogeneidad, y su puesta en marcha depende de los factores individuales de cada programa, hay algunos pasos o procedimientos que son comunes a todas las revisiones. Algunos de los más relevantes son los siguientes (Holdaway, 1988; Young, 1988):

— Determinar quién está solicitando la revisión y cuáles son sus perspectivas.

— Identificar los propósitos específicos de la revisión.

— Identificar las principales cuestiones sobre las que hay que centrar la atención.

— Limitar el campo de extensión de la revisión.

— Identificar el proceso que va a utilizarse y los períodos de tiempo que van a ser necesarios.

— Diseñar métodos para recoger y analizar datos numéricos, información y opiniones.

— Decidir sobre la composición de los comités de revisión y sobre la selección de sus miembros.

— Decidir sobre los criterios que van a ser usados en la valoración del programa y de las personas empleadas.

— Decidir sobre la publicidad de los resultados e informes de la revisión en el transcurso del tiempo.

— Decidir qué medidas deberán tomarse para llevar a cabo las recomendaciones.

— Establecer mecanismos para precisar el grado de acuerdo de la institución con las recomendaciones resultantes de la revisión, así como para evaluar si se están poniendo en práctica esas recomendaciones de la revisión.

Un caso típico de puesta en práctica de una revisión de programas puede constar de estas cuatro fases (Kells, 1988a):

— Un autoestudio preparado por los profesionales del programa bajo revisión.

— Una revisión externa basada en una visita de expertos del mismo campo científico que el programa revisado.

— Una revisión del autoestudio y del informe de los expertos por un grupo de profesionales de la propia institución, pero de otros programas.

— Una respuesta o un acuerdo negociado con la dirección de la institución, como consecuencia de los resultados de los tres primeros pasos.

Aunque el objetivo de una revisión de programa se extiende a la evaluación de todas las partes que constituyen el departamento (medios materiales, personal, organización, etc.), Roe *et al.* (1986) recomiendan separar la revisión del departamento (incluyendo cursos, docencia, medios, etc.) de la evaluación de los individuos, a los que se les debe dedicar una atención específica en revisiones individuales.

Siguiendo este criterio enunciado por Roe *et al.*, es también necesario distinguir la revisión de programas de otros tipos de revisiones que, aunque relacionadas con ella, no son estrictamente institucionales, por lo que no serán tratadas aquí a pesar de su indudable interés, ya que se salen del objetivo primario del presente estudio centrado en la evaluación de las instituciones. Son éstas tres:

Revisiones de resultados

Dedicadas a valorar el aprendizaje conseguido por los alumnos en su permanencia en la educación superior (Ewell, 1985; Banta, 1988).

Revisión de rendimiento docente

Estudio sobre el rendimiento individual del profesorado en su función docente (Roe *et al.*, 1986; Dallöf, 1990a; Dallöf, 1990b).

Revisión del rendimiento investigador

Valoración del rendimiento investigador y de las aportaciones del profesorado a la sociedad (Gibbons, 1985; Creswell, 1986; Rudd, 1988; Staropoli, 1990).

Una versión distinta de las revisiones de programas la constituyen los planes estratégicos, de muy reciente desarrollo en el campo de las instituciones de educación superior. Los planes estratégicos añaden a la revisión de los programas una clara orientación hacia las metas, un espíritu globalizador, están fuertemente ligados a la toma de decisiones y su principal finalidad es la de apoyar el cambio de la organización (Maassen y Vught, 1990).

V.4.2. Objetivos de la revisión de programas

El propósito más claro de la revisión de programas es la mejora y la revitalización de los programas que están funcionando en las instituciones; sin embargo, algunas otras misiones complementarias se les han atribuido (Conrad y Wilson, 1985; Roe *et al.*, 1986; Holdaway, 1988):

- Determinar la efectividad del programa y considerar sus posibles mejoras.
- Seguir los mandatos de algún organismo gubernamental.
- Proporcionar información a los responsables de la toma de decisiones sobre el funcionamiento de los programas, detectando los puntos débiles y facilitando la planificación de la universidad.
- Demostrar la calidad, productividad, necesidad y demanda del programa a la propia universidad, a la sociedad y a la Administración pública.
- Proporcionar datos comparativos con relación a otros departamentos del mismo área.
- Servir como fundamento para la asignación y reasignación de fondos, consiguiendo así una asignación óptima.

El último de los objetivos señalados es uno de los motivos que más influyen en la generalización de las revisiones de programas, hasta tal punto que, después de la mejora de los programas, es la razón más frecuente esgrimida para justificar la realización de estas revisiones (Conrad y Wilson, 1985).

Otro autor (Young, 1988) sintetiza así los beneficios que produce una revisión de programas:

- Mejora la asignación de recursos.
- Ayuda a la toma de decisiones difíciles.
- Fortalece los criterios de admisión.
- Clarifica las misiones institucionales.
- Estimula la cooperación interinstitucional.
- Promueve mejoras en la gestión.
- Proporciona criterios para realizar adecuadamente los recortes financieros que pueden ser necesarios.
- Sirve como un elemento de renovación de los programas en múltiples sentidos.

V.4.3. **Distintos enfoques a la revisión de programas**

Existen distintos modos de enfocar el problema de la revisión de un programa. La elección del más adecuado depende de cuáles sean las intenciones primarias que se persiguen con la evaluación. Aunque estos métodos no son excluyentes, y se pueden encontrar algún tipo de combinación entre ellos, cuatro modelos básicos se pueden definir (Conrad y Wilson, 1985):

A. *Modelo basado en los objetivos*

Es posiblemente el más usado. Su función básica es la de evaluar hasta qué punto el programa está alcanzando sus objetivos. Para ello identifica los objetivos del programa y los criterios de rendimiento preestablecidos; los compara con los datos recogidos sobre el funcionamiento del programa; y determina el grado de congruencia o discrepancia entre los objetivos y la realidad. En estos modelos suele incluirse un análisis de las razones del éxito o del fracaso del programa, de modo que la revisión tenga una función formativa y no sólo finalista. El proceso de una evaluación de este tipo puede ser resumido así:

- Clarificación de las metas y objetivos del programa bajo revisión.

- Identificación de las variables que afectan al rendimiento.
- Delimitación de los criterios bajo los que el programa debe ser evaluado.
- Recogida de datos.
- Comparación de los datos con los criterios previamente establecidos.
- Comunicación de los resultados.

B. *Modelos de funcionamiento*

Este tipo de evaluación centra su atención más en las actividades del programa, que en el logro de las metas que tiene establecidas. En ellos se valoran básicamente los efectos del programa, respondiendo a los requerimientos públicos de información, presentando los éxitos y fracasos del programa desde diversas perspectivas. La idea central de este tipo de revisión es la de recoger la información sobre el programa que puede ser de interés para toda la audiencia que puede estar interesada en su funcionamiento: administradores, profesores, estudiantes y otros grupos sociales. Una revisión de este tipo puede llevarse a cabo siguiendo los siguientes pasos, no necesariamente por orden:

- Entrevistar a los «clientes», al profesorado y a otros grupos interesados.
- Analizar las actividades del programa.
- Desentrañar los propósitos reales del programa.
- Recoger las preocupaciones respecto al programa de los diversos grupos interesados.
- Conceptualizar los problemas básicos.
- Seleccionar a las personas que han de intervenir en la revisión del programa.
- Recoger los datos sobre la realidad del programa.
- Presentar los resultados de la revisión.

C. *Modelo para la toma de decisiones*

Este tipo de revisión está básicamente motivado por la necesidad de reasignar recursos entre distintos departamentos, o la de decidir sobre la

creación o eliminación de algún programa. En este tipo de revisión ni las metas ni las opiniones de los usuarios son necesariamente básicas. El objetivo central de este tipo de revisión es el de preparar la información que facilite la toma de decisiones. Este informe deberá contener estas tres acciones:

- Delimitar los aspectos a evaluar, identificando las situaciones entre las que es posible decidir, y eligiendo los criterios que deben ser usados por los que han de tomar las decisiones.
- Obtener y procesar la información pertinente.
- Analizar los datos, y trasladar los resultados a los que han de decidir.

D. *Modelo del experto*

Este modelo, muy diferente de los anteriores, tiene como pieza básica la opinión de un experto externo al que se le encomienda el análisis del programa. Evidentemente el modelo implica que se puede encontrar una persona capaz de realizar esa misión evaluadora, lo que no siempre será posible. Aquí los criterios son, obviamente, los del propio evaluador, que es el que conduce la marcha de la evaluación en todos los aspectos. En algunos casos la evaluación de este tipo ha sido realizada por un pequeño equipo, pero utilizando los mismos criterios de autonomía que en el caso de un sólo experto.

V.4.4. **Criterios a utilizar en la revisión**

Todas las instituciones que inician una revisión de programas establecen algunos criterios previos para la evaluación. Conrad y Wilson (1985), enuncian una serie de puntos que deben formar parte de una revisión de programa. Según estos autores, en toda revisión de programa es necesario conocer en qué nivel se encuentran los siguientes aspectos del programa:

- *Calidad:*
 - Del profesorado.
 - De los estudiantes.
 - Del currículum.
 - De los servicios de apoyo: biblioteca, laboratorios, ordenadores, equipamiento general.
 - De los servicios administrativos.
 - De los recursos financieros.

— *Necesidad del programa:*

- Para la misión de la institución.
- Para la sociedad.

— *Demanda:*

- Demanda actual y demanda previsible de alumnos en el futuro.
- Demanda de graduados por la sociedad.

— *Costes:*

- Coste-efectividad.
- Costes y beneficios no pecuniarios.

Por su parte, Roe *et al.* (1986), relatando su experiencia en las universidades australianas, opinan que una revisión de programa debe centrarse en los cuatro puntos siguientes:

— La calidad de la docencia y del currículum. Lo apropiado y equilibrado de este último.

— La calidad y el equilibrio entre las actividades profesionales, de investigación y de desarrollo, dentro del departamento.

— Si los servicios, equipamiento, asignación de recursos, administración, personal, y el conjunto de las políticas institucionales están sirviendo de un modo adecuado al departamento y a los estudiantes.

— Si existe una adecuada relación del departamento con otros departamentos similares, con el resto de la institución y con otros organismos externos de interés para el programa.

Young (1988) indica las características que debe tener una revisión de programas para que sea efectiva, es decir, útil para el propio programa, para la institución y para la sociedad:

— El proceso de revisión debe ser periódico.

— La calidad del programa debe ser evaluada utilizando múltiples parámetros representativos de las diversas facetas de la calidad.

— La revisión debe dedicar especial atención al funcionamiento del programa, a cómo está articulado, a cuál es la respuesta del mercado, a cuál es la eficiencia de la gestión, a cómo están programados los servicios, etcétera.

— Los resultados deben permitir a los gestores de la institución universitaria el tener una visión global de la calidad de los programas, así como conocer cuáles son sus puntos débiles.

— Una consecuencia de la revisión de programas debe ser la reasignación de recursos.

— La revisión debe permitir descubrir cuáles son los requisitos imprescindibles para la calidad de un programa.

— Las recomendaciones de una revisión deben ser implementadas.

— La revisión debe conducir a un incremento en las relaciones y en la comunicación entre los miembros del departamento.

Las instituciones que realizan evaluaciones de programas cíclicamente suelen utilizar algunos criterios para evitar que todas las revisiones se realicen a la vez. Algunos de estos criterios son los siguientes:

— *Criterios explícitos*

Como un porcentaje del total de los programas cada año, o un número determinado, un área de estudio cada vez, o algún tipo de agrupación de departamentos por afinidades.

— *Indicadores cuantitativos*

De modo que se inicie la revisión por aquellos programas que más lo necesitan según algunos indicadores (coste, productividad, carga docente, etc.). En algunos casos, sólo se realizan las revisiones cuando algún indicador negativo da una señal de alerta.

— *Irregularmente*

Algunas instituciones sólo realizan revisiones de programas cuando se va a nombrar un nuevo director del departamento, o cuando aparecen problemas específicos sobre el número de alumnos o sobre la calidad.

Otro de los problemas que se plantea antes de la realización de una revisión de programas es la elección de las personas que van a llevar a cabo la revisión. En algunas instituciones ésta es una responsabilidad de un grupo de personas de la propia universidad, específicamente dedicadas a esa misión. En otras se

usan evaluadores externos a la institución. En ambos casos se pueden o no combinar con la participación de miembros del departamento bajo revisión. Esquemáticamente estas son las posibilidades más utilizadas en cuanto a los agentes evaluadores:

- *Autorrevisiones*
- Por el profesorado del programa.
- Por alumnos del programa (incluidos antiguos alumnos).
- Por personal de dentro de la institución:
 - Profesores de otras materias.
 - Individuos especializados en la realización de revisiones.
- Por personal de fuera de la institución:
 - Profesores del mismo tipo de programa en otras instituciones.
 - Evaluadores profesionales.
- *Revisión interna-externa.*
- Por grupos mixtos internos y externos.
- Revisiones paralelas realizadas simultáneamente por dos grupos, uno interno y otro externo.

Roe *et al.* (1986) recomiendan decididamente la utilización del enfoque mixto, en el que, tras un proceso de autoestudio, un equipo basado en miembros de otras instituciones realice una evaluación final. Esta doble evaluación permite contrastar la visión complementaria de unos y otros sobre la realidad del programa evaluado.

V.4.5. **Algunos problemas de la revisión de programas**

Algunos problemas son comunes a la mayoría de las revisiones de programas. Conrad y Wilson (1985) citan estos seis como los más significativos:

A. *Conseguir la concordancia entre propósitos múltiples*

Si los fines que se persiguen por la evaluación son múltiples, el hacerlos concordar en el mismo proceso de evaluación puede ser un grave problema.

Situaciones contradictorias pueden aparecer con facilidad. Entre las más frecuentes pueden estar:

— Combinar la reasignación de recursos con la mejora de la calidad del programa.

— Combinar los intereses de la propia institución con los de algún organismo gubernamental con responsabilidades sobre la institución.

B. *Elegir un modelo de evaluación.*

Es necesario elegir correctamente, para cada revisión en concreto, y en función de la finalidad perseguida, un modelo de evaluación, o una combinación de ellos. De esta elección depende muy fundamentalmente el resultado final del proceso.

C. *Elegir un criterio de calidad.*

Dado que se van a hacer estimaciones de la calidad de los programas es necesario adoptar un punto de vista sobre ese concepto. Como se trató en el capítulo II, cuatro son los enfoques básicos sobre este problema:

- El reputacional.
- El de los recursos disponibles.
- El de los resultados obtenidos.
- El del valor añadido.

Elegir una u otra de las anteriores acepciones de calidad es otra de las dificultades que han de superarse antes de iniciar una revisión de programas (George, 1982).

D. *Uso de evaluadores externos*

Una decisión que hay que tomar es la de si se van a usar evaluadores externos, ya sea de otras instituciones o de la misma institución pero de diferentes programas.

E. *Uso posterior de la evaluación*

Uno de los problemas más serios de una revisión de programa es el modo con el que otros pueden hacer uso de los esfuerzos de una revisión. Una de los

principales quejas es la de que las revisiones no son utilizadas para los fines concretos que se propusieron y quedan arrinconadas a perpetuidad. Cuatro razones pueden explicar la no utilización de las revisiones:

- La inercia de toda organización.
- La ineptitud de la revisión para definir medidas de acción que puedan ser llevadas a cabo.
- La posible incertidumbre previa acerca de la necesidad de hacer la evaluación.
- La existencia de múltiples fuentes de información dispares, actuando sobre los que han de tomar las decisiones.

Esta situación de indecisión puede ser evitada si se toman medidas como las siguientes:

- Clarificar los propósitos de la revisión.
- Involucrar a los que finalmente han de tomar decisiones, en todos los procesos de la evaluación.
- Maximizar la comunicación.
- Comprender el entorno político de la institución.
- Reconocer la subjetividad de la evaluación.
- Disponer de un servicio competente de investigación institucional que permita disponer de datos fiables.

F. *Evaluación de la revisión*

Cuando los resultados de la revisión son puestos en práctica, aparece otro problema que conviene tener en cuenta: el impacto que produce la revisión sobre el funcionamiento del programa. La primera pregunta que se puede uno plantear es si realmente los resultados de una revisión son beneficiosos, o al menos tan beneficiosos como para compensar los esfuerzos realizados. Todo parece indicar que las revisiones de programas son fructíferas cuando se realizan a nivel de departamento, aunque su utilidad es más moderada cuando se evalúan programas más extensos (Conrad y Wilson, 1985).

Para ayudar a resolver satisfactoriamente todos los inconvenientes mencionados, Conrad y Wilson (1985) hacen algunas recomendaciones. Según estos autores, para conseguir que las revisiones de programas sean todo lo útiles que sería deseable es aconsejable que se cumplan estos requisitos:

— El propósito de una revisión debe ser único, y debe estar ampliamente aceptado y comprendido por todos los implicados. Sólo se pueden hacer concordar propósitos diferentes cuando éstos se mueven en la misma dirección.

— Las instituciones deben crear un sistema regular para la revisión de todos sus programas.

— Las instituciones deben adaptar los modelos de revisiones ya existentes a las necesidades peculiares de ellas mismas.

— La revisión debe ser flexible, de modo que pueda acomodarse a los cambios que parezcan necesarios sobre la marcha, así como a las interacciones que aparezcan entre los que evalúan y los que están siendo evaluados.

— La selección de las unidades a revisar debe realizarse siguiendo un método sistemático.

— La mayoría de las cuestiones que están siendo evaluadas necesitan varios indicadores, y *nunca se deben simplificar las respuestas más allá de lo razonable.*

— Las evaluaciones de calidad deben incluir un amplio espectro de indicadores que recojan las múltiples facetas de ésta.

— No es imprescindible la utilización de evaluadores externos.

— Las instituciones deben ser conscientes y cautelosas con las conexiones entre la revisión de los programas y otras cuestiones de política interna de la institución.

— La revisión debe estar relacionada con la toma de decisiones, pero no exclusivamente realizada bajo ese punto de vista.

George (1982) hace una recomendación adicional: el que realiza la revisión debe tener presente la idea del principio de incertidumbre de Heisenberg, que trasladado de la física a la investigación institucional vendría a decir: «el querer

revisar los programas con demasiada precisión puede producir perturbaciones en ellos que no son deseables».

Kells (1988a) señala que es difícil encontrar revisiones de programas que sean realmente útiles. No porque no lo sean en sí, sino porque, mientras las instituciones como tales suelen estar preparadas para los autoestudios, a nivel departamental no existe ni la preparación humana ni las dotaciones que son habituales en las evaluaciones a nivel institucional. Además, los departamentos suelen tener un acceso menor a la información, y se encuentran más aislados de los centros de poder institucional. En cambio, el interés de los miembros de un departamento por las mejoras puede estar más profundamente sentido, lo que facilita enormemente la realización de la revisión. Kells indica también que donde mejor resultado han dado las revisiones de programas ha sido en aquellos procesos de evaluación que han contado con una dirección extremadamente buena, que han sido internamente motivados, y que han transcurrido en un ambiente «tranquilo», relativamente «privado» y despolitizado.

CAPITULO VI. VALORACION DE LA CALIDAD EN LAS UNIVERSIDADES ESTADOUNIDENSES

Just as Americans are interested in the top ten movies of all time, in the largest corporations in any year, and so on, we are interested in knowing which are the best colleges, the best graduates programs in economics. And those in the institutions of higher learning are also interested in knowing where they rate, less as a guide to how to do better than as discussion topic over cocktails.

Solmon, 1981: 8

VI.1. INTRODUCCION

La enorme afición que existe en los Estados Unidos a establecer comparaciones entre todo tipo de objetos, personas o instituciones ha generado desde hace ya bastantes años un amplio conjunto de estudios que intentan hacer evaluaciones comparativas de la calidad de las universidades del país. Ya en 1910 J. M. Cattell realizó una ordenación de las mejores universidades norteamericanas (Webster, 1985). Este trabajo tuvo continuación, hasta tal punto que, con anterioridad a 1950, se pueden citar 48 estudios dedicados a establecer algún tipo de ordenación jerárquica de las universidades o de los departamentos, en función de alguna valoración de la calidad (Webster, 1986). Como veremos, son incontables los estudios de este tipo realizados con posterioridad a esa fecha. Aparte de unos pocos estudios de carácter global que han evaluado una gran diversidad de disciplinas académicas al mismo tiempo, la mayor parte de estos estudios hacen referencia a una sola especialidad. Es curioso señalar que las disciplinas más estudiadas han sido las correspondientes a los departamentos de psicología, sociología y economía (Webster, 1985).

Desde un punto de vista teórico, como ya se discutió en el capítulo II, existe una indudable dificultad en el planteamiento de la evaluación de la calidad, tanto en el caso de las instituciones universitarias como en el de un programa de estudios determinado. Básicamente, dos son los problemas a los que se enfrentan estas evaluaciones (Blackburn y Lingenfelter, 1973):

— Definir el criterio de calidad.

— Cuantificar los criterios de modo que se puedan hacer comparaciones entre distintos programas.

Los distintos investigadores han resuelto ambos problemas de modos diferentes, lo que ha dado lugar a un amplio espectro de estudios que valoran la calidad de las universidades estadounidenses desde todos los puntos de vista, con todo tipo de técnicas, con propósitos diferentes y con resultados para todos los gustos. Si se utiliza como criterio clasificatorio la metodología básica utilizada para la evaluación, éstas pueden clasificarse en tres categorías:

— Estudios que utilizan encuestas de reputación.

— Estudios que usan medidas objetivas.

— Estudios basados en la correlación entre medidas objetivas y resultados de encuestas.

A cada uno de estos tres tipos de estudios se dedicará una sección en el presente capítulo, mientras que en la última sección se discutirán las ventajas e inconvenientes que presentan cada uno de los métodos. Antes de entrar en la descripción de estos estudios hay que citar algunos trabajos que hacen una revisión de las investigaciones sobre la valoración de la calidad en las universidades norteamericanas. Son éstos: Smith y Fiedler (1971), Blackburn y Lingenfelter (1973), Miller (1979), Lawrence y Green (1980), Conrard y Blackburn (1985a), Tan (1986) y Webster (1986). En ellos, y especialmente en los más recientes, se recoge una amplia información sobre el tema que nos ocupa.

VI.2. ESTUDIOS BASADOS EN ENCUESTAS

Estos estudios son llamados habitualmente estudios reputacionales o revisiones por colegas (*peer review*), y su característica esencial es la de utilizar encuestas, contestadas por miembros de la propia comunidad universitaria, como medio para valorar comparativamente el nivel de calidad de las universidades o, más frecuentemente, de los departamentos de determinadas disciplinas. El método consiste básicamente en la selección de uno o más criterios para evaluar la calidad, encuestar a un grupo de expertos para que valoren cada centro de acuerdo con esos criterios y agregar finalmente todas las respuestas para establecer unas clasificaciones jerárquicas de los departamentos de cada disciplina académica o de las instituciones globalmente. La

premisa que subyace bajo esta metodología es la que los expertos más conocida en una disciplina son los que están más capacitados para evaluar los programas académicos de su propia materia a lo ancho y largo del país (Hensley, 1980).

En este tipo de estudios lo que se está realmente evaluando es la reputación, esencialmente en el aspecto investigador del profesorado de los departamentos, por lo que otra idea subyacente en esta metodología es la de que la calidad de los profesores es la principal responsable de la calidad del departamento, hecho este que es evidentemente discutible y, de hecho, es fuertemente discutido; así, por ejemplo, Young *et al.* (1989) encuentran que la relación existente entre las estimaciones de la calidad de los individuos y las de los programas en conjunto son muy pequeñas. Una gran cantidad de autores consideran que utilizar este criterio es uno de los puntos más débiles de este método de evaluación. Aunque la mayoría de las evaluaciones de la calidad de las instituciones han utilizado al propio profesorado para responder a las encuestas, algún estudio ha utilizado otros sistemas, como es el caso de Wise *et al.* (1981), que utilizaron las respuestas de los antiguos alumnos como criterio para evaluar la calidad de los departamentos.

Se puede afirmar que los estudios reputacionales han dominado la escena de la evaluación de la calidad en las universidades norteamericanas, especialmente a partir del estudio organizado por el *American Council of Education* (Cartter, 1966), al que, «como discípulos siguiendo a un líder religioso» (Lawrence y Green, 1980: 6), todos los demás han intentado emular. Estos estudios, a pesar de las numerosas críticas que reciben, gozan de buena aceptación entre la comunidad científica más competitiva e incluso entre el público en general, las revistas de información general y los periódicos diarios. A pesar de que la validez de estos métodos de evaluación sea discutible, la mala utilización que a menudo se hace de los resultados de estos estudios es posiblemente más atribuible a esa popularidad, con cierta dosis de sensacionalismo, que a los propios estudios en sí (Lawrence y Green, 1980).

Los estudios reputacionales pueden clasificarse, según el tipo de programa al que se dedican, en tres categorías:

- Estudios sobre programas de postgrado (*graduate programs*), es decir, programas que conducen a la obtención del doctorado.
- Estudios sobre programas de pregrado (*undergraduate programs*), es decir, los que conducen a las titulaciones de *bachelor*, que podrían identificarse con las licenciaturas no profesionales españolas.

— Estudios sobre instituciones que conceden títulos profesionales, como derecho, medicina, educación, etc. (*professional schools*).

Dada la gran diferencia que existe en Estados Unidos entre los tres tipos de programas citados, los centros que se dedican a cada uno de ellos tienen peculiaridades muy específicas, por lo que las evaluaciones de calidad se realizan separadamente para cada uno de los tres tipos.

VI.2.1. Evaluación de programas de postgrado

Estos estudios están referidos exclusivamente a programas de postgrado en departamentos universitarios que conceden títulos de doctorado. Hay que tener en cuenta que en el complejo y diversificado mundo universitario norteamericano existen unas 3.000 instituciones de educación superior, de las cuales sólo unas 300 tienen programas de doctorado. Justamente, estas universidades, con gran dedicación a la investigación, son las más prestigiosas y las que concitan mayor interés para estas evaluaciones, pero no debe olvidarse que sólo constituyen una parte relativamente pequeña del total del sistema universitario norteamericano, y no deben generalizarse a otras instituciones las ideas sobre calidad que pueden extraerse de esas pocas. (Una clasificación de las universidades norteamericanas en función de los grados que conceden y de su dedicación a la investigación puede verse en Carnegie Foundation, 1987.)

Siguiendo un orden cronológico en la exposición, los más relevantes estudios de carácter reputacional sobre las instituciones que conceden doctorados son los siguientes:

A) *Estudios pioneros*

El primer intento significativo de medir la calidad de los departamentos universitarios se debe a Hughes (1925), quien ordenó 38 de las universidades más prestigiosas, teniendo en cuenta 20 especialidades académicas en las que se otorgaban doctorados. El único criterio utilizado era el del prestigio del profesorado, para lo que se recurría a la opinión de otros colegas de las mismas especialidades (Smith y Fiedler, 1971).

Una segunda encuesta realizada por el mismo autor con posterioridad (Hughes, 1934) ampliaba el estudio hasta clasificar 59 universidades, utilizando 35 especialidades. En este estudio se valoraba, además del profesorado, los medios de que disponían las universidades. Los estudios de Hughes

sientan las bases con las que serán construidos otros posteriores: utilización de expertos como evaluadores de la posición de cada centro en una escala prefijada, dedicación exclusiva a los centros con estudios de doctorado y atención tan sólo a unas pocas universidades punteras (Lawrence y Green, 1980); pero, sobre todo, se fija el criterio de tomar como medida de la calidad del programa la calidad (o mejor dicho, la reputación) del profesorado (Conrad y Blackburn, 1985a).

Bastantes años después, Keniston (1959) realizó otro estudio del mismo tipo para 25 universidades y 24 especialidades diferentes, usando a los directores de departamento como evaluadores. El resultado final es una ordenación jerárquica de los departamentos para cada uno de los 24 programas analizados, junto con una ordenación de las 20 instituciones que globalmente consiguen la más alta puntuación. Este estudio, así como los dos de Hughes, fueron fuertemente criticados por: lo pequeño de la muestra, el uso de pocos evaluadores y el sesgo numérico de éstos hacia unas pocas instituciones prestigiosas, muy tradicionales y ubicadas preferentemente en el noreste del país (Tan, 1986; Conrad y Blackburn, 1985a).

B) *Estudios promovidos por el ACE (American Council of Education)*

Algunas de las críticas anteriores sirvieron para mejorar el método seguido en nuevos estudios. Así lo hicieron los dos estudios organizados por el *American Council of Education*: Cartter (1966) y Roose y Andersen (1970). El estudio de Roose y Andersen es una mera actualización y ampliación del de Cartter, ya que utilizan la misma metodología, pero consideran un 55 por 100 más de programas.

El tamaño de la población considerada aumentó notablemente respecto a anteriores evaluaciones, aunque seguía reducida a las universidades más prestigiosas. Cartter consideró 106 universidades y 29 disciplinas distintas, mientras que Roose y Andersen ampliaron la evaluación a 130 universidades y 36 especialidades.

En ambos estudios los evaluadores fueron directores de departamentos, profesores distinguidos de las especialidades bajo escrutinio y una selección de jóvenes profesores de esas mismas especialidades. A cada uno de ellos se le pedía que clasificara los departamentos de su especialidad según dos criterios: calidad del programa (con seis niveles distintos, desde distinguido a insuficiente) y efectividad (cuatro niveles, desde extremadamente atractivo a nada atractivo). Para cada criterio y para cada especialidad se calcularon los valores medios, clasificándose los departamentos según la categoría media

que habían alcanzado. Los resultados para los departamentos clasificados en las dos primeras categorías se presentaron por orden jerárquico, mientras que en las restantes se incluían las instituciones solamente por orden alfabético. Hay que señalar que en ambos estudios los resultados son ordenaciones jerárquicas de los departamentos, y en ningún caso se presentan resultados agregados por universidades. Sin embargo, otros autores han agregado posteriormente los resultados, usando criterios personales de discutible validez y con fines poco relacionados con el interés científico en la calidad institucional.

En los dos estudios del ACE se dedica una destacada atención al análisis de las condiciones de validez estadística de la encuesta, así como a la comparación de los resultados con los que se obtienen en otros estudios (Somit y Tanenhaus, 1964; Bowker, 1964), que también realizaban ordenaciones jerárquicas de los departamentos, utilizando como criterio de evaluación algunas medidas objetivas de calidad, como puede ser el número de publicaciones del profesorado, el tamaño de la institución, el número de volúmenes de las bibliotecas, etc. Cartter encuentra los resultados tan consistentes, que afirma: «Parece que si se hubieran incluido suficientes factores para la construcción de un llamado índice objetivo..., los resultados de nuestro índice de evaluación subjetivo serían casi un duplicado de los suyos» (Cartter, 1966: 118). Una de las conclusiones de Cartter es que se necesitan al menos 50 evaluadores para llegar a resultados fiables.

El trabajo de Cartter, y posteriormente el de Roose y Andersen, recibieron fuertes críticas y numerosos elogios, pero, en cualquier caso, tuvieron una importante repercusión en el ámbito universitario norteamericano. Una de las mayores críticas que se formularon sobre la metodología de estas dos evaluaciones es la de que ignoran a un gran número de instituciones de educación superior, no teniendo en cuenta a la gran mayoría de las instituciones que no están en la cúspide (téngase presente que de las 3.000 instituciones de educación superior norteamericanas, sólo algo más de un centenar son evaluadas en estos estudios). Asimismo, se criticaba el que sólo se evaluaran los programas de doctorado y no otros programas profesionales de indudable importancia e interés, como pueden ser los de educación, medicina o negocios. Otra crítica importante es que estas evaluaciones no están midiendo realmente la calidad, sino sólo la reputación de los departamentos, lo que puede tener implicaciones negativas para la libre competitividad y para la creatividad, ya que por el sistema de evaluación empleado, sólo los departamentos más tradicionales y ortodoxos son valorados como de alta calidad (Dolan, 1976). Otro problema de utilizar la reputación como sinónimo de calidad, es que ésta puede estar muy relacionada con el tamaño de la

institución, ya que en los departamentos mayores existe una productividad total alta, lo que no es necesariamente sinónimo de calidad. Otro defecto achacado a estos estudios es el de que no contemplan otras medidas de la calidad distintas, y sin duda importantes, como pueden ser las medidas objetivas (Dolan, 1976; Webster, 1981). Otro motivo de controversia fue el panel de expertos utilizados para realizar la evaluación: existía un número desproporcionado de evaluadores que provenían justamente de las universidades de más prestigio, por lo que se daba el riesgo de que se produjera un efecto *alma mater*, es decir, una tendencia a valorar mejor las universidades de las que provenían los propios evaluadores (Lawrence y Green, 1980). En cualquier caso, estos estudios, por su rigor y relativa amplitud, causaron un fuerte impacto en la comunidad universitaria, y han servido de modelo para todos los posteriores.

Posteriormente a los estudios de la ACE se realizaron algunos otros en los que también se utilizaban encuestas de reputación como criterio de evaluación de programas de doctorado, aunque ninguno de ellos es tan amplio como aquéllos. Uno de estos estudios es el de Ladd y Lipset (Scully, 1979), que utiliza la misma metodología simple que los más antiguos de Hughes y Keniston: una amplia encuesta en la que los encuestados opinaban sobre cuáles eran los mejores departamentos en 19 disciplinas.

C) *Estudio del ETS (Educational Testing Service)*

Otro estudio, realmente muy interesante, es el promovido por el *Educational Testing Service* (Clark *et al.*, 1976; Hartnett *et al.*, 1978). Básicamente es un ensayo metodológico para diseñar un sistema de evaluación de la calidad en el que se combinen las encuestas de reputación con las medidas objetivas. Como tal ensayo, no es un estudio global, sino que está dedicado a sólo tres especialidades que cubren campos diferenciados y con características estructurales presumiblemente diferentes: química, historia y psicología.

Previamente al estudio, se realizó una encuesta dirigida a los directores de los departamentos de estas especialidades, para que opinaran sobre qué variables eran las que mejor determinaban, en su opinión, la calidad de un programa académico. Los resultados de esta encuesta previa, ordenados según la valoración media dada por los directores de departamento, fue la siguiente (Miller, 1977):

- Capacidad académica de los estudiantes.
- Logros, conocimientos y habilidades conseguidos por los alumnos al completar el programa.
- Apoyo financiero del programa.

- Biblioteca.
- Formación académica del profesorado.
- Propósito del programa.
- Laboratorios.
- Oferta de cursos.
- Política de admisión.
- Medidas de bienestar para el profesorado.

Tras esta encuesta previa, se tomó una muestra de 25 departamentos en cada una de las especialidades mencionadas. Se realizó una encuesta de reputación que se completó con datos sobre profesores, estudiantes y antiguos alumnos de los departamentos de la muestra, así como de algunas de las características del programa en cada uno de los departamentos. Las evaluaciones reputacionales fueron combinadas con otras informaciones objetivas hasta completar un total de 27 variables que se agrupaban en cinco apartados:

- Características del profesorado (nueve variables).
- Características de los estudiantes (seis variables).
- Recursos físicos y financieros (cuatro variables).
- Ambiente general (dos variables).
- Funcionamiento del programa (tres variables).
- Antiguos alumnos (tres variables).

El conjunto de variables seleccionadas fueron analizadas para estimar sus relaciones mutuas. Las variables se agruparon en dos conjuntos según su mayor relación con la docencia o con la investigación:

— Variables que son indicadores de investigación, como tamaño del departamento, reputación del profesorado, capacidad de los estudiantes, número de publicaciones del profesorado, recursos físicos y financieros, etcétera.

— Variables que son indicadores de experiencia educativa, y que están relacionados con el funcionamiento del proceso educativo y el ambiente académico, como son las relaciones interpersonales, la proporción de alumnos que leían tesis, etcétera.

En su análisis, los autores encuentran una alta correlación entre las variables que componen cada uno de los grupos, pero no entre variables procedentes de los dos diferentes grupos, lo que parece confirmar la tesis de que la reputación, obtenida básicamente por la investigación, y la calidad real del programa son cosas relativamente independientes (Clark *et al.*, 1976; Tan,

1986). Los resultados finales de este estudio no hacen referencia concreta a ninguna institución, ya que objetivo y su interés son puramente metodológicos.

Una interesante incorporación que hace este trabajo es un análisis de la correlación que existe entre las evaluaciones reputacionales realizadas por colegas, las evaluaciones basadas en medidas objetivas, las evaluaciones reputacionales realizadas por los estudiantes y las realizadas por los antiguos alumnos. Los resultados son realmente notables: no aparecen correlaciones importantes entre las evaluaciones de unos y otros, de tal modo que la calidad de la enseñanza valorada por los alumnos y ex alumnos, y la calidad de los profesores evaluada por sus colegas tienen correlaciones muy bajas e incluso negativas (Hartnett, *et al.*, 1978). Una vez más se patentiza la diferencia entre lo que podríamos llamar calidad investigadora, de lo que sería la calidad docente, que son las que, respectivamente, están valorando profesores y estudiantes.

Otro estudio que también utiliza una encuesta de reputación, en combinación con otras variables objetivas, es el de Bingham *et al.* (1981). En este trabajo se valora la calidad de los programas de Asuntos Urbanos de 31 departamentos universitarios, utilizando tres indicadores: los resultados de una encuesta a profesores de esta materia; el número de libros publicados por los profesores del programa y número de citaciones dedicadas a miembros de cada departamento. Es decir, tres medidas relacionadas con la reputación, la productividad y la calidad de la investigación. Además de establecer una ordenación jerárquica de los departamentos del área, este estudio concluye que no existe una correlación demasiado elevada entre la reputación y las medidas objetivas, lo que es achacado por los autores al *efecto halo*, es decir, al prestigio de la institución en la que se encuentra el departamento.

D) *El estudio de la CBARC (The Conference Board of the Associated Research Councils)*

El más reciente y amplio estudio sobre la calidad de los programas de doctorado es el promocionado por *The Conference Board of the Associated Research Councils*, publicado en cinco volúmenes (Jones, *et al.*, 1982). En él se pasa revista a 32 disciplinas de doctorado en 228 universidades, de modo que en total son revisados 2.699 departamentos. Una extensión sensiblemente superior a la de cualquier otro estudio realizado anteriormente sobre la calidad de las universidades.

El estudio trata de evitar las críticas que estudios anteriores habían provocado. Para ello se propuso usar complementariamente a la valoración reputa-

cional todo un conjunto de medidas objetivas. Para la elección de estas medidas se consideran tres criterios:

- Que la variable pueda ser relacionada con la calidad del programa.
- Que los correspondientes datos sean posibles de obtener.
- Que sean aplicables a la mayoría de las disciplinas objeto de revisión.

Con esos criterios se seleccionaron las siguientes 15 variables agrupadas en cuatro áreas:

1. *Tamaño del programa*

- Número de profesores en el programa en el correspondiente año.
- Número total de graduados en el programa durante los cinco años anteriores.
- Número de estudiantes matriculados en el año correspondiente.

2. *Características de los graduados*

- Número de graduados que han recibido alguna clase de beca o ayuda durante los cinco años anteriores.
- Número medio de años que tardan en completar los estudios.
- Proporción de graduados que tenían empleo ya previsto antes de acabar los estudios en los últimos cinco años.
- Proporción de graduados con empleo previsto, dentro del mundo académico, en los últimos cinco años.

3. *Resultados de la encuesta de reputación*

- Reputación del profesorado.
- Reputación de la efectividad del programa.
- Reputación de las mejoras del programa en los últimos cinco años.
- Familiaridad de los evaluadores con el programa.

4. *Tamaño de la biblioteca*

- Índice compuesto con datos de la biblioteca.

5. *Investigación*

- Gastos en investigación del departamento.
- Publicaciones del profesorado en los tres últimos años.
- Proporción de profesores con ayudas de investigación en los últimos tres años.
- Estimación de la influencia que tienen los artículos publicados por el profesorado.

Para cada una de estas variables se evaluaron todos los programas de la muestra. Los resultados fueron agrupados en cinco áreas: humanidades, ciencias sociales y del comportamiento, ciencias biológicas, ciencias físicas y matemáticas e ingeniería. Hay que resaltar que en el grupo de humanidades (arte, música, lenguas, literatura, etc.) no fueron usadas en la evaluación las variables correspondientes a investigación.

Este trabajo es, sin duda, el más riguroso de los realizados para evaluar la calidad de las universidades norteamericanas. Un hecho importante es que los resultados fueron presentados sin establecer ningún índice compuesto de calidad, sino simplemente presentando los resultados para cada una de las variables consideradas. Por tanto, cada departamento puede ser ordenado jerárquicamente según 15 criterios diferentes. Los autores no realizaron tampoco ninguna combinación de estas ordenaciones por departamentos, con el fin de establecer algún tipo de ordenación global de las universidades. Todas las presentaciones de resultados combinados para las universidades que se han realizado con posterioridad han sido hechas por otros autores, buscando más el resultado llamativo que el rigor científico.

Este estudio analiza también la validez de las variables reputacionales, estudiando las correlaciones entre todos los grupos de variables utilizados. Los resultados del análisis confirman la existencia de un alto grado de fiabilidad para los resultados de las encuestas reputacionales en todos los campos científicos revisados.

Este estudio también aborda algunas de las críticas habituales a los estudios reputacionales. El hecho de que los evaluadores tengan mayor o

menor familiaridad con los departamentos objeto de revisión es analizado, encontrándose que los mejores conocedores de cada programa tienden a valorar mejor a sus profesores. La proximidad geográfica, sin embargo, no resulta ser un problema grave que sesgue la evaluación en un sentido determinado. Tampoco parece relevante la mayor influencia de evaluadores originarios de ciertas universidades más prestigiosas. El estudio también considera la posible influencia del interés de los evaluadores con alguna orientación específica de los programas, pero no encuentra un efecto significativo. Se puede afirmar que este estudio, gracias a su mayor precisión metodológica y a su gran amplitud de datos, ha sido mucho más aceptado que los anteriores por la comunidad universitaria.

E) *Resultados agregados por instituciones*

A pesar de recomendaciones en contra de la mayoría de los autores de los anteriores estudios de que no debían hacerse agregaciones por instituciones, éstas han sido intentadas por diversos autores. Concretamente, Magoun (1966), Petrowski (1973), Morgan *et al.* (1986), Webster (1983) y Orr (1984) agregan datos de los informes del ACE (los tres primeros) y del CBARC (los dos últimos) para obtener, con diversos resultados según el tipo de criterio utilizado, ordenaciones jerárquicas de las universidades. Como muestra de algunos resultados obtenidos por varios de estos estudios y también como una concesión al «folklore del ranking», se presenta el cuadro VI.1, tomado de Webster (1983), y en el que se recogen resultados de encuestas reputacionales desde 1925 los más antiguos agregados por los propios autores de los estudios, y los más modernos agregados por Webster con criterios personales, sin duda, discutibles. Aun con la salvedad de la dudosa validez de estas ordenaciones, un hecho que sorprende es la notable estabilidad de la reputación de las instituciones más señeras: salvo una continua mejora de las instituciones californianas, la lista permanece casi invariable a lo largo de casi sesenta años.

VI.2.2. Evaluación de programas de pregrado (undergraduate)

La educación superior norteamericana, en su primer ciclo de cuatro años, no ha sido objeto de un estudio tan generalizado ni profundo como la de postgrado. Dada la gran cantidad, diversidad y multiplicidad de funciones que las instituciones universitarias de este nivel desempeñan, intentar hacer estudios reputacionales sobre la calidad de un sistema tan heterogéneo es una tarea realmente compleja. A pesar de todo, y dado el gran interés de la sociedad norteamericana por valorar este nivel educativo, también se han hecho algunos estudios sobre la calidad de estos centros. La mayoría de ellos

utilizan medidas objetivas (que serán tratados en la próxima sección), pero también los hay reputacionales. Entre éstos hay que citar un conjunto de guías

CUADRO VI.1.

Ordenación jerárquica de las 15 mejores universidades norteamericanas mejor evaluadas, según Webster (1983)

	J82	L79	R70	C66	K59	H34	H25
UC, Berkeley	1	2	1	1	2	1	9
Stanford	2	3	3	5	13	13	14
Harvard	3	1	2	2	1	1	2
Yale	3	4	5	6	4	7	4
MIT	5	9	15	21	—	16	19
Princeton	6	8	8	10	7	11	6
Chicago	7	6	7	9	6	5	1
Wisconsin, Mad.	8	7	6	3	8	4	5
Michigan	8	5	4	4	5	8	8
UCLA	8	15	10	14	14	—	—
Columbia	11	11	12	7	3	3	3
Cornell	11	12	11	11	9	6	10
Illinois, Urb.	13	10	9	8	10	11	11
Pennsylvania	14	15	14	15	11	14	12
Caltech	15	13	20	18	—	17	18

NOTAS:

— Los encabezamientos de las columnas indican el estudio realizado como base para la ordenación.

J82 = Jones *et al.*, 1982

L79 = Scully, 1979

R70 = Roose y Andersen, 1970

C66 = Carter, 1966

K59 = Keniston, 1959

H34 = Hughes, 1934

H25 = Hughes, 1925

— Los datos referentes a J82, R70 y C66 son reelaboraciones de Webster (1983), agregando ordenaciones realizadas originalmente para departamentos por los autores de los estudios que se toman como base.

populares y de reportajes en revistas de información general, que, con criterios de dudosa validez, cuando no absolutamente acientíficos (una guía popular de USA basa la calidad de los *Colleges* en función del número de antiguos alumnos que aparecen citados en el *Who's who in America*), cubren el vacío de estudios más rigurosos, satisfaciendo así el interés de un gran público interesado en dónde puede ir a estudiar o en dónde poder enviar a sus hijos para rentabilizar mejor, social y económicamente, la importante inversión que la educación universitaria representa en ese país.

Centrándonos en los estudios reputacionales de carácter científico se puede citar, en primer lugar, un estudio debido a Johnson (1978), en el que se usan encuestas enviadas a miembros de la comunidad académica de todo tipo de centros (no sólo a los que imparten doctorado), para conocer la opinión de los encuestados sobre cuáles son las universidades del país con mayor liderazgo, entendiendo por tal el prestigio junto con la capacidad de innovación y el de influencia en la comunidad. Lo más notable de resultados es que, mientras que los que responden desde las mayores universidades dan unas contestaciones muy homogéneas en torno a las 10 mejores universidades (California en Berkeley, California en Los Angeles, Columbia, Chicago, Harvard, Michigan, MIT, Princeton, Stanford y Yale), el grado de acuerdo disminuye según los encuestados pertenecen a instituciones que se van alejando de la cúspide jerárquica de la educación superior en USA.

Otro estudio sobre centros de pregrado es el promovido por el *Higher Education Research Institute* (Lawrence y Solmon, 1981; Solmon y Astin, 1981; Astin y Solmon, 1981), que trata de identificar los programas más reputados de pregrado en económicas, historia, inglés, química, sociología y biología. El método utilizado fue el de solicitar a los evaluadores que ordenaran los departamentos siguiendo seis criterios:

- La calidad global de la educación.
- El nivel de preparación de los estudiantes para graduarse o para ir a centros profesionales.
- El nivel de preparación de los estudiantes para el trabajo al acabar la graduación.
- El grado de entrega del profesorado a la enseñanza.
- Los logros académicos o profesionales del profesorado.
- Las innovaciones curriculares y pedagógicas.

Los autores dan la lista de aquéllos que han sido incluidos entre los 10 mejores en alguno de los seis criterios. También realizan un interesante estudio en el que se relacionan los resultados obtenidos de las observaciones subjetivas, con toda una serie de medidas objetivas de los centros evaluados. Los autores concluyen que, conocidos el tamaño relativo, el nivel de selectividad, el de gasto y algún otro factor adicional según el tipo de centro, puede ser predicha con bastante exactitud la valoración que va a recibir cada centro por

los encuestados. Basándose en este estudio, los autores afirman que deberían descartarse los costosos estudios reputacionales, aunque añaden que quizá su evidencia es insuficiente dados los pocos campos científicos objeto de revisión en este estudio.

VI.2.3. Evaluación de programas profesionales

Mientras que la mayoría de los estudios sobre calidad de las instituciones universitarias en USA se han referido a los programas de doctorado, sólo unos cuantos se han dedicado a los programas profesionales, ya sea utilizando técnicas reputacionales aisladas o combinadas con medidas objetivas. Los estudios de programas profesionales han sido promovidos por dos clases de organismos: los relacionados con el mundo académico y los propios grupos o asociaciones profesionales. Las guías populares y los trabajos periodísticos también se han unido recientemente a la evaluación de estos centros profesionales.

Un primer estudio sobre la evaluación de la calidad de los programas profesionales, usando técnicas reputacionales, es el debido a Marguiles y Blau (1973). En el que se evaluaban 17 programas profesionales: Arquitectura, Negocios, Educación, Ingeniería, Derecho, Bibliotecas, Enfermería, Farmacia, Trabajo Social, Odontología, Periodismo, Medicina, Optica, Salud Pública, Teología, Veterinaria e Ingeniería Forestal. El método seguido fue el de preguntar a los decanos de 1.180 escuelas profesionales que señalaran cuáles eran los programas más notables. Como el número de respuestas para cada programa era bajo (sólo el 36 por 100 de los encuestados respondieron), los autores repitieron el estudio (Blau y Marguiles, 1974-75), añadiendo las escuelas de Música, y consiguiendo que aumentaran considerablemente el número de respuestas (79 por 100), pese a lo cual los resultados no fueron muy diferentes de los anteriores. En este estudio se analizó la correlación existente entre la reputación de las escuelas y algunas medidas objetivas, como los recursos financieros o el tamaño de la biblioteca, encontrando que, para cada tipo de programa profesional, el nivel de correlación que se da es muy diferente.

Otro estudio del mismo tipo es el *Cartter Report* (1977). En él se usan como evaluadores, además de los decanos de las escuelas profesionales, a otros profesores de las escuelas. El estudio se reduce a sólo tres escuelas: Educación, Derecho y Negocios, y los resultados no difieren mucho de los obtenidos anteriormente por Marguiles y Blau.

Algunos otros estudios se refieren a especialidades concretas, como es el caso del trabajo de Cole y Lipton (1977) dedicado a las escuelas de Medicina.

Este es un interesante estudio desde el punto de vista metodológico, ya que se combina el uso de las respuestas reputacionales con todo un conjunto de medidas objetivas referidas a cada centro. Los autores encuentran que las variables objetivas que más se correlacionan con la calidad, tal y como es percibida por una encuesta reputacional, son: productividad investigadora, distinciones de profesorado, ayudas de investigación disponibles y tamaño del departamento. También es detectada la existencia de un cierto *efecto halo*, que afecta a las escuelas adscritas a las universidades más prestigiosas.

Dentro del campo de los estudios profesionales, son también muy frecuentes en la actualidad las encuestas realizadas por revistas profesionales o de amplia circulación pública, que, con periodicidad anual, hacen sus propias evaluaciones jerárquicas de los centros de estudios profesionales (Fraenkel, 1989). Cabe citar, por ejemplo, los estudios periodísticos anuales que publica *US News and World Report*, para escuelas de Negocios, Derecho, Medicina e Ingeniería (varios, 1990).

En el cuadro VI.2 se presentan algunos de los resultados mencionados en este apartado. Como se puede apreciar, quizá el rasgo más característico sea la tremenda estabilidad de los resultados, a pesar del tiempo transcurrido entre encuestas y de las diferencias de métodos y criterios utilizados en ellas. Realmente es difícil no sospechar la existencia de un *efecto halo* en torno a ciertas instituciones tradicionalmente prestigiosas.

VI.3. ESTUDIOS BASADOS EN MEDIDAS OBJETIVAS

Para evaluar la calidad de las universidades norteamericanas se han usado, además de estudios reputacionales, otros en los que se han ensayado distintas medidas o indicadores objetivos de la calidad. El hecho más frecuente es que la elección de las medidas objetivas utilizadas en cada caso sea una elección discrecional de los investigadores, más que el resultado de estudios previos sobre la relación entre calidad e indicadores (Conrad y Blackburn, 1985a). En consecuencia, han sido ensayados una gran cantidad de indicadores que se pueden agrupar en estas tres categorías:

— *Medidas sobre características del profesorado.* Habitualmente se consideran el número de publicaciones, el número de premios y distinciones, etcétera.

— *Medidas sobre características del alumnado.* La calidad de los estudiantes es evaluada generalmente por el resultado de los tests de aptitud que se realizan antes del ingreso en los centros de educación superior.

CUADRO VI.2.

Ordenaciones jerárquicas de algunas escuelas profesionales norteamericanas

EDUCACION		NEGOCIOS		
B74	C77	B74	C77	V90
Stanford	Stanford	Harvard	Stanford	Stanford
Columbia	Harvard	Stanford	Harvard	Harvard
Ohio Sta	Chicago	Chicago	MIT	Pennsylv.
Chicago	UCLA	Pennsylv.	Chicago	Northw.
Harvard	UC, Berkel.	Carnegie-M.	Carnegie-M.	MIT
DERECHO		INGENIERIA		
B74	C77	V90	B74	V90
Harvard	Harvard	Yale	MIT	MIT
Yale	Yale	Chicago	Stanford	Stanford
Michigan	Stanford	Stanford	Illinois	Illinois
Columbia	Michigan	Columbia	UC, Berkel.	Caltech
Chicago	Chicago	Harvard	Caltech	UC, Berke.

NOTAS:

- B74 = Blau y Marguiles, 1974.
- C77 = Cartter Report, 1977.
- V90 = Varios, 1990.

— *Medidas sobre características de las instituciones.* Son frecuentemente usadas el tamaño de la biblioteca, el número y la capacidad de los laboratorios, el gasto por estudiante, los salarios de los profesores, la relación profesor/alumno, etcétera.

En cualquier caso, el problema de establecer unos indicadores generales y simples para evaluar la calidad de las instituciones sigue estando sin resolver, por lo que el establecer un índice combinado de calidad cuya validez sea aceptable es, por el momento, una pura utopía. Los diferentes análisis de la calidad de las instituciones universitarias realizados mediante el uso de medidas objetivas, utilizan algún indicador de calidad o una combinación de ellos. En los siguientes apartados se hace referencia a algunos de estos trabajos, agrupados según sea el tipo básico de indicador utilizado.

VI.3.1. Estudios que usan indicadores relativos al profesorado

Los estudios más generalizados son posiblemente los que identifican

calidad de la institución con calidad del profesorado. Aun aceptando este principio, que es de por sí discutible, queda todavía por ver qué medidas de la calidad del profesorado deben ser utilizadas. Aunque se han usado también otros criterios, la mayoría de los autores utilizan la productividad investigadora de los profesores como medida evaluadora de su calidad y, en consecuencia, de la calidad de las instituciones. Algunos de los indicadores más utilizados son los siguientes:

— Número de publicaciones del profesorado en un cierto período de tiempo, normalmente de tres a seis años. En las publicaciones se suelen contabilizar:

– Todo tipo de publicaciones, con distintos criterios para los diversos tipos: libros, artículos, etc. Así lo hacen Somit y Tanenhaus (1964), Knudsen y Vaughan (1969) y Glenn y Villimez (1970).

– Sólo artículos en las revistas más significativas (Lewis, 1968; Morgan y Fitzgerald, 1978; Robey, 1979; Cox y Catt, 1977).

– Artículos en revistas con diferentes ponderaciones según la categoría científica de la revista (Moore, 1972).

– Número de páginas publicadas en revistas (Graves *et al.*, 1982; Hirsch *et al.*, 1984).

— Número de profesores participantes en consejos editoriales y como evaluadores en las principales revistas (Dean, 1976).

— Premios y honores conseguidos por el profesorado (Bowker, 1965).

— Número de citaciones según algún índice de citaciones de la materia objeto de evaluación (Davis y Papanek, 1984; Liu, 1978).

— Número y cantidad de las ayudas económicas recibidas por el profesorado para sus investigaciones (Guba y Clark, 1978).

La mayor parte de los estudios que utilizan medidas objetivas sobre el profesorado están referidos a departamentos de especialidades concretas o a escuelas profesionales. Agrupados según la disciplina académica que evalúan, algunos de los más representativos son los siguientes:

A) *Estudios globales sobre instituciones*

Uno de los primeros trabajos que usaron medidas objetivas para la evaluación de la calidad, en este caso de las instituciones como un todo, es el

de Bowker (1965). En este estudio el autor clasifica las instituciones que conceden títulos de doctorado, utilizando como criterio evaluador el número de profesores que han recibido ciertos premios y cierto tipo de becas prestigiosas. Como se puede apreciar en un criterio extremadamente simple que sólo puede servir para unas pocas universidades muy prestigiosas en las que se acumulan los profesores más afamados y laureados.

Muffo *et al.* (1987) proponen un sencillo método de evaluar las instituciones haciendo uso del número de veces que aparecen sus profesores en los índices de citación que se publican para todos los campos del conocimiento. Los autores lo ensayaron en una muestra de universidades norteamericanas. Estos índices de citación son un buen indicador de la productividad investigadora, tanto en cantidad como en calidad. (Aunque no sea el caso en las universidades norteamericanas, este indicador tiene el inconveniente en los países no anglófonos, de que los índices de citaciones que se publican están sesgados hacia la producción científica en inglés, algo que en el campo de las ciencias sociales y las humanidades es demasiado influyente.)

B) *Estudios sobre departamentos de ciencias políticas*

Somit y Tanenhaus (1964) evaluaron los departamentos de ciencias políticas usando el número de publicaciones, identificando, por tanto, calidad de los departamentos con cantidad de publicaciones de sus profesores. Es un sistema tan simple que los propios autores reconocen que una mayor carga docente o una mayor dificultad en el acceso a la bibliografía pueden disminuir la productividad cuantitativa de un departamento sin que ello signifique peor calidad.

Usando un criterio semejante, Morgan y Fitzgerald (1978) valoran la productividad del profesorado de los departamentos de ciencias políticas contabilizando el número de publicaciones en las revistas más importantes de la disciplina. Los autores establecen una ordenación jerárquica de estos departamentos, que comparan con los resultados reputacionales de Roose y Andersen (1970), encontrando diferencias importantes entre ambas ordenaciones. Como dicen los autores, se están valorando cosas diferentes, reputación y productividad, que no tienen por qué coincidir.

Un trabajo semejante es el que realizó Robey (1979). La calidad de los departamentos de ciencias políticas se identifica con la productividad medida a través del número de publicaciones en una serie de revistas de la especialidad. Al comparar los resultados de su ordenación con la encuesta de reputación de Cartter (1966) encuentra discrepancias, aunque todas las

universidades con mejor reputación se encuentran, al menos, entre las 30 más productivas.

C) *Departamentos de sociología*

Los estudios sobre calidad de los departamentos de sociología son posiblemente los más numerosos y los que mayor variedad de variables utilizan (Webster y Conrad, 1986). Algunos de los más representativos se reseñan a continuación.

Lewis (1968) evaluó los departamentos de sociología utilizando como criterio el número de publicaciones no sólo del profesorado, sino también de los alumnos de doctorado. El autor compara sus resultados con los reputacionales de Cartter (1966), encontrando una relación muy estrecha entre ambos.

Knudsen y Vaughan (1969) utilizaron como medidas objetivas para evaluar la calidad de los departamentos de sociología el número de publicaciones del profesorado (en total y *per cápita*), junto con el número de alumnos graduados en el programa. Sólo fueron contabilizadas las publicaciones en algunas revistas más importantes, asignando distintos pesos a las diferentes publicaciones según el siguiente criterio: 48 puntos a un libro de investigación, 24 a un libro de texto, 16 a un libro del que es editor, 16 a un artículo en *The American Sociological Review*, y puntuaciones menores a publicaciones en otras revistas según su categoría científica. Los autores encuentran similitudes con las evaluaciones reputacionales de Cartter (1966), pero sólo en el caso de las instituciones más destacadas.

Glenn y Villimez (1970) realizaron otro estudio semejante para los departamentos de sociología, pero aumentando el número de revistas contabilizadas y el período de tiempo considerado, variando asimismo los diferentes criterios para los pesos de cada tipo de publicación (en este caso determinados tras una encuesta previa entre miembros de la comunidad académica). Los resultados sólo están de acuerdo con los de Knudsen y Vaughan en los primeros puestos de la lista, lo que demuestra la gran dependencia de los criterios elegidos en los resultados finales, al menos para la gran mayoría de los departamentos evaluados.

D) *Departamentos de psicología*

Un estudio dedicado a evaluar la calidad de los departamentos de psicología fue realizado por Cox y Catt (1977). En él se contabilizaban los artículos

publicados por los profesores de los distintos departamentos en las revistas de psicología durante un período de seis años. Los departamentos fueron clasificados de acuerdo con este criterio, que a su vez se comparó con los resultados de Roose y Andersen (1970), encontrándose disparidades importantes en los resultados de ambos estudios.

E) *Escuelas de educación*

Guba y Clark (1978) realizaron un estudio sobre las escuelas de educación utilizando como variables objetivas seis medidas de la calidad investigadora del profesorado medida a través de su productividad en cuanto a publicaciones, participaciones en congresos y ayudas de investigaciones recibidas. El objetivo del trabajo era identificar las 25 instituciones más productivas en investigación dentro del campo de la educación, por lo que poco se puede objetar al método. Los autores no pretenden considerar estas valoraciones como índice de calidad de los centros.

F) *Departamentos de economía*

Son también numerosos los estudios sobre la calidad en los departamentos de economía. Citaremos sólo algunos de los más representativos.

Moore (1972) realizó una evaluación de los departamentos de economía, utilizando como variable objetiva el número de publicaciones de los profesores en las principales revistas de la especialidad, ponderadas de distinta manera, según el prestigio de la revista en que se publicaron. Utilizando métodos econométricos, el autor relaciona la ordenación de Roose y Andersen (1970) con sus propios resultados, encontrando diferencias significativas entre ambos, «lo que hace a uno preguntarse de qué modo, y con qué eficiencia se extiende (entre el mundo académico) la información relacionada con los programas de graduados» (Moore, 1972: 2).

Una curiosa alternativa a las medidas habituales sobre la calidad de los departamentos, es la presentada por Dean (1976). Este autor utiliza como criterio para valorar a los departamentos el número de evaluadores y editores de las principales revistas que pertenecen al profesorado del departamento.

La consistencia de las anteriores evaluaciones con las evaluaciones reputacionales de Roose y Andersen (1970), es estudiada por Stolen y Gnuschke (1977), quienes llegan a la conclusión de que sólo para los departamentos distinguidos son consistentes las distintas evaluaciones, mientras que para los

demás las cosas fluctúan mucho según el tipo de sistema utilizado. Una vez más, se pone de manifiesto la concordancia de los distintos métodos para las instituciones realmente sobresalientes, pero su debilidad de método para evaluar la calidad de un amplio abanico de instituciones, como podría ser todo un sistema universitario.

Graves *et al.* (1982, 1984) evalúan los departamentos de economía según el número de páginas publicadas en las 24 revistas más notables dentro de la especialidad. La ordenación jerárquica obtenida por este criterio la relacionan con una serie de variables relacionadas con los recursos disponibles, como salarios del profesorado, carga docente, razón secretarías/profesor, razón alumnos/profesor, calidad de los servicios de apoyo y existencia de personal de apoyo. Los resultados de este análisis permiten concluir que los salarios y el porcentaje de secretarías son las variables más relacionadas con la productividad investigadora, mientras que la carga docente es la más inhibitoria de esta productividad. Hirsch *et al.* (1984) realizaron una puesta al día del anterior trabajo e incluyeron también 40 universidades extranjeras. Hoga (1984) replicó a este trabajo matizando que había que distinguir entre las publicaciones de los profesores de los departamentos y las de los estudiantes u otras personas adscritas a ellos, ya que si se contabilizan juntas se producen algunos cambios importantes en las ordenaciones.

Otro estudio muy semejante a los anteriores es el de Davis y Papanek (1984), quienes utilizaron como criterio de la calidad de los departamentos el número de citaciones de sus profesores en las revistas de la especialidad.

G) *Departamentos de ingeniería mecánica*

Liu (1978) analizó la calidad de los departamentos de ingeniería mecánica valorando la calidad del profesorado según el número de citaciones en las revistas que cataloga el *Science Citation Index*.

VI.3.2. **Estudios que utilizan indicadores relativos al alumnado**

A pesar de la indudable importancia del alumnado en la calidad de las instituciones, fundamentalmente en el proceso docente, el uso de indicadores relativos al alumnado ha sido muy poco utilizados en la práctica. Como ya se dijo en el capítulo II, al hablar del concepto de calidad, la mejor medida de la calidad de las instituciones universitarias sería la que fuera capaz de valorar cómo mejoran las capacidades de los alumnos en el transcurso de su permanencia en la institución. Deberían ser las variables referidas a los alumnos las

que tuvieran una atención preferente en la evaluación de la calidad. Sin embargo, la calidad de los estudiantes no ha sido utilizada como demasiada frecuencia en Estados Unidos. Tres razones básicas se han alegado:

— Porque, realmente, la calidad de los estudiantes tiene mucho de medida reputacional en un sistema educativo como el norteamericano, donde los mejores alumnos asisten a las universidades más prestigiosas (Blackburn y Lingenfelter, 1973).

— Porque la calidad de los estudiantes que ingresan no garantiza un buen programa (Blackburn y Lingenfelter, 1973).

— Porque la calidad de los estudiantes que salen del sistema educativo es de muy difícil evaluación (Kuh, 1980).

A pesar de las dificultades, se han usado diferentes variables relacionadas con los alumnos como criterio para la evaluación de la calidad de las instituciones universitarias. Algunos de los utilizados por distintos investigadores son los siguientes:

— Proporción de alumnos del centro que acaban obteniendo un doctorado. Así, por ejemplo, Hall (1984) clasifica los centros de pregrado privados según la proporción de sus ex alumnos que acabaron consiguiendo un doctorado entre 1920 y 1980. Igualmente, Fuller (1986) establece una ordenación de las universidades, usando como medida la proporción de estudiantes de pregrado que acaba obteniendo un doctorado en cualquier universidad.

— Nivel de selectividad del centro (Astin y Henson, 1977) o centros preferidos por los alumnos más capacitados académicamente (Astin y Solmon, 1979).

— Carrera profesional de los ex alumnos del centro. Glower (1980) clasifica las escuelas de ingeniería según el éxito profesional de sus ex alumnos a partir de datos biográficos obtenidos de la publicación *Who's who in Engineering*.

— Producción científica de los ex alumnos de una institución. Clemente y Sturgis (1974) consideran la producción científica de los individuos que habían conseguido un doctorado en los departamentos de sociología, en los treinta años anteriores al estudio, como criterio para evaluar el nivel de calidad de estos departamentos.

Es evidente que existe una importante carencia de evaluaciones de la calidad de las instituciones que incluyan su principal elemento, los estudiantes,

y las que lo incluyen, como las reseñadas, son a través de índices demasiado parciales como para no dudar de su validez. Como señala Tan (1986), es necesario desarrollar nuevos índices mucho más directos, que permitan evaluar mejor la experiencia vivida por los estudiantes durante su permanencia en las universidades.

VI.3.3. Estudios que utilizan una combinación de indicadores

Es indudable que la utilización de una combinación de indicadores de diversos tipos (profesorado, alumnado y recursos) sería una aceptable manera de evaluar la calidad de las instituciones en todos sus aspectos (docente, investigador y de servicios). Sin embargo, pocos investigadores han utilizado este enfoque debido a una doble dificultad: en primer lugar, está el problema teórico de encontrar la combinación que mejor se aproxime al concepto de calidad de una institución; en segundo término, la incapacidad de medir las variables realmente importantes, que en muchos casos están relacionadas con los efectos a largo plazo del proceso educativo.

Enfoques mixtos han sido usados para analizar la relación entre diversas variables objetivas y las ordenaciones reputacionales, como se verá en la próxima sección. El estudio de *The Conference Board of the Associated Research Councils* (Jones *et al.*, 1982) es también de hecho un estudio mixto en el que se combinan variables objetivas con subjetivas, y dentro de aquéllas, variables de diversos tipos.

Los más conocidos estudios de carácter mixto que utilizan un conjunto amplio de características de las instituciones, del profesorado y del alumnado son *The Gourman Report* (Gourman, 1982, 1983), del que existen ediciones renovadas periódicamente. Estos populares estudios han sido muy criticados en los medios académicos, ya que la metodología utilizada para las evaluaciones no es satisfactoriamente explicada, pero, sin embargo, cuentan con una amplia difusión entre los «consumidores», posiblemente mucho más que estudios más rigurosos. Uno de los motivos de esta popularidad se debe, sin duda, a que clasifica toda clase de instituciones universitarias del país, e incluso otras extranjeras con más que dudosos criterios. A pesar de todo, como afirma Fraenkel (1989), ésta es la más completa, detallada y seguramente más fiable de todas las guías de este tipo.

Un trabajo más académico es el realizado por Brown (1967), que utilizó las siguientes ocho variables objetivas para evaluar la calidad de las universidades: proporción de profesores con doctorado, salario medio del profesorado, proporción de estudiantes que siguen estudios de postgrado, proporción de

estudiantes de postgrado, número de libros por estudiante, número total de profesores, coeficiente profesor/alumno y gasto total por estudiante. Con todas estas variables, Brown construyó un índice compuesto de la calidad de los centros universitarios.

VI.4. ESTUDIOS QUE RELACIONAN EL ENFOQUE REPUTACIONAL CON MEDIDAS OBJETIVAS

En las dos secciones anteriores se han considerado estudios que tratan de valorar la calidad de las universidades norteamericanas desde dos puntos de vista contrapuestos. Como es lógico suponer, se han realizado frecuentes análisis comparativos de ambos enfoques con la intención básica de determinar cuál es el conjunto de variables que mejor pueden servir para evaluar la calidad de las instituciones. Aunque estos estudios no son propiamente evaluaciones de la calidad de las universidades, son de un gran interés para clasificar la validez de los distintos métodos de evaluación de la calidad, permitiendo extraer consecuencias teóricas y empíricas importantes.

La metodología básica de casi todos estos estudios es la de establecer correlaciones entre alguna de las ordenaciones reputacionales citadas en la sección VI.2, con algún conjunto de variables objetivas, para intentar encontrar aquellas variables que mayor relación tienen con la calidad percibida.

Uno de los primeros estudios de este tipo lo realizó el propio Cartter (1986), utilizando los resultados reputacionales de su estudio y algunas medidas objetivas. Cartter encontró que las compensaciones económicas del profesorado y el tamaño de las bibliotecas eran las magnitudes más correlacionadas con los resultados reputacionales.

Con posterioridad a la publicación del primer estudio de Cartter (1966), son numerosos los estudios que intenten relacionar las medidas objetivas con los resultados reputacionales. Casi todos ellos realizan comparaciones con las ordenaciones obtenidas en las encuestas de la ACE (Cartter, 1966; Roose y Andersen, 1970). Algunos de los más representativos de estos estudios se resumen a continuación, siguiendo un orden cronológico:

A) *El National Science Board* (1969), analizando las características de los departamentos, junto con su valoración reputacional, identificó 15 variables relacionadas con la calidad de los departamentos. Son las siguientes:

- Número de títulos otorgados por el departamento.

- Fondos federales recibidos para investigación.
- Fondos recibidos para docencia y gastos generales.
- Origen académico de los becarios.
- Origen académico de los doctores.
- Nivel de selectividad para el ingreso a la institución.
- Grado de preferencia de los becarios por la institución.
- Estudiantes postdoctorales en ciencias e ingeniería que hay en la institución.
- Número de doctorados concedidos en proporción al número de profesores.
- Número de doctorados concedidos en proporción al número de estudiantes.
- Proporción entre el número total de doctores y el de graduados que reciben sus credenciales cada año.
- Salarios de los profesores.
- Proporción de *full professors* en el total de la plantilla docente.
- Relación profesor/alumno.
- Tamaño del departamento.

B) Oromaner (1970) encontró, en un análisis referido exclusivamente a los departamentos de sociología, que los más prestigiosos departamentos se caracterizaban por:

- Un mayor tamaño del departamento.
- Un profesorado relativamente joven.
- Un profesorado graduado a su vez en prestigiosas universidades.

C) Hagstrom (1971) analizó 125 departamentos de ciencias físicas y biológicas para encontrar que las variables más explicativas eran:

- La productividad del profesorado.
- Las características de los estudiantes.
- Los recursos dedicados a la investigación.
- El tamaño del departamento.

D) Abbot (1972) investigó la relación entre la reputación de los departamentos de sociología con dos grupos de variables objetivas: unas, institucionales, y otras, departamentales. Abbot encontró que las variables institucionales resultan ser más significativas para explicar la ordenación reputacional que las departamentales, confirmándose así la existencia de un *efecto halo* importante en los estudios reputacionales.

E) Abot y Barlow (1972), en un estudio referido a las instituciones como un

todo, no hallaron relación entre la reputación de las instituciones y variables como el salario del profesorado, el tamaño de la biblioteca, el número de profesores o los recursos disponibles, mientras que sí encontraron relación entre el prestigio y estas dos variables:

- La productividad de los profesores.
- El número de doctorados otorgados por la institución.

F) Elton y Rodgers (1973) analizaron, para los departamentos de ciencias físicas, la relación existente entre los resultados reputacionales del estudio de Cartter (1966) y una serie de variables objetivas. La más importante conclusión que extraen estos autores es la de la existencia de un importante *efecto halo* que contamina los estudios reputacionales.

G) Beyer y Snipper (1974) estudiaron, para diferentes campos académicos, la relación existente entre algunos indicadores objetivos y el resultado del estudio de Cartter (1966). El resultado más interesante que se obtiene de este análisis es el comprobar que las mismas medidas objetivas tienen diferente capacidad explicatoria según la disciplina académica de la que se trate, por lo que habría que concluir que no se pueden adoptar métodos simplistas para realizar evaluaciones objetivas, y que hay que ser muy cuidadoso con la elección de indicadores según el área que se intenta evaluar.

H) Morgan *et al.* (1976) compararon, para el caso de departamentos de ciencias políticas, los resultados del estudio reputacional de Roose y Andersen (1970) con varias medidas objetivas, llegando a la conclusión de que las más correlacionadas con la calidad percibida eran:

- El tamaño de la biblioteca.
- El gasto por estudiante.
- La razón profesor/alumno.
- El salario medio del profesorado.

I) Cox y Catt (1977) compararon la productividad de los departamentos de psicología con los resultados reputacionales de la encuesta de Roose y Andersen (1970), llegando a la conclusión de que existía muy poca relación entre ambas evaluaciones.

J) Hensley (1980) realizó un estudio sobre las correlaciones que existían entre los principales estudios reputacionales (agregando los resultados por instituciones) y tres medidas objetivas de las instituciones académicas: premios y honores conseguidos por el profesorado, productividad científica del

profesorado y fondos de investigación obtenidos. Hensley encontró que, por un lado, las correlaciones entre los propios estudios reputacionales eran muy altas, a pesar del tiempo transcurrido entre unos estudios y otros y, por otro, que las correlaciones entre medidas objetivas y reputacionales también eran relativamente altas, lo que hace concluir al autor que los métodos reputacionales son unos indicadores válidos de la calidad institucional.

K) Drew y Karpf (1981) analizaron la relación existente entre la productividad del profesorado de los departamentos de matemáticas, medida por el número de publicaciones, y la ordenación jerárquica de estos departamentos realizada por Cartter (1966). En este caso el autor llega a la conclusión de que la ordenación reputacional puede ser casi perfectamente predicha utilizando como predictor solamente el número de publicaciones del departamento.

L) Hogan (1981) realizó un estudio para los departamentos de economía, usando como medidas objetivas de la calidad las cinco siguientes: calidad de los estudiantes que entran en el programa (medida por la proporción de alumnos becarios), calidad del profesorado (medida por los resultados reputacionales del estudio de Cartter, 1966), tamaño del programa (medida por el número de doctores que finalizan el año y por la relación profesor/alumno), número de profesores y productividad del profesorado (medida por el número de publicaciones y por las páginas totales publicadas). Utilizando técnicas econométricas, el estudio permite concluir que las variables que más influyen sobre la calidad del programa (medida por la productividad investigadora de los que ya se han graduado en el programa) son:

— La productividad del profesorado (número de artículos y páginas totales).

— El nivel de calidad de los alumnos que son aceptados en el programa (proporción de alumnos becarios).

M) Conrad y Blackburn (1985b, 1986) realizaron un interesante estudio en el que relacionan la calidad estimada de los departamentos con un amplio número de variables objetivas. Para la evaluación de la calidad no se utilizaron los habituales estudios reputacionales, sino que se hizo uso de informes de expertos que habían realizado visitas de revisión a 45 departamentos universitarios de 15 universidades, en cinco áreas científicas diferentes. Las variables objetivas ensayadas, 73 en total, fueron agrupadas en cuatro apartados: las referentes al profesorado, al alumnado, al programa y a los recursos disponibles. Los resultados de este estudio son de gran interés; los más relevantes son los siguientes:

— La productividad del profesorado, medida por el número de publicaciones, explica el 28 por 100 de la varianza total, valor que, aunque importante, demuestra que es exagerado el peso que se le asigna a esta variable en otros estudios basados en las encuestas reputacionales, en los que esta productividad del profesorado es la más importante magnitud tenida en cuenta. Las ayudas a la investigación recibidas explican un 7 por 100 de la varianza, también bastante menos de lo que cabía esperar. Tanto la edad del profesorado como la proporción de éste que tiene contrato fijo (*tenure*), como la proporción de los que han tenido largas experiencias fuera de la universidad (negocios, política, industria, etc.) están relacionados negativamente con la calidad. La carga docente del profesorado tiene una moderada correlación negativa con la calidad, mientras que el nivel de formación académica del profesorado tiene una relación positiva.

— El tamaño del programa, medido por el número de estudiantes matriculados, explica el 32 por 100 de la varianza. Es decir, el tamaño es una característica importante para estimar la calidad de los departamentos. Por otra parte, la capacidad académica de los estudiantes aparece moderadamente correlacionada con la calidad del departamento.

— Las características propias del programa tienen una capacidad explicativa del 30 por 100. Las más altas correlaciones con la calidad se dan para una variable que indica la proporción de estudiantes de toda la institución que siguen estudios dentro del campo de las ciencias de la naturaleza.

— La única variable utilizada como índice de los medios disponibles es el número de libros en la biblioteca, y tan sólo explica un 2 por 100 de la varianza.

M) Saunier (1985) realizó un estudio en el que se establecen relaciones entre los resultados reputacionales de Jones *et al.* (1982) en seis tipos de departamentos (todos ellos de ciencias sociales y humanidades), con una serie de medidas objetivas de la calidad. Aunque los resultados varían por tipo de departamento, se puede afirmar que el tamaño del departamento, medido por el número de profesores y por el número de graduados, son las variables con mayor capacidad explicativa de la ordenación jerárquica reputacional. Otra variable que resulta relevante es el número de publicaciones del profesorado. Los gastos totales del departamento resultan tener una influencia muy débil sobre la calidad percibida. En un estudio posterior, King y Wolffe (1987) mejoran técnicamente el estudio de Saunier, pero llegan a conclusiones muy semejantes.

P) Fairweather (1988) analizó las relaciones entre algunas medidas objetivas de instituciones de ingeniería con los estudios reputacionales de Jones *et*

al. (1982). Es este un estudio técnicamente muy riguroso, que llega a la conclusión de que los índices reputacionales no están realmente midiendo la calidad de los departamentos, porque existe un importante efecto halo de las características de las instituciones sobre el prestigio de los departamentos.

R) En un trabajo cuya finalidad es encontrar el conjunto de variables que mejor predigan la calidad de los departamentos, Tan (1990) analiza un conjunto amplio de variables objetivas que son agrupadas en factores. Posteriormente compara las ordenaciones reputacionales de los departamentos de sociología, según el estudio de Jones *et al.* (1982), con las que resultarían de la aplicación de los distintos factores o de sus combinaciones. La principal innovación de este trabajo es el uso de análisis multivariante y la obtención de un factor compuesto como la mejor explicación para la calidad departamental.

S) Un estudio semejante al anterior, en el que se intenta también encontrar un conjunto de variables objetivas que puedan ser usadas como predictores de la calidad es el de Vinsonhaler y Vinsonhaler (1990). Este estudio proporciona también una combinación de factores como medida objetiva de la calidad para las instituciones que conceden doctorados. La calidad de los alumnos, los recursos dedicados a la investigación y el prestigio del profesorado, están entre las variables que con más peso influyen en el indicador combinado de la calidad.

VI.5. **DISCUSION SOBRE LOS DISTINTOS ENFOQUES**

Tanto los métodos reputacionales de evaluación de la calidad como los que utilizan medidas objetivas han sido siempre motivo de profunda controversia en los medios académicos norteamericanos. Se han discutido las ventajas e inconvenientes de cada enfoque, y se han hecho todas las comparaciones posibles entre ellos. Algunas de las conclusiones más relevantes de esta discusión se exponen a continuación.

Sin embargo, antes de entrar en la discusión entre cada uno de los dos enfoques, es necesario formular una reserva fundamental respecto a casi todos los estudios expuestos en las secciones anteriores, de una u otra metodología: el concepto de calidad está siempre sin especificar. La idea que pueden tener de la calidad varía para cada investigador, para cada disciplina evaluada y para cada departamento, por lo que resulta casi imposible establecer conclusiones generales sobre la calidad de las instituciones, observando los resultados de todos estos trabajos (Lawrence y Green, 1980).

VI.5.1. Crítica de métodos reputacionales

La validez de los métodos reputacionales para determinar la calidad de las instituciones ha sido objeto de intensa controversia durante más de medio siglo (Hensley, 1980). Sin embargo, es de justicia afirmar que las desafortunadas consecuencias de los estudios reputacionales son más atribuibles a «la competitividad de la comunidad universitaria, al sensacionalismo de los medios de masas y a la obsesión americana por conocer quién está en la cumbre, que a los fallos de los estudios por sí mismos» (Lawrence y Green, 1980: 14), cuyos autores, al menos los de los trabajos más importantes, han sido más rigurosos, y ha tenido más cuidado en las conclusiones que la mayor parte de los que han utilizado o comentado con posterioridad esos estudios.

Como es de suponer, una considerable parte de las críticas a estos métodos se ha centrado en su principal debilidad: la falta de definición del concepto de calidad que llevan implícitos (Tan, 1986). Realmente, lo que estos métodos están evaluando no es la calidad, sino algo que suele acompañarla, el prestigio. Pero es evidente que, aunque calidad y prestigio puedan caminar cercanas, no van necesariamente de la mano.

De todos modos, las críticas más frecuentes y concretas se refieren a los distintos sesgos que las evaluaciones reputacionales llevan implícitas. Algunos de los más citados por los autores que han discutido estos temas son los siguientes:

— Por su misma esencia de estudio subjetivo, la evaluación de la calidad por estos métodos se reduce a una mera colección de cotilleos académicos. Sólo los que conocen con cierta profundidad los programas están en condiciones de emitir una opinión, pero éstos son realmente pocos (Conrad y Blackburn, 1985a).

— No se puede evitar el *efecto halo* que sobre cada departamento tiene el prestigio global de la institución (Smith y Fiedler, 1971; Lawrence y Green, 1980). La existencia de este efecto se demuestra por la importante correlación entre la calidad percibida y toda una serie de variables institucionales que no están relacionadas directamente con el programa evaluado (Elton y Rodgers, 1973; Fairweather, 1988). Un caso extremo citado por Webster (1985: 70) es el de una encuesta que situaba los programas de pregraduados en Negocios de las universidades de Harvard, Stanford, Chicago, Northwestern y Columbia entre los 12 mejores del país, cuando ninguna de estas universidades tiene tal programa a nivel de pregraduados. Este efecto también fue encontrado en el estudio reputacional para centros de pregraduados realizado por Astin y Solmon (1981).

— Existe un retraso entre el funcionamiento real de un departamento y el prestigio, que llega después o que se mantiene aun cuando el departamento haya disminuido de calidad (Smit y Fiedler, 1971; Cox y Catt, 1977; Lawrence y Green, 1980). Como consecuencia, las variaciones en el tiempo son muy escasas y siempre están las mismas universidades en la cabeza de las listas (Astin, 1985).

— Puede darse una tendencia de los ex alumnos a supervalorar sus antiguas instituciones (*efecto alma mater*), con lo que se genera un cierto círculo cerrado (Lawrence y Green, 1980; Conrad y Blackburn, 1985a). Por otra parte, las instituciones que producen mayor número de doctores también tienen mayor número de evaluadores, y son justamente las que ocupan los primeros niveles de las ordenaciones jerárquicas, con lo que se potencia ese círculo cerrado (Blackburn y Lingenfelter, 1973; Tan, 1986). La correlación entre familiaridad con la institución y la valoración del profesorado en el estudio de Jones et al. (1982) llega a ser del 0,98, lo que prueba la existencia de este efecto (Astin, 1985).

— El tamaño de las instituciones es una de las variables que tienen mayor influencia en la valoración del prestigio (Oromaner, 1970; Smith y Fiedler, 1971; Hagstrom, 1971; Elton y Rodgers, 1973; Lawrence y Green, 1980; Saunier, 1985; Astin, 1985; King y Wolfle, 1987). Incluso cuando el tamaño se combina con medidas objetivas, como es el caso del estudio de Jones et al. (1982), sólo se consideran valores totales (no per cápita), como si se estuviera realmente más interesado en la cantidad que en la calidad. Pero es evidente que el tamaño no hace mejores a los departamentos, sólo los hace más visibles (Lewis, 1984).

— Sólo son evaluadas con cierta eficacia las instituciones más destacadas y que tienen profesorado con reputación investigadora, pero no la gran mayoría de instituciones del país (Petrowski *et al.*, 1973; Webster, 1981; Tan, 1986). La mayoría de las instituciones de educación superior norteamericanas quedan sin evaluar por este sistema reputacional (Conrad y Blackburn, 1985a).

— La técnica con la que se realice la agregación para toda la institución de los datos de los departamentos, influye muy sensiblemente en los resultados (Astin, 1985). Por otra parte, las universidades con pocos programas, aunque sean de alta calidad, quedan siempre colocadas en posición desfavorable (Webster, 1985).

— La utilización exclusiva de evaluadores originarios del mundo académico deja de lado a los usuarios del sistema universitario: empleadores y estudiantes (Lawrence y Green, 1980).

— Los estudios que establecen correlaciones entre medidas objetivas y ordenaciones reputacionales demuestran que se puede obtener una buena estimación de la calidad de las instituciones, usando simplemente algunas medidas objetivas de muy sencillo uso, que hacen innecesarios los costosos estudios reputacionales (Astin, 1985).

— Los métodos reputacionales no informan sobre la calidad de los programas, sino sólo de la posición relativa que tienen entre ellos (Conrad y Blackburn, 1985a).

— La utilización del departamento como unidad de evaluación deja al margen toda una serie de características institucionales de indudable importancia en la calidad de la educación (Conrad y Blackburn, 1985a).

— Los evaluadores suelen ser los jefes de departamento o los profesores más veteranos, lo que puede no ser representativo de la opinión mayoritaria del mundo académico (Blackburn y Lingenfelter, 1973).

A estas críticas metodológicas hay que añadir otras de tipo más conceptual, que también recogen diversos autores, en especial algunos de los más críticos a estos métodos (Tyler, 1972; Dolan, 1976; Lawrence y Green, 1980). Son éstas:

— Refuerzan el estatus e impiden la innovación, ya que las escuelas ideológicas juegan un papel importante en estas evaluaciones (Dolan, 1976).

— Sólo una variable está siendo utilizada para evaluar la calidad: el prestigio del profesorado en el campo de la investigación (Dolan, 1976; Lawrence y Green, 1980; Tan, 1986; Webster, 1986).

— La utilización posterior de los estudios reputacionales es contraproducente para el sistema universitario como un todo, pudiendo provocar el que las universidades actúen sólo en las direcciones que mejoren su posición jerárquica relativa (Dolan, 1976). Por otra parte, estas jerarquizaciones son usadas para la asignación de fondos, con lo que se están creando diferencias cada vez mayores entre las instituciones.

— Ni la calidad del proceso educativo ni otros servicios a la sociedad son tenidos en cuenta en este tipo de evaluaciones, con lo que esto supone de reduccionismo de las misiones de la universidad (Dolan, 1976). En ningún caso se está evaluando la calidad del proceso productivo que tiene lugar en los departamentos (Webster y Conrad, 1986). Estos métodos no informan para «nada de la calidad de la instrucción, del grado de humanidad, del grado en

que la inquietud académica es alimentada por la interacción entre estudiantes y profesores» (Hartnett *et al.*, 1978: 1311).

Un problema adicional de los estudios reputacionales es el de la técnica utilizada para recabar la información. Han sido utilizados varios mecanismos para encuestar, lo que, por sí solo, ya puede tener cierta influencia en los resultados finales de estos estudios (Kane y Lawler, 1978).

Por otra parte, contestando a las anteriores críticas, algunos autores encuentran en los métodos reputacionales las siguientes ventajas:

— La calidad de la educación superior es algo tan complejo que no puede ser objetivizado. Del mismo modo que se valora la calidad de una orquesta sinfónica, sólo la opinión de expertos es una aproximación válida a la evaluación de una universidad.

— Aunque sólo midan la reputación, hay que tener en cuenta que es un concepto muy conectado con el de la calidad de los programas (Tan, 1986).

— A pesar de todos sus defectos, los estudios reputacionales han ayudado a la comprensión de la calidad de los programas en las universidades y en las escuelas profesionales norteamericanas (Conrad y Blackburn, 1985a).

— El sesgo del panel de evaluadores ha sido eliminado en los últimos estudios reputacionales.

— En cualquier caso, la subjetividad no puede ser evitada, sea cual sea la técnica usada, y, como demuestran algunos estudios que establecen correlaciones entre medidas objetivas y reputacionales, la relación entre unas ordenaciones y otras es muy estrecha.

— Los colegas del mundo académico son los más adecuados para juzgar el trabajo de sus iguales (Webster, 1981).

— Aunque no es la única condición, la calidad del profesorado es una condición necesaria para la calidad del programa. Además en los últimos estudios reputacionales (Jones *et al.*, 1982) también han sido incluidas otras variables.

— Actualmente, el mundo académico conoce con mucha rapidez las variaciones en los más remotos departamentos, por lo que no se produce un retraso temporal importante (Webster, 1986).

Vistas sus ventajas y desventajas, Conrad y Blackburn (1985a) recomiendan los estudios reputacionales, porque, a pesar de todas sus deficiencias, pueden ayudar a conocer mejor la calidad de los programas, sobre todo si se tienen en cuenta las debidas rectificaciones metodológicas sobre las cuestiones que han sido más ampliamente criticadas.

V.5.2. **Crítica a los métodos basados en medidas objetivas**

A pesar de que cabría esperar de este tipo de evaluaciones unos resultados más convincentes, no ha sido así en la realidad, y las evaluaciones que usan medidas objetivas no dejan de tener serias deficiencias. Sin embargo, no han dado lugar a una controversia tan notable como los estudios reputacionales. La discusión se ha centrado principalmente en el tipo de indicadores utilizados en cada evaluación. Los métodos que han usado los autores de estos estudios, así como la elección de indicadores, responde más a las preferencias de cada investigador (y presumiblemente a la posibilidad o facilidad de obtener los datos necesarios), que a una fundamentación teórica sobre el concepto de calidad y de su mejor evaluación. La falta de razonamiento teórico sobre la calidad hace que los indicadores usados sean realmente de dudosa validez en muchos casos.

Algunas de las deficiencias más notables, que diversos autores han encontrado en estos métodos de evaluación de la calidad, son las siguientes:

— Algunas de las medidas objetivas utilizadas tienen fuertes componentes subjetivas, como es el caso de la valoración de premios, ayudas o distinciones, que pueden haber sido concedidos por razones no estrictamente objetivas. Por otro lado, la misma elección de los indicadores más adecuados también tiene dosis de subjetividad (Webster, 1986).

— En la mayor parte de este tipo de estudios se identifica la productividad del profesorado con la calidad del programa (Tan, 1986). Muchos investigadores encuentran que no son la misma cosa y, por tanto, no deben confundirse (Smith y Fiedler, 1971; Liu, 1978).

— La mayoría de los estudios no prestan atención ni a la efectividad del programa, ni a la calidad docente (Conrad y Blackburn, 1985a).

— La utilización del número de publicaciones o de citas del profesorado como índice de calidad favorece más a los individuos (y, en consecuencia, a los departamentos), que están trabajando en temas más populares, que a

aquellos otros que se dedican a temas más obtusos, aunque quizá más interesantes a largo plazo. Tampoco se suele tener en cuenta la calidad de la publicación, y cuando se hace es a través del índice de citas, cuya correcta estimación está llena de dificultades no siempre superables (Webster, 1986). Como idea general se podría afirmar que este tipo de medidas objetivas miden más la cantidad que la calidad. Asimismo, es también problemática la evaluación de los distintos tipos de publicaciones para conseguir un solo indicador agregado de productividad; es evidente que distintos sistemas pueden producir resultados muy diferentes (Webster y Conrad, 1986).

— Estos estudios tampoco tienen en cuenta las aportaciones de los profesores a todo tipo de actividades, académicas o extraacadémicas, de indudable importancia social (Webster y Conrad, 1986).

— Utilizar una medida, en principio tan adecuada para evaluar la calidad de las instituciones, como son los resultados de los alumnos después de acabar los estudios, tiene el problema del gran lapso de tiempo que transcurre entre la evaluación y la situación real que se está evaluando, con lo que se pierde efectividad en el objetivo de la evaluación, que debe ser la mejora de la calidad de las instituciones que se evalúan (Webster, 1986).

— El criterio de selectividad para el ingreso en el centro, como medida de calidad de la institución, no está teniendo en cuenta para nada el proceso educativo de la institución, por lo que es dudoso que pueda usarse como indicador de su nivel de calidad (Webster, 1986).

— Las medidas objetivas sobre los recursos disponibles en las instituciones no tienen en cuenta cómo están siendo utilizados esos recursos. Mientras que, por otra parte, no existe una clara evidencia de la relación directa entre recursos y calidad de la educación (Webster, 1986).

— Ninguno de estos estudios tiene en cuenta el *valor añadido* a los estudiantes y ni siquiera intenta valorarlo (Conrad y Blackburn, 1985a).

A pesar de todas sus deficiencias, hay que reconocer que los estudios basados en medidas objetivas han colaborado en gran medida a profundizar en el conocimiento del concepto de calidad en la educación superior (Tan, 1986). Estos estudios han ayudado a adoptar un enfoque multidimensional sobre la calidad de las instituciones universitarias, que ayuda a comprender los problemas de la educación superior (Conrad y Blackburn, 1985a). Además de estas importantes aportaciones, diversos autores han mencionado las siguientes ventajas:

— Las medidas objetivas tienen en principio más valor científico que las subjetivas (Tan, 1986), permitiendo la profundización metodológica.

— La utilización de los estudiantes (tanto de los que ingresan en la institución como de los que salen de ella) como medidas objetivas de la calidad, tiene la ventaja de que en todos los centros, y no sólo en los más prestigiosos, se pueden hacer evaluaciones (Webster, 1986).

— La productividad del profesorado, si se mide por el número de citaciones, valora claramente la importancia e influencia de las investigaciones de los departamentos, así como la actualidad del trabajo departamental (Webster, 1986).

V.5.3. **Requerimientos para una evaluación ideal**

A pesar de todos los defectos de los métodos de evaluación de la calidad de las universidades, estas evaluaciones son necesarias, dada la gran cantidad de instituciones universitarias que existen en Estados Unidos. Esto se hace especialmente cierto para las que no están en el vértice de la pirámide jerárquica (Lawrence y Green, 1980). En un mercado tan variado y complejo, las evaluaciones de la calidad son un método de informar a la opinión pública, a los usuarios del estado y funcionamiento real de las instituciones.

Los resultados de las comparaciones entre medidas objetivas y subjetivas no llevan a ninguna conclusión definitiva en ningún sentido (Webster y Conrad, 1986), por lo que las evaluaciones que utilizan ambos enfoques pueden ser las más recomendables en el momento actual.

Lawrence y Green (1980), como resumen de su estudio sobre la calidad de la educación superior, hacen las siguientes recomendaciones sobre el desarrollo futuro de estos tipos de estudios:

- Los estudios deben extenderse a todas las instituciones.
- Deben usarse indicadores múltiples.
- Las opiniones de los consumidores deben incorporarse al proceso evaluador.
- Deben intentarse evaluar los beneficios sociales e individuales de la educación.

— Se debe prestar mayor atención al aprendizaje y desarrollo de los estudiantes como objetivos de la educación superior.

— Debe considerarse la posibilidad de diferentes indicadores por programas, por sexo, por razas, etcétera.

Para Troutt (1981), por otra parte, todos los análisis de la calidad deberían cumplir tres requisitos:

— Emplear medidas de los logros intelectuales y de los resultados cognoscitivos.

— Recoger datos de instituciones de diversos tipos.

— Proporcionar información sobre el cambio de los estudiantes entre su entrada y algún momento posterior de su proceso educativo.

Por su parte, Webster (1986), resumiendo las ventajas y desventajas de cada método, recomienda los siguientes criterios para realizar una evaluación jerárquica de las universidades:

— Debe ser multidimensional, aunque está todavía por determinar qué características deben tenerse en cuenta para construir un indicador de calidad óptimo.

— Las medidas usadas deben tener en cuenta los resultados de la totalidad de los profesores o estudiantes, y no sólo a los de unos pocos más sobresalientes.

— Los estudios deben basarse siempre en medidas *per cápita* y no en cantidades agregadas, como ha sido habitual.

— Los estudios deberían no sólo comparar instituciones, sino establecer algún tipo de valoración respecto a unas medidas externas o reglas normalizadas.

— Deberían incluirse, además de departamentos, especialidades que sean especialmente significativas.

— Deberían incluirse las participaciones de departamentos auxiliares que intervienen en los programas.

— Es necesario incluir alguna medida de lo que los estudiantes aprenden en la universidad.

En cualquier caso, está por hacer la evaluación que mida el *valor añadido* en el sentido de Astin (1985), es decir, que valore el efecto global que la institución universitaria ejerce sobre el individuo que pasa por ella. Esta sería, sin lugar a dudas, la mejor medida de la calidad de una universidad.

CAPITULO VII. LA EVALUACION DE LA EDUCACION SUPERIOR EN DIVERSOS PAISES

The cultivation of quality, efficiency and enterprise (...) are not in themselves matters on which any rightminded person would dispute. We are all in favour of them, just as we are in favour of gastronomy and a thin waistline (...). When the controversy arises is their operational, legislative and financial consequences for the individual establishment that go to make up the nations' system of higher education.

Neave, 1988: 7.

VII.1. INTRODUCCION

En este capítulo se hace referencia a situaciones y experiencias concretas de la evaluación de la educación superior en un conjunto de países. La información disponible sobre estas experiencias es reducida no sólo porque no estén todavía demasiado generalizadas, sino también porque la participación en el «mercado internacional de la información» es desigual, siendo algo escasa la que proviene de países de habla no inglesa. No sólo es en Estados Unidos donde más atención se ha prestado a la evaluación de la educación superior, sino también el país que, con gran diferencia, más importación proporciona sobre este tema. Aunque en menor medida, también existe una amplia información sobre la evaluación institucional en el Reino Unido, Australia y Canadá. Cabe destacar muy especialmente la gran cantidad de aportaciones realizadas al estudio de la evaluación institucional en Holanda. Alguna información se puede también encontrar sobre la evaluación en otros países europeos, como Alemania, Francia, Suecia y Finlandia, pero poco más es posible.

Como consecuencia de esta situación, la información que se recoge en este capítulo es necesariamente parcial, ya que sólo recoge aquella parte de la información sobre la práctica de la evaluación de la educación superior que está publicada en los medios habituales de comunicación científica o que ha sido presentada en recientes congresos científicos sobre estos temas. La información que se proporciona en este capítulo no es tampoco sistemática, ya que se relatan experiencias concretas en diversos países, no siempre conexas.

nadas entre sí, pero no se pretende hacer una revisión exhaustiva de todas ellas.

El capítulo se ha dividido en tres secciones: tras esta introducción, la segunda sección está referida a tres países no europeos, de los que se dispone de suficiente información: USA, Australia y Canadá; la tercera sección se dedica a los pocos países de la Europa Occidental en los que algún sistema de evaluación institucional está actualmente en marcha.

Además de los países ya señalados, en algunos otros se han hecho también intentos de evaluación de la educación superior que merecen indicarse:

— Un estudio sobre la efectividad de las universidades de Indonesia realizado por Johnstone y Agustiar (1983).

— Una revisión de trabajos que evalúan la calidad en las universidades venezolanas realizado por Salcedo (1988).

— Una descripción de la organización y sistemas de control de la calidad en las universidades israelíes llevado a cabo por Iram (1987).

— Una metodología de evaluación institucional aplicada en instituciones de varios países latinoamericanos (Escotet, 1990a, 1990b).

VII.2. **ESTADOS UNIDOS, CANADÁ, AUSTRALIA**

En esta sección se recoge información sobre tres países en los que la evaluación de la educación superior ha tenido un amplio desarrollo en los últimos tiempos. Tienen también algo más en común: la existencia de sistemas de educación superior muy diversificados y plenamente autónomos. Algunos casos concretos de evaluación de sus instituciones de educación superior serán brevemente relatados para cada uno de estos países.

VII.2.1. **Estados Unidos**

Aunque la mayor parte de la información que se ha aportado en capítulos anteriores está referida a los Estados Unidos, en este apartado se tratarán específicamente algunos casos prácticos de métodos de evaluación de las instituciones y de los programas, con exclusión de los estudios que tratan de valorar la calidad de las instituciones mediante la utilización de comparaciones

y ordenaciones jerárquicas, tema al que ya se dedicó específicamente el capítulo anterior.

Es necesario señalar una vez más que el sistema de educación superior norteamericano está constituido por más de 3.000 instituciones con una gran diversidad de situaciones: pública y privadas; independientes y confesionales; controladas por los estados y no controladas; especializadas en investigación o no; concediendo doctorados o no; selectivas o abiertas, etc. (Carnegie Council, 1980; Carnegie Foundation, 1987; Kerr, 1990). Este sistema, aparentemente anárquico, es flexible, competitivo y despolitizado (Kells, 1986). Otra particularidad de este sistema es que la mayor parte del poder está en las presidencias de cada universidad, no como en los sistemas europeos, en los que el poder radica en los gobiernos, por un lado, y el profesorado (bajo la tradicional estructura de cátedras), por otro (Maassen, 1988). Veremos seguidamente algunas de las experiencias norteamericanas en la práctica de diferentes facetas de la evaluación institucional.

El sistema de evaluación institucional en los Estados Unidos

En Estados Unidos la educación en todos sus niveles es una competencia de los estados. Sin embargo, éstos no ejercieron ningún tipo de control sobre las instituciones de educación superior hasta hace relativamente muy poco tiempo. A principios de este siglo las instituciones de educación superior habían crecido y se habían diversificado enormemente. Fue entonces cuando empezó a verse la necesidad de evaluarlas con el fin de permitir trasvases de estudiantes entre ellas, pudiéndose valorar además la calidad de los diversos títulos profesionales. El sistema que se creó fue el de las acreditaciones (a las que ya se dedicó una amplia información en el capítulo V), un método basado en las revisiones realizadas por miembros de la comunidad académica. Las acreditaciones han cambiado a lo largo del tiempo; actualmente son un sistema complejo de autorrevisión y de revisión externa que tiene lugar cada diez años en la mayoría de las instituciones de educación superior. El sistema de acreditación es el dominante en el sistema educativo norteamericano. Sin embargo, en los años sesenta, como consecuencia del crecimiento del sistema universitario en alumnos y presupuestos, algunos estados empezaron a ejercer algún tipo de control sobre las instituciones, de modo que en la actualidad la mayoría de los estados tienen algún sistema de evaluación de programas en las instituciones que forman parte de sus respectivos sistemas estatales de educación superior (Young, 1988). La evaluación de la educación superior se ha convertido en una preocupación de una gran parte de los gobiernos estatales, de modo que, según una encuesta realizada por Boyer *et al.* (1987),

dos tercios de los estados ya han iniciado algún sistema de evaluación de la calidad de las instituciones públicas de educación superior.

Por otra parte, las propias instituciones universitarias iniciaron hace algunos años investigaciones sobre la eficiencia de su propio funcionamiento, revisiones de programas, planificaciones estratégicas, etc. Todo ello como consecuencia de la escasez de recursos, los costes crecientes y los cambios demográficos que obligaron a las universidades a prestar una mayor atención a los problemas de efectividad y eficiencia (Kells, 1986). El panorama de las evaluaciones de las instituciones en Estados Unidos se puede resumir en estos tres grupos:

— *Revisión y permisos de los estados*

Sólo nueve de los 50 estados están presentes de una forma activa en procesos de control de las instituciones de educación superior. El proceso varía desde una simple licencia inicial a las instituciones para que puedan iniciar su funcionamiento, a revisiones periódicas de todo el proceso institucional.

— *Acreditación por agencias regionales*

Virtualmente abarca a todas las instituciones del país. El proceso incluye una autorrevisión junto con una visita de estudio por parte de un equipo externo evaluador.

— *Evaluaciones y planificación iniciadas por las propias instituciones*

Una actividad de uso creciente, que abarca todo tipo de métodos y sistemas de evaluación.

Si pasamos del nivel institucional al de los programas, hay cuatro tipos de acciones que son frecuentes como métodos de evaluación (Kells, 1986):

— Pruebas estatales para validar títulos de carácter profesional.

— Revisiones cíclicas de los programas, unas veces obligadas desde los estados y otras solamente recomendadas.

— Acreditaciones especializadas para los programas de carácter profesional.

— Revisiones de programas promovidas y realizadas desde las propias instituciones.

El uso de indicadores de rendimiento

En los Estados Unidos la asignación de fondos a las universidades estatales está casi completamente determinada por la historia previa de la institución y por el número de matriculaciones en cada curso. Los criterios de asignación de fondos están basados en la actividad, pero no en la efectividad del proceso o en la calidad de los resultados (Bogue, 1982). Este sistema tiene la ventaja de que se garantiza la equidad y la objetividad en el reparto de fondos, pero tiene algunos inconvenientes como el que se enfatiza el crecimiento más que la calidad, no se estimulan las mejoras y se puede inducir una disminución de los niveles de exigencias a los alumnos para conseguir mantener altos niveles de matriculación.

Ultimamente se están poniendo en marcha procesos de asignación de recursos entre las universidades utilizando criterios más complejos basados en los indicadores de rendimiento. Dos ejemplos de la utilización de estos indicadores se citan a continuación.

El *Sistema Público de Educación Superior del Estado de Ohio* es un ejemplo (Coulter y Moore, 1987) de utilización de indicadores de rendimiento para la asignación de recursos en un sistema público. Este sistema desarrolló un programa selectivo de calidad en el que se asignaban recompensas especiales a aquellos centros en los que se cumplieran determinados requisitos. Para la asignación de estas recompensas (contratos especiales, becas, ayudas financieras, etc.) se exigía la puesta en marcha de toda una serie de indicadores que evaluaran los méritos de los programas que solicitaban las ayudas.

También el *Sistema Estatal de Tennessee* creó un sistema especial de asignación de fondos, que puede llegar a alcanzar hasta el 5 % del total del presupuesto de la institución, para aquéllas que demuestren haber alcanzado un aceptable nivel en estos cinco indicadores (Bogue, 1982; Banta *et al.*, 1986):

- Porcentaje de programas sometidos a acreditación que la consiguen.
- Porcentaje de programas que se han sometido a una revisión en los últimos cinco años o a un examen general de resultados en las materias que se imparten.
- *Valor añadido* por la educación proporcionada a los estudiantes, medido por la diferencia entre una prueba realizada al ingresar en la universidad y otra a la salida.
- Opiniones sobre la calidad recogidas de estudiantes, antiguos alumnos, empleadores y miembros de la comunidad.

— Implementación de un plan de mejoras en la institución como consecuencia de los resultados de las evaluaciones.

Autoestudios y revisiones de programa

Los autoestudios de las instituciones y las revisiones de los programas constituyen uno de los métodos habituales de investigación institucional en la educación superior en USA. Los autoestudios siguen un método normalizado por las propias agencias de acreditación, aunque pueden existir algunas variaciones en la puesta en práctica (O'Brien y Waters, 1983; Kells, 1988b). Sin embargo, la revisión de programas constituye una actividad más idiosincrática de cada institución o sistema educativo, de la que se pueden citar algunos ejemplos de interés. Dado el objetivo primordial de este estudio, se hará referencia solamente a casos de evaluación de programas realizados en sistemas universitarios de carácter estatal.

Un ejemplo de revisión de programas en una institución es el del *Chicago City-Wide College*, una institución pública que proporciona educación superior de dos años de duración, y que tenía una matriculación de 115.000 estudiantes en 1987 (Hart, 1988). En este centro se realiza una revisión sistemática de los programas, de modo que cada cinco años son revisados todos. La intención primaria de las revisiones es la mejora de los programas, aunque secundariamente proporcionan información a los gestores y se cumple un mandamiento legal del Estado de Illinois. Cada programa es revisado bajo tres perspectivas: necesidades del programa, calidad y costes. Los siguientes indicadores son utilizados para reflejar cada uno de estos aspectos:

- Características e información sobre el programa.
- Necesidades del programa:
 - Demanda del programa por parte de los estudiantes y de la comunidad.
 - Matriculaciones en los últimos tres años y perspectivas para los próximos dos.
 - Características de los estudiantes que se matriculan en el programa.
 - Intereses educativos y ocupacionales de los estudiantes.
 - Relación del programa con el mercado laboral, medido a través de información sobre la ocupación de los graduados y a través de las perspectivas sobre la demanda de empleo para los graduados de ese programa.

— Calidad del programa:

– Datos sobre retención de estudiantes, tasa de abandonos y de traslados de los últimos tres años, del momento presente y perspectivas para los próximos dos años.

— Datos sobre finalización de estudios y empleo posterior de los graduados para los mismos períodos de tiempo que en el caso anterior.

— Seguimiento del grado de satisfacción de los estudiantes y de la validez del programa con fines profesionales.

— Valoraciones sobre la calidad del profesorado, del currículum y de los servicios relacionados con el programa.

— Costes del programa:

– Desarrollo de una tabla de costes por unidad.

– Análisis de los factores subyacentes a cada uno de los costes.

– Comparaciones de estos costes con los costes medios de otras instituciones del Estado.

— Recomendaciones del programa:

– Recomendaciones relacionadas con los objetivos del programa, con las admisiones y con las características del programa.

– Recomendaciones de carácter administrativo.

– Recomendaciones de carácter docente y relacionadas con las unidades de apoyo.

— Planes de acción sobre las recomendaciones que han sido hechas con anterioridad.

Otro ejemplo de puesta en práctica de revisiones de programas es el de la *Universidad de California*, una institución pública con nueve campus autónomos entre sí, que inició en 1982 una sistemática revisión de sus programas en todos sus campos. La experiencia se enfrentó al principio a fuertes resistencias del profesorado, celoso de la autonomía de cada campus y del mantenimiento de la confidencialidad de los resultados. A pesar de este inicio, en la actualidad

es un mecanismo aceptado, y puede considerarse uno de los sistemas más eficaces que utiliza esta universidad para mejorar la calidad (Balderston, 1985).

El *Sistema Universitario Estatal de Florida* también se planteó, ya a finales de los años setenta, los problemas generales de la educación superior, que se incrementaban en el caso especial de este Estado a causa del fuerte crecimiento demográfico experimentado, unido a importantes cambios en las características de la población. En 1983 se iniciaron las revisiones de programas académicos para todo el sistema público de educación superior del Estado, constituido por nueve universidades de distintas tipologías. Las revisiones son cíclicas, de modo que, tras una revisión inicial, se realiza un nuevo seguimiento cinco años después. Las revisiones son presididas por una persona de prestigio de fuera del Estado y por otra de la propia universidad, que es distinta para cada programa revisado, mientras que el equipo está formado por representantes del Estado, de la disciplina objeto de estudio, del propio centro y por dos expertos de reconocido prestigio, uno nombrado por el Estado y otro por la propia institución. Para llevar a cabo la revisión, el equipo realiza una visita a la institución, de la que emite un informe sobre los programas. Este informe puede ser replicado posteriormente por el departamento. Toda la información pasa finalmente a un comité del Estado que, bien toma acciones en la dirección que recomiendan el equipo de la revisión o bien puede rechazar estas recomendaciones.

Evaluaciones de resultados

Otro de los aspectos de la evaluación de la educación superior, a los que se ha dedicado especial atención en los Estados Unidos, es el de la valoración de los resultados del proceso educativo, es decir, al valor añadido que el proceso educativo superior genera sobre quien pasa por él. Hay que citar en este sentido las investigaciones llevadas a cabo por el *National Center of Higher Education Management Systems* (NCHEMS) (Ewell, 1985), que están tratando de encontrar los mecanismos más adecuados para la evaluación de los resultados del proceso educativo.

Uno de los modelos de evaluación de resultados más interesantes es el realizado en la *Northeast Missouri State University*. En esta universidad se implementó, durante trece años, un estudio sobre el *valor añadido* a los estudiantes durante su permanencia en el centro (McClain y Krueger, 1985; McClain *et al.*, 1986). El método consistió básicamente en la realización de una prueba a todos los alumnos al ingresar en la universidad, y otra al finalizar, evaluando así el resultado del proceso de aprendizaje. En las pruebas no sólo se miden conocimientos, sino muy principalmente cualidades personales de

los individuos y su capacidad para enfrentarse a los problemas reales. En contraposición de otras medidas de la calidad, por este método se consiguen evaluar las ganancias reales de un individuo en conocimientos, capacidad analítica, destrezas, valores humanos y culturales, junto con otras cualidades del desarrollo personal.

Con un espíritu algo diferente se realizó un estudio de evaluación de los resultados educativos en la *University of Tennessee*, Knoxville (Banta, 1985; Banta *et al.*, 1986). En esta universidad los alumnos que finalizan los estudios deben realizar una prueba de evaluación de sus conocimientos; con los resultados de esta prueba la universidad valora los puntos débiles y fuertes del currículum. Se realizaron también encuestas sobre calidad de los programas académicos y de los servicios entre alumnos y antiguos alumnos (no sólo de los que han acabado los estudios, sino también de los que los han abandonado antes de finalizarlos). La evaluación se completó con revisiones por profesores de otras universidades, que realizaron visitas para hacer estimaciones de la calidad de los programas y de los servicios. La universidad está continuamente evaluando la efectividad del sistema con el objeto de alcanzar mejor los objetivos de la institución, de tal modo que son numerosas las medidas de tipo administrativo, docente y organizativo que han sido tomadas siguiendo las sugerencias que se desprenden de las evaluaciones realizadas.

VII.2.2. *Canadá*

La experiencia en autoestudios y revisiones de programas de los Estados Unidos es trasladable fácilmente al Canadá. Evaluaciones de diferentes tipos realizadas en la Universidad de Alberta, en la McGill University, en la Universidad de Guelph y en el Sistema Universitario de Quebec son algunos ejemplos que se van a referir a continuación sobre el uso de evaluaciones, básicamente autorreguladas, en las universidades canadienses.

La experiencia de la Universidad de Alberta

La *Universidad de Alberta* inició sus revisiones de programas ya en 1980, bajo los siguientes principios y objetivos declarados (Holdaway, 1988):

— La meta de la revisión debe ser la recogida de información que sirva como una base fiable para la planificación y la mejora de los programas.

— Los mecanismos que deben usarse en la revisión deben ser comprensibles, legibles, efectivos, eficientes, flexibles, creíbles y realistas.

— Se debe constituir un consejo asesor para las revisiones a nivel de la institución, cuyas misiones deben ser:

— Asesorar sobre el orden con el que se deben efectuar las revisiones y sobre la composición de los equipos.

— Asesorar a los equipos sobre los requerimientos de la revisión.

— Transmitir los informes recibidos.

— Mantener a la comunidad universitaria informada sobre los resultados.

— Cada unidad revisada debe realizar un autoestudio y emitir como resultado un informe.

— El equipo de la revisión debe estar constituido por dos miembros de la misma disciplina, pero de otras universidades, dos de la misma universidad, uno del departamento que es evaluado y otro de cualquier otra disciplina.

— Cada unidad de revisión deberá consultar a los estudiantes, a los graduados recientes, al personal académico y no académico, y a los posibles empleadores de los graduados, si se considera adecuado.

Este proceso de revisión de los programas en la Universidad de Alberta se realiza a un ritmo de 15 programas al año, de modo que se completa una revisión de todas las unidades en diez años. El proceso típico de una revisión consta de los siguientes pasos:

— Asesoramiento a la unidad que va a ser revisada de cómo hacer el autoestudio.

— Realización del autoestudio por parte de la unidad. Este proceso tarda entre seis y doce meses. Estos informes deben tratar la historia de la unidad, la situación actual, las cuestiones claves, una franca y equilibrada consideración sobre los puntos fuertes y débiles de la unidad, estrategias para cambios futuros y dotación para dichos cambios y el papel de los individuos y de los grupos. Se suministran autoestudios de otros departamentos como modelo. Se espera que el estudio involucre a la mayor parte de los individuos de la unidad, aunque esté dirigido por un pequeño grupo.

— Nombramiento de un equipo de revisión constituido por dos evaluadores externos y dos internos. Uno de los internos debe conocer bien el departa-

mento bajo revisión, mientras que el otro representa los intereses generales de la universidad. La lista de evaluadores es acordada por la unidad y el comité asesor de revisiones de la universidad.

— Un informe sobre cómo realizar la revisión y un ejemplar del autoestudio son entregados a cada miembro del equipo con suficiente antelación a la visita. En este informe se le pide al equipo que opine sobre los puntos fuertes y débiles del programa docente e investigador; haga una valoración del personal y de su carga de trabajo; estime lo adecuado de los recursos disponibles; valore la efectividad de la organización del departamento y sus relaciones con otros departamentos. Muy especialmente, se solicita al equipo evaluador que señale aquellos puntos que pueden ser mejorados en relación con los recursos disponibles.

— El equipo evaluador, en un plazo de dos meses, envía sus conclusiones al comité de evaluación de la universidad, quien envía una copia a la unidad revisada, de modo que esta unidad pueda emitir una respuesta a la evaluación.

— Con los tres informes (autoestudio del departamento, informe del equipo externo y respuesta del departamento) el comité resume los principales rasgos del programa y los discute con el jefe del departamento.

— El proceso finaliza con reuniones entre el comité de revisión de la universidad, el comité de planificación de la universidad, el jefe del departamento y el rector de la universidad. Como consecuencia de esta o estas reuniones son adoptadas medidas en el departamento que mejoren su calidad en las direcciones señaladas por el informe final.

— A los dieciocho meses de estas reuniones el departamento envía un informe de los progresos realizados al vicerrector de asuntos académicos.

Es de señalar que en algunas facultades con varios departamentos el proceso de revisión se realiza simultáneamente en todos ellos e incluso se presenta un estudio complementario de la facultad como un conjunto.

Holdaway (1988), relatando su experiencia en la Universidad de Alberta, recomienda estos siete puntos para que una revisión tenga éxito:

- Centrar la atención en la mejora de los programas.
- Nombrar un comité de revisión presidido por un profesor respetado, que no sea un administrador.

— Realizar tres informes: el autoestudio, la evaluación externa y la contestación del departamento.

— Tratar los informes del equipo externo con la apropiada confidencialidad, pero asegurarse que los procedimientos son abiertos y claros.

— Compensar económicamente a los departamentos que están realizando la revisión.

— Nombrar equipos mixtos de evaluadores internos y externos.

— Responsabilizar a un administrador de la tarea de seguir el progreso del departamento hacia las metas que se le recomendaron.

Otras experiencias canadienses

Ya hace más de una década que el *Centre d'animation, de développement et de reserche en education* (CADRE), de Montreal, organizó la institucionalización de autoevaluaciones entre las universidades de Quebec. Cada institución del sistema universitario de Quebec realiza su propia autoevaluación, que es pasada posteriormente al CADRE, cuyos expertos en investigación institucional la analizan y extraen las pertinentes conclusiones para todo el sistema (Calvert, 1981).

La *McGill University* inició un plan de revisiones de programas en 1982 con un ciclo de siete años, revisando una parte proporcional de los programas cada año. Tras la preparación de un autoestudio por el departamento, se obtienen dos informes sobre el programa, uno emitido por un equipo interno formado por estudiantes y miembros del departamento que está siendo revisado, y otro realizado por un equipo externo formado por miembros de otras universidades (Holdaway, 1988).

Jamieson (1988) relata la experiencia en autoestudio realizada por la Universidad de Guelph, en Ontario. El autoestudio, de carácter no estructurado, controla la calidad del proceso docente e investigador, revisa el cumplimiento de las metas institucionales y determina las necesidades de recursos necesarios para el buen funcionamiento de la institución. Una aplicación importante de este autoestudio es la de reasignar fondos desde los departamentos sobredotados a los menos dotados.

VII.2.3. **Australia**

En Australia existe una organización gubernamental responsable de la

educación superior: la *Commonwealth Tertiary Education Commission* (CTEC). Esta organización recomendó insistentemente a las universidades que pusieran en marcha sistemas de revisión de programas o de departamentos (Roe *et al.*, 1986). Existe también un instituto especializado en la investigación de los temas relacionados con la educación superior: el *Higher Education Research and Development Society of Australasia* (HERDSA), que se encarga de preparar metodológicamente las revisiones que desean hacer las universidades.

Miembros de HERDSA, con el apoyo de CTEC, desarrollaron un amplio proyecto para el diseño de un sistema de evaluación de la educación superior en las instituciones australianas. El proyecto fue llevado a cabo mediante reuniones de trabajo de jefes de departamento de distintas universidades que discutieron sobre tres temas básicos: evaluación de asignaturas y de docencia, evaluación de programas o de departamentos y evaluación de los profesores. Los resultados de este proyecto se presentaron en dos libros: el primero está dedicado a la evaluación de la docencia (Roe y McDonald, 1984), mientras que el segundo (Roe, McDonald y Moses, 1986) se dedica expresamente a la evaluación de departamentos y a la evaluación individual, tanto docente como investigadora. La influencia de los métodos de evaluación norteamericanos es patente en los resultados de este interesante proyecto metodológico.

Por otra parte, la CTEC realizó una encuesta pública sobre la manera de mejorar la eficiencia y la efectividad de las universidades australianas. Los resultados son presentados en un informe (Commonwealth Tertiary Education Commission, 1986), en el que se ofrece una amplia visión de la situación de las instituciones universitarias australianas desde el punto de vista del rendimiento. Entre las recomendaciones que se recogen en el informe para aumentar el rendimiento de las universidades citaremos las dos más directamente relacionadas con la evaluación de la calidad. Son éstas (Commonwealth Tertiary Education Commission, 1986: 29):

— «La AVCC y la ACDP (asociaciones de las universidades australianas) desarrollarán, de acuerdo con el CTEC, principios y procedimientos para la implementación de evaluaciones internas y regulares de los departamentos académicos y la extensión de este procedimiento a otras actividades (servicios y unidades administrativas) de las instituciones.

— Las instituciones de educación superior proporcionarán informes regulares al CTEC sobre los procedimientos usados para evaluar el rendimiento académico y su relación con la gestión institucional, de modo que la CTEC pueda realizar una valoración del rendimiento institucional en este área en su informe trienal.»

Como se puede apreciar, en Australia se está siguiendo un método básicamente autorregulado como sistema de controlar y mejorar la calidad de las instituciones universitarias, de modo que el control gubernamental sobre las instituciones de educación superior se limita a solicitar que los resultados de las autoevaluaciones le sean remitido cada tres años.

Por otra parte, la utilización de autoevaluaciones es algo que viene de antiguo en aquellas universidades: Calvert (1981) ya relata alguna experiencia de este tipo que se inició en los años setenta. Por último, señalar que el análisis de la efectividad de las instituciones de educación superior en Australia también ha sido objeto de estudio (Lysons y Ryder, 1988), siguiendo las mismas técnicas de medida de la efectividad que realizó Cameron (1978) para Estados Unidos.

VII.3. PAÍSES DE EUROPA OCCIDENTAL

La situación de la educación superior en Europa Occidental no es uniforme. De un modo muy general se puede decir que existen dos tipos de sistemas universitarios: los que han gozado de amplia autonomía en el pasado y ahora se mueven hacia sistemas más regulados y controlados por el Gobierno (el caso del Reino Unido) y los que originalmente eran sistemas centralizados bajo control gubernamental, pero que se dirigen ahora hacia situaciones de mayor autonomía (Suecia, Holanda o España son casos representativos de esta situación). Como consecuencia de esta doble tendencia, es posible que en un próximo futuro se alcance una cierta situación de homogeneidad, fruto de la confluencia en un punto más o menos común de sistemas con orígenes opuestos.

La información que se recoge en esta sección está referida a unos pocos países que tienen establecido algún modelo sistemático de evaluación institucional. Se presta atención muy especialmente a dos países, Reino Unido y Holanda, en los que ha existido un mayor interés por el problema de la evaluación del sistema universitario. En cualquier caso, dada la diversidad de sistemas universitarios, no se pueden sacar conclusiones únicas sobre la situación de la evaluación institucional en los distintos países de la Europa Occidental, salvo quizá una que acuñó Neave (1988): que en todos ellos hay un auténtico despertar de un «estado evaluador».

VII.3.1. Visión general de estos países

Los sistemas nacionales de educación superior europeos estuvieron sometidos en el pasado reciente a una gran demanda social que multiplicó el

número de estudiantes universitarios. Estos sistemas tuvieron que enfrentarse al reto de fundar nuevas universidades, construir nuevos edificios, contratar personal y elevar cada vez más los gastos de todo tipo, sin que diera tiempo a designar nuevas estructuras y formas de organización que adaptaran una universidad de élite a una de masas que funcionara correctamente.

Como consecuencia de ese crecimiento tan rápido, la calidad de la educación quizá no recibió siempre la debida atención. Alcanzada en la actualidad una cierta situación de estabilidad en las universidades europeas, la preocupación por la calidad ha despertado con fuerza. Las razones no son exclusivamente académicas, sino que, unido a la posible devaluación de los niveles de calidad en las universidades, tanto en investigación como en docencia, está el problema de la competencia y de la supervivencia tecnológica en el mercado mundial, para cuya solución parece necesaria una educación superior de alta calidad (Hüfner y Rau, 1987). Recientemente, la educación superior está siendo vista en muchos países como la herramienta para entrar en un nuevo período de desarrollo económico, para lo que se necesita una educación superior creativa y de calidad, que produzca científicos y técnicos capaces de generar los avances tecnológicos necesarios y profesionales capaces de gestionar esos avances (Vught, 1988).

Otro motivo del interés actual por la educación superior en Europa hay que buscarlo en las altas tasas de desempleo juvenil que se alcanzaron en la última década, incluso entre titulados universitarios. Ante esta situación se volvieron las miradas hacia las universidades, pidiéndoles una educación más efectiva que ayudara a los jóvenes a desenvolverse mejor en el mercado laboral.

Por otra parte, durante los años setenta y ochenta la mayoría de los sistemas de educación superior en Europa se tuvieron que enfrentar con nuevas situaciones originadas por un cambio de las estrategias de los gobiernos hacia la educación universitaria. Este cambio no es independiente de la subida al poder de gobiernos conservadores en una gran parte de estos países. La filosofía del «valor del dinero» de estos gobiernos acabó con la política de financiamiento incondicional de las universidades que se había practicado en épocas anteriores. Como consecuencia, el control de la calidad de la investigación y de la docencia se ha convertido en un tema clave de la política educativa para los gobiernos de muchos países europeos (Goedegebuure *et al.*, 1990).

Como se puede apreciar, son varios los motivos que inciden en la actual preocupación por la calidad de las universidades en Europa Occidental. Resumiendo, se podría decir que es la consecuencia de tres hechos fundamen-

tales: la necesidad de limitar el gasto social y de gestionarlo más eficazmente; la necesidad de orientar a los estudiantes hacia los estudios con mejores posibilidades de empleo, y la necesidad de desarrollar estudios de calidad que posibiliten una educación adecuada a una economía de alta tecnología (Neave, 1986).

Para conseguir este aumento de la calidad en las universidades europeas se están siguiendo dos tipos de estrategias, una de más control y planificación estatal, y otra que justo hace lo contrario: crear condiciones de mayor autonomía, de modo que en las instituciones se cree el clima necesario para desarrollar la calidad y la creatividad. En esta última postura el Gobierno no desaparece de la escena, sino que actúa influenciando a las instituciones por medio de recompensas y sanciones (Vught, 1988). En esta estrategia, que puede ser llamada de autorregulación, se pone el énfasis en la capacidad autorregulatoria de las instituciones dentro de un marco general proporcionado por el Gobierno.

En un sistema autorregulado las instituciones están sometidas a un control de calidad *ex post*, recompensándolas por su nivel de rendimiento, e incluso penalizando la carencia del adecuado nivel de rendimiento. Las instituciones se hacen también responsables de sus propios asuntos financieros, deben atraer a los clientes y adaptarse mejor a su entorno, es decir, «las instituciones deben buscar su propio nicho en el mercado» (Maassen, 1988: 1).

En varios países europeos se están dando pasos en el sentido de la desregularización de las instituciones universitarias (Holanda, Alemania, Francia, Suecia y España, entre otros). Curiosamente, en muchos de estos países los políticos y administradores de las universidades vuelven sus miradas hacia el sistema universitario norteamericano, al que ven como flexible, eficiente, competitivo, diversificado y variado, como un modelo a seguir para resolver los problemas de la educación superior europea (Kells, 1986; Hufner y Rau, 1987).

El nacimiento de un «estado evaluador» generalizado en las universidades europeas puede considerarse el mayor cambio que está teniendo lugar dentro de las relaciones entre educación superior, Gobierno y sociedad en la Europa Occidental en los últimos años. Aunque el desarrollo de esta situación no es uniforme, mientras que hay países que ya han avanzado un cierto camino: Finlandia, Francia, Holanda, Suecia y el Reino Unido (Neave, 1988), en otros se están iniciando los primeros pasos, pero, en cualquier caso, la evaluación de la calidad en sus diferentes dimensiones y configuraciones (autoevaluación, revisiones por colegas, indicadores de rendimiento, etc.) es en estos momentos un tema de plena actualidad en las universidades europeas (Kells, 1990).

VII.3.2. **Holanda**

Sin lugar a dudas, Holanda es uno de los países europeos donde mayor interés y dedicación ha suscitado el problema de la calidad de la educación superior y, por tanto, una de las mejores fuentes de información sobre los problemas de su evaluación. En este apartado se recogen diversos aspectos de la evaluación de la calidad en el sistema universitario holandés.

El sistema de educación superior holandés

La educación superior en Holanda está constituida por un sistema público sostenido por el Estado central, y que consta de tres sectores: el universitario, el profesional y una universidad a distancia (Frackmann y Maassen, 1987). Hay 13 universidades con un total de 170.000 alumnos en el curso 87-88, y 85 instituciones de educación superior profesional (Maassen, 1989).

Los principios que definen la educación superior holandesa en la postguerra son éstos (Bormans *et al.*, 1987; Maassen, 1987):

— Igualdad de oportunidades y acceso abierto a todos los estudiantes que lo deseen y estén cualificados.

— Igualdad entre todas las universidades para su desarrollo, bajo la suposición de que todas tienen el mismo nivel de calidad.

Estos principios se han reflejado en algunas características del sistema holandés de educación superior (Bormans *et al.*, 1987):

— Acceso libre a cualquier área de estudio con la excepción de medicina.

— No existencia de ningún tipo de barreras académicas de entrada a la educación superior, como exámenes o pruebas.

— Crecimiento desmedido del personal funcionario en las universidades, dentro de una política de personal basada más en la seguridad en el empleo que en criterios de calidad o de diferenciación.

— La política interna de las instituciones ha sido un reflejo de la política gubernamental en lo que respecta a justicia distributiva y a equidad.

— El Estado proporciona el 85 % del presupuesto de las universidades. El criterio de asignación ha estado basado exclusivamente en el número de alumnos de cada institución, y en la razón profesor/alumno.

— Carencia de un sistema de control de calidad que valorara el proceso educativo.

En los finales de los años setenta empiezan a darse algunos hechos importantes en Holanda, semejantes a los que acontecían en otros países desarrollados (Bormans et al., 1987):

— Estancamiento de los presupuestos y del número de alumnos.

— Incremento del desempleo de los graduados universitarios.

— Preocupación por la falta de relación entre el mundo del trabajo y las instituciones de educación superior.

— Proporción de abandonos demasiado alta, y duración excesiva del período necesario para completar los estudios.

— Preocupación creciente por la calidad y por la carencia de una política de selección.

— Un excesivo grado de especialización y compartimentalización de las actividades académicas.

— Un personal que, disfrutando de una situación de funcionario, no siempre satisfacía las demandas que se le hacían respecto a calidad, creatividad, productividad y destrezas pedagógicas.

— Carencia de una política científica nacional adecuada.

Como consecuencia de esta situación, en los primeros años ochenta se llevó a cabo un importante proceso de reorganización que ha afectado a todas las universidades. Algunas de las medidas adoptadas fueron las siguientes (Bormans et al., 1987; Frackmann y Maassen, 1987):

— Introducción de un sistema de dos ciclos de docencia. El primero, de cuatro años de duración, abierto a todo el mundo, y el segundo, de carácter selectivo, orientado a la investigación y a la especialización profesional.

— Cierre de algunos departamentos y facultades que no se consideraban imprescindibles y que generaban un gasto importante.

— Introducción de formas de financiación a las instituciones condicionadas a la investigación realizada.

— Reorganización de las responsabilidades de enseñanza e investigación dentro de las universidades.

— Reorganización de los institutos de educación superior de tipo profesional, reduciéndolos, de casi 400 que eran, a menos de 90.

— Se crearon también criterios laborales y académicos asociados con cada categoría profesional dentro de las universidades, de modo que se crearan las condiciones para la entrada de jóvenes investigadores capacitados.

El informe HOAK (Educación superior: autonomía y calidad)

En 1985 el Gobierno holandés hizo público un documento, en el que definía su nueva política respecto a las universidades. En este documento, conocido como HOAK (*Educación superior: autonomía y calidad*), se introducían planes para desarrollar la autonomía de las instituciones, a la vez que se proponían métodos para aumentar el control de calidad de las instituciones. Básicamente, la idea del informe es la de cambiar el control *ex ante* de las instituciones por un control *ex post*, de tal modo que la intervención del Estado debería reducirse a aquellos casos en los que por alguna razón se produzcan efectos contraproducentes (Spee, 1990). Cuatro son los objetivos básicos que se proponen en este documento (Frackmann y Maassen, 1987):

— Dirigir desde el Estado la educación superior de un modo más global, dejando mucha más capacidad de autorregulación en manos de las instituciones.

— Enfatizar la propia responsabilidad de las instituciones.

— Diseñar un sistema de control de calidad adecuado.

— Desarrollar unos mecanismos de cooperación más adecuados.

Como contrapartida al aumento de la autonomía de las instituciones, parece también necesario desarrollar sistemas eficaces de autocontrol del funcionamiento de cada institución. El sistema de control de la calidad de la educación superior propuesto desde el Ministerio de Educación en el HOAK puede resumirse así (Frackmann y Maassen, 1987; Maassen, 1987; Prins, 1990):

— Las propias instituciones de educación superior deben desarrollar sistemas de autoevaluación a los tres niveles de funcionamiento: departamen-

to, facultad e institución. Se deben desarrollar sistemas de indicadores de rendimiento que faciliten la información entre instituciones y Gobierno.

— Los «consumidores» de la educación superior (graduados, empleadores, servicios) deben estar involucrados en la evaluación del sistema educativo, especialmente en las áreas con más carácter profesional. Se propone la creación de comités mixtos de evaluación formados por representantes de las universidades y de la sociedad.

— El Estado por su parte instituye un sistema de inspección de la educación superior que será suplementario de las autoevaluaciones y que publicará sus resultados cada dos años. Las evaluaciones gubernamentales serán realizadas en visitas *in situ* de los inspectores, que utilizarán como guía de su revisión las autoevaluaciones realizadas por la propia institución.

Se espera que las instituciones por sí mismas tomen medidas para corregir los defectos señalados para la evaluación. El Gobierno, por su parte, puede también adoptar medidas para estimular esas correcciones, tanto positivamente como mediante la utilización de sanciones, que pueden ir desde la advertencia pública hasta recortes en la financiación (Frackmann y Maassen, 1987).

El Plan para la Educación Superior y la Investigación

Con la experiencia recabada desde la publicación del HOAK, se preparó un plan de reforma de la educación superior que se convirtió en una ley que, desde la primavera de 1990, establece nuevas normas para la educación superior holandesa. La autonomía de las universidades se ha incrementado, consiguiendo mayor poder para el desarrollo de la docencia y de la investigación, para fijar criterios de admisión de estudiantes y para la asignación interna de fondos. El poder del Gobierno queda reducido a determinar el nivel total de fondos para la institución, a restringir el número de plazas en los casos de fuertes desequilibrios, y a diseñar las líneas maestras del sistema de evaluación que deben ser implementados por las instituciones (Spee, 1990).

La asignación global de fondos para cada universidad se hará en función del número de alumnos y de su rendimiento, para lo que se propone un sistema basado en el número de créditos conseguidos por los estudiantes de la institución durante ese año.

Para la valoración de la calidad se establecen unos comités externos que visitan cada universidad para evaluar la docencia, y un organismo nacional de evaluación, también externo, que evalúa la investigación. Disciplinas que sean

evaluadas negativamente repetidas veces no recibirán fondos del Estado (Spee, 1990).

El sistema de control de la calidad

El sistema de control de la calidad en la educación superior holandesa es el resultado de acuerdos entre las instituciones y el ministerio. La reacción de las instituciones a las últimas propuestas del Gobierno ha sido la de ponerse a trabajar para desarrollar un verdadero sistema de control de calidad (Maassen, 1989). Dos tipos de controles independientes están funcionando ya en la actualidad: sobre la investigación y sobre la docencia.

La investigación es evaluada por medio de organismos externos e independientes de las instituciones y del ministerio (academias científicas, organizaciones profesionales, etc.). La evaluación es por disciplinas, lo que significa que todos los proyectos de investigación de la misma materia son evaluados el mismo año, y no son vueltos a evaluar hasta pasados cinco años (Spee, 1990). Los resultados de esta evaluación de la investigación son importantes, porque de ellos depende la asignación de recursos a los departamentos en cuanto a gastos de investigación se refiere. Una de las consecuencias más llamativas del sistema de control de la investigación es que, desde su implantación en 1982, la producción de las universidades holandesas, en cuanto a número de publicaciones, se ha incrementado en un 50 % (Maassen, 1989).

La docencia es evaluada por medio de un comité que visita las instituciones. Los comités estuvieron formados en 1989 (primer año de su funcionamiento no experimental) por profesores (25 % de eméritos, 10 % de pedagogos, 15 % de otros), profesores extranjeros (20 %) y representantes de asociaciones profesionales y de los empleadores (30 %) (Vroeijenstijn, 1990).

La evaluación se realiza también por disciplinas académicas. El proceso de evaluación se basa, en primer lugar, en el autoestudio realizado por los propios departamentos, pero el comité recoge información por sí mismo, especialmente de los alumnos. Con toda esta información emite un informe que es enviado a la propia universidad cuyo departamento ha sido evaluado. La experiencia presente muestra la carencia de criterios comunes para la evaluación, por lo que parece deseable el desarrollo de indicadores de rendimiento cualitativos y cuantitativos (Spee, 1990).

Hay que señalar que en Holanda existe desde 1986 un organismo llamado el *Inspectorado de la Educación Superior* que organiza todo el proceso de evaluaciones de la calidad de la investigación y de la docencia; sin embargo,

este organismo no participa directamente en ninguno de los procesos de evaluación (Maassen, 1989; Bresters y Kalkwijk, 1990).

Además de los comités que rutinariamente controlan la docencia y la investigación, existen otros llamados comités de reconocimiento, formados por científicos de una misma disciplina, y que, nombrados por el Ministerio para ciertas situaciones problemáticas, tratan de encontrar los puntos débiles de la disciplina o prever su futuro, con el fin de tomar medidas concretas (Westerheijden, 1990).

Para una más detallada descripción del proceso y de los resultados de las evaluaciones externas en las universidades holandesas, pueden consultarse los artículos de Vroeijenstijn y Acherman (1990); Acherman (1990) y Vroeijenstijn (1990).

Modelos de evaluación institucional

Un modelo de evaluación institucional desarrollado en Holanda, como consecuencia de un nuevo espíritu de autocontrol de las instituciones, es el *modelo AMOS*, que diseñó la Universidad Libre de Amsterdam (Van Os *et al.*, 1987). Este modelo está centrado en el proceso de aprendizaje. Se establecen unos criterios básicos para el flujo de estudiantes a través del proceso educativo, de modo que las disciplinas que se salen de estos criterios son inmediatamente puestas bajo investigación de una comisión constituida por miembros de la institución.

Un interesante programa de autoevaluación es el diseñado por el *Center for Higher Education Policy Studies* (CHEPS, 1990). Se trata de un paquete informatizado para llevar a cabo el proceso de autoevaluación de la calidad de las instituciones. El paquete está integrado por 14 módulos que recorren otros tantos aspectos esenciales de la calidad institucional.

El uso de indicadores de rendimiento

Como consecuencia de las nuevas relaciones entre Gobierno y universidades, se ha propuesto una serie de indicadores de rendimiento que, aceptados por ambas partes, permitan establecer un lenguaje común entre ellas. Aunque el sistema dominante en Holanda sea el de las revisiones por colegas, el uso de los indicadores de rendimiento, como medidas más objetivas de funcionamiento y rendimiento, se está propugnando como complemento de la información que proporcionan las revisiones (Paardekooper y Spee, 1990). Los

indicadores que inicialmente fueron propuestos por parte gubernamental son, clasificados por tipologías, los siguientes (Bormans *et al.*, 1987):

- Los que valoran la calidad docente e investigadora:
 - Evaluación media de los estudiantes.
 - Tasa de éxito de los estudiantes.
 - Tasa de abandono por nivel.
 - Duración media de los estudios.
 - Índice de citas.
 - Índice de publicaciones en revistas clave.
 - Número de disertaciones doctorales.
 - Trabajos en congresos.
 - Pertenencias a comités científicos.
 - Premios académicos.
 - Evaluación por colegas.
 - Participación en los fondos de investigación de la universidad.
- Los que valoran la funcionalidad:
 - Porcentaje de graduados que encuentran empleo.
 - Tiempo medio entre graduación y empleo.
 - Salario medio después de x años.
 - Evaluación de los empleadores.
 - Evaluación media de los estudiantes.
 - Ayudas financieras del exterior.
 - Utilización de asistentes de investigación.

- Los que valoran la eficiencia:
 - Coste por estudiante y por graduado.
 - Número de estudiantes y graduados en relación con el número de profesores.
 - Relación profesor/alumno.
 - Coste por publicación.
- Los que valoran las relaciones con el medio exterior:
 - Profesores invitados.
 - Cátedras honorarias.
 - Prácticas en empresas.
 - Proporción del presupuesto de investigación conseguido en contratos externos.
 - Patentes y otras transferencias tecnológicas.
 - Cooperación con investigadores de empresas.

Estos indicadores, propuestos en principio por el Gobierno, fueron duramente criticados, lo que motivó que se creara un comité organizado por el VSNU, con el fin de proyectar un conjunto de indicadores válidos y aceptados por todos. Los resultados de esta comisión son un amplio conjunto de 70 indicadores cuantitativos y no-cuantitativos, que recogen información sobre cada uno de los aspectos interesantes de la vida académica en relación con la calidad [pueden verse en Spee (1990), que a su vez los toma del documento original en holandés publicado por el Ministerio de Educación de aquel país].

Algunas conclusiones sobre la evaluación institucional en Holanda

El sistema de control de calidad de las universidades holandesas que está siendo implantado en la actualidad ha sabido evitar una serie de errores en lo que posiblemente radica su éxito. Según Kells (1989), son estos los errores que no se han cometido:

- Adoptar como un todo un sistema traído de otro lugar.

- Construirlo todo de nuevo.
- Adoptar medidas de rendimiento como base del sistema.
- Esperar hasta que todo estuviera diseñado para empezar a proceder a algún nivel.
- Confiar solamente en la evaluación externa.
- Hacerlo todo a escondidas.
- Exigir perfección inmediatamente.

No sólo se evitaron esos errores, sino que también fueron tomadas algunas decisiones acertadas (Maassen, 1989):

- Poner el control del sistema de autorregulación en manos del VSNU (organismo comparable al Consejo de Universidades español).
- Usar un sistema integrado.
- Poner un énfases creciente en el autoestudio.
- Adoptar un sistema inicial de revisiones por departamentos.

De acuerdo con esta gradual introducción de la evaluación, las evaluaciones externas han sido introducidas cuando había ya suficiente experiencia con la evaluación interna (Maassen, 1989).

La actual política del Gobierno holandés respecto a la educación superior es juzgada por Maassen y Van Vught (1988), quienes consideran el carácter híbrido de la situación actual, en la que la autonomía no está exenta de control. Según estos autores, aunque las instituciones son autónomas, están todavía demasiado influidas por el Estado, lo que produce una serie de desventajas y anomalías que dificultan el buen funcionamiento del sistema, aunque, en cualquier caso, la situación es notoriamente mejor que la anterior.

El problema que parece plantearse en un futuro inmediato es el de hasta qué punto podrán convivir el uso de los indicadores de rendimiento con las revisiones de los programas realizadas por otros profesores. Los primeros no son plenamente aceptados por los profesores como sistema de evaluación de la calidad, mientras que las segundas no parecen útiles para valoraciones de la calidad en gran escala (Westerheijden, 1990a).

Para finalizar esta visión general de la situación de la evaluación de la calidad en Holanda, hemos de señalar que sólo se ha hablado del sector universitario, y no del sector profesional. Este sector, que tiene un alumnado más numeroso que las universidades, ha sufrido unas transformaciones semejantes a éstas, con diferencias de matices o de organización del proceso evaluador. Así, por ejemplo, estas instituciones, que son más numerosas y pequeñas que las universidades, son revisadas como un todo y no por programas. Una completa descripción de la evaluación en este sector puede encontrarse en Rullmann (1990).

VII.3.3. Reino Unido

La instauración de un sistema de evaluación de la educación superior en el Reino Unido es un largo proceso histórico que se inició a principios de siglo, pero que no ha llegado todavía a unos mecanismos aceptados por todos. Las discusiones sobre la calidad y su evaluación han sido, y siguen siéndolo en la actualidad, algo muy vivo (Silver, 1986). Dado que no existe un consenso universal sobre las metas de una educación superior caracterizada por la pluralidad y diversidad, es muy difícil establecer planes y estrategias conjuntas que permitan desarrollar sistemas de evaluación ampliamente aceptados (Billing, 1986).

El sistema de evaluación institucional británico es, junto con el holandés, el que más se ha desarrollado en los últimos tiempos en Europa, pero, a diferencia del holandés, el británico hace mayor hincapié en el uso de indicadores de rendimiento que en los métodos de evaluación autorregulada (Goedegebuure *et al.*, 1990). Pero para entender este proceso de evaluación es necesario, primero, hacer algunas consideraciones sobre el sistema educativo superior británico.

Características del sistema de educación superior británico

El Reino Unido ha tenido siempre unas universidades prestigiosas y autónomas, cada una de las cuales es responsable de su propio funcionamiento, reclutamiento del profesorado y de los empleados. En estas universidades, como consecuencia de la política que se desprendía del Informe Robbins (Robbins Report, 1963), el crecimiento de la educación superior en los años sesenta y setenta fue seguido de un incremento paralelo de los recursos, por lo que no se produjo una reducción drástica de la calidad de la educación superior, como sucedió en otros países europeos (Neave, 1986). Sin embargo, el Estado puso en marcha en los años sesenta un sistema paralelo de

educación superior, el llamado sector público, controlado por el propio Gobierno y que es el que ha absorbido la mayor parte del incremento del sistema educativo superior.

Las universidades británicas están asociadas en el *Committee of Vice-Chancellors and Principals* (CVCP), un organismo formado por representantes de todas las universidades. Aunque el Departamento de Educación y Ciencia fija los gastos totales de las universidades, así como las líneas maestras para ese gasto, un organismo independiente constituido por miembros de la comunidad académica (el *Universikty Grants Committee*: UGC) era el responsable de repartir ese presupuesto y de hacer las evaluaciones pertinentes para ese reparto. A partir de 1989 esta institución dio paso a la *Universities Funding Council* (UFC), con semejantes propósitos, pero con carácter más centralizador. En el sector público los organismos equivalentes son el *Council for National Academic Awards* (CNAA) y el *Polytechnics and Colleges Foundign Council* (PCFC). Mientras que las universidades son autónomas, y sólo recientemente han recibido presiones para ser controladas, el sector público ha sido tutelado desde su nacimiento por la CNAA, organismo al que no pertenecen las instituciones, sino que está de hecho creado para controlarlas (Goedegebuure *et al.*, 1990).

Las instituciones de educación superior dependen en dos terceras partes de los fondos que reparten los organismos citados. La asignación de recursos entre las universidades se realiza por períodos de tres años, las correspondientes a la docencia se hacen según el número de alumnos de cada institución, mientras que los presupuestos para la investigación se aprueban según el nivel con el que se ha evaluado la investigación de la institución (siguiendo un método de revisión por colegas). Las asignaciones se hacen en bloques para cada universidad, de modo que éstas son responsables posteriormente de la distribución del presupuesto entre sus departamentos y servicios (Frackmann, 1987). En el sector público, el *National Advisory Body* distribuía las dotaciones en función de los números de alumnos, mientras que los fondos para investigación funcionaban siempre por contratos (Starapoli, 1987); sin embargo, desde 1989 el régimen se asemeja al de las universidades.

Cambios en la educación superior británica en los años ochenta

Mientras que en algunos países se están poniendo en marcha sistemas de descentralización para aumentar la calidad de la enseñanza universitaria, en el Reino Unido se está siguiendo el camino opuesto con el objetivo de aumentar el control gubernamental sobre el sistema (Bauer, 1988). Y es que, como afirma Sizer (1990), el hecho más importante que le ha ocurrido a la educación

superior británica en los últimos tiempos ha sido la elección de Thatcher como primer ministro.

Consideraciones de tipo político y no educativo condujeron a un cambio de actitud hacia la educación superior, de modo que pasó de estar marcada por la demanda social, a estarlo por los recursos financieros. El gasto público en educación declinó fuertemente y se pasó a una política de fuerte control económico. La equidad en el acceso pasó a ser un valor en baja, en contra de lo que es habitual en las universidades americanas y otras europeas (Moodie, 1988). Este mismo autor señala que: «El hecho de que la mayor parte del Gobierno, de sus asesores y de los parlamentarios que le apoyan sean graduados de Oxford, produce un especial aroma: una mezcla de meritocracia y de elitismo social» (Moodie, 1988: 6).

Como consecuencia de los cambios políticos, se avivaron las discusiones sobre los indicadores de rendimiento que habían estado paradas durante algún tiempo. Su uso permitiría a las instituciones educativas el demostrar a la sociedad, y especialmente, al gobierno, que estaban proporcionando «cosas de valor a cambio de su dinero» (Goedegebuure *et al.*, 1990).

En 1984, como consecuencia de discusiones entre Gobierno y universidades, la CVCP nombró un comité que elaboró un informe para promocionar la eficiencia en las universidades. Este comité investigó el funcionamiento de cinco universidades británicas y elaboró un informe conocido como Informe Jarrat (Jarrat, 1985), en el que se recomienda un cambio drástico en el funcionamiento de las universidades, promoviendo la utilización de técnicas de gestión y el uso de indicadores de rendimiento para adaptarse a las restricciones económicas que desde el Gobierno se impusieron a las universidades. En el informe se citan entre otros objetivos el de: «Desarrollar un conjunto de indicadores de rendimiento, cubriendo inputs y outputs, diseñados para uso dentro de las universidades y para hacer comparaciones entre instituciones» (Lockwood, 1986: 12). El *Informe Jarrat* es una manifestación de la política general del Gobierno conservador británico, basado más en consideraciones políticas que científicas, sobre el modo en el que deben gestionarse las universidades. El informe fue fuertemente criticado y provocó muchas discusiones (Scott, 1986).

Siguiendo las indicaciones del *Informe Jarrat*, un grupo de trabajo conjunto de la CVCP y de la UGP propusieron un conjunto de indicadores en 1986, que fue adoptado por las distintas instancias implicadas. Unas recomendaciones semejantes fueron hechas por el sector público. Un documento gubernamental, el libro blanco *Higher Education: Meeting the Challenge* planeaba la in-

troducción de los indicadores de rendimiento en la política educativa superior (Goedegebuure *et al.*, 1990). La *Education Reform Act*, de 1988, instituyó legalmente todos estos principios y ligaba definitivamente la calidad de las universidades a su eficiencia (McVicar, 1990).

El control de la calidad en la educación superior británica

Tradicionalmente la calidad docente de las universidades ha estado garantizada por examinadores externos, nombrados por las mismas universidades, que controlan la calidad de la formación de los graduados que salen de ellas. La calidad de los alumnos que entran en las universidades está controlada por los mecanismos de selección de las propias universidades (Goedegebuure *et al.*, 1990). Este sistema sigue en vigor en la actualidad; sin embargo, ha sido creado recientemente por la CVCP un nuevo organismo, la *Academic Audit Unit*, con el fin de asistir a las universidades en sus responsabilidades para el control de la calidad de la docencia. Realmente es un organismo que inspecciona las universidades por «invitación negociada» de ellas (Young, 1990).

El control de la calidad de la investigación en las universidades está en manos de quienes asignan los correspondientes fondos, el UFC, que, mediante revisiones por colegas, pero con criterios no explícitos, valora la investigación de las universidades. Para el sector público existe un organismo gubernamental, el *Research Council*, que evalúa la investigación y asigna los correspondientes fondos de acuerdo con las evaluaciones (Young, 1990).

Para el sector público se creó un organismo especial de control de la docencia: el *Her Majesty's Inspectorate*. Está formado por inspectores (funcionarios no pertenecientes al mundo académico) que visitan las instituciones para controlar la docencia, incluso entrando en las clases. Sus recomendaciones sirven para que el PCFP haga reasignaciones de recursos entre estos centros.

El *Council for National Academic Awards* también controla la calidad de la educación superior del sistema público de educación superior. Este organismo es el que concede las titulaciones a los estudiantes de la gran mayoría de estos centros (Brennan, 1986, 1990). Su principal misión es la de aprobar los cursos que se imparten en estos centros. Para ello realiza revisiones de los programas de estudios cada diez años.

Durante los años ochenta se ha ido extendiendo el uso de indicadores de rendimiento de las instituciones de educación superior, recogándose datos de las instituciones que son publicados para intentar establecer comparaciones

entre las distintas instituciones. El numeroso grupo de indicadores propuestos por los comités nombrados *ad hoc* pueden verse en Cullen (1987) y en Sizer (1990). Lo más significativo de estas nuevas tendencias es el intento manifiesto de introducir la competitividad entre las instituciones de educación superior. Las contradicciones que subyacen en estos planes están sin resolver, y no parece claro cuál será su futuro (McVicar, 1990).

Hay que señalar, como conclusión, que no da la impresión de que el sistema de evaluación de la calidad de las universidades británicas sea demasiado sólido. Como concluye un autor británico refiriéndose a los métodos de evaluación en el Reino Unido: «No existen estudios convincentes sobre la evaluación de las necesidades, y los que hay no llegan a conclusiones sobre los efectos y la efectividad de tales evaluaciones. En todos los niveles existe más oscuridad que luz y más conflicto que consenso» (Billing, 1986: 86). Para otros, sin embargo, el problema está en que la educación superior británica está dominada por la propia profesión académica, en contra de lo que sucede en USA (Wright, 1989). Como alternativa, Wright propone la apertura a las fuerzas del mercado en vez de la instauración de sistemas de control instituidos por los propios profesionales.

VII.3.4. **Finlandia**

En Finlandia existen 17 universidades públicas, con un organismo de coordinación, el *Consejo de Educación Superior*, que actúa como cuerpo asesor del ministerio. Otro organismo, la *Academia de Finlandia* es la encargada de promover y controlar la investigación en las universidades (Jäppinen, 1987).

En los últimos años la educación superior en Finlandia ha sufrido cambios importantes. Un sistema centralizado empezó a transformarse en otro más autónomo y más encaminado hacia la gestión por objetivos (Goedegebuure y Maassen, 1990). El proceso en marcha está transformando el antiguo sistema de control por un sistema de evaluación pública (Hölttä, 1990).

En 1985 un grupo de trabajo que había sido creado para establecer un modelo de evaluación de la educación superior llegó, entre otras, a las siguientes conclusiones:

- La necesidad de evaluar el rendimiento de la educación superior con el fin de mejorar la calidad y subsanar las deficiencias que existían.
- Los resultados de la evaluación deberían ser usados para la toma de

decisiones internas dentro de cada universidad, o dentro de sus departamentos cuando sea respecto a actividades de éstos.

— La necesidad de establecer un sistema de evaluación continua tanto de la investigación como de la calidad de la docencia.

Como consecuencia de este informe, a partir de 1986, coincidiendo con un plan para el incremento anual de la inversión en las universidades del 15 % durante cuatro años consecutivos, fue puesto en marcha un sistema de utilización de indicadores de rendimiento. Estos indicadores están referidos tanto a los medios con que cuentan las instituciones como a la productividad docente e investigadora. Respecto a su uso, una normativa estatal establece desde aquel año estas tres normas para todas las instituciones universitarias del país (Hüfner y Rau, 1987):

— Todas las instituciones de educación superior deben desarrollar un modelo normalizado de evaluación en orden a ofrecer información que sea comparable, acerca de los resultados docente e investigadores, así como de los costes implicados en ellos.

— Todas las instituciones de educación superior deben entregar regularmente un informe de sus actividades. Estos informes servirán para elaborar los informes que el Gobierno realizará cada cuatro años sobre el estado global de la educación superior.

— Los resultados docentes e investigadores serán tenidos en cuenta para las asignaciones de recursos por parte del Gobierno a las instituciones.

Por otra parte, las facultades están poniendo en marcha en estos momentos sistemas de evaluación institucional que les sirvan para la asignación de fondos a sus propios departamentos. Este es un proceso que está actualmente generalizándose, aunque con distinta rapidez, según el tipo de facultad (Hölttä, 1990).

En resumen, en Finlandia existen unas universidades con autonomía creciente, pero que a su vez se ven obligadas a suministrar al Gobierno información continua sobre su funcionamiento, quien, por otra parte, proporciona los medios financieros de acuerdo con el rendimiento de cada institución, y curiosamente, para los tiempos que corren en la educación superior europea, incrementándose anualmente.

VII.3.5. **Alemania**

En la época de mayor expansión de la educación superior en Alemania,

como en otros países europeos, la demanda social ha sido la que ha dirigido la política educativa. El incremento del número de alumnos ha superado las asignaciones de recursos al sistema, de modo que se ha ido produciendo un paulatino deterioro de la situación general de las universidades, a pesar de lo cual no se han tomado las medidas adecuadas, sino que ante la hipótesis de una disminución de alumnos en un futuro próximo se ha utilizado la política de «hacer un túnel a través de la joroba», manteniendo los presupuestos hasta que disminuya el número de alumnos (Frackmann y Maassen, 1987).

En Alemania hay una preocupación por aumentar la competencia y la diferenciación entre las instituciones de educación superior, lo que lleva consigo una búsqueda de indicadores de rendimiento que sean capaces de reflejar las diferencias y de proporcionar un sistema para la asignación de recursos de acuerdo con esos rendimientos. Las razones de esa nueva situación hay que buscarlas en el cambio de actitud política hacia la educación superior, a la precariedad de los presupuestos de los *länder*, a la necesidad de una mayor preparación de cara al desarrollo económico y tecnológico y al esperado decrecimiento del número de alumnos (Hüfner, 1987).

Por parte de las universidades parece existir una cierta reserva a poner en marcha sistemas de indicadores cuantitativos para las universidades. La idea dominante es la de que los indicadores sólo se deben usar en situaciones muy específicas, mientras que los sistemas cualitativos de evaluación deben de ser predominantes (Frackmann, 1987). La utilización de indicadores de rendimiento se reduce exclusivamente a la de aquéllos que son necesarios para cubrir las demandas estatales de datos cuantitativos para la planificación. Sin embargo, está teniendo lugar un incremento en la utilización de las autoevaluaciones por parte de las propias universidades, especialmente en el campo de la evaluación de la actividad investigadora (Hüfner, 1987).

Dado que la toma de medidas hacia una mayor competitividad entre las universidades requiere el acuerdo y la cooperación entre 11 gobiernos regionales, es poco presumible que se adopten medidas en ese sentido en un futuro inmediato (Hüfner y Rau, 1987). Por otro lado, las contradicciones originadas por el deseo gubernamental de una mayor competitividad entre las instituciones y de un mayor control a través de los indicadores de rendimiento, no está ayudando a una actitud de colaboración entre académicos y gobiernos regionales (Hüfner y Rau, 1987).

En este clima no es de extrañar que algunas universidades que han realizado estudios prospectivos sobre sí mismas (las universidades de Frankfurt y Ulm, por ejemplo) han utilizado empresas externas de consultoría, dada

la carencia de metodologías y mecanismos internos de evaluación (Frackmann y Maassen, 1987).

VII.3.6. Suecia

En Suecia existen 31 instituciones de educación superior, todas ellas públicas. En 1977 se produjo una importante reforma descentralizadora, que trasladó la responsabilidad sobre las universidades del Gobierno central a los gobiernos locales y regionales. Igualmente, hubo un cambio en el sistema presupuestario de asignación de fondos a las universidades, de modo que los fondos globales a cada universidad se le asignan de acuerdo con su productividad docente e investigadora. Como consecuencia de esta reforma, la gestión de las universidades sufrió importantes cambios, por lo que la *Junta Nacional de Universidades* sueca inició un programa de autoevaluación de los centros para realizar el seguimiento del proceso de reforma. No existe en Suecia un control externo sobre el rendimiento de las universidades, por lo que la autoevaluación se convierte en el principal método de control de las instituciones sobre sí mismas. «La autoevaluación debe ser entendida como un asunto estrictamente interno de las instituciones que intenta proporcionar una base adecuada para la toma de decisiones por las mismas instituciones» (Furumark, 1981: 209).

Como consecuencia del proceso de reforma de la educación superior se inició un proyecto para el diseño de un método de autoevaluación para las universidades suecas. El primer paso fue el de recoger información teórica y empírica sobre los métodos aplicados en Estados Unidos, adaptándolos posteriormente a la realidad sueca. El segundo paso fue el de promover la idea de la necesidad de la autoevaluación entre las instituciones universitarias, organizando cursillos de preparación para su personal. Posteriormente, con distintos ritmos, según las universidades, las autoevaluaciones se fueron imponiendo en todas las instituciones a nivel global y a nivel departamental.

A pesar de la dedicación y del esfuerzo realizados, la autoevaluación ha tenido menor éxito a nivel de departamentos, posiblemente debido a que la actividad evaluadora estuvo rodeada de un ambiente general de recortes financieros que hacía que la autoevaluación fuera vista como una amenaza por la mayoría de los miembros de la comunidad académica (Bauer, 1988). Por esa razón, actualmente se están complementando las evaluaciones internas con otras externas realizadas por organismos con responsabilidad en algún área científica o técnica concreta.

El interés de la actividad investigadora en Suecia en el campo de la

evaluación se dirige en la actualidad hacia temas como son los grupos de trabajo para el desarrollo de métodos de organizar responsabilidades entre los distintos niveles de las instituciones universitarias; el desarrollo de un sistema de indicadores de calidad; la evaluación de programas desde el punto de vista de utilidad para el empleo, y el desarrollo de indicadores de rendimiento para describir globalmente el sistema universitario (Bauer, 1988).

En la actualidad, aunque la autonomía se ha incrementado, también lo han hecho las presiones desde la sociedad y desde el Parlamento para controlar el proceso educativo superior. La idea subyacente de esta actitud es la de que la actual situación de autogobierno de las universidades exige en contrapartida una mayor dosis de responsabilidad por parte de las propias instituciones (Bauer, 1988).

En general, puede afirmarse que el sistema sueco de evaluación está dominado por la idea de la autoevaluación, sin que se descarte el uso de indicadores para valorar el rendimiento del sistema globalmente, ni la posible utilización de evaluadores externos cuando es necesario.

VII.3.7. **Francia**

La preocupación en Francia por la evaluación de su sistema de educación superior se concretó en 1985 cuando se creó el *Comité National d'Evaluation*. Este es un organismo autónomo, independiente de las autoridades académicas y administrativas, diseñado para hacer juicios sobre la calidad de la investigación y de la docencia no sólo de las universidades globalmente, sino también de otros centros educativos superiores y de centros de investigación (Staropoli, 1986). El comité está constituido por 15 miembros destacados de la comunidad académica y algunas personalidades de otros organismos del Estado, dedicados a las actividades del comité a tiempo parcial, y no meramente como asistentes a reuniones esporádicas.

El método operativo de este comité es el de nombrar un pequeño equipo de expertos que visitan cada institución, examinan todos sus componentes *in situ* e informan directamente al comité. Las universidades no han de hacer ningún informe previo para el equipo, sino que han de facilitar toda la información disponible para que el equipo pueda conocer todos los datos que desee. Todos los aspectos de la universidad deben ser objeto de evaluación. En las evaluaciones realizadas hasta la fecha han sido utilizados expertos en las principales áreas, incluidos extranjeros y representantes de la industria, que visitan cada universidad durante un corto período de tiempo, siguiendo una metodología preparada de antemano por el secretariado del comité. Los datos recogidos

por estos observadores son posteriormente analizados en ese secretariado, quien, basándose en ellos, elabora un informe final sobre la institución (Starapoli, 1987).

Por el momento, las consecuencias de la evaluación no se reflejan en la asignación de recursos. Las evaluaciones están permitiendo la construcción de un conjunto de indicadores que puedan permitir en el futuro establecer ordenaciones entre las instituciones bajo criterios que sean aceptados por la comunidad académica.

El *Comité National d'Evaluation*, según su secretario general, André Starapoli, ha tratado desde el principio de trabajar en armoniosa colaboración con el estamento científico, y se puede afirmar tras la experiencia acumulada que el mundo científico y académico lo apoyan plenamente (Starapoli, 1989).

CAPITULO VIII: CONCLUSIONES Y PROPUESTAS PARA EL SISTEMA UNIVERSITARIO ESPAÑOL

Me parece muy oportuno empezar ahora este proceso de búsqueda de la metodología de evaluación en nuestro país (...) Es evidente que en un asunto como éste las cosas deben plantearse con pragmatismo y generosidad. No es fácil autoevaluarse, pero debemos abordar el proceso como una herramienta útil para el futuro y encontrar la mejor manera de diseñarlo y los mejores ejecutores para ponerlo en marcha.

Crespo, 1989: 129.

VIII.1. INTRODUCCION

A lo largo de los siete capítulos anteriores se ha ido recopilando una gran cantidad de información sobre la teoría y práctica de la evaluación de las instituciones de educación superior en el mundo. Para la mayoría de los temas que se han tratado en esos capítulos existen buenas y completas revisiones. El presentar aquí un panorama completo de la realidad de la evaluación institucional en el mundo tiene como objetivo esencial ofrecer a aquéllos que por sus responsabilidades en el sistema universitario español o en las instituciones universitarias, o a los que simplemente tengan interés en este tema la posibilidad de disponer de un panorama general actualizado sobre las disponibilidades metodológicas para abordar el problema de la evaluación de las instituciones universitarias españolas. Obviamente, como es frecuente cuando se presenta una visión demasiado amplia, la que aquí se ha ofrecido puede adolecer en algún momento de excesivamente simplificada y esquemática, pero en compensación se proporciona una extensa y reciente bibliografía a la que poder recurrir para profundizar en cada tema.

Dado el objetivo final de este estudio, en este último capítulo se intentan dos cosas:

1. Hacer una recapitulación de las ideas básicas que se han ido exponiendo con anterioridad.
2. Plantear la aplicabilidad de esas ideas a la realidad universitaria española.

Ni que decir tiene que la segunda cuestión que se propone sólo será un esbozo de las posibilidades existentes de implantación de un sistema de evaluación institucional en nuestras universidades. El autor no es un experto, ni en la práctica de la evaluación, ni en la política universitaria, por lo que las ideas o sugerencias que se hacen sobre la implantación de un sistema coherente de evaluación institucional en las universidades españolas sólo tienen como finalidad servir de base para una discusión amplia y profunda por parte de todos los que estamos implicados en la educación superior, pero muy especialmente por aquéllos que representan a las universidades, y que finalmente serán los que habrán de decidir sobre el método de evaluación más adecuado para ellas.

A pesar de lo dicho, hay dos ideas que el autor sí se atreve a afirmar con rotundidad:

1. La necesidad de instaurar un sistema de evaluación de las universidades españolas, con toda la urgencia que permita una serena reflexión y discusión sobre metodologías, experiencias y posibilidades reales de implantación. Utilizando palabras del director general de Enseñanza Superior en octubre de 1989: «...hoy nadie puede cuestionar la necesidad y exigencia de la evaluación. En la actualidad todo se evalúa y todo se audita: si hay instituciones privadas, que por razones evidentes, se autoexigen esa evaluación, con mayor razón las instituciones públicas, que existen gracias al dinero de los contribuyentes, deben estar sometidas también a una evaluación de los resultados, a fin de que la sociedad pueda entender y aceptar el sentido de su contribución económica a las mismas» (Blas Arítio, 1989: 159).

2. La necesidad de que la implantación de un sistema de evaluación de las universidades sea introducido, de tal modo que no sólo se evite el rechazo del mundo académico (es natural que no agrade ser evaluado), sino que se cuente con su colaboración de buen grado. Para ello será necesario, primero: convencer de las ventajas de la evaluación para el buen funcionamiento del sistema universitario y, a continuación, implantar paulatina y prudentemente un sistema que haya sido discutido y aceptado mayoritariamente. Tomando también palabras textuales del mismo director general: «Si bien es cierto que el argumento de la evidencia complejidad y, en algunos casos, imprecisión de los parámetros evaluadores no puede servir de coartada para no realizar nunca la evaluación, no deja de ser cierto que dicha complejidad, sobre todo por los efectos socioacadémicos que produce, nos obliga a una enorme cautela y rigurosidad en la metodología...» (Blas Arítio, 1989: 159).

El presente capítulo se ha estructurado en cuatro secciones: tras esta introducción, se hace una síntesis de las ideas básicas sobre evaluación

institucional que se han tratado a lo largo de anteriores capítulos; en la tercera parte se hacen algunas reflexiones sobre el sistema universitario español y se resumen algunas experiencias relacionadas con la evaluación que se han realizado en este país; por último, se presenta el boceto de un posible sistema de evaluación para el sistema universitario español.

Conviene recordar para finalizar que, como en el resto de este estudio, sólo trataremos el tema de la evaluación institucional, es decir, de la evaluación de las instituciones como un todo, de sus departamentos o de sus programas, pero no de los individuos, sean alumnos o profesores, aisladamente de su pertenencia a las anteriores instituciones.

VIII.2. **ALGUNAS CONCLUSIONES GENERALES SOBRE LA EVALUACION INSTITUCIONAL**

En los anteriores capítulos se ha hecho un recorrido teórico y empírico sobre diversas facetas de la evaluación de la calidad y del rendimiento en las instituciones universitarias. La gran variedad de sistemas de educación superior en el mundo, con una gran diversidad interna dentro de muchos de ellos, con tradiciones distintas, e inmersos en la actualidad en diferentes contextos políticos, hace que los métodos de evaluación de las instituciones universitarias de esos sistemas sean también muy variados. El problema de resumir en pocas líneas toda la complejidad de las situaciones que se pueden encontrar no es sencillo. Pero además existe otro factor adicional que incrementa la dificultad de resumir coherentemente las experiencias de esas evaluaciones: excetuan-do Estados Unidos, con una cierta tradición evaluadora, en el resto de países la evaluación de las instituciones universitarias es algo muy reciente, de tal modo que los primeros ensayos evaluadores, en los pocos países que ya se han realizado, son casi todos de la década de los ochenta, y, lógicamente, no están suficientemente experimentados como para poder ofrecernos unos resultados claros y concluyentes.

A pesar de estas limitaciones, revisamos a continuación las ideas clave que sería necesario tener presentes antes de plantearse el diseño de un sistema de evaluación institucional.

VIII.2.1. **El problema de la calidad de la educación superior.**

Un principio que debe presidir todo intento de evaluación de las instituciones universitarias es el de que el objetivo básico de la evaluación es la mejora

de la calidad de la institución. Sólo cumplido este objetivo básico, se deben buscar otros objetivos adicionales, y también deseables, como mejorar la eficiencia, recompensar el rendimiento o penalizar la ineficiencia.

A pesar de que la mejora de la calidad es el objetivo buscado por tantas instituciones universitarias en todo el mundo, no existe ningún acuerdo sobre un concepto tan difuso como el de calidad de las universidades, de su docencia, de su investigación o de los servicios que presta a la sociedad. Sin embargo, todos hablan, en los últimos tiempos más que nunca, de la necesidad de una educación de calidad, de unas universidades de calidad. El clima de desarrollo tecnológico que se vive en la actualidad, unido al creciente interés por la eficiencia y efectividad de los servicios, tiene mucho que ver con esa preocupación creciente por la calidad de las universidades.

A pesar de la falta de definición del concepto de calidad en las instituciones de educación superior, se han hecho múltiples intentos de evaluarla. Dos posiciones básicas pueden encontrarse:

— La que considera que la calidad puede ser medida a través de un conjunto de variables o indicadores objetivos.

— La que cree que la calidad de algo tan complejo como una universidad se valora mejor a través de las opiniones de expertos que sean conocedores de las instituciones que se están evaluando.

Aunque son conceptos diferentes, la preferencia por el uso de variables objetivas para la evaluación de la calidad suele llevar aparejada en la práctica el uso de los indicadores de rendimiento, es decir, medidas, generalmente cuantitativas, de cómo están funcionando los distintos aspectos de una institución universitaria. La preferencia por la segunda postura conduce a la utilización de evaluaciones de la calidad realizadas por colegas, de la misma institución o de otras, formando equipos nombrados *ad hoc* para cada revisión o formando comités de evaluación para todas las instituciones homogéneas de un país.

Ambas posturas no son irreconocibles, y existe una tendencia en la actualidad a la utilización de sistemas combinados de ambas como el método más adecuado para realizar una evaluación de la calidad que tenga en cuenta el mayor número de facetas implicadas.

VIII.2.2. **La evaluación institucional**

A lo largo de todo este estudio se ha estado tratando casi exclusivamente

el tema de la evaluación institucional, pero recordemos el concepto: una evaluación institucional es una valoración de cómo una institución universitaria está alcanzando las metas preestablecidas. Esta valoración se extiende a todas las componentes de la institución: departamentos, facultades o escuelas, universidad como un todo. La evaluación institucional está alcanzando en la actualidad una amplia difusión y un profundo desarrollo metodológico como consecuencia de la creciente preocupación por la calidad de la educación superior, pero también por el notable aumento de los costes totales de las universidades. Para la mayoría de expertos que tratan este tema, el objetivo esencial de la evaluación institucional debe ser la mejora de la calidad de la institución, y sólo secundariamente debería utilizarse como ayuda para la gestión y la toma de decisiones.

Existen multitud de enfoques para abordar el problema de la evaluación institucional. Veamos algunos de los más significativos:

— Puede estar centrada en las magnitudes contables y de control o dedicarse especialmente al análisis de los aspectos organizativos con el objetivo de su mejora.

— Puede estar promovida por las propias instituciones o puede ser consecuencia de algún mandato legal o imperativo gubernamental.

— Puede ser interna (autoevaluación), utilizar evaluadores externos o ser una combinación de ambas. Las autoevaluaciones consisten esencialmente en la toma interna de datos y en la recogida de las opiniones, juicios y valoraciones de todos los implicados en la institución: profesores, administradores, estudiantes, antiguos alumnos y empleadores.

Todos y cada uno de los tipos de evaluación citados encuentran ejemplificación en algunos de los países o instituciones que han sido revisados en capítulos anteriores.

Un problema importante, desde un punto de vista operativo, es el de qué nivel institucional debe ser objeto de la evaluación: la institución como un todo, los departamentos o los programas (en nuestro caso las facultades que imparten un plan de estudios) o los individuos (profesores y empleados). La respuesta es sencilla: todos. Otro problema distinto es el de por dónde se debe empezar al instaurar un sistema de evaluación *ex novo*.

Habitualmente suele considerarse el departamento como la pieza fundamental de las universidades, y, por tanto, el punto de partida de la evaluación.

Sin embargo, no se puede excluir el efecto que la institución como un todo tiene sobre el funcionamiento de cada departamento o sobre la docencia impartida por cada facultad. Para cada una de las diferentes situaciones posibles se encuentran también experiencias reales:

— Evaluaciones globales para todas las instancias de una institución universitaria se realizan en Estados Unidos, Canadá, Francia y Australia. En este último país existe un modelo de revisión sistemática de instituciones, departamentos e individuos. La evaluación global de la institución puede estar relacionada con la asignación de recursos, como sucede en el Reino Unido para los presupuestos de investigación o en algunos sistemas estatales estadounidenses en los que existe un método de asignación en función del rendimiento.

— Cuando se valora la investigación es usual utilizar como unidad de evaluación al departamento, mientras que para evaluar los aspectos docentes suele considerarse la unidad que se responsabiliza de los programas docentes. En otros países pueden ser también los departamentos, aunque en España serían los centros (facultades y escuelas) las unidades más adecuadas.

VIII.2.3. **Clasificación de las universidades por criterios de calidad**

El utilizar algún criterio de calidad, medidas objetivas u opiniones de expertos, para establecer ordenaciones de calidad entre instituciones universitarias es una práctica habitual exclusivamente en los Estados Unidos, en donde desde siempre existe un sistema de instituciones universitarias absolutamente autónomas, diversificadas y, en consecuencia, competitivas. Aun así, sólo una proporción pequeña de ellas (las más distinguidas) entran a formar parte habitual de estas ordenaciones jerárquicas. Por otra parte, los mecanismos utilizados para la realización de estas ordenaciones han sido fuertemente criticados por sus defectos metodológicos. Aun aceptando la validez y utilidad de tales ordenaciones, es necesario insistir en que sólo son posibles en sistemas universitarios en los que, al menos durante un período suficientemente largo, las instituciones universitarias hayan gozado de autonomía y de la necesaria igualdad en las condiciones de partida, que les haya permitido diversificarse y diferenciarse unas de otras según méritos propios. Esta situación no se da actualmente en ningún país europeo (salvo quizá el Reino Unido), y de hecho en ninguno de ellos se ha intentado por el momento realizar algún tipo de ordenación jerárquica entre universidades o departamentos, al menos siguiendo criterios razonablemente rigurosos. La situación en España es semejante a la de otros países: con una autonomía casi recién estrenada, y con unas diferencias de recursos entre instituciones todavía importantes, sólo

tras las oportunas reasignaciones de recursos que crearan un estado de relativa igualdad en las universidades, y transcurrido un período de tiempo razonable, sería el momento de poder empezar a pensar en criterios de ordenación entre instituciones. Es necesario reasignar recursos «...para conseguir que en un plazo razonable las universidades puedan sentirse en igualdad de condiciones para competir por su desigualdad» (Blas Arriño, 1989: 159).

VIII.2.4. **Evaluaciones autorreguladas y evaluaciones de control**

Desde el punto de vista de quién es el agente que realiza la evaluación institucional se encuentran dos tipos básicos de evaluación, que en la práctica pueden encontrarse con algún grado de combinación entre sí. Otras veces la combinación de criterios se encuentra en la utilización de uno para la evaluación de los aspectos docentes y otro distinto para los investigadores. A riesgo de ser demasiado simplificadores, éstas son las alternativas:

— Las evaluaciones autorreguladas en las que la parte activa de todo el proceso de evaluación está controlado por las propias instituciones, bien individualmente (caso de las universidades norteamericanas), bien por una asociación de universidades (caso de las universidades británicas, suecas, australianas y holandesas). La realización de estas evaluaciones es a veces una iniciativa propia de las instituciones (universidades norteamericanas y británicas) o puede ser consecuencia de un imperativo legal (universidades holandesas, finesas y suecas). La puesta en práctica de la evaluación externos puede hacerse usando equipos de evaluación a la institución (caso estadounidense, holandés y británico), internos (caso sueco o finés) o mixtos (caso canadiense). La revisión del equipo externo puede basarse en un autoestudio realizado previamente por la institución (universidades norteamericanas y holandesas) o utilizar otros criterios (universidades británicas).

— Frente a las evaluaciones autorreguladas se encuentran aquéllas que son promovidas y realizadas por organismos externos a las instituciones educativas. Dos casos típicos, aunque notablemente diferentes, son el sistema de evaluación de la educación superior en Francia y el del sector público de la educación superior británica. En el caso francés la evaluación es realizada por un comité nacional autónomo en el que no participan las universidades como tales, pero tampoco el Gobierno o las autoridades académicas. En cambio, el sector público británico es controlado por organismos netamente gubernamentales, sin ninguna participación de las instituciones.

Se puede afirmar que la gran mayoría de los expertos (académicos y

gubernamentales) reconocen las ventajas de los métodos autorregulados de evaluación frente a los otros. Las intromisiones externas en las instituciones de educación superior son, como está sucediendo en el Reino Unido, un foco constante de conflicto. Una situación intermedia puede ser la francesa, en la que la «intromisión» es realizada por un equipo «neutral».

Sobre las posibilidades de estos métodos en España, Castell (1989) opina: «En el contexto español, yo francamente sería crítico de una evaluación directa de las universidades por parte de la Administración (...) Por lo tanto, la evaluación se puede descentralizar llevándola a las propias universidades, y, sobre todo, remitirla hacia el mercado y a criterios relativamente objetivos, con una estimación estrictamente técnica por organismos competentes desde el ministerio, pero como depositarios casi-notariales de evaluaciones definidas por criterios externos a la Administración» (Castell, 1989: 69).

VIII.2.5. **Indicadores de rendimiento y revisiones por colegas (peer review)**

Desde el punto de vista de cuál es el criterio de medida para realizar la evaluación, éstas se pueden fundamentar en dos sistemas alternativos que, como en otros casos, pueden encontrarse mezclados con diferentes intensidades. Son los dos siguientes:

— Uso de revisiones por colegas (*peer review*). Esta metodología se fundamenta en la hipótesis de que el sistema educativo es tan complejo que los que están inmersos en él son los más capacitados para opinar sobre la calidad o el buen funcionamiento de las instituciones universitarias. Estas evaluaciones pueden ser realizadas por colegas de otras instituciones, o de la propia institución, e incluso por equipos de revisores académicos constituidos permanentemente con este objetivo. Distintas combinaciones de los tres sistemas se dan con mayor o menor frecuencia en casi todos los países en los que se realizan evaluaciones.

— Uso de indicadores de rendimiento. Son indicadores generalmente cuantitativos, que permiten conocer con mayor operatividad el funcionamiento de las instituciones y su eficacia. Los indicadores de rendimiento resaltan los aspectos institucionales más relacionados con la gestión y la toma de decisiones. La principal acusación que les hacen sus detractores es que pueden ser demasiado simplificadoros de la realidad. Tras la formación de comités *ad hoc* para hacer propuestas de indicadores, y tras amplias discusiones, la utilización de estos indicadores es ya un imperativo legal en el Reino Unido, en Finlandia,

y próximamente lo será en Holanda. En Estados Unidos son también utilizados por los sistemas públicos de educación superior de algunos estados.

La utilización de un enfoque u otro ha sido causa de controversia, especialmente en los últimos cinco años. En los dos países europeos más desarrollados en la aplicación de la evaluación institucional, Holanda y el Reino Unido, se ha hecho hincapié justamente en cada uno de los dos enfoques; mientras que en Holanda se ha desarrollado un sistema basado en las revisiones autorreguladas realizadas por colegas, en el Reino Unido, por causas no ajenas a la política gubernamental (la política thacheriana de las tres EEE: economía, eficiencia, efectividad), el centro del interés ha estado en la implantación de los indicadores de rendimiento. Realmente la dicotomía no es tan absoluta, porque también en el Reino Unido se han usado desde siempre las revisiones por colegas, tanto para la docencia como para la investigación de las universidades, y en Holanda el Gobierno ha llegado a un acuerdo negociado con la asociación de universidades holandesas para el uso de un conjunto de indicadores en aquellas universidades en un próximo futuro.

Parece razonable aceptar que un sistema de evaluación que haga uso de ambos criterios, indicadores y revisiones por colegas, puede ser el más adecuado. La ventaja del uso de los indicadores para resolver problemas como el de establecer unos criterios de asignación racional de recursos parece incuestionable. Por otro lado, los métodos de revisiones por colegas tienen, entre muchas otras, una ventaja esencial que posibilita dar los primeros pasos en la instauración de un sistema de evaluación institucional: su mayor aceptabilidad por los individuos y grupos que van a ser evaluados.

El diseño de un sistema *ex novo* de evaluación institucional debería, según opinión generalizada de los expertos en el tema, iniciarse por un sistema autorregulado que hiciera uso preferente de la evaluación por equipos de colegas, y sólo posteriormente se debería ir completando con la introducción de nuevos elementos evaluadores.

VIII.3. PARTICULARIDADES DEL SISTEMA UNIVERSITARIO ESPAÑOL

La instauración de un sistema de evaluación institucional efectivo exige obviamente el conocimiento de las realidades peculiares de cada sistema universitario. En esta sección recogemos algunas características del sistema universitario español que, aunque conocidas por todos, es necesario remarcarlas por su relación con la evaluación institucional. También haremos una breve revisión de algunas experiencias relacionadas con la evaluación institucional que se han realizado en este país.

VIII.3.1. **Algunas características del sistema en relación con la evaluación**

Cualquier enfoque sobre la evaluación de las universidades está directamente relacionado con su grado de autonomía. Todos los sistemas de evaluación implantados están basados en la existencia de universidades autónomas, ya que sólo a instituciones con un margen de maniobra importante cabe exigirles responsabilidades sobre sus actuaciones; en las instituciones reguladas desde una autoridad central, sólo cabe controlar que las normas impuestas son cumplidas correctamente. «Efectivamente, la pregunta por la calidad sólo tiene sentido si se parte del principio de que las universidades son y deben ser distintas y de que pueden y deben competir entre sí. Ello sólo es posible si, aparte de aceptar un principio de autonomía universitaria y, por lo tanto, de autogobierno, se sientan las bases para que el conjunto de todas las universidades, en el juego espontáneo e interactivo de acciones y reacciones, pueda generar calidad diferenciada» (Lamo, 1989: 87). En España la LRU estableció un sistema de amplia autonomía que, sin embargo, no ha sido suficientemente desarrollado. Veamos lo que opinan dos observadores extranjeros, dos expertos en investigación institucional, Goedegebuure y Maassen (1990), sobre el problema de la autonomía universitaria en España: «En España, la tendencia hacia la descentralización de la toma de decisiones desde el Ministerio de Educación a las instituciones de educación superior ha sido muy pronunciada. La descentralización no es un fenómeno limitado a la educación superior, sino que es detectable en toda la sociedad. Especialmente, el cambio de una sociedad estatalmente centralizada a una sociedad descentralizada con regiones autónomas ha creado un importante dilema para la educación superior española. Las administraciones regionales quieren operar tan independientemente como sea posible e intentan extender sus propias estructuras burocráticas, mientras que las autoridades del Estado no quieren abandonar su poder. Como resultado, las universidades no son todavía tan autónomas como suponía la Ley de Reforma Universitaria de 1983. Ellas se han de enfrentar a las autoridades del Estado central, así como a las autoridades regionales.»

«La LRU pretendía promocionar la descentralización y la diferenciación (...) La LRU concedía a las universidades más autonomía, entre otras cosas, para hacerles posible apoyar los intentos de España de integrarse entre las sociedades industriales avanzadas. Aunque es demasiado pronto para sacar conclusiones definitivas, puede decirse que, entre otras cosas, como consecuencia de los problemas mencionados arriba, hay importantes barreras en el camino de las universidades hacia mayores niveles de autonomía. Las decisiones de cada universidad han de ser confrontadas con el Ministerio de Educación, con el Consejo de Universidades, con un gobierno regional dotado de su

propio parlamento y de órganos ejecutivos, y con un Consejo Social, con una mayoría no académica, que supervisa los asuntos económicos».

«Las cuatro relaciones mencionadas arriba pueden ser consideradas como cuatro de las más obvias relaciones que pueden influir en el éxito de los cambios en la educación superior española. Con respecto a la primera relación, puede argumentarse que el actual equilibrio debe cambiar hacia un aumento en la autonomía de las instituciones. El Gobierno, en cualquier caso, no quiere abandonar su poder. Ese parece ser el primer dilema español: el Gobierno promueve la autonomía universitaria, pero no quiere descentralizar la toma de decisiones (...) El segundo dilema consiste en la actitud de las autoridades regionales que parecen estar más interesadas en extender sus propias estructuras burocráticas que en descentralizar el poder a las universidades (...) Se supone que el Consejo de Universidades debe proteger a las universidades y ayudarlas a desarrollar su autonomía. Dado que el presidente del Consejo es el ministro y los presidentes de dos de sus principales comités son altos cargos del ministerio, no es muy verosímil que en las circunstancias presentes esta relación conduzca a un equilibrio satisfactorio hacia la meta de incrementar la autonomía. Este es el tercer dilema de la educación superior española (...) Dado que tres quintos de los miembros del Consejo Social provienen de fuera de la comunidad académica, esta responsabilidad (la del Consejo de aprobar el presupuesto) representa una amenaza para la autonomía de las universidades. Este puede ser considerado como el cuarto dilema al que se enfrentan las universidades en España. Estos cuatro dilemas influyen fuertemente en las posibilidades de éxito del proceso de cambio hacia el incremento de la autonomía universitaria en España.» (Goedegebuure y Maassen, 1990: 6-7, 19-20).

Una peculiaridad del sistema universitario español, que lo diferencia notablemente del de otros países del entorno, es el de que su situación actual todavía es expansiva. Como ya se dijo en el capítulo I, por razones demográficas, pero también económicas, mientras que en otros países desarrollados se ha alcanzado una situación de relativa estabilidad numérica en sus sistemas universitarios, en España estamos todavía en momentos de plena expansión en cuanto al número de estudiantes universitarios, y posiblemente lo estaremos hasta finales de siglo. En esta situación, la dificultad de evaluar un sistema, de por sí tan complejo como el universitario, evidentemente se complica (Montero, 1989).

Como también se comentó en el capítulo I, otro serio problema de la educación superior en España es el de su escasa dotación económica. Como idea de la posición relativa de la situación española frente a la de otros países europeos se pueden ofrecer los siguientes datos: en el año 1983, el gasto

público por estudiante en España era de 814 dólares, frente a 5.538 en el Reino Unido, 2.618 en Alemania (RF), 2.709 en Suecia, 1.664 en Italia o 737 en Yugoslavia (Mora, 1988). A pesar de los indudables esfuerzos presupuestarios (en los cuatro años siguientes hasta 1986 creció la financiación pública total a las universidades en un 23 % en términos reales), dado el continuo incremento de estudiantes, el gasto público por estudiante en 1986 era levemente inferior al de 1983 (Pedró, 1988). Indudablemente, al hablar de evaluación de la calidad de nuestras universidades, no se pueden olvidar estas cifras.

VIII.3.2. **Algunas experiencias relacionadas con la evaluación institucional**

Aunque no se han realizado en España investigaciones sistemáticas en el campo de la evaluación institucional, existen toda una serie de experiencias que es necesario recordar por las aportaciones que hacen, aunque sea lateralmente, al tema que nos ocupa. Sin pretender que ésta sea una revisión exhaustiva, comentaremos a continuación algunas experiencias interesantes.

El sistema de retribuciones como resultado de una evaluación docente e investigadora del profesorado

Evidentemente la primera experiencia evaluadora que hay que reseñar es el sistema actual de evaluación de las actividades docentes e investigadoras del profesorado, con repercusiones en las retribuciones; la evaluación docente realizada por las propias instituciones, la investigadora a cargo de una comisión de carácter estatal.

Las opiniones del profesorado sobre las evaluaciones

Para la implementación de un modelo de evaluación parece necesario conocer las opiniones de los que van a ser evaluados. Diferentes estudios realizados en las universidades de Barcelona y Oviedo (Rodríguez, 1989; De Miguel y Rodríguez, 1990) permiten conocer algunas de las opiniones del profesorado respecto a la evaluación. Recogiendo estas opiniones, los autores de estos estudios proponen una evaluación ponderada de las componentes docente, investigadora y de servicios del profesorado, según estos criterios:

1. 50 por 100 para la docencia, evaluada por:
 - Estudiantes, 50 por 100.
 - Colegas, 30 por 100.

- Uno mismo, 20 por 100.
- 2. 40 por 100 para la investigación, evaluada por:
 - Expertos externos, 50 por 100.
 - Colegas, 25 por 100.
 - Uno mismo, 25 por 100.
- 3. 10 por 100 para servicios públicos y administrativos valorados por colegas y/o expertos.

Otro estudio en el que se investigan las opiniones del profesorado sobre la evaluación es el realizado en la Universidad de Valencia (Jornet *et al.*, 1989). Algunos resultados destacados sobre estas opiniones son los siguientes:

— Una proporción alta del profesorado está de acuerdo con la idea de la evaluación como un elemento de promoción personal y de control sobre sus actividades.

— Existe un acuerdo bastante unánime acerca de la importancia de la valoración de la docencia y de la investigación, algo menos sobre el valor de las actividades administrativas, y bastante menos sobre la de los servicios a la comunidad y las actividades profesionales.

— Que los alumnos intervengan en la evaluación docente es aceptado por un 67 por 100 de los encuestados. Como manifestación del clima poco cordial que existe dentro de los departamentos universitarios, un 45 por 100 se manifiesta en contra de la idea de ser evaluados por los colegas del propio departamento. La evaluación por comisiones de colegas democráticamente elegidos es aceptada por un 48 por 100 del profesorado, aunque un 42 por 100 de los catedráticos se oponen a ella. La autoevaluación es aceptada por el 58 por 100 del profesorado.

— Una gran mayoría del profesorado está de acuerdo en la evaluación de su actividad, como una práctica continua a lo largo de su vida académica.

Evaluaciones de la docencia

La evaluación de la calidad de la educación universitaria, o la de los aspectos docentes del profesorado es, con mucho, la actividad evaluadora

que más se ha desarrollado en España. Son abundantes los estudios, procedentes en gran parte de los ICEs, que han tratado este problema. El desarrollo de los estatutos de las universidades ha creado la necesidad de encontrar respuestas al problema de la evaluación, que en la práctica, al menos por el momento, se está enfocando casi exclusivamente hacia la evaluación de la docencia, en la mayoría de los casos basada en encuestas realizadas a los alumnos. Muchas de estas propuestas evaluadoras están todavía en fase de proyecto o de ensayo, por lo que es difícil sacar conclusiones definitivas de ellas. Como ejemplos de estos enfoques pueden citarse, entre otros, los recientes trabajos de Escudero (1989), Fernández Pérez (1989), Jornet *et al.* (1989), Tejedor *et al.* (1989).

La utilización de indicadores cuantitativos

Al contrario que la anterior, la experiencia en el uso de indicadores cuantitativos para evaluar algún aspecto de las universidades españolas es realmente escaso. Con el objetivo de analizar comparativamente la asignación de recursos se ha hecho algún estudio que valoraba la relación entre la docencia total impartida por las universidades y su correspondiente dotación de profesorado. Utilizando criterios algo diferentes entre sí, cabe citar los trabajos de Moltó y Oroval (1982), y Velasco y Velázquez (1983), en los que se establecía una ordenación de universidades en función de sus necesidades de profesorado.

Refiriéndose a la evaluación de la actividad investigadora, Goberna y Pastor (1988) hacían un ensayo de evaluación de esta actividad por facultades, utilizando como criterio las publicaciones realizadas por sus miembros.

Un modelo de evaluación del rendimiento educativo

Escudero (1986) ofrece diversos criterios de evaluación del rendimiento de las universidades a partir de la información que proporcionan los alumnos, ya que éstos son para el autor «la principal fuente de información para evaluar el rendimiento de la universidad» (Escudero, 1986: 191). Estos son los indicadores propuestos:

1. Tasa de finalización de estudios.
2. Tasa de abandonos.
3. Duración real de los estudios en relación a las previsiones curriculares.

4. Calificaciones obtenidas.

5. Valoraciones de:

— La satisfacción global con la carrera.

— Del grado de interés hacia los alumnos por parte de profesores e institución.

— De la adecuación de la docencia a las capacidades de los alumnos y del nivel de exigencias en los exámenes.

— De la organización y ambiente académico.

Un modelo de evaluación de la educación superior española

Fernández Sánchez *et al.* (1989) presentan el marco teórico-estructural para un modelo interactivo de evaluación de la educación superior española. El modelo se basa en el análisis conjunto de cuatro grupos de variables presuntamente indicadoras de la calidad de las instituciones universitarias: variables de infraestructura del centro, variables relativas al profesorado, variables relativas al alumnado y variables relativas a las relaciones internas y al clima del centro. El objetivo del estudio es el de analizar la validez de las variables definidas como instrumentos para la evaluación de la calidad en las instituciones universitarias españolas. Ignoramos si se ha llegado ya a resultados definitivos en este proyecto.

Criterios para una evaluación de las universidades españolas

Tras un seminario sobre criterios de calidad en las universidades, Pulido (1989) resumía las discusiones de la reunión y proponía, entre otras, estas conclusiones sobre los rasgos esenciales de una evaluación:

1. Sobre las características de la evaluación:

— La evaluación debe estar referida a centros, o departamentos, dado que a ese nivel es más factible realizar valoraciones homogéneas. Ello no excluye evaluaciones sobre eficacia de la gestión universitaria en su conjunto.

— El sistema de evaluación debe ser continuo, permanentemente actualizado, y de implantación progresiva.

— Se deben usar enfoques y criterios múltiples, y su puesta en marcha debe ser un proceso cuidadoso.

— Se deberían usar evaluaciones alternativas complementarias, realizadas por instituciones privadas y por organismos oficiales.

2. Sobre los criterios de evaluación:

— Sería conveniente armonizar criterios cuantitativos y cualitativos.

— Se ha de evaluar tanto la docencia como la investigación.

VIII.4. **UNA PROPUESTA DE EVALUACIÓN INSTITUCIONAL PARA LAS UNIVERSIDADES ESPAÑOLAS**

Hacer una propuesta de evaluación institucional para las universidades españolas desde una posición individual es, sin duda alguna, una osadía. Pero como ya se señaló al inicio del capítulo, se trata simplemente de esbozar unas líneas maestras que recojan las ideas más interesantes que se han ido exponiendo a lo largo del estudio, recogidas a su vez de la experiencia de otros países, y que puedan servir de pie para iniciar la discusión.

Trataremos de marcar las líneas básicas de una evaluación institucional que podríamos llamar «ideal», pero que dada la realidad actual de las universidades españolas y la inexperiencia evaluadora, no necesariamente habría de ponerse en marcha al unísono, sino que preferiblemente debería iniciarse con pasos cortos pero seguros. Veamos algunas de las características esenciales que podría tener este proceso evaluador.

VIII.4.1. **Algunos requisitos generales de la evaluación**

La evaluación institucional de las universidades españolas debe ser un proceso absolutamente respetuoso con la autonomía universitaria; es más, sería deseable que la puesta en marcha de un sistema continuo de evaluación institucional llevara aparejado el que las universidades alcanzaran mayores dosis de autonomía, especialmente en cuestiones de tipo económico. Sin embargo, el proceso evaluador ha de hacerse con unos criterios de homogeneidad que permitan utilizar los resultados para hacer asignaciones de recursos más racionales y justas entre todas las componentes del sistema público de educación superior español. En estas circunstancias parece obvio que el Consejo de Universidades, en cuanto representante conjunto de todas las

universidades, debería ser el organismo que se responsabilice de la gestión o puesta en marcha del proceso evaluador, lo que no quiere decir lógicamente que haya de ser el Consejo como tal el que realice las evaluaciones. Otra alternativa viable sería la de crear un ente específico para la gestión de la evaluación de las universidades, independiente de cualquier organismo gubernamental, que se hiciera cargo del seguimiento de todo el proceso, tal y como sucede en Francia. En cualquier caso, este no debería ser un proceso controlado desde el Ministerio de Educación, aunque éste deberá estimular su implantación, e incluso exigirla en una fase posterior.

La evaluación de las instituciones universitarias debe cubrir dos objetivos básicos: la mejora de la calidad de la institución y la gestión más eficaz. Para ello parece razonable adoptar dos tipos de enfoques: el de las evaluaciones autorreguladas de tipo básicamente cualitativo, y el uso complementario de indicadores de rendimiento cuantitativos y cualitativos.

Es esencial para el éxito del proceso el que éste sea visto y aceptado por todos los implicados como un método para mejorar la calidad y el rendimiento de las propias instituciones. Por tanto el sistema de evaluación que finalmente se implante debe ser el resultado de la recogida de información sobre los criterios que deben utilizarse, de la discusión amplia entre los interesados, y, finalmente, de la decisión que tomen los legítimos representantes de las universidades que se van a autosometer al proceso de evaluación.

La evaluación debe llegar a todas las instancias de la vida académica; departamentos, centros y universidad. Realmente la situación ideal sería la de evaluar una por una cada universidad, globalmente y de una sola vez, desde los departamentos hasta el rectorado. Sin embargo, se hace difícil pensar cómo sería posible realizar de una sola vez una evaluación global de todas y cada una de las partes de una de nuestras macrouiversidades. Por esta razón parece conveniente que el sistema de evaluación que se adopte, al menos en una primera fase de implantación y de adquisición de experiencia, considere separadamente unidades académicas menores y homogéneas (departamentos de un mismo tipo, facultades de la misma clase), para todas las universidades a la vez. Por la misma razón, parece razonable evaluar separadamente investigación, docencia y otros servicios, aunque sean distintas facetas de una misma realidad. En cualquier caso, la evaluación de los individuos (no solamente del profesorado, al que, no se sabe muy bien por qué razones, suele considerársele objeto exclusivo de la evaluación) se deberá hacer considerándolos parte integrante de cada una de las instancias a las que pertenecen, según el aspecto concreto que se esté evaluando.

La construcción de un sistema de evaluación transversal por unidades homogéneas tiene la desventaja de que dificulta una imagen general de la institución, y, por tanto, la posibilidad de una política global de mejoras. Por eso, cada cierto tiempo, cerrado un ciclo de revisión de todas sus partes, las universidades deberían realizar un resumen general de las distintas evaluaciones que proporcionará una imagen de conjunto de la institución. Por otra parte, este sistema de evaluación transversal tiene la ventaja de permitir con mayor facilidad la construcción de un sistema de indicadores de rendimiento homogéneos que faciliten el análisis de la efectividad y de la eficiencia de cada institución y del sistema universitario en su conjunto.

Es evidente que la existencia de una práctica de evaluación institucional de carácter uniforme para todas las universidades públicas, no debe impedir que éstas desarrollen cuantos sistemas internos de evaluación consideren oportunos, ya que la práctica continuada de estas experiencias es una de las bases más firmes para la mejora de la calidad de las instituciones. Es más, la existencia de un gabinete dedicado a la evaluación sistemática de la institución debería ser una práctica habitual de nuestras universidades, como lo está siendo en las universidades norteamericanas.

VIII.4.2. **La evaluación de la investigación**

La evaluación de la actividad investigadora debe realizarse con criterios diferenciados por departamentos, pero comunes para todos los departamentos del mismo tipo en distintas universidades. Parece lógico que el departamento sea la unidad evaluada, y, dentro de él, las diferentes líneas de investigación que estén constituidas. Llegando a la evaluación hasta el nivel de líneas o equipos de investigación, parece innecesario, e incluso contraproducente, llegar a valoraciones individuales, que pueden desanimar la ya escasa colaboración científica entre los miembros de los departamentos.

La evaluación investigadora deberá tener en cuenta cantidad y calidad de la producción científica. Para ello parece adecuado la existencia de una comisión de expertos a nivel estatal nombrada *ex profeso* para cada tipo de departamento, y que pueda realizar una evaluación con los mismos criterios para todos ellos. Esta comisión debería evaluar basándose en los informes y autoevaluaciones enviados por los propios equipos de investigación y por los departamentos. Los resultados finales de la evaluación deberían darse no sólo por departamentos globalmente, sino también especificando los resultados por líneas de investigación. El informe emitido debería calificar no sólo negativa o positivamente, sino que, en este último caso, debería diferenciar entre un nivel aceptable y otro más notable.

Sería deseable que tanto los departamentos como los equipos o líneas de investigación recibieran una asignación de fondos para la investigación que tuviera en cuenta los resultados de la evaluación. Una evaluación negativa debería ser objeto de investigación por parte de la universidad, e incluso de algún tipo de penalización si las causas así lo aconsejaran (no promoción de sus profesores, no concesión de fondos, etc.). En cualquier caso los resultados de la evaluación deben ser totalmente públicos y, de alguna manera, publicados.

Las comisiones de expertos que realizaran estas evaluaciones deberían estar constituidas por personas con una buena experiencia investigadora, algunos del campo que se está evaluando, otros de un campo próximo, pero alguno de un área absolutamente distante, para evitar tendencias corporativistas. La participación de algún investigador extranjero sería también deseable. Por poner un ejemplo, la evaluación de los departamentos de Economía Aplicada podría realizarse por una comisión de cinco miembros, dos de Economía Aplicada, dos de otros departamentos económicos distintos, y otro de un departamento de física, por poner otro ejemplo, todos ellos con una buena reputación investigadora, aunque no necesariamente ilustre. El nombramiento de las comisiones debería hacerse mediante una elección restringida entre personas previamente seleccionadas y dispuestas a intervenir en estos procesos. Los criterios generales del funcionamiento de las comisiones, así como la infraestructura necesaria, serían misiones del organismo gestor del proceso de evaluación, en nuestra opinión, el Consejo de Universidades. Para establecer los criterios generales de la evaluación se deberá recurrir a los conocimientos disponibles sobre la valoración de la investigación que existen a nivel internacional.

Este proceso debería hacerse cíclicamente, de modo que cada cinco años (o un período de tiempo de este orden) todos los departamentos universitarios fueran revisados. Quizá fuera conveniente revisar cada año departamentos homogéneos (con docencia en los mismos centros) para facilitar una visión de conjunto, en un mismo instante de tiempo, de cada facultad o escuela.

VIII.4.3. **La evaluación de la actividad docente**

A la evaluación docente es a la única que se le ha dedicado hasta el momento alguna atención en nuestras universidades, pero casi siempre centrada en la figura del profesor y no en otras cuestiones que son de indudable importancia, como el plan de estudios, la organización interna del centro docente, los servicios educativos que se ofrecen a los alumnos, etcétera. Dado

el actual funcionamiento de la docencia en nuestras universidades, parece que la facultad o escuela es la unidad que debería ser evaluada. Evaluar los departamentos parece más problemático, ya que su peso en la organización total del currículum de una carrera universitaria no es excesivo, exceptuando, claro está, los programas del tercer ciclo. De hecho, los alumnos, agentes importantes de esta evaluación docente, difícilmente diferenciarían entre unos departamentos y otros: es la formación que reciben en la facultad o escuela la que ellos pueden valorar.

La evaluación de las facultades y escuelas podría hacerse también cíclicamente por períodos de cinco años, de modo que cada universidad revise una parte proporcional de sus centros cada año. Podría coincidir con la revisión de las actividades investigadoras de los departamentos afines integrados en un centro, con lo que se completaría anualmente una perspectiva global de cada centro.

Básicamente el proceso de evaluación de cada centro debería consistir en un informe de autoevaluación realizado por cada unidad docente, por cada departamento y por el centro como un todo, siguiendo una sistemática que habría de ser común para todos los centros del mismo tipo, y bastante semejante para todos los centros del país. Esta sistemática común la debería elaborar, y explicar en cada centro antes de cada evaluación, el Consejo de Universidades, como gestor de todo el proceso de evaluación.

La autoevaluación deberá recoger las opiniones sobre el funcionamiento del centro (aspectos docentes, pero también organizativos, estructurales y de recursos) de todas las partes implicadas en el proceso educativo, incluido el personal de servicios. Además de los profesores, los alumnos deberán ser consultados no sólo para que opinen sobre el profesorado, sino también sobre la organización docente, los planes de estudio, los servicios que reciben, el clima académico en el que se desenvuelven y tantas otras cosas que son necesarias para la formación de los estudiantes. Sería muy interesante que se recogiera información de los antiguos alumnos (a los que lamentablemente las universidades españolas no prestan la más mínima atención bajo ningún aspecto), así como de los empleadores locales.

Como segundo paso las universidades deberían nombrar una comisión que evaluara al centro, tras una corta visita, utilizando como base el propio informe de autoevaluación. Esta comisión podría consultar personalmente a algunos de los implicados (profesores, alumnos y personal de servicios, pero también antiguos alumnos y empleadores), para finalmente emitir un informe con los resultados de su evaluación. Este informe sería remitido a los responsables del

centro para que hicieran las alegaciones oportunas, y finalmente toda esta información pasaría a la universidad, que debería tomar las medidas para corregir los defectos señalados por el informe y hacerlo público, de un modo resumido, para conocimiento del mundo académico y de la sociedad.

La comisión evaluadora debería estar formada por profesores de otras universidades y un representante de la propia universidad, aunque de otra facultad diferente a la evaluada. Los representantes de otras universidades deberían ser mayoritariamente, pero no exclusivamente, del mismo tipo de facultad (quizá fuese adecuado incluir en todas las comisiones un experto de ciencias de la educación). La presencia de algún evaluador extranjero sería recomendable.

VIII.4.4. **La evaluación global de las instituciones**

Siguiendo las pautas anteriores, en un período de cinco años se habría realizado una revisión completa de todos los departamentos y centros de una universidad. Podría ser el momento de realizar una revisión global de la institución que se centrara preferentemente en sus aspectos organizativos, pero que resumiera también las evaluaciones docentes e investigadora de cada una de sus componentes. Una comisión formada por miembros de otras universidades, no sólo profesores, sino también algún gestor, podrían hacer esa revisión basándose en el informe autoevaluado realizado por la propia universidad. Esta evaluación, además de analizar el funcionamiento institucional como en las otras instancias evaluadas, podría utilizar indicadores de rendimiento para resumir la información global de toda la institución, y permitir las comparaciones entre distintas instituciones, todo ello con las debidas precauciones con que deben tomarse estos indicadores.

VIII.4.5. **La posibilidad del uso de indicadores de rendimiento**

En nuestra opinión la evaluación de las instituciones universitarias debe basarse en la utilización de autoevaluaciones combinadas con evaluaciones de comisiones externas que revaliden, si es el caso, las primeras. Sin embargo, con la sucesiva utilización de estas fórmulas, puede llegar el momento en el que se disponga de la suficiente experiencia evaluadora y del necesario conocimiento de la realidad de nuestras instituciones, como para que fuera posible establecer unos indicadores de rendimiento que fueran lo suficientemente amplios como para no simplificar excesivamente la realidad de las instituciones, y que permitieran, por otra parte, una mejor evaluación de la eficiencia y ayudaran a la toma de decisiones globales sobre el sistema universitario.

Transcurridos algunos años desde que se iniciara el proceso de evaluación, recogiendo toda la experiencia acumulada, sería el momento oportuno para establecer (tras las propuestas y discusiones necesarias) un conjunto de indicadores de rendimiento adecuados para sintetizar la realidad de las universidades españolas.

VIII.4.6. **Cómo iniciar el proceso de la evaluación**

«En la práctica resulta difícil moverse en la dimensión de la eficiencia y, por el momento, lo más razonable es la aplicación y mejora de estrategias de autoevaluación desde múltiples y diversas perspectivas, de manera que un análisis posterior de resultados de todas ellas nos acerque a un diagnóstico comprensible global» (Escudero, 1986: 200). De acuerdo con estas ideas, los primeros pasos deberían estar encaminados al aprendizaje y puesta en práctica de las técnicas de autoevaluación, para pasar a continuación a la utilización de equipos de revisión externos, y sólo finalmente al uso de indicadores de rendimiento.

En una primera fase, un primer curso académico como máximo, se deberían realizar ensayos de autoevaluaciones en diferentes tipos de departamentos y de facultades, de modo que se puedan descubrir las deficiencias del sistema de evaluación en diversas situaciones. A continuación se podría empezar formalmente el proceso de evaluación con el nombramiento de las correspondientes comisiones que evaluaran aproximadamente un quinto de los departamentos de todo el país y un quinto de los centros de cada universidad, para en cinco años cerrar el proceso con una evaluación global de la universidad, antes de volver a iniciarlo de nuevo. Sólo pasada esta primera fase de evaluaciones institucionales, sería el momento oportuno de replantearse el modelo de evaluación, de iniciar la utilización de indicadores de rendimiento, o incluso la de crear situaciones que puedan estimular una cierta dosis de competitividad entre diferentes instituciones.

Aunque desde el primer momento se debería establecer alguna ligazón entre los resultados de la evaluación y la asignación de recursos, esta conexión deberá ser sumamente cuidadosa y paulatina, entre otras cosas porque no debe perjudicar a aquellas instituciones menos brillantes, precisamente porque tal vez disponen inicialmente de recursos insuficientes.

Otra cuestión de vital importancia para el éxito del proceso evaluador es el seguimiento, desde el primer momento, del cumplimiento con el que cada institución responde a las recomendaciones que le han hecho las respectivas

comisiones de evaluación. Las universidades deberían ejercer ese control sobre sus propios departamentos y centros a través de su gabinete de evaluación y de investigación institucional que debería estar constituido en cada una de ellas.

VIII.4.7. **Unas sugerencias finales**

Hasta aquí el boceto de propuesta que hacemos como resumen de nuestro estudio. En realidad creemos que existe un paso previo a los enunciados con anterioridad: la declaración de nuestras universidades de la voluntad política de iniciar un proceso de autoevaluación que persiga estos dos objetivos esenciales:

— Situar a las universidades españolas en los niveles de calidad que serán necesarios para competir en un mundo científica y tecnológicamente avanzados.

— Demostrar a la sociedad que su importante inversión en educación superior está siendo bien utilizada.

Manifestada la voluntad de la evaluación, el segundo paso a dar sería el de crear una comisión amplia, pero operativa, de algunos expertos en evaluación y de representantes de las universidades, que hiciera una propuesta de sistema de evaluación para nuestras universidades mucho más completo que el que aquí se ha hecho. Esta propuesta podría pasarse a discusión en los departamentos universitarios, de modo que se recogiera la opinión de todos los implicados antes de hacer una propuesta definitiva. En cualquier caso la propuesta no debería ser excesivamente rígida, de modo que pueda ir reformándose con la experiencia que se vaya acumulando en el transcurso de su aplicación.

Una última recomendación: el proceso debe ser todo lo cuidadoso y paulatino que sea necesario para garantizar su éxito, pero debería iniciarse sin más dilación.

REFERENCIAS BIBLIOGRAFICAS

- ABBOT, W. F. (1972): «University and Departmental Determinants of Prestige of Sociology Departments». *American Sociologist*, 7, Nov., 14-15.
- ABBOTT, W. F., y BARLOW, H. M. (1972): «Stratification Theory and Organizational Rank. Resources, Functions, and Organizational Prestige in American Universities». *Pacific Sociological Quarterly*, 15, 4, 401-424.
- ACHERMAN, H. A. (1990): «Evaluation de la qualité par un controle collegial. Un nouveau domaine de cooperation interuniversitaire». *Gestion de l'enseignement superieur*, 2, 2, 203-218.
- ANTIA, J. M. (1976): «Critical success factors in polytechnic performance». *Educational Administration*, 5, 1, 14-32.
- ASHBY, E. (1963): «Introduction: Decision-Making in the Academic World». *Sociological Studies in British University Education. Monograph*, no. 7, Keele, UK, University of Keele.
- ASTIN, A. W. (1977): *Four Critical Years: Effects of College on Beliefs, Attitudes and Knowledge*. San Francisco, Jossey-Bass.
- (1985): *Achieving Educational Excellence*. San Francisco. Jossey-Bass.
- (1989): «Moral Messages of the University». *Educational Record*, 70, 2, 22-25.
- ASTIN, A. W., y HENSON, J. W. (1977): «New Measures of College Selectivity». *Research in Higher Education*, 6, 1-8.
- ASTIN, A. W., y SOLMON, L. C. (1979): «Measuring Academic Quality: An Interim Report». *Change*, 11, 48-51.
- (1981): «Are Reputational Ratings Needed to Measure Quality?». *Change*, Oct., 14-19.
- BALDERSTON, F. E. (1985): «Academic Program Review and the Determination of Universities Priorities». *International Journal of Institutional Management in Higher Education*, 9, 3, 237-248.
- BALL, C. (1985): *Fitness for Purpose*. Guildford, UK, Society for Research into Higher Education.
- BALL, R., y HALWACHI, J. (1987): «Performance indicators in higher education». *Higher Education*, 16, 393-405.
- BANTA, T. W. (ed.) (1988): *Implementing Outcomes Assessment: Promise and Perils*. New Directions for Institutional Research, 59, San Francisco, Jossey-Bass.
- BANTA, T. W. (1985): «Use the Outcomes Information at the University of Tennessee, Knoxville». En Ewell, P. T. (ed.): *Assessing Educational Outcomes*. New Directions for Institutional Research, 47, San Francisco, Jossey-Bass.

- BANTA, T. W.; FISCHER, H. S., y MINKEL, C. W. (1986): «Assessment of Institutional Effectiveness at the University of Tennessee, Knoxville». *International Journal of Institutional Management in Higher Education*, 10, 3, 262-271.
- BARNETT, R. A. (1987): «The maintenance of quality in the public sector of UK higher education». *Higher Education*, 16, 279-301.
- BAUER, M. (1988): «Evaluation in Swedish Higher Education: recent trends and the outlines of a model». *European Journal of Education*, 23, 1/2, 25-36.
- BAUMOL, W. (1967): «Macroeconomics of unbalanced growth: The anatomy of urban crisis». *American Economic Review*, 57, 3, 415-426.
- BEYER, J. M., y SNIPPER, R. (1974): «Objective Versus Subjective Indicators of Quality in Graduate Education». *Sociology of Education*, 47, 541-557.
- BILLING, D. (1986): «Judging Institutions». En Moodie, G. C. (ed.): *Standards and Criteria in Higher Education*. Guilford: Society for Research into Higher Education.
- BINGHAM, R. D.; HENRY, G. T., y BLAIR, J. P. (1981): «Urban Affairs Programs. A Ranking in US Colleges and Universities». *Urban Affairs Quarterly*, 16, 3, 355-368.
- BIRCH, D. W., y CALVERT, J. R. (1977): «Performance indicators in higher education: A comparative study». *Educational Administration*, 5, 2, 15-29.
- BLACKBURN, R. T., y LINGENFELTER, P. E. (1973): *Assessing Quality in Doctoral Programs: Criteria and Correlates of Excellence*. Ann Arbor, Michigan, Center for the Study of Higher Education, ED 078728.
- BLAS ARITIO, F. A. (1989): «Discurso de clausura». En Ministerio de Educación y Ciencia: *Hacia una clasificación de las universidades según criterios de calidad*. Madrid, Ministerio de Educación y Ciencia y Fundación Universidad-Empresa.
- BLAU, P. M., y MARGULIES, R. Z. (1974): «The Reputations of American Professional Schools». *Change*, 6, Winter, 42-47.
- BLAUG, M. (1968): «The productivity of universities». *Minerva*, 6, 2, 398-407.
- BOGUE, E. G. (1982): «Allocation of Public Funds on Instructional Performance/Quality Indicators». *International Journal of Institutional Management in Higher Education*, 6, 1, 37-43.
- BORMANS, M. J.; BROUWER, R.; IN'T VELD, R. J., y MERTENS, F. J. (1987): «The role of performance indicators in improving the dialogue between government and universities». *International Journal of Management in Higher Education*, 11, 2, 181-194.
- BOWEN, H. R. (1977): *Investment in Learning: The Individual and Social Values of American Higher Education*. San Francisco, Jossey-Bass.
- (1979): «Goals, Outcomes, and Academic Evaluation». En: *Evaluating Educational Quality: A Conference Summary*. Washington, Council of Postsecondary Accreditation.
- (1981): «Cost differences: The Amazing Disparity Among Institutions of Higher Education in Educational Cost Per Student». *Change*, En/Feb., 21-27.
- BOWKER, A. H. (1965): «Quality and quantity in higher education». *Journal of the American Statistical Association*, 60, 309, 1-15.
- BOYER, C. M.; EWELL, P. T.; FINNEY, J. E., y MINGLE, J. R. (1987): «Assessment and Outcomes Measurement. A View from the States». *AAHE Bulletin*, 39, 7, 8-12.

- BOYER, E. L. (1987): «American higher education: the tide and the undertow». *International Journal of Institutional Management of higher Education*, 11, 1, 5-12.
- BRENNAN, J. (1986): «Peer review and partnership: changing patterns of validation in the public sector of higher education in the United Kingdom». *International Journal of Institutional Management in Higher Education*, 10, 2, 150-156.
- (1990): «Quality Assessment in the Public Sector in Great Britain». En Goedegebuure, L. C. J.; Maassen, P. A. M., y Westerheijden, D. F. (eds.): *Peer Review and Performance Indicators*. Utrecht, NL, Uitgeverij Lemma.
- BRESTERS, D. W., y KALKWIJK, J. P. Th. (1990): «The Role of Inspectorate of Higher Education». En Goedegebuure, L. C. J.; Maassen, P. A. M., y Westerheijden, D. F. (eds.): *Peer Review and Performance Indicators*. Utrecht, NL, Uitgeverij Lemma.
- BRIDGER, G. (1989): «Attitudinal Surveys in Institutional Effectiveness». En Nichols, J. O.: *Institutional Effectiveness and Outcomes Assessment Implementation on Campus: A Practitioner's Handbook*. Nueva York, Agathon Press.
- BROWN, D. G. (1967): *The Mobile Professors*. Washington, The American Council of Education.
- CALVERT, J. R. (1979): «Relative performance». En Billing, D. (ed.): *Indicators of performance*. Guildford, UK, Society for Research into Higher Education.
- (1981): «Workshop report. Institutional performance assessment under conditions of changing needs». *International Journal of Institutional Management in Higher Education*, 5, 3, 237-249.
- CAMERON, K. S. (1978): «Measuring Organizational Effectiveness in Institutions of Higher Education». *Administrative Science Quarterly*, 23, 4, 604-632.
- (1983): «Strategic Responses to Conditions of Decline». *Journal of Higher Education*, 54, 4, 359-380.
- (1984): «Assessing Institutional Ineffectiveness: A Strategy for Improvement». En Scott, R. A. (ed.): *Determining the Effectiveness of Campus Services*, New Directions for Institutional Research, 50. San Francisco, Jossey-Bass.
- CARNEGIE COMMISSION ON HIGHER EDUCATION (1973): *The Purposes and the Performance of Higher Education in the United States. Approaching the year 2000*. New York, McGraw-Hill.
- CARNEGIE COUNCIL (1980): *Three Thousand Futures*. San Francisco, Jossey-Bass.
- CARNEGIE FOUNDATION (1987): *A Classification of Institutions of Higher Education. 1987 Edition*. Princeton, The Carnegie Foundation for the Advancement of Teaching.
- CARTER, A. M. (1966): *An Assessment of Quality in Graduate Education*. Washington, American Council in Education.
- CARTTER REPORT (1977): «The Cartter Report on the Leading Schools of Education, Law and Business». *Change*, 9, Feb. 44-48.
- CASSIN, E.; BOTTERO, J., y VERCOUTTER, J. (eds.) (1983): *Los imperios del antiguo oriente*. Historia Universal Siglo Veintiuno, vol. 2, 14 edición, Madrid, Siglo XXI.
- CASTELLS, M. (1989): «Los sistemas de evaluación de la calidad de las universidades: reflexiones sobre la experiencia norteamericana, con España como punto de referencia».

En Ministerio de Educación y Ciencia: *Hacia una clasificación de las universidades según criterios de calidad*. Madrid, Ministerio de Educación y Ciencia y Fundación Universidad-Empresa.

- CERI (1982): *The University and the Community*, París, OCDE-Centre for Educational Research and Innovation.
- CLARK, M. J.; HARTNETT, R. T., y BAIRD, L. L. (1976): *Assessing Dimensions of Quality in Doctoral Education: A Technical Report of a National Study in Three Fields*. Princeton, Educational Testing Service.
- CLEMENTE, F., y STURGIS, R. B. (1974): «Quality of Department of Doctoral Training and Research Productivity». *Sociology of Education*, 47, 287-299.
- COLE, J. R., y LIPTON, J. A. (1977): «The Reputations of American Medical Schools». *Social Forces*, 55, 3, 662-684.
- COMMONWEALTH TERTIARY EDUCATION COMMISSION (1986): *Review of efficiency and effectiveness in higher education*. Camberra, Australian Government Publishing Service.
- CONRAD, C. F., y BLACKBURN, R. T. (1985a): «Program Quality in Higher Education: A Review and Critique of Literature and Research». En Smart, J. C. (ed.): *Higher Education: Handbook of Theory and Research*, Vol. 1, New York, Agaton Press.
- (1985b): «Correlates of Departmental Quality in Regional Colleges and Universities». *American Educational Research Journal*, 22, 2, 279-295.
- (1986): «Current views of departmental quality: An empirical examination». *The Review of Higher Education*, 9, 3, 249-266.
- CONRAD, C. F., y WILSON, R. F. (1985): *Academic Program Review*. ASHE-ERIC Higher Education Report, 5, Washington: Association for the Study of Higher Education.
- COULTER, W. B., y MOORE, A. H. (1987): «Utilization of performance indicators for financing at state level in the United States: The Ohio case». *International Journal of Management in Higher Education*, 11, 2, 195-207.
- COWAN, J. (1985): «Effectiveness and efficiency in higher education». *Higher Education*, 14, 235-239.
- COX, W. M., y CATT, V. (1977): «Productivity Ratings of Graduate Programs in Psychology Based on Publication in the Journals of the American Psychological Association». *American Psychologist*, 32, 10, 793-813.
- CRAVEN, E. C. (1980): «Evaluating Program Performance». En Jedamus, P.; Peterson, M. W., et al. (eds.): *Improving Academic Management*. San Francisco, Jossey-Bass.
- CRESPO, A. (1989): «Posibilidades y consecuencias de una ordenación de universidades en España». En Ministerio de Educación y Ciencia: *Hacia una clasificación de las universidades según criterios de calidad*, Madrid, Ministerio de Educación y Ciencia y Fundación Universidad-Empresa.
- CRESWELL, J. W. (ed.) (1986): *Measuring Faculty Research Performance*. New Directions for Institutional Research, 50. San Francisco, Jossey-Bass.
- CUENIN, S. (1987): «The use of performance indicators in universities: an international survey». *International Journal of Management in Higher Education*, 11, 2, 117-139.

- CULLEN, B. D. (1987): «Performance indicators in UK higher education. Progress and prospects». *International Journal of Management in Higher Education*, 11, 2, 171-180.
- CUTHBERT, R. E., y BIRCH, D. W. (1979): «Limitations of the student-staff ratio as a resource control». En Billing, D. (ed.): *Indicators of performance*, Guildford, UK, Society for Research into Higher Education.
- CHAMBERS, R. H. (1984): «Enhancing Campus Quality Through Self-Study». En Scott, R. A. (ed.): *Determining the Effectiveness of Campus Services*, New Directions for Institutional Research, 50, San Francisco, Josse-Bass.
- CHEPS (1990): *The CHEPS Quality Assessment Data Package*. Enchede, NL, CHEPS.
- CHURCH, C. H. (1988): «The Qualities of Validation». *Studies in Higher Education*, 13, 1, 27-43.
- DAHLLÖF, U. (1990b): «Towards a new model for the evaluation of teaching: An interactive process-centred approach». En el Report of the IMHE Study Group on Evaluation in Higher Education. Informe presentado a la *10th General Conference of Member Institutions* del Programme on Institutional Management in Higher Education, París, OCDE, septiembre.
- DAVIS, P., y PAPANÉK, G. (1984): «Faculty Ratings of Major Economic Departments by Citations». *American Economic Review*, 74, 225-230.
- DEAN, J. W. (1976): «An alternative rating system for university economics departments». *Economic Inquiry*, 14, 1, 146-153.
- DE MIGUEL, M., y RODRIGUEZ, S. (1990): «Evaluating university system: a Spanish faculty point of view». Ponencia presentada en el *12th Annual Forum of the European Association for Institutional Research*, Lyon, Septiembre.
- DOCHY, F. J. R. C.; SEGERS, M. S. R., y WYNAND, H. F. M. (1990): «Selecting Performance Indicators». En Goedegebuure, L. C. J.; Maassen, P. A. M., y Westerheijden, D. F. (eds.): *Peer Review and Performance Indicators*. Utrecht, NL, Uitgeverij Lemma.
- CHURCH, C. H. (1988): «The Qualities of Validation». *Studies in Higher Education*, 13, 1, 27-43.
- DAHLLÖF, U. (1990b): «Towards a new model for the evaluation of teaching: An interactive process-centred approach». En el Report of the IMHE Study Group on Evaluation in Higher Education. Informe Presentado a la *10th General Conference of Member Institutions* del Programme on Institutional Management in Higher Education, París, OCDE, septiembre.
- DAVIS, P., y PAPANÉK, G. (1984): «Faculty Ratings of Major Economic Departments by Citations». *American Economic Review*, 74, 225-230.
- DEAN, J. W. (1976): «An alternative rating system for university economics departments». *Economic Inquiry*, 14, 1, 146-153.
- DE MIGUEL, M., y RODRIGUEZ, S. (1990): «Evaluating university system: a Spanish faculty point of view». Ponencia presentada en el *12th Annual Forum of The European Association for Institutional Research*, Lyon, Septiembre.
- DOCHY, F. J. R. C.; SEGERS, M. S. R., y WIJNEN, W. H. F. W. (1990): «Selecting Performance Indicators». En Goedegebuure, L. C. J.; Maassen, P. A. M., y Westerheijden, D. F. (eds.): *Peer Review and Performance Indicators*, Utrecht, NL, Uitgeverij Lemma.

- DOLAN, W. P. (1976): *The Ranking Game: The Power of the Academic Elite*, Lincoln, Nebraska, The Nebraska Curriculum Development Center.
- DREW, D. E., y KARPFF, R. (1981): «Ranking Academic Departments: Empirical Findings and a Theoretical Perspective». *Research in Higher Education*, 14, 4, 305-320.
- ELTON, C. F., y RODGERS, S. A. (1973): «The departmental rating game: Measure of quantity or quality?». *Higher Education*, 2, 439-446.
- ELTON, L. (1986): «Quality in higher education: Nature and purpose». *Studies in Higher Education*, 11, 1, 83-84.
- EICHER, J. C. (1988): «Financiación de la enseñanza superior». En Grao, J. (ed.): *Planificación educativa y mercado de trabajo*, Madrid, Narcea.
- ESCOTET, M. A. (1990a): *Evaluación Institucional Universitaria*, Buenos Aires, Losada.
- (1990b): «Experimental and Quasi-experimental Designs for Institutional Evaluation». Ponencia presentada en el *12th Annual Forum of the European Association for Institutional Research*, Lyon, Septiembre.
- ESCODERO ESCORZA, T. (1986): «Algunos criterios y evidencias del rendimiento universitario». En Latiesa, M. (ed.): *Educación y rendimiento académico en la universidad*. Madrid, CIDE.
- (1989): «Experiencias evaluativas en la Universidad de Zaragoza». En *Consideraciones Metodológicas sobre la Evaluación y Mejora de la Docencia Universitaria*. Informes de Investigación Evaluativa núm. 1, Valencia, Servei de Formació Permanent, Universitat de Valencia.
- EWELL, P. T. (1985): *Assessing Educational Outcomes*. New Directions for Institutional Research, 47, San Francisco, Jossey-Bass.
- FAIRWEATHER, J. S. (1988): «Reputational Quality of Academic Programs: The Institutional Halo». *Research in Higher Education*, 28, 4, 345-355.
- FEASLBY, C. E. (1980): *Program Evaluation*. Washington, American Association for Higher Education.
- FERNANDEZ PEREZ, M. (1989): *Así enseña nuestra universidad*. Madrid, edición del autor.
- FERNANDEZ SANCHEZ, J.; LILLO, B., y MARTINEZ ARIAS, M. R. (1989): «Proyecto de evaluación de la Educación Superior Española». En: *Consideraciones Metodológicas sobre la Evaluación y Mejora de la Docencia Universitaria*. Informes de Investigación Evaluativa núm. 1, Valencia, Servei de Formació Permanent, Universitat de Valencia.
- FINDLAY, P. (1990): «Developments in the Performance Indicator Debate in the United Kingdom». En Goedegebuure, L. C. J.; Maassen, P. A. M., y Westerheijden, D. F. (eds.): *Peer Review and Performance Indicators*, Utrecht, NL, Uitgeverij Lemma.
- FRACKMANN, E. (1987): «Lessons to be learnt from a decade of discussions on performance indicators». *International Journal of Institutional Management of Higher Education*. 11, 2, 1-162.
- FRACKMANN, E., y MAASSEN, P. A. M. (1987): «The Meaning of Institutional Self-Evaluation within the Framework of a Self-Regulation System». Ponencia presentada en el *27 Annual Forum of the Association for Institutional Research*, Kansas City, Missouri, 3-6 de mayo, ED 293429.

- FRAENKEL, P. (1989): «La experiencia de Estados Unidos». En Ministerio de Educación y Ciencia: *Hacia una clasificación de las universidades según criterios de calidad*. Madrid, Ministerio de Educación y Ciencia y Fundación Universidad-Empresa.
- FULLER, C. H. (1986): «Ph. D. Recipients: Where Did They Go To College?». *Change*, 18, 6, 42-51.
- FURUMARK, A. M. (1981): «Institutional Self-evaluation in Sweden». *International Journal of Institutional Management of Higher Education*, 5, 3, 207-216.
- GARDNER, D. E. (1977): «Five Evaluation Frameworks: Implications for Decision-making in Higher Education». *Journal of Higher Education*, 48, 5, 571-592.
- GEORGE, M. D. (1982): «Assessing Program Quality». En Wilson, R. (ed.): *Designing Academic Program Review. New Directions for Higher Education*, 37. San Francisco, Jossey-Bass.
- GIBBONS, M. (1985): «Methods for the Evaluation of Research». *International Journal of Institutional Management of Higher Education*, 9, 1, 79-85.
- GLENN, N. D., y VILLIMEZ, W. (1970): «The Productivity of Sociologists in 45 American Universities». *American Sociologist*, 5, Agt., 244-252.
- GLOWER, D. D. (1980): «A Rational Method for Ranking Engineering Programs». *Engineering Education*, 70, 788-794.
- GOBERNA, M. A., y PASTOR, J. T. (1988): «Sobre el rendimiento investigador de la universidad española». *Revista de Educación*, 286, 281-303.
- GOEDEGEBUURE, L. C. J., y MAASSEN, P. A. M. (1990): «Caught in the Middle? The (im)possibilities of fundamental system change». Ponencia presentada en el *12th Annual Forum of the European Association for Institutional Research*, Lyon, Septiembre.
- GOEDEGEBUURE, L. C. J.; MAASSEN, P. A. M., y WESTERHEIJDEN, D. F. (1990): «Quality Assessment in Higher Education». En Goedegebuure, L. C. J.; Maassen, P. A. M., y Westerheijden, D. F. (eds.): *Peer Review and Performance Indicators*. Utrecht, NL, Uitgeverij Lemma.
- GOURMAN, J. (1982): *The Gourman Report: A Rating of Undergraduate Programs in American and International Universities*. Los Angeles, The National Education Standards.
- (1983): *The Gourman Report: A Rating of Graduate and Professional Programs in American and International Universities*. Los Angeles, The National Education Standards.
- GRAVES, P. E.; MARCHAND, J. R., y THOMPSON, R. (1982): «Economics Departmental Rankings: Research Incentives, Constraints, and Efficiency». *American Economic Review*, 72, 5, 1131-1141.
- (1984): «Economics Departmental Rankings: Reply and Errata». *American Economic Review*, 74, 4, 834-836.
- GUBA, E. G., y CLARK, D. L. (1978): «Levels of R&D Productivity in Schools of Educations». *Change*, May, 3-9.
- HAGSTROM, W. O. (1971): «Inputs, Outputs, and Prestige of University Science Departments». *Sociology of Education*, 44, 375-397.
- HALL, A. E. (1984): «Starting at the Beginning: The Baccalaureate Origins of Doctorate Recipients, 1920-1980». *Change*, April, 40-43.

- HARCLEROAD, F. F., y DICKEY, F. G. (1975): *Educational Auditing and Voluntary Institutional Accrediting*. Washington, American Association for Higher Education. ED 102919.
- HARRIS, J. (1978): «Critical Characteristics of an Accreditable Institution, Basic Purposes of Accreditation, and Nontraditional Forms of Most Concern». En: *Research Report 2* of the Project to Develop Evaluative Criteria and Procedures for the Accreditation of Nontraditional Education. Washington, Council of Postsecondary Accreditation.
- HART, S. Y. (1988): «Designing a systematic program review process». En Kells, H. R., y Van Vught, F. A. (eds.): *Self-regulation, self-study and program review in higher education*. Culemborg, NL, Lemma.
- HARTNETT, R. T.; CLARK, M. J., y BAIRD, L. L. (1978): «Reputational Ratings of Doctoral Programs». *Science*, 199, 1310-1314.
- HAVEMAN, R. H., y WOLFE, B. L. (1984): «Schooling and Economic Well-Being: The Role of Nonmarket Effects». *Journal of Human Resources*, 19, 3, 377-407.
- HENSLEY, O. D. (1980): «How valid are scholars departmental rating models for assessing the quality of graduate programs?». *Research in Higher Education*, 12, 3, 255-269.
- HEVERON, E. D. (1987): «Boosting Academic Reputations: A Study of University Departments». *The Review of Higher Education*, 11, 2, 177-197.
- HIRCH, B.; AUSTIN, R.; BROOKS, J., y MOORE, B. (1984): «Economics Departmental Ranking: Comment». *American Economic Review*, 74, 4, 822-826.
- HOGAN, T. D. (1981): «Faculty research activity and the quality of graduate training». *Journal of Human Resources*, 16, 3, 400-415.
- (1984): «Economics Departmental Ranking: Comment». *American Economic Review*, 74, 4, 827-833.
- HOLDAWAY, E. A. (1988): «Institutional self-study and the improvement of quality». En Kells, H. R., y Van Vught, F. A. (eds.): *Self-regulation, self-study and program review in higher education*. Culemborg, NL, Lemma.
- HÖLTTÄ, S. (1990): «The Relationship between the State and Universities in Finland-Evaluation of Performance». Ponencia presentada en el *12th Annual Forum of the European Association for Institutional Research*, Lyon, Septiembre.
- HOPKINS, D. S. P., y MASSY, W. F. (1981): *Planning Models for Colleges and Universities*. Stanford, CA, Stanford UP
- HÜFNER, K. (1987): «Differentiation and Competition in Higher Education: recent trends in the Federal Republic of Germany». *European Journal of Education*, 22, 2, 133-143.
- HÜFNER, K., y RAU, E. (1987): «Measuring performance in higher education: problems and perspectives». *Higher Education in Europe*, 12, 4, 5-13.
- HUGHES, R. M. (1925): *A Study of the Graduate School in America*. Oxford, OH, Miami University Press.
- (1934): «Report of the Committee on Graduate Instruction». *Educational Record*, 15, 192-234.
- IN'T VELD, R. (1990): «The dynamics of evaluation: threats and opportunities». En el Report of the IMHE Study Group on Evaluation in Higher Education, presentado a la *10th General Conference of Member Institutions* del Programme on Institutional Management in Higher Education, París, OCDE, septiembre.

- IRAM, Y. (1987): «Quality and Control in Higher Education in Israel». *European Journal of Education*, 22, 2, 145-159.
- JACOBSON, H. K. (1978): «Framework for evaluation: Indicators of effort, performance, effects». En Jacobson, H. K. (ed.): *Evaluating Advancement Programs*, New Directions for Institutional Advancement 1, San Francisco, Jossey-Bass.
- JAMIESON, D. M. (1988): «Self-Study and its role in strategic planning». En Kells, H. R., y Van Vught, F. A. (eds.): *Self-regulation, self-study and program review in higher education*, Culemborg, NL, Lemma.
- JAPPINEN, A. (1987): «The current situation regarding the development and use of performance indicators in Finland». *International Journal of Management in Higher Education*, 11, 2, 163-170.
- JARRATT, A. (1985): *Efficiency Studies in Universities*. Londres, Committee of Vice Chancellors and Principals.
- JOHNES, J., y TAYLOR, J. (1987): «Degree quality: an investigation into differences between UK universities». *Higher Education*, 16, 581-602.
- JOHNES, G.; TAYLOR, J., y FERGUSON, G. (1987): «The employability of new graduates: a study of differences between UK universities». *Applied Economics*, 19, 695-710.
- JOHNSON, F. C., y CHRISTAL, M. E. (1985): *Preparing for Self-Study*. The AIR Professional File, 22. Tallahassee, FL, The Association for Institutional Research, ED 263861.
- JOHNSON, F. C. (1988): «Departamental self-study and its effect on autonomy». En Kells, H. R., y Van Vught, F. A. (eds.): *Self-regulation, self-study and program review in higher education*. Culemborg, NL, Lemma.
- JOHNSON, R. R. (1978): «Leadership Among American Colleges». *Change*, 10, nov., 50-51.
- JOHNSTONE, J. N., y AGUSTIAR (1983): «An analysis of the perceptions teaching staff hold towards factors useful for evaluating an institution of higher education». *Higher Education*, 12, 215-229.
- JONES, L. V.; LINDZEY, G., y COGGESHALL, P. E. (eds.) (1982): *An Assessment of Research-Doctorate Programs in the United States*. (5 vol.: *Biological Science, Humanities, Social and Behavioral Science, Mathematical and Physical Science, Engineering*). Washington, National Academic Press.
- JORNET, J.; VILLANUEVA, P.; SUAREZ, J., y ALFARO, I. (1989): «Proyecto de implantación de un sistema de evaluación del profesorado en la Universitat de Valencia». En: *Consideraciones Metodológicas sobre la Evaluación y Mejora de la Docencia Universitaria*. Informes de Investigación Evaluativa núm. 1. Valencia: Servei de Formació Permanent, Universitat de Valencia.
- KANE, J. S., y LAWLER, E. E. (1978): «Methods of Peer Assessment». *Psychological Bulletin*, 85, 3, 555-586.
- KELLS, H. R. (1981): «Some Theoretical and Practical Suggestions for Institutional Management». En Miller, R. I. (ed.): *Institutional Assessment for Self-improvement*. New Directions for Institutional Research 29. San Francisco, Jossey-Bass.
- (1986): «The second irony: the system of institutional evaluation of higher education in the United States». *International Journal of Institutional Management in Higher Education*, 10, 2, 140-149.

- (1988a): «Perspectives on university self-assessment for Western Europe». En Kells, H. R., y Van Vught, F. A. (eds.): *Self-regulation, self-study and program review in higher education*. Culemborg, NL, Lemma.
- (1988b): *Self-study Processes. A Guide for Postsecondary and Similar Service-Oriented Institutions and Programs*. 3.^a edición. Nueva York, American Council of Education.
- (1989): «The future of self-regulation in Dutch higher education». Publicación núm. 65 de las series *Higher Education Policy Studies*, Enchede, NL, CHEPS.
- (1990): «Quality Assessment in European Higher Education. A commentary». En Goedegebuure, L. C. J.; Maassen, P. A. M., y Westerheijden, D. F. (eds.): *Peer Review and Performance Indicators*. Utrecht, NL, Uitgeverij, Lemma.
- KELLS, H. R., y VAN VUGHT, F. A. (1988): «Theoretical and Practical Aspects of a Self-Regulation and Quality Control System for Dutch Higher Education». *Tijdschrift voor Hoger Onderwijs*, 6, 1, 15-20.
- KENISTON, H. (1959): *Graduate Study and Research in the Arts and Sciences at the University of Pennsylvania*. Philadelphia, University of Pennsylvania Press.
- KERR, C. (1990): «The American mixture of higher education in perspective: four dimensions». *Higher Education*, 19, 1-19.
- KING, S., y WOLFLE, L. M. (1987): «A latent-variable causal model of faculty reputational ratings». *Research in Higher Education*, 27, 2, 99-106.
- KNUDSEN, D. D., y VAUGHEN, T. R. (1969): «Quality in Graduate Education: A Reevaluation of the Rankings of Sociology Departments in the Cartter Report». *American Sociologist*, 4, Feb., 12-19.
- KOGAN, M. (1986): «The evaluation of higher education. An introductory review». *International Journal of Institutional Management in Higher Education*, 10, 2, 125-139.
- KUH, G. D. (1981): *Indices of Quality in the Undergraduate Experience*. Washington, American Association for Higher Education.
- LAMO DE ESPINOSA, E. (1989): «Evaluación de la calidad en la enseñanza». En Ministerio de Educación y Ciencia: *Hacia una clasificación de las universidades según criterios de calidad*. Madrid, Ministerio de Educación y Ciencia y Fundación Universidad-Empresa.
- LAWRENCE, J. K., y GREEN, K. C. (1980): *A Question of Quality: The Higher Education Rating Game*. Washington, American Association for Higher Education.
- LAWRENCE, J. K., y SOLMON, L. C. (1981): «Reputational Ratings and Undergraduate Education Assessment». En Miller, R. I. (ed.): *Institutional Assessment for Self-improvement*. New Directions for Institutional Research, 29, San Francisco, Jossey-Bass.
- LESLIE, L. L., y BRINKMAN, P. T. (1988): *The Economic Value of Higher Education*. Nueva York, American Council on Education, MacMillan Pub, Co.
- LEVIN, H. M. (1989): «Raising productivity in Higher Education». Papers of *The Pew Higher Education Research Program*, University of Pennsylvania.
- LEWIS, L. (1968): «On Subjective and Objective Rankings of Sociology Departments». *American Sociologist*, 3, May, 129-131.
- (1984): «Graduate Education: Evaluating the Evaluators. A Review Essay». *The Sociological Quarterly*, 25, 125-134.

- LIEBERMAN, M. (1987): «Peer Review and Faculty Self-Government: A dissenting View». En Shere, W., y Duhamel, R. (eds.): *Academic Futures. Prospects for Post-Secondary Education*. Toronto, The Ontario Institute for Studies in Education.
- LINDSAY, A. W. (1981): «Assessing institutional performance in higher education: a managerial perspective». *Higher Education*, 10, 687-706.
- (1982): «Institutional Performance in Higher Education: The Efficiency Dimension». *Review of Educational Research*, 52, 2, 175-199.
- LINDSAY, A. W., y BAYLEY, M. (1980): «A convex polytope technique for analyzing the performance of universities». *Socio-Economic Planning Science*, 14, 33-44.
- LIU, H. C. (1978): «Faculty Citation and Quality of Graduate Engineering Departments». *Engineering Education*, 68, 7, 739-741.
- LOCKWOOD, G. (1986): «Efficiency studies in the British universities». *International Journal of Institutional Management in Higher Education*, 10, 1, 5-13.
- LYSONS, A., y RYDER, P. (1988): «An empirical test of Cameron's dimensions of effectiveness: Implication for Australian tertiary institutions». *Higher Education*, 17, 323-332.
- MAASSEN, P. A. M. (1987): «Quality Control in Dutch Higher Education: internal versus external evaluation». *European Journal of Education*, 22, 2, 161-170.
- (1988): «The Relationship between Institutional Research and Quality Control/Self-Assessment in European Higher Education». Publicación núm. 73 de las series *Higher Education Policy Studies*. Enchede, NL, CHEPS.
- (1989): «Quality Assessment in Higher Education. The Dutch experience». En McVicar, M. (ed.): *Performance indicators and Quality Control in Higher Education*. Portsmouth, UK, Portsmouth Polytechnic.
- MAASSEN, P. A. M., y VAN VUGHT, F. A. (1988): «An Intriguing Janus-head: the two faces of the new governmental strategy for higher education in the Netherlands». *European Journal of Education*, 23, 1/2, 65-76.
- «Strategic Planning». Publicación núm. 91 de las series *Higher Education Policy Studies*. Enchede, NL, CHEPS.
- MAGOUN, H. W. (1966): «The Cartter Report on Quality in Graduate Education». *Journal of Higher Education*, 37, 481-492.
- MARGULIES, R. Z., y BLAU, P. M. (1973): «America's Leading Professional Schools». *Change*, 5, nov., 21-27.
- McCLAIN, C. J., y KRUEGER, D. W. (1985): «Using Outcomes Assessment: A Case Study in Institutional Change». En Ewell, P. T. (ed.): *Assessing Educational Outcomes*. New Directions for Institutional Research, 47, San Francisco, Jossey-Bass.
- McCLAIN, C. J.; KRUEGER, D. W., y TAYLOR, T. (1986): «Northeast Missouri State University's Value Added Assessment Program: A Model for Educational Accountability». *International Journal of Institutional Management in Higher Education*, 10, 3, 252-261.
- McVICAR, M. (1990): «Quality Assessment in British Higher Education». En Goedegebuure, L. C. J.; Maassen, P. A. M., y Westerheijden, D. F. (eds.): *Peer Review and Performance Indicators*. Utrecht, NL, Uitgeverij Lemma.
- MERTENS, P. (1979): «Comparative indicators for German universities». *International Journal of Institutional Management in Higher Education*, 3, 1, 155-168.

- MILLER, R. I. (1979): *The Assessment of College Performance*. San Francisco, Jossey-Bass.
- (1980): «Appraising Institutional Performance». En Jedamus, P.; Peterson, M. W., et al. (eds.): *Improving Academic Management*. San Francisco, Jossey-Bass.
- (1986): «Evaluating institutional quality: some ways and some problems». *International Journal of Institutional Management in Higher Education*, 10, 3, 241-251.
- MILLARD, R. M. (1983): «The Accrediting Association: Ensuring the Quality of Programs and Institutions». *Change*, 15, 4, 32-36.
- MONTERO MARTIN, S. (1989): «Valoración según proyectos de investigación». En Ministerio de Educación y Ciencia: *Hacia una clasificación de las universidades según criterios de calidad*. Madrid, Ministerio de Educación y Ciencia y Fundación Universidad-Empresa.
- MOLTO, T., y OROVAL, E. (1982): «La enseñanza superior en España: una comparación entre objetivos y resultados». *Revista de Educación*, 270, 250-259.
- MOODIE, G. C. (1988): «The Debates about Higher Education Quality in Britain and USA». *Studies in Higher Education*, 13, 1, 5-13.
- MOORE, W. J. (1973): «The relative quality of graduate programs in economics, 1958-1972: Who published and who perished». *Western Economic Journal*, 11, 1, 1-23.
- MORA RUIZ, J. G. (1987): «Un modelo de análisis de la demanda de educación superior». Tesis doctoral, Valencia, Department d'Economia Aplicada. Universitat de Valencia.
- (1988): «El gasto público en la enseñanza universitaria española». *Nuestro Tiempo*, 408, 46-55.
- (1989): *La demanda de educación superior*. Madrid, Consejo de Universidades.
- MORGAN, D. R., y FITZGERALD, M. R. (1977): «Recognition and productivity among American Political Science departments». *Western Political Quarterly*, 30, 3, 342-350.
- MORGAN, D. R.; KEARNEY, R. C., y REGENS, J. L. (1976): «Assessing quality among graduate institutions of higher education in the United States». *Social Science Quarterly*, 57, 671-675.
- MUFFO, J. A.; MEAD, S. V., y BAYER, A. E. (1987): «Using faculty publication rates for comparing "peer" institutions». *Research in Higher Education*, 27, 2, 163-174.
- NATIONAL SCIENCE BOARD (1969): *Graduate Education: Parameters for Public Policy*. Washington, The National Science Foundation.
- NEAVE, G. (1986): «The All-Seeing Eye of the Prince in Western Europe». En Moodie, G. C. (ed.): *Standards and Criteria in Higher Education*. Guilford, Society for Research into Higher Education.
- (1988): «On the Cultivation of Quality, Efficiency and Enterprise: an overview of recent trends in higher education in Western Europe, 1986-1988». *European Journal of Education*, 23, 1/2, 7-23.
- NICHOLS, J. O. (1989): *Institutional Effectiveness and Outcomes Assessment Implementation on Campus: A Practitioner's Handbook*. Nueva York: Agathon Press.
- O'BRIEN, G. M. St. L., y WATERS, G. L. (1983): «A Thematic Approach to the Institutional Self-Study: A Model for University Evaluation and Planning». *The North Central Association Quarterly*, 58, Summer, 13-18.

- OROMANER, M. J. (1970): «A Note on Analytical Properties and Prestige of Sociology Departments». *American Sociologist*, 5, Agst, 240-244.
- ORR, D. (1984): «An Economist Looks at College Rankings». *Change*, Mar., 45-49.
- OS, W. V.; DRENTH, P. J. D., y BERNAERT, G. F. (1987): «AMOS: an evaluation model for institutions of higher education». *European Journal of Education*, 22, 2, 171-181.
- PAARDEKOOPER, C. M. M., y SPEE, A. A. J. (1990): «A Government Perspective on Quality Assessment in Dutch Higher Education». En Goedegebuure, L. C. J.; Maassen, P. A. M., y Westerheijden, D. F. (eds.): *Peer Review and Performance Indicators*. Utrecht, NL, Uitgeverij Lemma.
- PEDRO, F. (1988): «Higher Education in Spain: setting the conditions for an Evaluative State». *European Journal of Education*, 23, 1/2, 125-139.
- PETROWSKI, W. R.; BROWN, E. L., y DUFFY, J. (1973): «National Universities and the ACE Ratings». *Journal of Higher Education*, 44, 7, 495-513.
- PIRSIG, R. M. (1974): *Zen and the art of motorcycle maintenance*. New York, William Morrow & Co.
- POLLITT, C. (1987): «The Politics of Performance Assessment: lessons for higher education?». *Studies in Higher Education*, 12, 1, 87-98.
- PREMFORS, R. (1986): «Evaluating basic units: seven fundamental questions». *International Journal of Institutional Management in Higher Education*, 10, 2, 169-174.
- PRINS, J. (1990): «Students are judging teachers». Ponencia presentada en el *12th Annual Forum of The European Association for Institutional Research*, Lyon, septiembre.
- PULIDO, A. (1989): «Conclusiones del seminario». En Ministerio de Educación y Ciencia: *Hacia una clasificación de las universidades según criterios de calidad*. Madrid, Ministerio de Educación y Ciencia y Fundación Universidad-Empresa.
- ROBEY, J. S. (1979): «Political Science Departments: Reputations Versus Productivity». *PS*, 12, 2, 202-209.
- ROBBINS REPORT (1963): *On Higer Education*. Report of the Committee appointed by the Prime Minister under the Chairmanship of Lord Robbins, London, HMSO.
- RODRIGUEZ, S. (1989): «La Evaluación del Profesorado Universitario. Análisis de la opinión del profesorado de la Universidad de Barcelona». En: *Consideraciones Metodológicas sobre la Evaluación y Mejora de la Docencia Universitaria*. Informes de Investigación Evaluativa, núm. 1, Valencia, Servei de Formació Permanent, Universitat de Valencia.
- ROE, E., y McDONALD, R. (1984): *Informed professional judgment: A guide to evaluation in post-secondary education*. St. Lucía, Australia, University of Queensland Press.
- ROE, E.; McDONALD, R., y MOSES, I. (1986): *Reviewing Academic Performance*. St. Lucía, Australia, University of Queensland Press.
- ROGERS, B. H., y GENTEMANN, K. M. (1989): «The value of institutional research in the assessment of institutional effectiveness». *Research in Higher Education*, 30, 3, 345-355.
- ROMNEY, L. C. (1978): *Measures of institutional goal achievement*. Boulder, Co, National Center for Higher Education Management System.
- ROMNEY, L. C.; BOGEN, G., y MICEK, S. S. (1979): «Assessing institutional performance: the importance of being careful». *International Journal of Institutional Management in Higher Education*, 3, 1, 79-89.

- ROOSE, K. D., y ANDERSEN, C. J. (1970): *A Rating of Graduate Programs*. Washington, American Council on Education.
- RUDD, E. (1988): «The evaluation of the Quality of Research». *Studies in Higher Education*, 13, 1, 45-57.
- RULLMANN, P. M. M. (1990): «A Non-university Perspective on Quality Assessment in the Netherlands». En Goedegebuure, L. C. J.; Maassen, P. A. M., y Westerheijden, D. F. (eds.): *Peer Review and Performance Indicators*. Utrecht, NL, Uitgeverij Lemma.
- SALCEDO, H. (1988): «Quality indicators in Venezuelan higher education». *Studies in Educational Evaluation*, 14, 25-35.
- SAUNIER, M. E. (1985): «Objective measures as predictors of reputational ratings». *Research in Higher Education*, 23, 3, 227-244.
- SCOTT, P. (1986): The Jarratt Report: a commentary. *International Journal of Institutional Management in Higher Education*, 10, 1, 15-21.
- SCOTT, R. A. (1983): «The State Commission: Who Controls the Process?». *Change*, 15, 4, 32-37.
- SCULLY, M. G. (1979): «The Well-Known Universities Lead in Rating of Faculties Reputations». *Chronicle of Higher Education*, 19, 17, January 15, 6-7.
- SHATTOCK, M. L. (1990): «The evaluation of the universities contribution to society». En el Report of the IMHE Study Group on Evaluation in Higher Education, presentado a la *10th General Conference of Member Institutions* del Programme on Institutional Management in Higher Education, Paris, OCDE, septiembre.
- SILVER, H., y SILVER, P. (1986): «The Escaping Answer». En Moodie, G. C. (ed.): *Standards and Criteria in Higher Education*. Guilford, Society for Research into Higher Education.
- SIZER, J. (1979a): «Assessing institutional performance: An overview». *International Journal of Institutional Management in Higher Education*, 3, 1, 49-77.
- (1979b): «Indicators in times of financial stringency, contraction and changing needs». En Billing, D. (ed.): *Indicators of performance*. Guildford, UK, Society for Research into Higher Education.
- (1981): «European Perspectives Suggest Other Criteria». En Miller, R. I. (ed.): *Institutional Assessment for Self-improvement*. New Directions for Institutional Research, 29, San Francisco, Jossey-Bass.
- (1982): «Assessing institutional performance and progress». En Wagner, L. (ed.): *Agenda for institutional change in higher education*. Guildford, UK, Society for Research in Higher Education.
- (1990): «Funding Councils and Performance Indicators in Quality Assessment in the United Kingdom». En Goedegebuure, L. C. J.; Maassen, P. A. M., y Westerheijden, D. F. (eds.): *Peer Review and Performance Indicators*. Utrecht, NL, Uitgeverij Lemma.
- SLOVACECK, S. P. (1988): «Strategic planning and self-study». En Kells, H. R., y Van Vught, F. A. (eds.): *Self-regulation, self-study and program review in higher education*. Culemborg, NL, Lemma.
- SMITH, R., y FIEDLER, F. E. (1971): «The measurement of scholarly work: a critical review of the literature». *Educational Record*, 52, 3, 225-232.

- SOLMON, L. C. (1981): «A Multidimensional Approach to Quality». En Stauffer, T. M. (ed.): *Quality-Higher Education's Principal Challenge*. Washington, American Council on Education.
- SOLMON, L. C., y ASTIN, A. W. (1981): «Departments Without Distinguished Graduate Programs». *Change*, Sept., 23-28.
- SOMIT, A., y TANENHAUS, J. (1964): *American Political Science: A Profile of a Discipline*. Nueva York: Atherton Press.
- SPEE, A. A. J. (1990): «What is new in quality assessment in The Netherlands». Ponencia presentada en el *12th Annual Forum of the European Association for Institutional Research*, Lyon, septiembre.
- STAROPOLI, A. (1986): «Le Comité National d'Evaluation. An innovation in French higher education». *International Journal of Institutional Management in Higher Education*, 10, 2, 157-161.
- (1987): «The Comité National d'Evaluation: preliminary results of a French experiment». *European Journal of Education*, 22, 2, 123-131.
- (1989): «La experiencia francesa». En Ministerio de Educación y Ciencia: *Hacia una clasificación de las universidades según criterios de calidad*. Madrid, Ministerio de Educación y Ciencia y Fundación Universidad-Empresa.
- STOLEN, J. D., y GNUSCHKE, J. E. (1977): «Reflections on alternative rating systems for university economics departments». *Economic Inquiry*, 15, 2, 277-282.
- TAN, D. L. (1986): «The Assessment of Quality in Higher Education: A Critical Review of the Literature and Research». *Research in Higher Education*, 24, 3, 223-265.
- (1990): «Is there a better way to measure quality of programs? Ponencia presentada en *12th Annual Forum of the European Association for Institutional Research*. Lyon, septiembre.
- TAYLOR, J., y JONES, J. (1989): «An Evaluation of Performance Indicators based upon the First Destination of University Graduates». *Studies in Higher Education*, 14, 2, 201-217.
- TEJEDOR, F. J.; JATO, E., y MINGUEZ, C. (1989): «Evaluación del Profesorado Universitario por los alumnos en la Universidad de Santiago». En: *Consideraciones Metodológicas sobre la Evaluación y Mejora de la Docencia Universitaria*. Informes de Investigación Evaluativa núm. 1. Valencia, Servei de Formació Permanent, Universitat de Valencia.
- TRITSCHLER, D. (1981): «Strategies for Assessing Performance at Your Own Institution». En Miller, R. I. (ed.): *Institutional Assessment for Self-improvement*. New Directions for Institutional Research, 29. San Francisco: Jossey-Bass.
- TROUTT, W. E. (1981): «Relationships between Regional Accrediting Standards and Educational Quality». En Miller, R. I. (ed.): *Institutional Assessment for Self-improvement*. New Directions for Institutional Research, 29, San Francisco: Jossey-Bass.
- TYLER, F. B. (1972): «Knowledgeable respondent: private club or public service?». *American Psychologist*, 27, 3, 191-196.
- UNESCO (1989): *Anuario estadístico*, París, UNESCO.
- VARIOS (1990): «America's best graduate and professional schools». *US News & World Report*, Mar. 19, 46-79.

- VELASCO, C., y VELAZQUEZ, M. P. (1983): «Notas y disfunciones en la distribución del profesorado universitario en España: una aproximación». *Revista Española de Investigaciones Sociológicas*, 24, 143-161.
- VINSONHALER, J. C., y VISONHALER, J. F. (1990): «A Problem Oriented Data Base for Monitoring and Improving Academic Quality». Ponencia presentada en el *12th Annual Forum of the European Association for Institutional Research*, Lyon, septiembre.
- VROEIENSTIJN, T. I. (1990): «Self-regulation Based on Self-assessment and Peer Review. Experience in Dutch universities with external quality assessment». Ponencia presentada en el *12th Annual Forum of the European Association for Institutional Research*, Lyon, septiembre.
- VROEIENSTIJN, T. I., y ACHERMAN, H. (1990): «Control Oriented versus Improvement Oriented Quality Assessment». En Goedegebuure, L. C. J.; Maassen, P. A. M., y Westerheijden, D. F. (eds.): *Peer Review and Performance Indicators*. Utrecht, NL, Uitgeverij Lemma.
- VUGHT, F. A. V. (1988): «A new autonomy in European higher education». En Kells, H. R., y Van Vught, F. A. (eds.): *Self-regulation, self-study and program review in higher education*. Culemborg, NL, Lemma.
- WEBSTER, D. S. (1981): «Advantages and disadvantages of Methods of Assessing Quality». *Change*, Oct., 20-24.
- (1983): «America's Highest Ranked Graduate Schools, 1925-1982». *Change*, May/Jun, 14-24.
- (1985): «Institutional effectiveness using scholarly peer assessment as major criteria». *The Review of Higher Education*, 9, 1, 67-82.
- (1986): *Academic quality rankings of American colleges and universities*. Springfield, Illinois, Charles C. Thomas.
- WEBSTER, D. S., y CONRAD, C. F. (1986): «Using Faculty Research Performance for Academic Quality Rankings». En Creswell, J. W. (ed.): *Measuring Faculty Research Performance*. New Directions for Institutional Research, 50, San Francisco: Jossey-Bass.
- WEERT, E. (1990): «A macro-analysis of quality assessment in higher education». *Higher Education*, 19, 57-72.
- WESTERHEIJDEN, D. F. (1990a): «Peers, Performance and Power. Quality Assessment in the Netherlands». En Goedegebuure, L. C. J.; Maassen, P. A. M., y Westerheijden, D. F. (eds.): *Peer Review and Performance Indicators*. Utrecht, NL, Uitgeverij Lemma.
- (1990b): «University Evaluation and its Political Context. Quality Management and Decision Making in Higher Education». Ponencia presentada en el *Coloquio Internacional: La Pedagogia Universitaria*, Barcelona, octubre.
- WILLIAMS, G. (1986): «The Missing Bottom Line». En Moodie, G. C. (ed.): *Standards and Criteria in Higher Education*. Guilford, Society for Research into Higher Education.
- WISE, S. L.; HENGSTLER, D. D., y BRASKAMP, L. A. (1981): «Alumni Ratings as an Indicator of Departmental Quality». *Journal of Educational Psychology*, 73, 1, 71-77.
- WRIGHT, P. W. G. (1989): «Who defines quality in higher education? Reflections on the role of professional power in determining conceptions of quality in English higher education». *Higher Education*, 18, 149-165.

- YORKE, D. M. (1986): «Indicadores del logro institucional. Algunas consideraciones teóricas y empíricas». En Latiesa, M. (ed.): *Educación y rendimiento académico en la universidad*. Madrid, CIDE.
- YOUNG, D. (1990): «The Academic Audit Unit». En Goedegebuure, L. C. J.; Maassen, P. A. M., y Westerheijden, D. F. (eds.): *Peer Review and Performance Indicators*. Utrecht, NL, Uitgeverij Lemma.
- YOUNG, D. L. (1988): «Program review in Florida: a universal model for government and universities». En Kells, H. R., y Van Vught, F. A. (eds.): *Self-regulation, self-study and program review in higher education*. Culemborg, NL: Lemma.
- YOUNG, D. L.; BLACKBURN, R. T., y CAMERON, K. S. (1989): «Leadership, Student Effort, and Departmental Program Quality: An Exploration of Quality across levels of Analysis». *The Review of Higher Education*, 12, 3, 265-277.

BIBLIOGRAFÍAS

- HALSTEAD, K. (ed.) (1988): *Higher Education. Bibliography Yearbook 1987*. Washington, Research Associates of Washington.
- HÜFNER, K.; HUMMEL, T., y RAU, E. (1986): *Efficiency in Higher Education: An Annotated Bibliography*. Berlin, Freie Universität Berlin.

