

LA EVALUACION DE LAS INSTITUCIONES UNIVERSITARIAS



EDICION DE:
Mario de Miguel
José-Ginés Mora
Sebastián Rodríguez



Consejo de Universidades
Secretaría General

LA EVALUACION DE LAS INSTITUCIONES UNIVERSITARIAS

**Una selección de textos preparada
por:**

MARIO DE MIGUEL
JOSE-GINES MORA
SEBASTIAN RODRIGUEZ

CONSEJO DE UNIVERSIDADES
SECRETARIA GENERAL
1991

© CONSEJO DE UNIVERSIDADES
Secretaría General
NIPO: 176-91-022-7
ISBN: 84-369-2028-7
Depósito Legal: M-41766-1991
Imprime: REGLETA
Juan de Olías, 11-13
28020 Madrid

INDICE

	<u>Páginas</u>
Agradecimientos	7
Acerca de los autores	9
Introducción	11
NEAVE, Guy: Sobre el desarrollo de la calidad, la eficiencia y la iniciativa: un resumen de las tendencias recientes de la enseñanza superior en Europa Occidental, 1986-1988	17
RODRIGUEZ, Sebastián: Calidad universitaria: un enfoque institucional y multidimensional	39
MORA RUIZ, José-Ginés: La evaluación institucional: una perspectiva general	73
VROEIJENSTIJN, Ton I. y ACHERMAN, Hans, Evaluación de la calidad basada en el control frente a la evaluación de la calidad basada en la mejora	93
WESTERHEIJDEN, Don F.: La evaluación de la universidad y su contexto político: gestión de la calidad y toma de decisiones en la educación superior	121
VAN VUGHT, Frans: ¿Cómo evaluar la calidad de la educación superior en Europa Occidental?	151
TAN, David L.: Evaluación de la calidad de la enseñanza superior: una revisión de la bibliografía y la investigación	165
SPEE, Arnold A.J.: Sobre la comunicación de la calidad para la mejora de la investigación y la educación superior en los Países Bajos	213
Mc DONALD, Rod y ROE, Ernest: Evaluación de los departamentos	239
HOLDAWAY, Edward A.: Autoestudio institucional y mejora de la calidad	265

SIZER, John: Comites de financiación e indicadores de rendimiento en la evaluación de la calidad en el Reino Unido	285
DOCHY, Filip J.R.C., SEGERS, Mien S.R. y WIJNEN, Wynand H.F.W.: Selección de indicadores de rendimiento: una propuesta con resultado de la investigación	317
DE MIGUEL, Mario: Utilización de indicadores en la evaluación de la docencia universitaria	341
CUENIN, Serge: Un examen de las políticas europeas en materia de evaluación de la enseñanza superior	371
FRACKMANN, Edgar: Lecciones que deben aprenderse de una década de discusiones sobre indicadores de rendimiento	399
KELLS, Herbert R.: La insuficiencia de los indicadores de rendimiento para la educación superior: la necesidad de un modelo más global y formativo	423
DE GREVE, J. P. y FRIJDAL; A.: La evaluación de la investigación científica. Un método combinado de análisis de perfil	443

AGRADECIMIENTOS

Esta obra ha sido posible gracias al permiso desinteresado para la traducción y publicación de los trabajos originales que han otorgado:

Las organizaciones:

- Center of Higher Education Policies Studies de la Universidad de Twente en los Países Bajos.
- O C D E, París.
- Higher Education Research and Development Society of Australasia, Australia.

Y los señores:

- S. Cuenin, E.A. Holdaway, G. Neave, A.A.J. Spee, D. Tan, F. van Vught y D. F. Westerheijden.

El agradecimiento hay que hacerlo extensivo a los diferentes traductores que han pasado al castellano los artículos originalmente escritos en inglés, con excepción del texto del Sr. Cuenin traducido del francés.

Una subvención del Consejo de Universidades ha hecho factible la realización de este proyecto.

ACERCA DE LOS AUTORES

- H. ACHERMAN. Director del Departamento de Personal de la Universidad de Amsterdam, Países Bajos.
- S. CUENIN. Investigador del CNRS (Centre National de la Recherche Scientifique), Francia.
- J.P. DE GREVE. Profesor de la Universidad Libre de Bruselas, Bélgica.
- M. DE MIGUEL. Catedrático de Métodos de Investigación y Diagnóstico en Educación de la Universidad de Oviedo, España.
- F.J.R.C. DOCHY. Director de Investigación de la Universidad Abierta de Heerlen y Profesor Asociado de la Universidad de Maastrich, Países Bajos.
- E. FRACKMANN. Miembro del Hochschul-Information System GMBH, Hannover, Alemania.
- A. FRIJDAL. Miembro del Instituto Universitario Europeo, Florencia, Italia.
- E.A. HOLDAWAY. Profesor de Administración Educativa de la Universidad de Alberta, Canadá.
- H.R. KELLS. Profesor del Danbury Park Management Center, Chelmsford, Inglaterra.
- R. McDONALD. Director del Centro de Recursos y Servicios Educativos de la Universidad de Murdoch, Perth, Australia.
- J.G. MORA. Profesor Titular de Economía Aplicada de la Universidad de Valencia, España.
- G. NEAVE. Profesor de Educación en la Universidad de Londres y Director de Investigación de la IAU (International Association of Universities), Inglaterra.
- S. RODRIGUEZ. Director del Gabinete de Evaluación e Innovación Universitaria de la Universidad de Barcelona, España.
- E. ROE. Director del Instituto de Educación Superior de la Universidad de Queensland, Brisbane, Australia.

- M.S.R. SEGERS. Profesor Asociado de la Universidad de Maastrich, Países Bajos.
- J. SIZER. Profesor de Gestión Financiera de la Universidad de Tecnología de Loughborough, Inglaterra.
- A.A.J. SPEE. Director delegado de Política de Becas y Ayuda Financiera a los estudiantes del Ministerio de Educación y Ciencia, Países Bajos.
- D. TAN. Profesor de Educación Superior en la Universidad de Oklahoma, Norman, Estados Unidos.
- T.I. VROEIENSTIJN. Consejero de la Asociación para la Cooperación de las Universidades Holandesas, Países Bajos.
- F. VAN VUGHT. Profesor y Director del CHEPS (Center Higher Education Policy Studies) de la Universidad de Twente, Países Bajos.
- D. F. WESTERHEIJDEN. Investigador del CHEPS (Center Higher Education Policy Studies) de la Universidad de Twente, Países Bajos.
- W.H.F.W. WIJNEN. Profesor de Investigación en Educación Superior de la Universidad de Maastrich, Países Bajos.

INTRODUCCION

Este libro es una recopilación de artículos sobre evaluación institucional. La selección se ha hecho preferentemente entre los textos más actuales que sobre este tema han aparecido en revistas y libros, o que han sido presentados en congresos recientes. Como toda selección es limitada y se han tenido que dejar de lado aportaciones de indudable interés realizadas por diversos autores. Los que han sido seleccionados creemos que representan adecuadamente los distintos puntos de vista sobre la evaluación institucional, sobre las ideas básicas en torno a ella, sobre las dificultades para su puesta en práctica y sobre las experiencias existentes en los países en los que ha sido ya recorrido algún camino en este sentido.

La importancia de la evaluación de las instituciones de educación superior en los momentos actuales resulta obvia. En los países más avanzados la calidad de esta educación se ha convertido en un tema prioritario para las propias instituciones universitarias y para los ministerios de educación en esos países. Especialmente en los países anglosajones y de la Europa noroccidental se están poniendo en marcha un amplio conjunto de ensayos evaluadores con el objetivo de mejorar la calidad de sus instituciones de educación superior. Casi todos los artículos de este libro hacen referencia a esas experiencias.

El interés por la evaluación parece estar también despertando en España. Sería deseable que ese interés estuviera ligado a la preocupación por la calidad del servicio que prestan las universidades, por la calidad en los programas docentes, por la calidad de la investigación llevada a cabo por los equipos de trabajo en los departamentos, y por la calidad de la vida universitaria en todos sus aspectos. Sería lamentable que el interés por la evaluación se redujera exclusivamente al interés por unas pequeñas recompensas económicas individuales. Esperamos que estas lecturas ayuden a adquirir esa perspectiva más amplia.

El objetivo de esta recopilación es facilitar a las personas interesa-

das en este tema, pero que no son especialistas en estas cuestiones (cabe suponer que la mayoría de los miembros de la comunidad universitaria deberían incluirse en esta categoría), una lectura que les permita conocer algunas de las ideas básicas sobre la evaluación institucional que circulan entre los investigadores dedicados específicamente a estos problemas.

Los diecisiete artículos que se han seleccionado pueden ser agrupados a *grosso modo* en cinco apartados: los de carácter general que plantean visiones amplias y conceptuales sobre la evaluación institucional (Neave, Rodríguez, Mora, Vroeiijensstijn y Acherman, Westerheijden), los dedicados a relatar la situación de la evaluación institucional en algunos países (van Vught, Tan, Spee), los dedicados específicamente a los autoestudios y revisiones internas (Holdaway, Mc Donald), los que trata sobre el uso de indicadores de rendimiento (Sizer, Dochy *et al.*, De Miguel, Cuenin, Frackmann, Kells) y finalmente uno que trata sobre la evaluación de la investigación (De Greve y Frijdal).

El primer artículo, escrito por Guy NEAVE, constituye un análisis crítico de las directrices políticas asumidas por los principales países de la Europa Occidental entre los años 1986 y 1988 en relación con la evaluación de la educación superior. A título de síntesis se podría decir que el trabajo del autor se estructura en dos partes. En la primera analiza (lo que, en palabras del autor, se denomina *estado evaluativo*) los nuevos planteamientos de las relaciones entre educación superior y poder gubernamental, que consideran la evaluación institucional como una estrategia que permite coordinar las prioridades establecidas a nivel político con la autonomía de las universidades. En la segunda comenta las medidas tomadas al respecto en algunos países europeos, y plantea la necesidad de reconducir las relaciones entre educación superior y sociedad como única forma de encarar el futuro.

Sebastián RODRIGUEZ introduce el tema de la calidad de la educación universitaria, pasando revista a puntos como: el concepto de educación superior, los aspectos de moda y política de la calidad, así como a la autonomía como presupuesto fundamental de la misma. A continuación el autor defiende la tesis de la relatividad y multidimensionalidad del concepto de calidad en relación a los objetivos y actores del propio sistema y marco universitario, así como a las conceptualizaciones de la calidad, tanto desde la perspectiva de la *reputación de las instituciones*

como desde una perspectiva *objetiva* o del enfoque de la calidad como *desarrollo del potencial* de los agentes de la propia institución.

En su artículo José-Ginés MORA presenta una perspectiva general sobre la evaluación institucional, indicándose tipos de evaluaciones y características de cada una de ellas, con especial referencia a los autoestudios y a los indicadores de rendimiento. Se trata de una *primera aproximación* a los conceptos y rasgos esenciales que caracterizan la evaluación institucional y que puede interesar especialmente a los no especialistas para adquirir una visión general que les permita situarse dentro del tema.

Ton I. VROEIJENSTIJN y Hans ACHERMAN exponen, desde su perspectiva de la Asociación para la Cooperación de las Universidades Holandesas (VSNU), el modelo de evaluación de la calidad de la educación universitaria que ha sido implantado en Holanda y que esta basado en la revisión por colegas (*peer review*). Son analizados los fundamentos y características de los autoestudios como vía de la autorregulación frente al control externo, así como el papel a desempeñar por las propias universidades como organización. Especial interés reviste el análisis crítico del modelo de *indicadores de rendimiento*, así como la síntesis de guía para el autoestudio institucional que se ofrece.

Don F. WESTERHEIJDEN analiza, en primer lugar, los elementos que componen el marco de referencia de una *teoría de la acción política* que le servirá, en una segunda parte, para contextualizar los planteamientos o enfoques de la evaluación de la calidad en las instituciones universitarias en relación a los diferentes agentes y niveles de la organización universitaria, tanto a nivel interno como en sus relaciones con el marco político exterior. En una tercera parte, se exponen las aplicaciones concretas para la práctica real de la evaluación de la calidad, defendiéndose la tesis de que sin la *expectativa de consecuencias reales, los incentivos para organizar la evaluación de la calidad serán inexistentes; con la expectativa de consecuencias reales, la evaluación se convierte en un juego de poder.*

Frans Van VUGHT presenta un modelo de evaluación institucional que sigue las ideas básicas de la evaluación en EE.UU. y Canadá tras pasadas a una situación como la europea caracterizada por la existencia de sistemas de educación superior básicamente públicos. Este modelo adaptado es el que esta siendo puesto en práctica con aparente

éxito por las universidades holandesas. La experiencia holandesa, por sus características y por su relativa similitud con el contexto español puede ser de gran interés. En este mismo sentido son importantes las reflexiones críticas que el autor hace sobre el funcionamiento del modelo.

El artículo de David L. TAN constituye una revisión muy completa de los estudios sobre evaluación de la calidad de la educación superior en EE.UU. Tras una introducción dedicada a poner de relieve las dificultades para operativizar el concepto de calidad, el trabajo se estructura en tres apartados que responden a los tres tipos de metodologías más utilizadas en la investigación sobre el tema: estudios sobre reputación (en la traducción hemos utilizado el termino acreditación), investigaciones basadas en indicadores objetivos y evaluaciones a partir de correlatos. El autor realiza una revisión crítica de la investigación efectuada partiendo de estos tres tipos de metodologías, poniendo de relieve los puntos fuertes y débiles de los principales trabajos, e identificando los asuntos que quedan por resolver.

Arnold A.J. SPEE describe, en primer término, el nuevo marco jurídico de la universidad holandesa por lo que respecta a: planificación, investigación, docencia, admisiones y financiación. En segundo lugar, analiza —desde la posición del Ministerio de Educación y Ciencia holandés— el sistema de evaluación y control de la calidad de la docencia y la investigación de las universidades holandesas, así como su puesta en práctica. Finalmente reflexiona sobre el proceso de comunicación de los resultados de la propia evaluación, aspecto de particular importancia ya que en dicha comunicación se determina parte del impacto del control de calidad.

El trabajo de Rod McDONALD y Ernest ROE constituye una guía muy interesante y práctica sobre como se debe proceder en la evaluación de un departamento o de cualquier otra unidad universitaria. Tras una introducción en la que se definen y justifican este tipo de evaluaciones, los autores analizan detalladamente los aspectos que han de tenerse en cuenta para la planificación y realización de una evaluación departamental, especificando las normas prácticas que deben orientar las etapas del proceso. Finalmente añaden una serie de sugerencias y recomendaciones que deben ser tenidos en cuenta con el fin de crear y mantener dentro de las instituciones un clima favorable a este tipo de evaluaciones.

Edward A. HOLDAWAY describe en su artículo la experiencia evaluadora puesta en práctica en la Universidad de Alberta en Canadá. El interés esencial de este artículo está en la descripción detallada del proceso de autoevaluación llevado a cabo en aquella universidad, modelo que, salvando las distancias sociopolíticas que rodean ambos sistemas universitarios, podría servir como ejemplo para un proceso semejante en la universidad española.

John SIZER analiza inicialmente el complejo marco político y económico en el que se desarrolla el modelo de indicadores de rendimiento de las universidades británicas. A continuación ofrece un exhaustivo análisis del proceso de aplicación de los indicadores seleccionados para los distintos aspectos de la educación superior —investigación, docencia y servicios— con un especial énfasis en las repercusiones financieras de tal proceso. Así mismo se exponen las dificultades de desarrollar determinados indicadores.

En su artículo Filip DOCHY, Mien SEGERS y Wynand WIJNEN resumen la investigación que sobre indicadores de rendimiento llevaron a cabo en el marco de las universidades holandesas. Los autores analizan el concepto de *indicador de rendimiento*, así como las condiciones para asegurar la calidad de las instituciones. En una segunda parte se ofrece una síntesis de los indicadores considerados como válidos desde un modelo *input-proceso-output* y referidos tanto a docencia como a investigación y servicios. Finalmente se ofrecen una serie de consideraciones sobre su aplicación y uso.

El trabajo de Mario DE MIGUEL plantea la necesidad de superar los esquemas reduccionistas que habitualmente se utilizan en relación con la docencia universitaria, elaborando un marco conceptual que permita una visión comprensiva de todos los factores que intervienen en la misma a nivel institucional, así como las competencias relativas a las diferentes unidades académicas que intervienen en su gestión. En una segunda parte el autor comenta la utilidad que aportan los *indicadores de rendimiento* como estrategia para la evaluación, así como los requisitos que se deben tener en cuenta en su elaboración y utilización. Finalmente, después de efectuar una descripción detallada de los principales indicadores que se pueden utilizar para evaluar cada uno de los factores de contexto, input, proceso y producto de la función docente, propone un modelo concreto de evaluación en base a indicadores que conjuga una visión global de todos los componentes de la docencia, a la vez

que permite efectuar evaluaciones concretas tomando como referencia cada una de las unidades de gestión (departamentos, centros, profesores, etc.).

Ante las restricciones presupuestarias de la última década, Serge CUENIN sugiere la necesidad de revisar las políticas de intervención gubernamental en la educación superior, postulando que los indicadores de rendimiento constituyen una buena estrategia para canalizar en el futuro las relaciones entre los gobiernos y las universidades. Partiendo de este supuesto, en una segunda parte, define el concepto de *indicadores de rendimiento*, establece diversas clasificaciones de los mismos y efectúa un breve análisis de los más utilizados en materia de enseñanza. La parte final del trabajo se dedica a comentar de forma crítica su utilidad en dos grandes ámbitos de aplicación: para tomar decisiones dentro de los centros y para regular la comunicación entre éstos y otros organismos externos. El trabajo concluye analizando algunas experiencias concretas en universidades europeas.

Edgar FRACKMANN plantea una discusión teórica sobre la utilidad y conveniencia del uso de los indicadores de rendimiento, basándose en la experiencia de su uso en los últimos años en Europa. Aunque con reservas, este autor se muestra partidario de su utilización en el contexto de los países europeos, señalando los criterios que deberían tenerse en cuenta para evitar los problemas que puede conllevar el uso de indicadores de rendimiento.

Herb KELLS presenta en su artículo un enérgico alegato a favor de las evaluaciones basadas en los autoestudios y en las revisiones por colegas, frente a los que hacen hincapié en el uso de indicadores de rendimiento. Las ventajas de un sistema frente a otro son detalladas en su artículo. La utilización de *rankings* entre universidades es también objeto de dura crítica por parte de este autor.

Por último se recoge un artículo escrito por DE GREVE y FRIJDAL en el que se propone un sistema de evaluación de la investigación de los departamentos universitarios en la que se tiene en cuenta, no solo el rendimiento del equipo de investigación según medidas objetivas, sino también las opiniones de comisiones de colegas que han de seguir unos criterios determinados para evaluar el perfil de la investigación.

SOBRE EL DESARROLLO DE LA CALIDAD, LA EFICIENCIA Y LA INICIATIVA: UN RESUMEN DE LAS TENDENCIAS RECIENTES DE LA ENSEÑANZA SUPERIOR EN EUROPA OCCIDENTAL, 1986-88 *

GUY NEAVE

Universidad de Londres. Reino Unido

Introducción

Hace muchos años, antes del concilio Vaticano II, varias iglesias, denominaciones y sectas protestantes se unieron en un espíritu de ecumenismo. Lo que las había dividido en el pasado era ahora fuente de gozo y unidad. Para celebrar el acontecimiento y la nueva teología apareció un «Libro de Oración» de todas las denominaciones. Como un bromista señaló, complació a todo el mundo, ¡pero no a Dios!

Se puede decir lo mismo de las reformas que en los dos últimos años han iluminado el mundo de la educación superior. El desarrollo de la calidad, la eficiencia y la iniciativa —la versión actual de la Nueva Teología Gubernamental— no pueden discutirse. Todos estamos a favor de ellas, igual que estamos a favor de la gastronomía y de una cintura estrecha. Como principios abstractos o como puntos de retórica, no se les puede atacar. Donde surge la controversia es en las consecuencias operativas, legislativas y financieras para las instituciones que componen el sistema de enseñanza superior de una nación. Es en este punto donde el espíritu de acuerdo fraternal empieza a desaparecer y donde la nueva ética social, aunque agrada a todo el mundo, no siempre es placentero para los más afectados por ella.

Ciertamente, en los últimos dos años se han reanudado las protestas de estudiantes contra aspectos específicos de la política gubernamental

* Este artículo es una traducción del original aparecido en la revista *European Journal of Education*, 23, 1/2, 7-23, 1988.

—contra la selectividad en España y Francia— y, en el caso de Francia, con una protesta tan masiva que el Ministro encargado de la educación superior se vio obligado a dimitir. En Bélgica y Portugal tuvieron lugar protestas más esporádicas contra la subida de las tasas de matrícula (las manifestaciones en Francia y España también estaban en contra de esta subida). Aunque las acciones estudiantiles fueron espectaculares, no han servido para detener el cambio, cuyos aspectos principales se han clarificado y precisado más en los dos últimos años.

En estos dos últimos años se ha hecho evidente que las políticas que se desarrollaron como respuestas empíricas y a corto plazo a dificultades financieras al principio de esta década, ahora han asumido un empuje estratégico a largo plazo. Este empuje toma diversas formas, pero esencialmente se basa, de una parte, en una nueva relación entre educación superior y gobierno, y, de otro, entre educación superior y sociedad. En cuanto a la primera se puede interpretar como la aparición del concepto «Evaluative State»¹, y, la otra, como un intento de introducir una forma particular de «Ética competitiva», definida externamente como la principal fuerza motriz del desarrollo institucional y, en consecuencia, del sistema de enseñanza superior.

La aparición del «Evaluative State»

En cierto sentido, la noción de «Evaluative State» no es nuevo desde que los gobiernos asumieron la tarea de garantizar las finanzas de la educación superior, así como de definir el marco legal y administrativo dentro del cual evolucionaron estas instituciones. La evaluación ha sido un aspecto importante de esa responsabilidad, aunque a menudo implícito. Sin embargo, debería distinguirse entre evaluación para mantenimiento del sistema y evaluación para el cambio estratégico. La primera es una función rutinaria, asociada al presupuesto anual. En aquellos países donde la distribución de los fondos es responsabilidad directa de la autoridad central —como en Francia, España, Holanda y Suecia— esa evaluación es responsabilidad del Ministro Nacional de Educación, o, en el caso de Suecia, del Consejo Nacional de Universidades y Facultades.

1. N.T.: El concepto de "Evaluative State" puede ser interpretado como un estado, principio o corriente de evaluación que regula las relaciones entre los gobiernos y las instituciones de educación superior.

En resumen, la asignación de recursos, bien sea de dinero o personal, requiere el análisis de las necesidades del sistema para mantener la calidad de la enseñanza superior que permita cumplir los objetivos establecidos por los gobiernos, bien sea en términos de número de estudiantes, distribuidos en diferentes disciplinas, o de las prioridades de investigación establecidas a nivel nacional. El proceso de evaluación rutinaria no se limita a los Ministerios de Educación, porque aunque el ministro de Educación o el Comité de Ayudas a las Universidades pueda hacer propuestas, a menudo es el Ministerio de Economía o del Tesoro quien dispone. Por ello, los mecanismos del «Evaluative State» en el dominio particular de educación superior no se limita simplemente a los aspectos en los que la administración central tiene una responsabilidad directa.

Otros organismos pueden también realizar evaluaciones desde su propia perspectiva. En Suecia el *Audit Bureau* juega un papel significativo, aunque menos visible, en la evaluación de la eficiencia financiera de la educación superior en su conjunto y de las instituciones individuales. En Francia el *Cour des Comptes* —la Agencia de Auditoría Nacional— cumple una función similar. Sus informes anuales, aunque no disponibles para el público, proporcionan un relato detallado y a veces fascinante de cómo las instituciones individuales han usado el dinero que se les ha confiado. Hay una especie de intersección horizontal entre las diferentes responsabilidades de la administración central, cada una de las cuales tiene su propia perspectiva sobre la evaluación de la educación superior. Existe, además, en Francia, una intersección vertical entre el ministro de Educación y el *Commissariat Général du Plan*. Este último organismo tiene la responsabilidad principal de preparar el «Plan de cuatro años». A grosso modo, el «Plan» une los diferentes sectores de la economía nacional y les otorga prioridades, estima las tasas probables de crecimiento y establece, de este modo, las asignaciones del presupuesto para llevarlo a cabo. Desde 1981, la educación en general y la educación superior en particular han recibido bastante atención, especialmente en el área de la investigación y de postgrado (OCDE, 1986).

El *Commissariat Général du Plan* francés se encuentra en la intersección entre la «evaluación rutinaria» y lo que se ha denominado «evaluación estratégica». Proporciona una perspectiva global a la planificación nacional por sectores. Aunque sus previsiones y estimaciones no son obligatorias, facilita, sin embargo, una serie de puntos o metas a alcanzar en áreas concretas de la economía nacional que pueden ser tenidas en cuenta por los diversos Ministerios a la hora de distribuir los recursos.

La evaluación estratégica, como su nombre indica, tiene el propósito de establecer metas a largo plazo para el sistema de educación superior y estimar, en términos monetarios, los recursos, equipos y personal necesarios para llevarlas a cabo. No todos los países, cuando realizan este tipo de evaluación, establecen una temporalización del tiempo necesario para lograr cada uno de estos objetivos específicos (Wittrock, 1985). La función global de «evaluación estratégica» es evaluar el rendimiento previo en una dimensión particular de la política nacional sobre la que se pretende efectuar cambios. El enfoque de este tipo de evaluación se centra en temas amplios: acceso de estudiantes a la Universidad y la incorporación a la enseñanza superior de grupos hasta ahora excluidos; la distribución espacial de los centros de educación superior, características de las instituciones de educación superior a lo largo de la nación; la estructura, duración y equilibrio entre las diferentes disciplinas; políticas nacionales de personal, etc.

Ejemplos típicos de «evaluación estratégica» en el área de política de acceso sería, en Gran Bretaña, el informe Robbins de 1963, que sentó el principio de que debería haber plazas disponibles en la educación superior para todos los alumnos cualificados que deseaban aprovechar esta oportunidad. La Comisión de 1955 en Suecia tenía un propósito estratégico similar, pero referido a un marco mucho más amplio. Se proponía el desarrollo cuantitativo de la educación superior, la enseñanza y el aprendizaje, la investigación y la administración (Premfors, 1983), aunque se centró especialmente sobre el tema de «las reservas de los talentos». En Francia, la aparición de un nuevo sector de enseñanza superior, paralelo a la Universidad pero relacionado más específicamente con las habilidades exigidas a los mandos intermedios en industria y comercio, fue el resultado de las deliberaciones del así llamado *Comités des 18*. Sus consideraciones dieron lugar a los Institutos Universitarios de Tecnología (Lamoure, 1984). El contenido fundamental del que se ocupó la Comisión Sueca de 1968 fueron los temas relativos a la distribución espacial de los centros de la educación superior, así como las nuevas condiciones de acceso del alumnado y la estructura de los estudios de licenciatura. Su análisis de la realidad reafirmó, a su vez, el principio de que la educación superior se apoya sobre un «modelo comprensivo» unitario. Hay otros ejemplos, por supuesto. Aunque estos ejemplos no se parezcan demasiado, su elección puede ser discutida: tienen algo en común. La evaluación estratégica, así como la «rutinaria», se realizan ambas a priori (van Vught, 1988). Es decir, establecen metas y asignan recursos con la esperanza de que estas metas y recursos influirán

en el proceso de enseñanza, el aprendizaje, las aspiraciones individuales de los sujetos y el establecimiento de metas institucionales, de modo que el *producto* —es decir, el número de estudiantes que tienen acceso a este nivel de enseñanza y sus calificaciones (Neave, 1987— se correspondan con las previsiones iniciales.

Desde este punto de vista, la evaluación estratégica pretende remediar los defectos de un sistema estable revisando los objetivos periódicamente, que es el proceso a seguir en todo sistema que se quiera reformar. Por ello, la evaluación estratégica sólo funciona cuando hay crisis en el orden establecido. Desde el punto de vista de las reformas recomendadas —a las que hemos aludido anteriormente—, este tipo de evaluación constituye el comienzo de todo el proceso para implantar dichas reformas. En cambio, la aparición del «Evaluative State» está unido a dos grandes cambios en el momento, propósito y lugar de la evaluación tanto en el proceso de creación de un programa sobre la enseñanza superior como de la «política de adhesión» a ese programa. El primero supone la unión de la «evaluación rutinaria» y la «evaluación estratégica». El segundo conlleva un claro cambio hacia una evaluación «a posteriori».

Una evaluación «a posteriori» pretende averiguar hasta qué punto se han cumplido las metas globales a través de la evaluación del «producto». No pretende, como la evaluación «a priori», estimar el producto que puede alcanzarse uniendo los objetivos y los recursos. Supone que habrá mayores probabilidades de alcanzar los objetivos si la asignación de recursos se hace de acuerdo con el grado de eficacia que desarrolla una institución. Una evaluación *a posteriori* trabaja controlando el *producto*, no controlando el *proceso*. Este cambio de control del proceso o control del producto, como un modo de «dirigir» la enseñanza superior hacia «prioridades nacionales», constituye uno de los desarrollos más importantes en la política sobre educación superior que se ha experimentado desde que surgió el fenómeno de la «enseñanza superior para todos». Primero, porque representa un cambio de enfoque, de una preocupación centrada sobre políticas de igualdad y justicia social temas predominantes cuando la educación superior pasó de ser elitista a ser de masas por aspectos del «input» relativos a las relaciones entre educación superior, sociedad y economía. Segundo, porque concentrándose en el producto —en la adecuación del «output» a la luz de las exigencias de la economía nacional— y, por tanto, sobre la calida del «output», establece una nueva finalidad de la enseñanza superior sin relacionarla con las demandas individuales, sino con las necesidades de «mercado». Tercer-

ro, porque este tipo de evaluación proporciona un poderoso instrumento para que la política estatal «pueda regular» la respuesta institucional individual.

Quizás sería exagerado considerar los sistemas de control *a posteriori*, como un medio de contrarrestar los problemas de implementación que han obstaculizado muchas reformas de la educación superior en Europa occidental durante los años 70 (Bauer, 1988). Puede que una consecuencia de los cambios actuales en Gran Bretaña, Holanda, Suecia y Finlandia sea limitar el número de los aspectos «vetados» mediante la unión de la evaluación estratégica con la rutinaria, y a su vez ambas con el concepto de financiación condicionada. Esto dará lugar a otros *effets pervers* todavía por descubrir. Sin embargo, la aparición de nuevos modos de control y coordinación motivan la pregunta ¿por qué?

Justificación del «Evaluative State»

La aparición del «Evaluative State», como cualquier otro cambio significativo en materia de política educativa, no es obra de un momento. Menos aún es el resultado de una planificación clara, anticipación lúcida y la omnisciencia de la administración. Por el contrario, su aparición se debe más a la acumulación de pequeñas respuestas a las circunstancias algunas económicas, otras ideológicas. Además, como tal evaluación se encuentra inmersa en la práctica diaria, las actitudes y los valores, aunque sólo es percibido parcialmente, o dentro del contexto de los conceptos y modelos descriptivos anteriores. Por último, no es irrelevante señalar que el fenómeno del «Evaluative State», aunque es una clave para comprender los desarrollos actuales en educación superior, no se limita a este campo. Se extiende a otras áreas de la política social, como la seguridad social y la sanidad, en la medida de que estos, obviamente, dependan del estado. Resumiendo, este desarrollo, crucial en la política actual sobre educación superior en Francia, Holanda, Suecia y Gran Bretaña, no se centra exclusivamente en esta institución. No obstante, como en nuestro caso, si no está relacionado de un forma u otra con la educación superior, a menudo puede parecer como si lo estuviera. La educación superior es simplemente una de las instituciones sociales directamente afectadas por este desarrollo.

¿Qué se entiende, entonces, por «Evaluative State»? Existen muchas respuestas, todas ideológicas. Hay varias teorías, que cohabitan y que,

a menudo, se confunden. Por un lado, «Evaluative State» se percibe como una alternativa a la regulación mediante autorizaciones burocráticas. Al relacionar la evaluación con el output de los sistemas de educación superior —esto se discute en Holanda y en Gran Bretaña—, se puede abandonar el control sobre la manera en que las instituciones ejecutan las recomendaciones estatales. Al cambiar el énfasis del control de calidad a la etapa de output se genera una mayor libertad institucional. Esta apertura o, en términos de Durand Prinborgne, ideología neoliberal, también se encuentra en Francia.

Por otro lado, un margen mayor de maniobra a nivel de instituciones individuales no implica una descentralización. En ciertos casos, por ejemplo en Francia y España, la descentralización del control y el aumento de la gestión a nivel regional se consideran como parte de una misma estrategia global. Por el contrario, en Suecia, el movimiento hacia «Evaluative State» concede menor importancia a la gestión a nivel regional. Los desarrollos actuales suecos sugieren la posibilidad de mantener un control central sobre la delimitación de objetivos, y al mismo tiempo dar mayor libertad a las instituciones para escoger la trayectoria que mejor se adapte a sus circunstancias. Esta postura, aparentemente contradictoria, sugiere que con la aparición del «Evaluative State» no se puede mantener la dicotomía maniquea que han adoptado los educadores comparativos desde 1930: sistemas centralizados versus sistemas descentralizados.

El desarrollo actual del «Evaluative State» se puede justificar porque alivia la inercia aparente y las objeciones sin sentido que, según muchos, son las facetas típicas del control central. También se puede justificar el argumento opuesto, la necesidad de establecer responsabilidades claras para asegurar las prioridades nacionales o las demandas del gobierno en cuanto «consumidor» de los servicios de la enseñanza superior. Así, en sistemas basados en la descentralización, el «Evaluative State» aparece como un paso hacia un mayor control central, y en aquellos basados en el centralismo se percibe como el inicio de una mayor flexibilidad y, por tanto, más descentralización.

Lo que quiere decir esta aparente contradicción es que ninguno de estos dos sistemas es apropiado para la situación en la que se halla la educación superior. Tampoco han sido capaces de resolver el hasta ahora «acertijo administrativo insoluble» sobre cuánto tiempo puede quedar, lo que Clark ha denominado el «nivel medio» de coordinación a

nivel institucional (Clark, 1983) y, simultáneamente, sostener la coherencia global de un sistema estatal de educación superior. El «Evaluative State» constituye entonces una racionalización y redistribución de funciones entre el centro y la periferia, de modo que el centro mantiene un control estratégico global por medio de menos decisiones políticas, aunque más precisas, explícitas «en las declaraciones oficiales, el establecimiento de metas del sistema y la operativización de los criterios relativos a la calidad del «output». Estos elementos forman el marco general en el que se maneja el control estratégico y al que se confía la coordinación a nivel medio con la tarea de llevar a cabo los medios necesarios para su logro.

No causa, por tanto, sorpresa que en países donde se ha definido previamente un marco similar a nivel medio —como en el Reino Unido— el desarrollo de tal marco se considera más restrictivo que liberalizador. En suma, «Evaluative State» es una réplica a la afirmación que hace diez años realizó el sociólogo francés Michel Crozier: *«No se puede cambiar el mundo por medio de decretos. Sin embargo, se puede cambiar mediante la aplicación prudente del arte del experto en econometría y de las técnicas —quizá también— del contable. Esto no quiere decir que la racionalización de los sistemas de dirección y control de la política sobre la educación superior signifique el desprecio total del «estado», hecho que complacería a neoliberales y a criptomarxistas. Por el contrario, supone que el «estado» retira su control sobre los detalles para centrarse más significativamente sobre un «perfil» estratégico efectivo.*

Ciertamente, el proceso de redefinir el marco global y la mejora de instrumentos de la política para «dirigir» la educación superior por control remoto puede presentarse como una especie de «privatización». La suposición que subyace a esta tesis es que la educación superior, al menos en Gran Bretaña, está sujeta a consideraciones similares a las que han llevado a la venta masiva de bienes nacionales a un sector particular de la nación, a saber, a una «democracia» poseedora de acciones. Esta última es una nación que confunde —quizá a sabiendas— el concepto de posesión con el concepto de control. Se puede poseer algo sin tener la menor influencia sobre el destino o resultado de lo que uno posee, como muchos pequeños accionistas se encontraron en octubre de 1987. Otros, menos caritativos, han interpretado este mismo proceso como «pérdida de capital activo», es decir, un proceso para llevarse las partes beneficiosas de una empresa y vender el resto a precio de ganga a los amigos o parientes (Becher y Kogan, 1988).

Esto no refleja la situación exactamente. Primero porque en Gran Bretaña las Universidades ni son corporaciones públicas ni son parte del estado central. Tampoco son organismos enteramente privados. Desde este punto de vista, el estado no tiene poder para vender lo que no posee. Pero tiene la libertad para ejercer presión financiera de manera que los departamentos se ven obligados a cerrar —como ya lo han hecho— y así llevan a las Universidades a buscar apoyo financiero al margen de los fondos estatales. El que esta política sea o no realista es otro asunto. La industria británica no estaba dispuesta a crear y pagar íntegramente los sistemas de formación para sus propios empleados, actitud que se ha visto recompensada dado que ha obligado al gobierno central a asumir parte del costo de entrenamiento, principalmente a través de la *Manpower Services Commission*. El que la industria esté dispuesta a financiar a la educación superior lo que el estado no está dispuesto a hacer constituye un acto de fe, frente al desconcierto de los cambistas en el templo —o la ciudad—, que ahora tiene todas las dimensiones de un credo.

El caso británico —aunque presentado desde dentro y justificado con referencia a la teoría económica de la oferta que algunas veces invoca los «derechos del consumidor»— es un ejemplo muy claro del proceso de aparición del «Evaluative State», como también de los mecanismos por los cuales opera. Es también el ejemplo más claro porque —a diferencia de otros mecanismos de control de los sistemas centralizados, que tienden a ser jurídicos por naturaleza y se explicitan en cientos de disposiciones, decretos, órdenes ministeriales con la fuerza de la ley y las excepciones a la misma— la coordinación en el sistema británico no se establece sobre una base legal de arriba-abajo o de la autoridad definida administrativamente. Más bien trabaja con un sistema de negociación pragmática mediante un organismo intermediario, el Comité de Ayudas a la Universidad (UGC), que actúa en favor de las Universidades presentando las necesidades al gobierno de forma colectiva. Los últimos dos o tres años han contemplado un cambio extraordinario que no sólo ha terminado con este *entente cordial*, sino que por medio del trabajo de varias comisiones de investigación se ha establecido una serie de criterios externos para evaluar la eficiencia formal, así como un conjunto de normas estatales para determinar la viabilidad de conjunto de disciplinas. En el momento de escribir estas líneas, se trata de ubicar esto dentro del marco legal contenido en la *Education Bill* de 1988. Esta última publicación (Neave, 1988a) no tiene precedentes en el Reino Unido, ya que coloca a la enseñanza superior dentro de un marco legislativo global y, particular-

mente, al sector de Universidades. Un proceso similar se ha realizado, por ejemplo, en Francia con la *Lois d'Oriention* de 1968 y 1984, la *Hochschulrahmengesetz* de 1976 y 1983 en la República Federal de Alemania, o el Acta de Reforma de Educación Superior de Suecia de 1977.

Por supuesto, en el Reino Unido hay diferentes circunstancias y detalles que lo separan del resto. Sin embargo, desde una perspectiva «menos británica», las recomendaciones del informe Jarratt sobre el desarrollo de una serie de criterios de input y output para evaluar la eficiencia de Universidades concretas, son esencialmente parte de un proceso similar sobre elaboración de criterios no jurídicos con el fin de evaluar el rendimiento del sistema en Holanda, Francia, la República Federal de Alemania y Suecia. El hecho de que Gran Bretaña se vea obligada a ampliar la «reglamentación» de la educación superior refleja que previamente las bases de la relación entre gobierno y educación superior eran insuficientes para permitir que el modelo británico de «Evaluative State» se implantara en la educación superior y, al mismo tiempo, fuera bien aceptado (Jarrat Report, 1985).

Este proceso de «reglamentación» de la educación superior británica contrasta, sin embargo, con los movimientos en Holanda, Francia, Suecia y Finlandia, que intentan una nueva flexibilidad en los sistemas de gestión y evaluación, aligerando el peso del control legal y, en su lugar, transfiriendo lo que hasta ahora estaba en medios legales a una serie de indicadores, criterios y objetivos. Estos objetivos son apoyados mediante un sistema de incentivos, recompensas o sanciones dependiendo de si las instituciones concretas han excedido las normas nacionales en cuanto al número de graduados en Holanda o han mostrado iniciativas para realizar contratos con la industria o servicios locales y regionales en Francia. Existen dos aspectos de la «desburocratización» del control de la educación superior en Europa. El primero proviene en parte de consideraciones políticas, mientras que el segundo tiene que ver con la necesidad de unir la evaluación a posteriori a la política de gestión.

El propósito básico de la reforma introducida por los gobiernos francés, holandés y sueco por vía jurídica era implementar un cambio rápido de forma homogénea a través de los sectores de la educación superior. Dado que los decretos legislativos tenían que pasar a través del proceso legislativo, así servían, a menudo, para politizar la cuestión. En el caso francés, los debates que acompañaron a la que iba a ser *Loi d'Oriention* de 1984 sobrepasaron todos los récords en el número de enmiendas

desde 1943. De hecho, el proceso legislativo se convierte así en un obstáculo para la reforma, al polarizar opiniones dentro y fuera de la Universidad, sobre todo cuando se trata de detallar la reforma. Tiene ciertas ventajas cuando se reparte la responsabilidad de la decisión desde arriba hasta el nivel de la institución —en relación a unos cuantos objetivos globales que se pretenden alcanzar—. Primero, el debate sale del centro y se concentra en las implicaciones que tiene para cada una de las Universidades concretas; segundo, al crear una mayor «amplitud de maniobra» se estimula más la iniciativa, a fin de que cada Universidad concreta supervise su propio esfuerzo en lugar de buscar guía y ayuda del gobierno central; y tercero, un sistema como éste dará pie a que las Universidades —al igual que todo el sistema de educación superior— sintonicen mejor con las necesidades y las circunstancias locales o regionales, en respuesta a cambios —tanto nacionales como regionales— en el mercado de trabajo. En resumen, la autonomía se considera en Francia, Suecia y Finlandia como crucial para crear las condiciones necesarias para una «reforma en marcha», es decir, para un ajuste institucional ante circunstancias variables que, por su propia naturaleza, no siempre se pueden predecir y cumplimentar con suficiente rapidez partiendo de la legislación estatal.

Existen variaciones sobre este tema. En Finlandia, por ejemplo, la «amplitud de maniobra» aparece claramente en los estudios de licenciatura. Sin embargo, hay una mayor coordinación central a nivel de estudios de postgrado e investigación. Algo similar sucede en Holanda, donde se presta bastante atención a la planificación nacional del sistema de investigación durante el postgrado.

Procedimientos de Evaluación

No obstante, si el mecanismo de ajuste al cambio se basa menos en el proceso legislativo y es, más bien, parte de una operación técnica autoafirmativa efectuada dentro de un marco definido por la legislación, son necesarias estructuras específicas para unir la evaluación de instituciones a la dirección de la política global. La aparición de nuevas estructuras con esta responsabilidad o la redefinición de organismos ya establecidos es muy clara y, durante los dos últimos años, ya han aparecido en Europa occidental. Lo principal de ellas es el *Comité National d'Evaluation*, creado en 1984 en Francia para evaluar el rendimiento de las Universidades. Este organismo, que actúa como una agencia de evaluación

nacional, está formado por profesores destacados. Rinde un informe directamente al presidente de la República, no al ministro de Educación Superior e Investigación —detalle muy importante, que simboliza que la evaluación global de las Universidades del país está más allá de partidismos.

El *Consejo de Universidades* en España realiza funciones similares, principalmente en la evaluación del currículum y en la asignación de cargos nacionales. No obstante, el desarrollo del «Evaluative State» en España ha sido menos importante que en Francia. Sin embargo, el Consejo de Universidades español —organizado, en parte, según el modelo del Comité de Ayudas a la Universidad de Gran Bretaña— puede ilustrar la tendencia actual a crear organismo especializados e intermediarios fuera del Ministerio, con la responsabilidad de coordinar la educación superior y las exigencias del mercado de trabajo. Las relaciones del Consejo de Universidades con el Ministerio de Educación se establecen a través de su presidente, que es el propio ministro.

En Holanda el desarrollo del «Evaluative State» ha supuesto más la redefinición del estatus y funciones de organismos ya establecidos, específicamente de la Inspección General, que la creación de nuevos organismos. Anteriormente, la responsabilidad de la inspección holandesa, además de las escuelas y la formación del profesorado, era la educación profesional superior (*Hoger Beroeps Onderwijs*). En 1985, debido a la reorganización de este organismo, su responsabilidad incluyó a las Universidades, al mismo tiempo que se redefinió su posición en el organigrama del gobierno central. En la actualidad rinde sus informes directamente al ministro. Desde el punto de vista de las Universidades, esta inspección está en un proceso de reorganización y establecer lazos con las instituciones universitarias.

En Gran Bretaña el proceso de creación de mecanismos evaluativos es más difuso y no está concentrado en ningún organismo. Se ha fortalecido el papel evaluativo de agencias que ya no poseen esta función —a saber los consejos de investigación— así como que el Comité de Ayudas a la Universidad asuma un papel más directivo, aunque éste último va a ser reemplazado en breve por un Comité de Fondos para las Universidades (UFC) y, en la educación superior no universitaria, por el Consejo de Fondos para *Colleges* y Politécnicos (PCFC). No obstante hay dos desarrollos que evidencian el crecimiento del «Evaluative State» en Gran Bretaña. El primero es el requisito establecido por el Comité

de Ayudas a la Universidad en 1986, según el cual todas las Universidades debían dar informes detallados de su planificación de 1987 a 1992. También tenían que incluir en estos informes estimaciones sobre el nivel y la excelencia de los resultados académicos, así como medidas posibles para reestructurar sus disciplinas según las prioridades nacionales y las posibilidades de apoyo del mundo empresarial. El segundo es la introducción de la «legislación sobre contratos de enseñanza», un sistema por el que las instituciones compiten a nivel nacional por contratos del gobierno para enseñar a un número determinado de estudiantes. Los contratos se dan al mejor postor. Esta es una nueva versión del «Evaluative State» que parece distinguir a Gran Bretaña del resto de Europa. Aunque el sistema de contratos también está implantado en Francia y Holanda, en estas ocupa un papel marginal. En el caso holandés, los contratos de enseñanza se limitan a servicios prestados a la región. En el caso francés, se circunscribe a determinados aspectos de estudios de postgrado (Bijleveld, 1987).

El «Evaluative State» toma forma de diferentes maneras. Algunas implican el establecimiento de agencias nacionales cuyo único propósito es examinar el estado presente y el potencial futuro de la educación superior evaluando desde arriba las instituciones. Este modelo aparece claramente en Francia, aunque en Alemania también se ha prestado mucho interés al trabajo del *Comité National d'Evaluation* francés. Otras naciones, como España y Holanda, tratan de fortalecer esta función con servicios ya existentes. En Holanda, como muestran Van Vught y Maassen, hay también un proceso paralelo que va de abajo-arriba. Este proceso gira en torno al desarrollo de programas estratégicos a nivel institucional, dentro del marco de nuevos procedimientos de planificación entre gobierno e instituciones. En Gran Bretaña se constata un proceso similar de dos niveles: primero, el reforzamiento de la función evaluadora entre agencias de fondos, tanto en el campo de la investigación como en la distribución de recursos; segundo, el desarrollo de una autoevaluación institucional de abajo-arriba, aunque inicialmente haya surgido de arriba-abajo, explícita en diversos programas expuestos por cada institución al Comité de Ayudas a la Universidad.

Avances como éstos pueden interpretarse como parte de un empuje más general orientado a clarificar el modelo de gestión de la enseñanza superior en el que subyace el concepto de «Evaluative State», así como también a poner de relieve el cambio de una evaluación centrada sobre el «input» y el «proceso» a una evaluación sobre el «producto». Existen

diferencias sobre las estructuras y canales encargados de los aspectos técnicos de la «Evaluative State», así como respecto a cargo de quiénes debe estar esta operación a nivel central. Este último punto es un aspecto muy importante dado que, como señala Bauer, el grado de confianza que las unidades de base tienen sobre este proceso depende de que la evaluación esté en manos de «un cuerpo profesional de evaluadores». La cuestión es: ¿se debe confiar la «Evaluative State» a la opinión de expertos? o, por el contrario, ¿son más fiables los puntos de vista y las opiniones de funcionarios profesionales? La respuesta depende mucho de si el proceso técnico de la evaluación va a tener una función más política, en la que se trate de consensuar las distintas opciones que se pongan de relieve a través de la evaluación, o si la evaluación se va a utilizar para someter a profesores recalcitrantes a lo que los gobiernos consideran «la nueva realidad». Ambas orientaciones implican, de algún modo, la renegociación entre la educación superior y la sociedad: la primera, centrando la evaluación en agencias semi-independientes y, la segunda, en estructuras dentro de la administración. En Francia parece que está surgiendo el primer proceso y en Holanda y Gran Bretaña el segundo; aunque en ambos casos el control estatal tiende a disminuir en la medida que está surgiendo la «autoevaluación» a nivel institucional (Neave, 1988b). Considerando de nuevo las áreas a negociar, al unir la asignación de recursos a una evaluación a posteriori, se cambia también el «foco de tensión» entre las necesidades definidas por los profesores y las prioridades estipuladas por el gobierno. No obstante, al cambiar el enfoque político por el evaluativo, se corre el riesgo de que todos los aspectos técnicos implícitos en todo sistema de evaluación se transformen en uno de los conflictos y obstáculos más serios, a menos que el proceso se considere como legítimo por los evaluados.

Liderazgo institucional

La creación de sistemas de evaluación en los centros, la utilización de indicadores de rendimiento y la elaboración de criterios para operativizar la calidad requieren cambios a nivel institucional. Por ello la aparición del «Evaluative State» no se limita a la relación entre educación superior y sociedad. Al contrario, demanda nuevos tipos de información y da lugar a nuevas estructuras y modelos de autoridad para buscar y utilizar esa información a nivel institucional. A ello se añaden mayores responsabilidades a nivel de estructuras de gestión intermedias (especialmente en sistemas donde históricamente el modelo ha sido un siste-

ma de administración muy centralizado). Ahí están los contratos sobre investigación y docencia firmados en Francia —hasta hace poco las Universidades no tenían personalidad jurídica y dependían del Ministerio para estos asuntos—, la colocación de estudiantes en empresas y la evaluación interna de sus propias instituciones (OCDE, 1986). En Holanda, una de las principales áreas donde aumentó la responsabilidad de las instituciones fue en el poder establecer nuevos cursos sin la autorización del gobierno central. Todo el «tiempo perdido» implica un mayor esfuerzo de las instituciones para orientar su propio desarrollo, que indudablemente tendrá efecto en las estructuras de autoridad y toma de decisiones. Un liderazgo fuerte parece ser el requisito de una gestión orientada hacia metas, principio reconocido por sistemas tan diferentes política e históricamente como Finlandia, Gran Bretaña, Holanda y Francia. Otro asunto es la forma que esto pueda tomar.

Históricamente la mayoría de los sistemas de educación superior en Europa occidental han evolucionado hacia un liderazgo institucional basado en la persona del Rector o, en ciertos «Lander» de Alemania y en Francia, en el presidente de la Universidad. En contraste con los presidentes o vicescancelleres americanos e ingleses, el puesto del líder tiende a ser de corta duración —salvo excepciones, debido a personalidades muy concretas—, y es más un puesto honorífico para representar a la Universidad oficialmente que una autoridad que asume el poder ejecutivo. El liderazgo institucional, en estas situaciones, se ejerce a nivel de las facultades por los catedráticos responsables de las mismas, entre los que se escoge por elección al Rector (Neave, 1988,a).

Aunque se desconoce cómo evolucionará el modelo tradicional de liderazgo, tenemos algunos indicios. Uno de estos, contenido en el *Informe Jarratt* de Gran Bretaña, reclama la transformación del papel del vicescanciller británico, sugiriendo que debe dejar de estar en la jerarquía académica para asumir un estatus semejante al modelo americano de presidente de la Universidad. Esta redefinición de estatus implica dos cambios. Primero, coloca al vicescanciller a la cabeza en un sistema de «gestión lineal» en lugar de ser emanación del «collegium» académico; segundo, atribuye la responsabilidad de coordinación a una administración que está por encima de los profesores, en lugar de una administración en una posición subalterna. El grado de concentración del poder en la cima —sobre todo en relación con la planificación a largo plazo— se puede evidenciar por la aparición gradual en muchas Universidades británicas de una especie de gabinete —una forma de «personal del

cuartel general»—, conocido como «comité estratégico» o «comité de política y recursos», cuya responsabilidad y funciones tiene más que ver con la organización empresarial que con la Universidad tradicional. Si la evolución constatada en Gran Bretaña prefigura el futuro, la aparición del «Evaluative State» hará que a nivel institucional se concentren y centralicen el poder y la autoridad más que nunca.

Las propuestas de la comisión francesa «*Demain l'université*», cuyo primer informe apareció en 1988, son similares: «el aumento de comités técnicos y especializados asesorando al presidente de la Universidad y una reducción significativa de los organismos que se ocupan de asuntos generales de la Universidad. La necesidad de una mayor coordinación interna a nivel institucional ha sido tema de debate en el sector público de las Universidades belgas. Aunque se reconoce el principio de autonomía, las consecuencias administrativas y organizativas y los medios a través de los cuales podría efectuarse están menos claros.

Una propuesta, formulada el verano de 1987, sugería, además del Consejo de Administración, la creación de un comité permanente para unir representantes de todos los sectores de la comunidad académica, profesores, personal técnico, estudiantes y otros representantes con intereses legítimos. No obstante, ni la propuesta francesa ni belga han llegado tan lejos como la británica a la hora de asumir completamente el concepto de organización empresarial. Además, cualquiera que sea el resultado, se requiere legislación nacional para aplicarla, no pudiendo hacerse de forma individual. Aun así, ciertas áreas de responsabilidad financiera, como las matrículas de estudiantes, los contratos de investigación y, en el caso de Francia, las sumas procedentes del Impuesto de Aprendizajes requieren una gestión y una planificación diferentes a las que se llevaban a cabo por la administración central, cuando el principal objetivo de esta administración era vigilar si se observaban las normas gubernamentales. Sin embargo, al igual que en Gran Bretaña, cualquier cambio en las tareas que debe asumir la institución —la forma de efectuar la planificación para satisfacer las prioridades nacionales y las demandas locales es una de ellas— también implica un equilibrio de poder entre el presidente, la administración y el profesorado.

Este problema surge en los sistemas de enseñanza superior que están altamente politizados como en Europa occidental. Sin aumentar el poder y el número del personal administrativo, es poco probable que las instituciones sean capaces de construir nuevas relaciones con su región

o con las empresas. Por otro lado, si esta misión se confía a los profesores, lo cual puede ser una solución a corto plazo, aparece una clase híbrida de consultores semiprofesionales que van a tener problemas para dedicarse a la docencia, a la investigación o a ambas. Resumiendo, el dilema de muchas Universidades europeas, que en el pasado han disfrutado de poder a nivel de Facultad o UFR, es si resulta políticamente rentable concentrar el poder en un presidente o dividirlo entre tres: los que se dedican a la investigación, los que se dedican a la docencia y aquéllos cuya ocupación es salir fuera o buscar contratos para recaudar fondos.

Existen modelos de gestión distintos a los modelos lineales que emergen en algunas Universidades Británicas. La pregunta es ¿cuál se adapta mejor a la Europa Occidental? o ¿es mejor que cada país adopte el que más se adecúe a sus circunstancias particulares? Si el propósito del autogobierno —que no siempre es lo mismo que autonomía— es optimizar las relaciones con las empresas a nivel nacional o regional, entonces es probable que los cambios del entorno tengan influencia en el sistema de gestión que se adopte, o que al menos introduzcan variaciones locales. No sabemos si esto reemplazará la identidad «profesional» por la identidad «institucional» como sugiere Holttá. Cualquiera de las dos formas que cambie en la gestión y en las estructuras de gobierno de las Universidades no puede reducirse a una tarea técnica o puramente administrativa. Estas estructuras incorporan valores académicos esenciales —colegialidad, disciplina, etc.—a, que no se adaptan fácilmente a los modelos empresariales que abogan sus defensores. La cuestión no es un asunto de profesores que defienden la «libertad de enseñar y aprender», sino que se trata de si la «Evaluative State» será capaz de incorporar la racionalidad de la gestión empresarial desde lo alto del sistema de educación superior hasta las unidades más básicas (Neave y Rhoades, 1987).

La Ética del «Evaluative State»

Los sistemas de administración, control y evaluación, aunque parecen técnicos, son también expresiones de creencias y valores, ya que, en realidad, son los fines que se persiguen, mientras que las estructuras y modus operandi son los medios. Algunos pueden reducirse a fines técnicos. Por ejemplo, aumentar la eficiencia de recursos físicos, financieros o humanos; aumentar el número de graduados o crear la estruc-

tura de los estudios y las especialidades de acuerdo con lo que necesita el bien común. A menudo son los fines los que justifican la aparición del «Evaluative State». Aunque existe un gran debate actualmente sobre estos aspectos, todo lo que se puede decir acerca de ellos es que no son nuevos. Han constituido la preocupación constante de los políticos desde que la educación superior dejó de ser elitista y, en algunos casos, incluso antes (Premfors, 1983). Se puede argumentar que, cualquiera que fuese la intención de los gobiernos antes de la aparición del «Evaluative State», no existió voluntad para llevarlos a la práctica o, finalmente, para hacerlos cumplir con suficiente rapidez. Por ello, el «Evaluative State» constituye un medio de llevar a cabo fines antiguos todavía sin cumplir.

Aparece entonces la pregunta: «¿Por qué ha aparecido tan rápidamente el «Evaluative State» en los últimos cinco o siete años?» Se pueden dar dos respuestas: primera, debido a factores generales que implican un cambio estructural de carácter económico y demográfico. Segundo, a causa de un cambio en la ética social. En relación a lo primero normalmente se habla de la importancia creciente de una economía basada en la alta tecnología, y de la necesidad urgente de elevar las cualificaciones profesionales de la población. No todos los países consideran esta tarea exclusiva de la educación superior. En el Reino Unido, por ejemplo, se cree que este ajuste es responsabilidad del sistema de orientación y formación profesional. Otros lo ven como una responsabilidad de la Educación Superior, que debe definir mejor las «salidas profesionales» poniendo un mayor énfasis en carreras de ciclo medio. Las propuestas recientes de la comisión «*Demaine l'université*» en Francia son reveladores. Su sugerencia principal es la creación de Colegios Universitarios que acogerían el 40 % de los estudiantes admitidos en la Universidad. Otra variante de este proceso se encuentra en Holanda, donde la creación de amplios sectores de enseñanzas que se imparten en la Universidad y los centros de Educación Superior Profesional posibilitan a los estudiantes cambiar de un tipo de estudio a otro según las variaciones en el mercado de trabajo o en función de las nuevas profesiones que surgen a través de combinaciones de habilidades.

Existe otro aspecto a tener en cuenta en la formación profesional, aunque más a largo plazo. Se trata de que las profesiones importantes para las necesidades básicas de un país son cada vez más escasas, a excepción de Irlanda y España. Con la disminución de la tasa de natalidad, Europa es una sociedad vieja y el número de jóvenes de los que

depende la riqueza de las naciones está disminuyendo. El énfasis que ponen la mayoría de los gobiernos en que la educación superior sea más eficiente en transformar a los jóvenes que acceden a las instituciones en graduados calificados está variando debido a esta consideración. Ante el problema del envejecimiento de la población existen dos opciones: o se limita la proporción de jóvenes que pueden acceder a los estudios superiores o aumentan las «habilidades productivas» de los sujetos que cursan estudios superiores. En el caso de Francia se emplean las dos opciones. Otra opción seguida en la actualidad por casi todos los países europeos es prolongar el período de alta productividad entre los grupos de mayor edad por medio de la educación permanente. Aunque esto es una salida importante, sólo podrá servir durante los próximos veinte años. Después, se sentirán los efectos de la caída de la tasa de natalidad, incluso a nivel de educación permanente.

Por último, surge la cuestión del conflicto de necesidades entre los distintos sectores de la economía, entre los que se ocupan de la «creación de riqueza» y los que se encargan de facilitar los medios para crear riqueza a largo plazo, especialmente dentro del sistema educativo mismo. Sin un aumento de personas cualificadas es difícil que el output de la educación superior en algunos países —por ejemplo, Francia y Holanda— sea suficiente en diez años para cumplir las exigencias de alta tecnología, demandadas por el sector «privado» y las necesidades del sistema educativo para satisfacer las demandas de los jóvenes con habilidades y motivación ante los continuos desafíos de la educación, la enseñanza, la investigación, la formación y el desarrollo de una nación conlleva.

Es aquí donde aparece el asunto de la ética del «Evaluative State». El fin estratégico del «Evaluative State», aunque parezca técnico, resulta ser una redefinición de valores sociales. Frases como «abierto al mercado», «fomento de la iniciativa individual y empresarial», «instaurar una “ética competitiva” entre individuos e instituciones»..., son la expresión de una nueva fuerza motriz que pretende la renovación empresarial y del país. Sería descortés comparar esto al eslogan emitido hace 150 años por el rey burgués Louis Philippe: «*Enrichissez vous*», consigna que obedecieron con prontitud y provecho algunos de los personajes más insulsos de la *Comédie Humaine* de Balzac.

La competitividad y la excelencia no son exclusivos de las empresas ni de la Universidad. Por el contrario, residen en ambos mundos, aun-

que a menudo los interpretan de forma diferente. Sin embargo, la aparición de lo que Van Vught y Maassen denominan «Darwinismo social», como prueba de fuego para instituciones e individuos, no carece de peligros. Se puede justificar esto por el hecho de que la sociedad no es capaz de predecir qué tipo de habilidades se requerirán en el futuro. Si se sensibiliza a los estudiantes e individuos en cuanto a la situación del «mercado» de trabajo, la responsabilidad de prepararse para el cambio es del individuo. Aquellos que no estén dispuestos a realizar tal ajuste deben sufrir por su decisión. El peligro surge por el énfasis que ponen los gobiernos al usar esta ideología para aumentar el número de graduados en las áreas de ciencias y tecnologías, viendo la misión de la Universidad casi exclusivamente en estos términos. El objetivo llevado a cabo en Gran Bretaña de aumentar el número de estudiantes de ciencias del 48 % del total al 52 % es un indicativo de este tipo de política.

Pero esta perspectiva, al igual que algunos criterios utilizados para evaluar la eficiencia, tiende a olvidar que la Universidad no es exclusivamente el proveedor de mano de obra de las empresas y que, en muchos sistemas europeos de educación superior, sólo una minoría estudia ciencia y tecnología. Las otras áreas —denominadas «áreas de valores»— no son menos importantes en la formación de ciudadanos conscientes y responsables en cuyo sentido de justicia social y solidaridad descansa la estabilidad política de las naciones, particularmente cuando, como resultado del cambio industrial, están aumentando las diferencias salariales entre los de arriba y los de abajo.

Conclusión

La aparición del «Evaluative State» es uno de los principales cambios en la relación entre la enseñanza superior, los gobiernos y la sociedad en la Europa occidental. Desde un punto de vista histórico, puede considerarse como uno de los pasos más significativos en el intento de ir más allá de los modelos basados en el control administrativo, cuyos orígenes se pueden remontar a más de 70 años y, en caso de Francia y España, al período napoleónico (Sáez, 1987). El empuje hacia el «Evaluative State» no es uniforme en los diferentes países. En la actualidad, los más avanzados parecen ser Finlandia, Holanda, Suecia y el Reino Unido. Aun así, es posible detectar desarrollos significativos en Francia y, en menor medida, en España, que si se perfilan aún más pueden situar a ambos países en este camino más firmemente.

Otro asunto bastante diferente es considerar este proceso como una «revolución», aunque si las revoluciones implican la percepción de las prácticas y procedimientos establecidos de forma totalmente diferente —que se justifican por diferentes causas—, entonces es posible calificar el «Evaluative State» como un proceso «revolucionario».

La mayoría de las áreas importantes en la educación superior quedan afectadas por este cambio: la financiación, el presupuesto, el nivel de supervisión de la administración central, el potencial para realizar cambios profundos en los modelos de autoridad, la responsabilidad y liderazgo académico, y la misma autonomía académica, por citar los más obvios. Estos elementos tienen que ver con aspectos de la sociedad que influyen en la educación superior, pero el «Evaluative State» tiene otro propósito. Este conlleva una estrategia indirecta que implica una movilización social muy amplia para encarar el desafío del cambio tecnológico e introducir la racionalidad organizacional necesaria en aquellas instituciones cruciales para efectuar el ajuste social que se deriva de tales cambios. Algunos han interpretado esto de varias maneras: por ejemplo, a través de la privatización de la educación superior. Otros dan una expresión más concreta a este fenómeno y buscan modelos que actúen de inspiración y guía en este proceso, modelos que, por otra parte, también posean un dinamismo importante y apropiado para las circunstancias actuales.

Las referencias al «modelo de Universidad» de Estados Unidos son múltiples. Algunos gobiernos lo mencionan para limitar el número de Universidades que pueden investigar —propuesta reciente del Consejo Asesor Británico al Consejo de Investigación—. Otros resaltan su habilidad para conseguir fondos de benefactores, fundaciones y empresas privadas; mientras que un tercer grupo, especialmente el gobierno francés, establece como meta que el 75 % de estudiantes alcancen el bachillerato porque un número similar de estudiantes se graduó en América en las escuelas secundarias.

Aunque estos aspectos aisladamente puede verse como distorsionados y descontextualizados, juntos forman lo que se denomina «*change-ments des mentalités*». El modelo de la Universidad americana o la privatización, en este contexto, forman parte de la voluntad política que normalmente subyace a la aparición del «Evaluative State». Son emblemas de una dinámica social en la que las metas de las instituciones se moldean en función de la preocupación que deberían tener presente los

individuos al «consumir» sus «productos». El consumo de conocimiento debería ser, pues, «guiado por el mercado», cuando no «determinado por el mercado». Las implicaciones de este «*changement des mentalités*» en términos del proceso de elección de ocupación individual son muy interesantes porque sugieren que a los estudiantes se les pregunte qué tipo de estudios desean realizar en función del trabajo que piensan desempeñar al finalizar sus estudios, mientras que actualmente el criterio principal de los sistemas escolares europeos es seleccionar a sus estudiantes en base al rendimiento alcanzado previamente (Duner, 1977).

Cualquiera que sea la implicación —asunto que merece una mayor investigación comparativa— el «Evaluative State» representa un nuevo empuje para reformar la enseñanza superior tan significativo en las dimensiones que atañe como la decisión tomada hace un cuarto de siglo de posibilitar el acceso a este nivel educativo al mayor número de estudiantes. Como hemos dicho, existen factores coyunturales, a corto plazo, que han acelerado su nacimiento: oportunidad financiera, cambio económico, necesidad de movilizar recursos humanos a nivel nacional e internacional (por ejemplo, los programas ERASMUS y COMETT), etc. Aunque hay otras consideraciones más a largo plazo. La expansión de la educación superior, tanto en sectores universitarios como no universitarios, ha dado lugar a un sistema global de tal complejidad que las formas históricas de control, de establecimiento de metas, y de ajuste a los cambios ya no son capaces de cumplir su función con suficiente rapidez o precisión. Desde este punto de vista, la aparición del «Evaluative State» se puede interpretar como otro paso en el largo proceso a través del cual las organizaciones de la Europa occidental tienden a ajustarse al fenómeno de una enseñanza superior para todos.

CALIDAD UNIVERSITARIA: UN ENFOQUE INSTITUCIONAL Y MULTIDIMENSIONAL *

SEBASTIAN RODRIGUEZ ESPINAR
Universidad de Barcelona, España

1. Algunas cuestiones previas

Antes de abordar de forma directa el tema de la calidad universitaria es necesario precisar el marco de referencia desde el que vamos a realizar nuestro análisis. En primer lugar, nuestra escasa tradición en el debate sobre la calidad de la universidad, al menos desde la perspectiva en que se plantea a nivel internacional, unida a las propias características de nuestro sistema universitario, obligan a establecer previamente ciertos considerandos que nos permitan contextualizar la reflexión. En segundo término, entender que la calidad de la Educación universitaria se centra exclusivamente en la calidad de la enseñanza impartida en el aula, considerando al profesor como el agente único, nos llevaría a un reduccionismo tal que entraría en confrontación con la perspectiva institucional que toma en consideración la globalidad del sistema.

1.1. Educación universitaria o Educación superior

No sería aventurado afirmar que la práctica totalidad de nuestra sociedad tiene identificada a la educación universitaria como el nivel exclusivo de educación postsecundaria. Sin duda alguna, la *Ley General de Educación* de 1970 sentó el marco jurídico para que la sociedad haya producido tal identificación: bajo la égida de la

* El presente trabajo toma como núcleo la ponencia que bajo el título *Dimensiones de la calidad universitaria* el autor desarrolló en el *I Congreso Internacional sobre Calidad de la Educación Universitaria* (I.C.E. Universidad de Cádiz, Pto. Sta. María, marzo 1991).

Universidad se colocaron la casi totalidad de los *títulos* de nivel postsecundario, si bien es cierto que existen una serie de elementos de diferenciación (sobre todo en la estructura del profesorado) entre las Escuelas Universitarias y las Facultades o Escuelas Superiores. Esta situación contrasta con la existente en otros países europeos en donde el sistema de formación profesional tiene el carácter de educación superior pero no universitaria. Sin embargo, la *Ley de Reforma Universitaria* considera la existencia de Centros de educación superior no integrados en la Universidad:

«Los Centros docentes de educación superior que por la naturaleza de las enseñanzas que imparten o los títulos y diplomas que estén autorizados a expedir no se integren, o no proceda su integración, en una Universidad...» (disposición Quinta Adicional. L.R.U., 1983).

La realidad nos dice que a lo largo de los últimos años el proceso no ha sido el de potenciar este tipo de centros, sino que se ha dado, muy al contrario, la petición constante de integración de los existentes en una universidad.

En consecuencia, podemos afirmar que nuestro sistema universitario es un sistema global o único de enseñanza superior. Ahora bien, la reforma de la Enseñanza profesional que se contempla en la *Ley de Ordenación General del Sistema Educativo* de 1990 contempla la puesta en marcha de módulos de tercer grado de formación profesional, con requisitos de titulación de acceso (Bachillerato) idénticos a los de la Enseñanza universitaria. Al final de esta década, momento álgido de la expansión de esta nueva titulación, nos encontraremos, de modo encubierto, con un sistema dual de Enseñanza superior que exigirá la revisión de la propia Enseñanza universitaria.

Como señala Husén (1988), en una época en que las disposiciones educativas después de la escuela secundaria se consideran un derecho civil, sería más adecuado hablar de educación *postsecundaria* o *terciaria*. Husén cita la propuesta taxonómica de Martin Trow que distingue tres tipos de enseñanza superior en función de los grupos de edad que acceden: *elitista-selectivo* (sólo acceden el 5 % del grupo de edad —era la situación en la universidad tradicional—), *de masas* (con un porcentaje de acceso entre el 15 y el 50 %) y *universal*, cuando la mayoría de adultos ingresan de un modo y otro en este tipo de educación.

Si tenemos en cuenta todo lo anterior, no sería incorrecto concluir que nuestro sistema de educación universitaria es de masas y sin diferenciación, fundamentalmente por lo que respecta al acceso, entre los estudios de estricta orientación profesionalizadora (Escuelas) y aquellos que añaden las dimensiones de orientación disciplinar y de investigación (Facultades).

1.2. *La moda y cultura de la calidad*

Atlbach (1986) afirma que los norteamericanos son muy dados a buscar soluciones rápidas a problemas complejos: todo el mundo es experto en educación, puesto que todos la han conocido (en diferentes niveles), y manifiesta que el interés público por la calidad de la educación tiene una milla de ancho y una pulgada de hondo. Las instituciones educativas cambian lentamente y hoy no están *sincronizadas* con las orientaciones generales de la sociedad, debiendo ser conscientes de que los cambios de política educativa tienen que recorrer un largo y complicado camino hasta que se hacen realidad.

Hemos de reconocer que en la última década, y por factores externos a la universidad, el tema de la calidad ha cobrado especial relevancia: era necesario evaluar en qué medida era posible reducir ciertos costes universitarios dada su escasa contribución a la calidad de la educación universitaria. Inicialmente el motivo no ha sido una loable preocupación por la calidad, sino la presión ejercida por la reducción presupuestaria que se produjo en las universidades. Así, las universidades europeas se han tenido que enfrentar con una situación relativamente nueva: la evaluación institucional.

En el contexto estadounidense, la tradición evaluativa, ligada fundamentalmente a la acreditación de programas o a las *imágenes de reputación*, unida a la peculiar estructura de su sistema de Enseñanza universitaria, representa un caso especial. La fuerza del desarrollo de su sistema y modelos de evaluación institucional es necesario *canalizarla* a la hora de la aplicación en nuestro contexto. De lo contrario, corremos el riesgo de ser arrastrados fuera de este planeta.

Otro aspecto de interés es el de la correspondencia entre calidad de la infraestructura del edificio universitario y la calidad global de éste. El profesor Fernández Pérez (1990), con su proverbial discurso, pone el

dedo en la llaga de la calidad de la docencia universitaria y hace la certera reflexión de que la preocupación por la calidad de la educación universitaria no es sino la lógica consecuencia de estar en vías de solucionar los graves problemas de infraestructura que se han padecido en una etapa de masificación universitaria.

No es de extrañar, aunque sí nos debe preocupar, que nuestra *Ley de Reforma Universitaria* (1983), a diferencia de la reciente aprobada *Ley de Ordenación General del Sistema Educativo* (1990), no se distingue por prestar atención al tema de la calidad universitaria ¹. Por

1. La *Ley de Reforma Universitaria* no hace mención específica del tema de la calidad de las instituciones universitarias. Por lo que respecta a la evaluación institucional, es sorprendente que sólo tome en consideración al profesor (art. 45.3):

«Los Estatutos de la Universidad dispondrán los procedimientos para la evaluación periódica del rendimiento docente y científico del profesorado, que será tenido en cuenta en los concursos a que aluden los artículos treinta y cinco a treinta y nueve, a efectos de su continuidad y promoción. [El art. 8.4 del R.D. 1427/1986 *vinculó este punto al simple hecho de que los informes llegasen a tiempo.*]

[Una Ley que algunos la han llamado del profesorado universitario —hecho por el gobierno y diputados universitarios—: el Título Quinto (Del Profesorado) contiene el 27 % del articulado (16/59) y más del 35 % de la extensión material del mismo. Así mismo, ocupa el 75 % de las Disposiciones Adicionales y Transitorias (9/12), con un 50 % de la extensión material de las mismas.]

La LOGSE, por el contrario, dedica su Título Cuarto a la Calidad de la Enseñanza. Su art. 55 dice:

«Los poderes públicos observarán una atención prioritaria al conjunto de factores que favorecen la calidad y mejora de la enseñanza en especial a:

- a) La cualificación y formación del profesorado.
- b) La programación docente.
- c) Los recursos educativos y la función directiva.
- d) La innovación y la investigación educativa.
- e) La orientación educativa y profesional.
- f) La inspección educativa.
- g) La evaluación del sistema educativo.

Ya en la *Ley General de Educación* se hacía expresa referencia al tema en su art. 11.5: «La valoración del rendimiento de los Centros se hará en función de:

- el rendimiento promedio del alumnado en su vida académica y profesional;
- la titulación académica del profesorado;
- la relación numérica alumno-profesor;
- la disponibilidad y utilización de medios y métodos modernos de enseñanza;
- las instalaciones y actividades docentes, culturales y deportivas;
- el número e importancia de las materias facultativas;
- los servicios de orientación pedagógica y profesional;

tanto, y por lo que hace referencia a nuestro contexto universitario, deberemos reconocer la ausencia de una *cultura y experiencias de análisis propias* del fenómeno de la calidad universitaria. No habrá de extrañar que con frecuencia nos adentremos en aspectos que puedan parecer lejanos. Si queremos asomarnos al *escenario* de la calidad y su evaluación, no debemos olvidar que, con anterioridad, otros han fijado ciertas reglas del juego. Lo peligroso puede ser aceptarlas sin la previa reflexión sobre su adecuación a nuestra realidad.

1.3. *La dimensión política: la autonomía como presupuesto de la calidad*

En el último tramo de la década de los ochenta, varios sistemas universitarios occidentales han estrenado un nuevo marco jurídico (Suecia, Holanda, Gran Bretaña). Una de sus notas características ha sido el establecimiento de un modelo de relación de las universidades con la autoridad gubernamental basado en la autonomía de aquéllas. Un argumento fundamental para la adopción de este marco autonómico ha sido el de considerar la autonomía como un factor decisivo para el logro de los objetivos de calidad y progreso de la educación superior de dichos países (Vroeijenstijn y Acherman, 1990) ².

Ahora bien, con excesiva frecuencia el planteamiento de la autonomía ha estado ligado al modelo de financiación de la Universidad por parte del Estado, así como al cumplimiento de una serie de requisitos de la política gubernamental: oferta del número de plazas, criterios de admisión, oferta de programas educativos, procedimientos de selección y promoción del profesorado, salarios, etc. En definitiva, y como apunta Neave (1987), se está en el marco de una *autonomía condicionada*, o autonomía vigilada si queremos utilizar un adjetivo más familiar a nuestro proceso político, y, como manifiesta Mora (1991), analizar la calidad supone, en muchos casos, entrar en la obtención y repartición de fon-

— la formación y experiencia del equipo directivo del Centro;
— las relaciones del Centro con la comunidad».

2. De especial interés resulta el análisis del marco jurídico del sistema holandés: una ley gestada a partir de 1985 (*Ministerio de Educación y Ciencia*, 1985) y que fue remitida al parlamento en 1989. Su puesta en vigor permitirá la puesta en marcha de toda una serie de decisiones autónomas en relación al número de estudiantes y su acceso, tasas, tipos de estudios, etc., así como el compromiso de un sistema de evaluación de la calidad universitaria, especificado en la propia ley (Spee, 1990).

dos. En consecuencia, no sólo se plantea un problema conceptual, sino político, ya que va a afectar al funcionamiento global de la institución.

Las autoridades políticas, basándose en determinados estudios o informes que ponen de manifiesto importantes deficiencias (fallos de calidad) del sistema educativo, consideran, como señala Trow (1987), que las universidades se han convertido en algo excesivamente importante como para dejarlas solas. Por ello se emprenden *determinadas campañas* que reclaman mayor control sobre ellas. Como ejemplo sirvan los diferentes informes sobre la crisis de la Educación Superior en los EE.UU. aparecidos en la década de los ochenta.

A pesar de toda esta problemática, no puede negarse la evidencia de una preocupación política por la calidad de las instituciones de enseñanza superior. Así se pone de manifiesto en el *Programa sobre la Gestión de los Establecimientos de Enseñanza Superior* que se desarrolla bajo el patrocinio de la O.C.D.E., aunque con escasa participación de universidades españolas.

Sin entrar en un análisis detallado, la *L.R.U.* toma en consideración la filosofía de la autonomía universitaria como factor de mejora del sistema, cuando considera que la autonomía generará diversificación, y que esta supuesta variabilidad nos llevará al incremento de la calidad.

«...el sistema de universidades que resulta de la aplicación progresiva de esta Ley se caracterizará por una diversificación entre las Universidades, que estimulará, sin duda, la competencia entre las mismas para alcanzar niveles más altos de calidad y excelencia, si bien se garantiza una calidad mínima homogénea para todas las Universidades nacionales.»

(Preámbulo, LRU, 1983).

Claro está que es necesario precisar los ámbitos o áreas de autonomía. El art. 3.2 especifica (se señalan los más significativos):

«La autonomía de las Universidades comprende:

...c) La elaboración, aprobación y gestión de sus presupuestos y la administración de sus bienes.

...d) La selección, formación y promoción del personal docente e investigador y de administración y servicios, así como la determinación en que ha de desarrollar sus actividades. *[A partir de los resultados de la aplicación del Título Quinto.]*

f) La elaboración y aprobación de planes de estudio e investigación. *[A partir de la autonomía que deja el art. 28.1.]*

g) *La creación de estructuras específicas que actúen como soporte de la investigación y la docencia.*

h) *La admisión, régimen de permanencia y verificación de conocimientos de los estudiantes. [A partir de lo dispuesto por el art. 26.1.]* (Los [] son añadidos).*

No es necesario extender el comentario para concluir que la autonomía universitaria de la *L.R.U.* también es *condicionada*, observándose que los factores de mayor variabilidad en el conjunto de las universidades norteamericanas quedan *fijados* en nuestro sistema (acceso alumnos, tasas, selección y salarios, profesores, oferta de estudios). No estaría de más introducir este dato a la hora de que los *expertos* elaboren los modelos de productividad y *ranking* de nuestras universidades. Por otra parte, la peculiar composición del Consejo de Universidades queda patente en el artículo 24, en el que se señala, entre otras cosas, que el trabajo de los Rectores de las Universidades es ocuparse de las cuestiones *académicas*: así se pone de manifiesto en el art. 24.4 que recoge la composición y funciones de sus Comisiones:

«Las Comisiones serán dos: una de Coordinación y Planificación y otra Académica.

a) La Comisión de Coordinación y Planificación, cuyo presidente será el del Consejo de Universidades, estará constituida por los responsables de enseñanza universitaria en los Consejos de Gobierno de las Comunidades Autónomas que hayan asumido competencias en materia de enseñanza superior y por aquellos miembros del Consejo de Universidades que el Presidente designe.

b) La Comisión Académica, cuyo Presidente será el del Consejo de Universidades o el miembro del mismo en quien delegue, estará constituida por los Rectores de las Universidades públicas y aquellos miembros del Consejo de Universidades que el Presidente designe...»

Tendremos ocasión de volver sobre algún otro aspecto específico de la autonomía universitaria que ponga de manifiesto la dificultad de homologar nuestra realidad universitaria con la existente en otros países. Por tanto, a la hora de hablar de calidad de nuestra universidad deberemos tener en mente, entre otros puntos, el marco político (autonomía vs. centralismo), el marco social (educación de masas vs. educación de élites), el marco institucional-organizativo universitario (abierto vs. cerrado), así como la escasa experiencia de análisis del fenómeno de la calidad universitaria.

1.4. *Beneficios y calidad de la educación superior*

Parece obvio afirmar que la preocupación por la calidad de algo está en consonancia con su relevancia y utilidad, es decir, de los beneficios, de toda índole y naturaleza, que produce. Por tanto, sería irrelevante ocuparse de la calidad universitaria, si estas instituciones no fuesen *beneficiosas* tanto para el individuo como para la sociedad. La duda sobre los beneficios de la educación es siempre una duda encubierta que se centra fundamentalmente en la eficacia y eficiencia del sistema. El problema estriba en determinar el contenido sobre el cual aplicar el análisis de beneficios. En definitiva, subyace el problema de los objetivos de la Enseñanza Superior, tema sobre el que en buena medida gira la problemática conceptualización de la calidad universitaria.

Mora (1991, pp. 28-31) esquematiza la tipología de beneficios de la Enseñanza Superior y recoge las aportaciones de Shattock (1990) en el Informe del Grupo de Evaluación de la Enseñanza Superior de la O.C.-D.E. A nivel individual, tres tipos de beneficios podemos considerar: *educativos* (de desarrollo y capacitación), *marginales* (se desprenden del prestigio de las instituciones que favorecen múltiples oportunidades) y *existenciales* (cúmulo de experiencias personales e intransferibles que el sujeto *vive*). A nivel social, las universidades realizan una gran contribución al *progreso social* (igualdad de oportunidades, educación permanente, influencias sobre otros niveles educativos, impacto sobre el entorno), al *crecimiento económico* (desarrollo tecnológico y formación profesional), y al *desarrollo cultural* (aportaciones a la ciencia, arte...).

Esta perspectiva nos obliga a reflexionar sobre el hecho de que la preocupación y ocupación por la calidad de las universidades no puede entenderse como ejercicio intelectual de los universitarios o como una dialéctica de confrontación entre poderes, aunque ésta exista, sino que ha de enmarcarse en la óptica del servicio público que han de prestar las universidades en el país. Así lo define nuestro marco jurídico.

2. **Calidad universitaria: concepto y dimensiones**

Plantear una aproximación conceptual al término calidad de la educación universitaria es adentrarse por senderos de indefinición y contro-

versia; no ya sólo por la inconcreción o relatividad del término calidad, reflejada en la ya conocida y utilizada cita del filósofo R. Pirsig ³, sino por la dificultad de encontrar consenso en cuanto a la definición del *producto* de la educación universitaria. Por otra parte, y como señalan Os *et al.* (1987), refugiarse en la dificultad de la definición no es sino evitar una discusión abierta sobre temas más concretos. Por ello, y con la advertencia que no daremos solución a la imprecisión, vaguedad o multisemia del concepto de *calidad de la educación universitaria*, nuestro objetivo estriba en aportar unas ideas y reflexiones que permitan al lector disponer de unos elementos que, al contribuir al desarrollo de su propio marco conceptual, le faciliten abordar un tratamiento más complejo del tema. Si esto aconteciera, podría decirse que estas páginas participarían del concepto más genérico de calidad educativa: conseguir que la persona sea capaz de acometer tareas cada vez más complejas.

La propia *Enciclopedia of Educational Research* señala que toda definición de *calidad de la educación* será poco precisa debido a las múltiples acepciones de la misma (calidad en el proceso, en el producto, en la consecución de metas, en la adecuación de las acciones a un fin, en la eficiencia, etc.). Los siguientes testimonios, extraídos de nuestro propio contexto, ahondan en la misma dirección:

«La noción de "calidad de la enseñanza" constituye uno de esos lugares críticos de toda investigación de carácter empírico en el que lo que se desea medir o estimar queda radicalmente definido "a priori" por el instrumento de medida o estimación por el que se opta» (Fernández Pérez, 1990, p. 17).

Comenta el profesor M. Fernández que si ya no es posible cerrar la noción de enseñanza de calidad en un colectivo de profesores que trabajan codo a codo, cómo podrá cerrarse en el amplio y heterogéneo contexto de la universidad española.

3. «Quality... you know what it is, yet you don't know what it is. But that's self-contradictory. But some things are better than others. That is they have more quality. But when you try to say what quality is, apart from the things that have it, it all goes poof! There's nothing to talk about. But if you can't say what quality is, how do you know that it even exists? If no one knows what it is, then for all practical purposes it doesn't exist at all. But for all practical purposes it does exist. What else are the grades based upon? ...Obviously, some things are better than others... but what's the *betterness*... So round and round you go, spinning mental wheels and nowhere finding any place to get traction. What the hell is Quality? What is it?» (Pirsig, 1974, p. 179).

Santos (1990) señala que *calidad de enseñanza* es un tópico que se maneja con pretendida univocidad. El problema aparece en el momento de precisar en qué consiste la calidad:

«No puede entenderse la calidad de la enseñanza de espaldas a las exigencias constitucionales, ideológicas y técnicas, que parten de una concepción de la Universidad, atenta a la reconstrucción del conocimiento científico, a la investigación de carácter básico y a la preparación de personas que, desde procedencias desiguales y expectativas diferentes, quieren realizar unos aprendizajes encaminados a la capacitación profesional y al enriquecimiento personal y social» (pp. 49-50).

Por otra parte, afirma el profesor Santos, es necesario plantear una reflexión acerca de si se puede establecer y asumir la igualdad: *calidad de enseñanza = calidad de aprendizaje*. Con excesivo optimismo se afirma la relación de calidad entre enseñanza y aprendizaje. Asimismo, atender a un criterio de calidad cifrado en resultados académicos externos a la misma universidad no es muy adecuado.

«Decir que una Universidad es mejor que otra porque ha obtenido mejores calificaciones en pruebas homologadas exteriores es una afirmación, cuanto menos, aventurada» (pp. 64-65).

Para ello habrá que tener en cuenta las condiciones de partida, los medios, el contexto, etc. En definitiva, como apunta De Weert (1990), la calidad no puede hacer exclusiva referencia a los logros y resultados, sino también a las entradas y a los procesos.

2.1. *La calidad y el rendimiento de los elementos de la organización*

Kurz *et al.* (1989), adoptando el marco de referencia del rendimiento de las organizaciones en el ámbito de la salud, plantean una triple perspectiva desde la que consideran que puede conceptualizarse el rendimiento y, por tanto, la calidad de una organización.

- *Sistema racional*: Enfatiza el logro de objetivos, poniendo su atención en dos aspectos: la calidad del producto —eficacia— y en la economía de la producción —eficiencia—. Dentro del continuo de este sistema, y simplificando, tal vez pudiera hablarse de dos extremos en la aproximación teórica al concepto de calidad universitaria:

a) *El modelo de producción industrial*. En ella se asume que el proceso educativo es un proceso semejante al de producción industrial, no

habiendo razones para establecer diferencias significativas entre una universidad y una fábrica (Church, 1988). La calidad universitaria sería semejante a la calidad de los profesionales producidos para satisfacer las necesidades técnicas del sector productivo y de servicios. Se asume que pueden especificarse con claridad las especificaciones del *producto*, así como contar con un fundamentado modelo explicativo del proceso de transformación de los *inputs* (alumnos) en *outputs* (graduados).

b) *El modelo de desarrollo global* de los elementos personales de la institución universitaria. En ella se asume que la calidad de una universidad estaría determinada por el *valor añadido* que consigue generar en sus alumnos, profesores y personal de administración y servicios, en cuanto a su progreso y desarrollo global como personas. Aunque esta vía parece ofrecer una positiva prospectiva (Astin, 1985; Miller, 1986), participa de parte de la problemática planteada en el modo de producción: ¿cuáles son los ingredientes o dimensiones que definen el desarrollo personal?

Para Kurz *et al.* (1989), la ambigüedad en cuanto a los estándares de deseabilidad del *producto universitario*, así como el parcial conocimiento que se tiene de las relaciones causa-efecto en el *proceso de producción*, hace que la perspectiva racional sea de dudosa aplicación; sobre todo cuando se adopta como exclusiva o no se tiene en cuenta el *momento* de desarrollo en que se encuentra la organización.

● *Sistema natural*: Examina el rendimiento en relación al mantenimiento del sistema y a su soporte social. La universidad, afirman los citados autores, participa de un *modelo orgánico de organización*, caracterizado por:

- Alta concentración de especialistas.
- Baja centralización en las estrategias de toma de decisión.
- Baja estratificación en cuanto a niveles retributivos (este punto puede presentar grandes diferencias según los sistemas nacionales).
- Elevado volumen de comunicación.
- Alta comunicación horizontal.
- Equipos multidisciplinares.
- Interdependencia entre la investigación y la producción.

Este tipo de modelo organizativo se da, en parte, cuando la organización tiene que afrontar la demanda creciente de mayor calidad e individualización en el producto o servicio que presta. Para dar respuesta a esta demanda deberá incrementar la complejidad de las tareas que realizan sus miembros, reclamándose una mayor cohesión de los mismos. En la Universidad, muchos de sus cometidos tienen estas notas (sobre todo en universidades con cierto nivel de desarrollo): el trabajo en equipo, mantenimiento de servicios, constante revisión de programas reclaman una dinámica institucional que difícilmente puede tener estándares prefijados de calidad. Por lo tanto, la *indefinición* y lo *implícito* serán aspectos con los que habrá que contar.

- *Sistema abierto*: Preconiza que la adquisición de recursos por una determinada *unidad de la organización* (departamento, facultad...) mejora su posición de poder, incrementando su influencia dentro de la organización, a la vez que produce efectos positivos en la posición de la organización en su contexto. Pfeffer y Salancik (1978) aportan evidencia empírica de la relación entre adquisición de recursos y el poder de los departamentos o grupos. Por tanto, y atendiendo al continuo del proceso de desarrollo en el que los sistemas abiertos coinciden con niveles iniciales del mismo, podrá considerarse la calidad desde la perspectiva de la simple adquisición de recursos, estableciéndose como sinónimo de incremento de calidad el simple incremento de medios. Los riesgos de esta perspectiva son obvios: la posible espiral del nivel de satisfacción por la dominación dentro de la organización a expensas de la atención a los usuarios externos de la misma (alumnos, sociedad).

Como puede desprenderse de lo expuesto, cada uno de los sistemas puede asociarse a un determinado grado de desarrollo de la organización abierto-inicial, natural-medio y racional-superior. Cada uno de ellos puede aportarnos elementos de interés que son integrados en un concepto dinámico de la calidad universitaria. No es aventurado, por tanto, tomar un marco de referencia adoptado por Westerheijden (1990a) y que formulamos en la siguiente proposición:

«La calidad universitaria es un concepto relativo y multidimensional en relación a los objetivos y actores del sistema universitario. Su análisis debe hacerse en el contexto de los procesos sociales y políticos en que interactúan objetivos y actores.»

2.2. *La calidad desde los enfoques evolutivos*

Desde la perspectiva de la universidad como un sistema organizativo (racional, natural o abierto) el logro de la calidad (rendimiento) está determinado tanto por las acciones individuales como por la interacción de los individuos con sus unidades organizativas. Por tanto, es totalmente reduccionista aproximarse a la evaluación de la calidad a través de la evaluación individual de los agentes. Es necesario un análisis del rendimiento institucional, así como de los factores organizativos y ambientales que contextualicen posibles y posteriores análisis individuales.

Ahora bien, en el momento que se acomete el análisis funcional de una institución, es decir, la adecuación del funcionamiento de la misma a los objetivos, no sólo nos encontramos con divergencias operativas de los objetivos asumidos, sino también con la falta de consenso en cuanto a los indicadores de logro de los mismos. Por esta razón, hemos de avanzar en la concreción del marco último definitorio de la calidad. Esta concreción sólo es posible al adoptar un determinado modelo o enfoque evaluativo de la calidad de la educación universitaria en un momento y contexto determinado. La siguiente proposición cierra el marco en el que abordaremos las múltiples notas que pueden componer, según las perspectivas adoptadas, el concepto objeto de estudio.

«No se dá ni podrá darse una única conceptualización de la calidad universitaria. El único posible nivel de concreción será el que se adopte a partir del enfoque evaluativo desde el que se evalúe dicha calidad.»

El conjunto de los trabajos recogidos en esta publicación, a través de su aproximación evaluativa de la calidad, constituyen piezas importantes para componer una amplia y precisa imagen del concepto de calidad de las instituciones universitarias, pese a que entre ellos puedan darse ciertos elementos de comunidad. Mora (1991), referenciando los trabajos de Georges (1982) y Astin (1985), resume en cinco enfoques evaluativos la conceptualización de la calidad universitaria:

- La calidad como reputación.
- La calidad como disponibilidad de recursos.
- La calidad a través de los resultados.
- La calidad por el contenido.
- La calidad por el valor añadido.

3. Relatividad y multidimensionalidad de la calidad universitaria

La primera proposición señalada toma en consideración un doble análisis: el de los objetivos de la universidad y el del marco organizativo y sus respectivos agentes. La diversidad de objetivos de carácter institucional, unida a la diversidad de intereses de los diferentes *estamentos* universitarios (desde las autoridades gubernamentales al último nivel de los servicios administrativos, pasando por los distintos niveles académicos), conforman una compleja red en donde se diluye o queda atrapado cualquier concepto unívoco de calidad. Aproximémonos a esta complejidad.

3.1. Calidad en relación a los objetivos

Desde esta perspectiva puede afirmarse que el elemento característico de la calidad es la proporción o magnitud de los objetivos alcanzados o, dicho de otra forma, el grado en que un determinado *producto* hace lo que se ha propuesto que haga (Westerheijden, 1990a). En consecuencia, la calidad es un *concepto relativo* que sólo puede definirse operacionalmente en relación a un conjunto de objetivos previamente definidos. En el caso que nos ocupa, los objetivos de la educación universitaria.

El análisis de los objetivos de la educación universitaria pone de manifiesto, entre otros puntos, los siguientes:

- No se explicitan o se da un elevado nivel de inconcreción. Así aparece desprenderse del análisis de los marcos jurídicos en diferentes países occidentales (Pollit, 1990). Si aceptamos la similitud entre fines y objetivos, nuestro marco jurídico señalaba:

«La Educación Universitaria tiene por finalidad:

1. Completar la formación integral de la juventud, preparar a los profesionales que requiera el país y atender al perfeccionamiento en ejercicio de los mismos.
2. Fomentar el progreso cultural, desarrollar la investigación en todos los niveles con libre objetividad y formar a científicos y educadores.
3. Contribuir al perfeccionamiento del sistema educativo nacional, así como al desarrollo social y económico del país» (art. 30 de *Ley General de Educación, 1970*),

habiéndose producido escasas o nulas modificaciones:

«Son funciones de la Universidad al servicio de la sociedad:

- a) La creación, desarrollo, transmisión y crítica de la ciencia, de la técnica y de la cultura.
- b) La preparación para el ejercicio de actividades profesionales que exijan la aplicación de conocimientos y métodos científicos o para la creación artística.
- c) El apoyo científico y técnico al desarrollo cultural, social y económico tanto nacional como de las Comunidades Autónomas» (art. 1.2, *Ley de Reforma Universitaria, 1983*).

Mora (1991, pp. 25-27 recoge el listado de metas/objetivos de diferentes sistemas de enseñanza superior formulados en trabajos como el *Robbins Report, 1963*), *The Purposes and the Performance of Higher Education in the United States* (Carnegie Comisión, 1973) o el informe *The University and the Community* del Centre for Educational Research and Innovation (O.C.D.E., 1982). No aportan especificaciones significativas: las declaraciones gubernamentales, en ocasiones, se parecen como las gotas de agua, aunque los sistemas educativos estén en muy diferentes niveles de desarrollo.

- Los ámbitos o dimensiones de estos fines u objetivos se identifican fácilmente. Medina (1985) en su análisis de los fines formativos de la Universidad, a través de distintos pensadores y sistemas, señala la constante aparición de los fines en relación a: *la formación cultural y humana, la enseñanza o transmisión crítica de la ciencia, la investigación y la formación profesional*. Lo que no aparece es la asignación de igual importancia o, incluso, la posibilidad de compatibilizarlos. Así, el profesor Morente caracterizaba cuatro tipos de Universidad: *docente, educativo, científico y profesional*, afirmando:

«Los tres primeros se determinan esencialmente en función del concepto que se tenga del sentido y el valor de la ciencia. El cuarto es una perversión del ideal universitario, causado por el abuso del poder central del Estado» (cit. en Medina, 1985, p. 355).

- La diversa naturaleza de objetivos asumidos o impuestos a la Universidad es un tema claro (Cave *et al.*, 1988; De Weert, 1990), así como el hecho de que tradicionalmente han sido agrupados en tres dimensiones: docencia, investigación y servicios públicos (extensión universita-

ria). Si bien con desigual atención y consecuencias para cada una de ellas: difuso planteamiento de la calidad de la docencia, atención a la cantidad de la investigación y escasa atención a los objetivos de extensión y servicios universitarios.

3.2. *Calidad en relación a los actores*

El sistema universitario es complejo y en el mismo existen múltiples niveles —de muy diferente naturaleza y jerarquía— en donde interactúan diferentes agentes o actores con puntos de vista y objetivos diferentes, y con capacidad de decisión en diferentes niveles y ámbitos de competencias.

La naturaleza de la propia organización universitaria pone de manifiesto que no es viable plantear la hipótesis de un consenso o adhesión de un nivel o estamento de los agentes de esa organización a la propuesta de objetivos de otro. No será difícil constatarlo en temas como:

- Procesos de selección del profesorado.
- Criterios de admisión de alumnos.
- Volumen y distribución de los recursos.
- Naturaleza del control presupuestario.
- Promoción y apoyo a la investigación.
- Roles y funciones de los órganos de gestión, etc.

Podría afirmarse que en el análisis de la calidad de una institución universitaria habrá de tenerse en cuenta el concepto que de la misma sustentan o se desprende de:

a) Los estamentos académico y administrativo. No sería ofensivo para nadie el señalar la real pugna de poder (por otra parte lógica y comprensible) que subyace entre estos dos estamentos y en sus diferentes niveles (Rectorado-Gerencia... Decano-Jefe de negociado, etc.). La falta de preparación para la gestión de los cargos académicos, así como la lógica transitoriedad de los mismos, hace realidad la conocida anécdota del ministro y el conserje: ¡otro que viene mandando; el pobre

no sabe que él se irá y yo me quedaré! Por otra parte, la complejidad de la gestión, derivada de la relativa autonomía de decisiones, pone en evidencia el peso significativo del sistema de administración y servicios de las universidades.

b) La organización jerarquizada del sistema social y político deja su impronta en el marco universitario, planteando el eterno dilema de autoridad gubernamental frente a la autoridad universitaria (Ministerio, Comunidad Autónoma, Universidad). Dentro de la propia Universidad vuelve a repetirse el esquema (Rectorado, Facultades, Departamentos).

c) El sistema de financiación de la Universidad. Todos admitimos la presión ejercida por quien controla los fondos. Si en la universidad privada el usuario directo puede ser un agente activo en la definición de la calidad, en la universidad pública es el gobierno quien actúa como portavoz del mismo. El hecho de que entre el 72 %-73 % esté la media de la financiación gubernamental de las universidades españolas, unido al articulado de la *LRU* que regula la política presupuestaria de las universidades, pone de manifiesto la posición preponderante de las autoridades políticas en la definición *indirecta* de la calidad. Si afirmo lo de indirecta es debido a que no se puede hablar de una directa o inmediata asunción de los planteamientos políticos por el propio sistema universitario. Véase si no lo acontecido con el contenido de las pruebas del acceso a los cuerpos de profesores ordinarios. Sin embargo, sí que tendremos que admitir que la cuestión de la financiación ha sido el *arma* utilizada para promover una calidad universitaria basada en la eficiencia, sobre todo a partir de los primeros años de la pasada década.

d) La dimensión científica basada en las disciplinas (Matemáticas, Medicina o Sociología, más que en las Areas de Conocimiento) presenta diferentes niveles de homogeneidad y, consiguientemente, de poder, no ya sólo en el contexto global de un sistema universitario nacional, sino dentro de una determinada universidad. Como señala Spaapen (1989) esta dimensión aparece en la universidad predominantemente al nivel de departamento o grupos de investigación. En nuestra realidad, tal vez no sea fácil en estos momentos determinar en qué nivel deberíamos colocar esa dimensión científica: ¿Facultad, Area de conocimiento, Departamento? La prueba evidente de esta indefinición podemos tenerla tanto en el proceso de elaboración de las propuestas de nuevos planes de estudio, como en el *modelo* de evaluación de la actividad investigadora.

e) Los alumnos, en tantas ocasiones voz silenciosa o actores pasivos, deben ser protagonistas en el proceso de conceptualización de la calidad universitaria, o, mejor dicho, de la explicitación de su propia dimensión de calidad. No habrá de extrañarnos, simplificando, la existencia de discrepancias entre objetivos del profesorado y del alumnado con respecto a los requisitos y desarrollo de la actividad académica. El problema en estos momentos es la *aparente comunión de ideales*: sólo mueve a la acción aspectos muy puntuales, estando ausente un serio y sereno debate sobre la dirección de la Universidad en esta década. No sería de extrañar que profesionalización y desarrollo personal y cultural fuesen los dos grandes objetivos propuestos por estos *actores*.

f) Finalmente tomemos en consideración a los empleadores. Este punto es controvertido y su análisis detallado nos sobrepasaría en estos momentos. Ahora bien, sí que considero importante tener presente la reflexión del profesor Carabaña (1984):

«Los mismos que hace veinte años propugnaban el desarrollo del *capital humano* como premisa del desarrollo económico, insisten hoy en deshacerse de la gran parte de ese capital humano para el que desesperan de encontrar empleo productivo. *Se trata, claramente, de una opción ideológica, disfrazada de exigencia económica con la amenaza de otro milenio apocalíptico más. Sea cual sea el futuro, los pueblos estarán mejor preparados para enfrentarlo si cuentan en abundancia con los conocimientos y los hábitos de reflexión y crítica que, por mal que lo hagan, cultivan hoy las instituciones escolares*» (p. 45).

Por otra parte, no deberemos olvidar dos aspectos de interés: el papel de empleador que el sector público tiene en relación a los graduados universitarios; en determinadas áreas es empleador casi único. El segundo aspecto hace referencia a la concepción o cultura empresarial española sobre el qué, quién y cómo de la formación específica para el puesto de trabajo. Cuán difícil resulta convencer de la importancia de una mente bien formada frente a una mente bien llena.

Examinemos ahora los objetivos de la universidad a la luz de la perspectiva de su interacción con los diversos actores del sistema universitario.

4. **Objetivos del sistema universitario: la perspectiva de los actores**

La perspectiva de análisis que vamos a adoptar no va a seguir la ortodoxa compartimentación de Docencia, Investigación y Servicios, sino que vamos a plantearnos un marco de relaciones entre objetivos y agentes universitarios que sobrepase esa trilogía, aunque la misma esté presente en cada uno de los objetivos sometidos a examen.

4.1. *Educación universitaria de masas. La igualdad de oportunidades*

Una serie de hechos (Vázquez, 1985) ponen de manifiesto el cambio estructural —cualitativo y cuantitativo— de la población de acceso a la universidad que han producido su masificación en los países occidentales. Fenómenos como los del acceso de la mujer, la incorporación de grupos minoritarios (multiraciales, inmigrantes, mayores de edad), el *baby boom* (ya pasado), el descenso de la mortalidad estudiantil (no siempre se han establecido los correctos mecanismos de selectividad), el incremento del porcentaje de alumnos que continúan sus estudios de segundo ciclo, junto con factores como los del establecimiento de turnos nocturnos o la puesta en marcha de la *universidad abierta*, han marcado la fisonomía de esa educación superior de masas. Menos consistencia puede tener el argumento aducido por algunos políticos de que al necesitar los alumnos un excesivo número de años para concluir sus estudios (debido a su bajo rendimiento), las universidades se llenan de repetidores.

Ahora bien, sí es un hecho innegable el incremento absoluto de los estudiantes de enseñanza superior. No existe el mismo consenso al establecer el diagnóstico sobre los resultados obtenidos en cuanto a la política de igualdad de oportunidades. Valgan estos tres testimonios:

«Los políticos en favor de la igualdad de oportunidades... no intentaron resolver la contradicción fundamental que existe entre la búsqueda de la igualdad en educación y la persistencia de la desigual distribución del trabajo, la riqueza y el poder» (Kallen, 1987; pp. 229-230).

«Se llegó a la conclusión de que el sistema educativo no puede actuar como sustituto de las reformas sociales y económicas. No es posible tener en la educación más igualdad que la que existe en la sociedad en general» (Husén, 1988, p. 136)

Por otra parte, la contención de la demanda de enseñanza superior es un tema controvertido, pese que por parte de algunos estudiosos españoles la posición a adoptar no tiene duda:

«La contención de la oferta universitaria no puede ser completa ni indefinida. La razón de ello es muy simple: la demanda social de educación debe ser considerada a largo plazo como *incontenible*, porque no depende sólo del sistema ocupacional. Su impulso radica tanto en oportunidades ocupacionales, presentes o esperadas, como en aspiraciones de cultura y de estatus social» (Carabaña, 1984, citando a V. Pérez Díaz, 1981: *Universidad y empleo*).

Hace ya tres lustros que el prestigioso profesor sueco Husén (recop. 1988), en la Tercera Conferencia Internacional sobre Educación Superior, al tratar el dilema entre igualdad y excelencia de la educación superior, dijo:

«La educación superior orientada y relacionada con la investigación está a punto, en algunas instituciones, de dar paso a una educación superior orientada más profesionalmente, lo que en la educación de los no licenciados adopta las formas de un "campo de instrucción"» (p. 132).

«El crecimiento de la universidad en función tanto de la cantidad de matriculación durante la transición del sistema elitista al masivo, como de la diversidad de funciones, ha provocado tensiones manifiestas dentro de la institución. La tensión más evidente es la que existe entre la educación universitaria y para la investigación por un lado y la formación profesionalmente orientada por los no licenciados por el otro» (p. 133).

Se plantea, por tanto, la tensión entre el sistema de investigación y el de formación de la universidad de masas. Es en éste último donde la calidad de la docencia universitaria ha adquirido su máxima expresión. La salida parece estar en adoptar un punto de vista que permita el desarrollo hacia una multifuncionalidad de la enseñanza superior, asumiendo las consecuencias que se derivan en el acceso a la misma. En este sentido, el Conseller de Educació de la Generalitat de Catalunya, profesor Laporte (1990), afirma:

«Creo que ahora es absolutamente inútil oponerse a la masificación a no ser que en este país se establezcan también vías de enseñanza post-secundaria no universitarias... en estas circunstancias sólo nos queda un remedio, que es la diversificación total de la Universidad. Es decir, la Universidad si ha de atender a todos ha de ser diversa, no todas las universidades han de ser iguales entre ellas ni todos los centros o estudios que se imparten en la Universidad han de tener el mismo nivel de exigencia» (pp. 14-15).

Es difícil vaticinar las posibilidades que tiene esta idea de ser operativa y quedar reflejada en el diseño del mapa universitario de Catalunya. Presiones o aspiraciones legítimas de muy diversa índole pueden alterar la racionalidad de las decisiones. Tampoco debe olvidarse la *filosofía* que subyace en determinados planteamientos de la educación superior⁴.

4.2. *Educación para la profesionalización*

Como apunta Husén (1988), en una época de desequilibrio entre puestos de trabajo y demandantes de empleo, es fácil concebir la Universidad básicamente como una institución para la formación profesional específica y olvidar los valores educativos que no necesariamente están vinculados a la dimensión profesionalizadora. Sin embargo, es necesario señalar que la inmensa mayoría de los puestos de trabajo de nivel superior reclaman competencias específicas que corren a cargo del empleador. En los puestos de trabajo de alta cualificación no sólo son necesarias capacidades de carácter técnico, sino que se reclaman aptitudes para la toma de decisiones e iniciativa para abordar situaciones complejas. En consecuencia, y para este último tipo de competencias, unos programas de estudio que atiendan a una formación más

4. Una de las propuestas de reforma educativa elaboradas en los Estados Unidos en los primeros años de los ochenta y publicada en castellano (Grupo Paideia: *Manifiesto Educativo*; Adler, 1986), se señala que en las instituciones de enseñanza superior: escuelas y colegios universitarios, universidades y escuelas técnicas las finalidades del primer ciclo universitario deberían ser:

- La preparación para una profesión que requiere un conocimiento especializado y una capacitación técnica. Determinadas profesiones requerirán, por supuesto, estudios ulteriores. La amplia oferta de programas de carácter técnico no debería ser óbice para la exigencia obligada de cursar un programa *menor* de carácter liberal y humanístico que elevaría el nivel cultural general obtenido en la enseñanza básica obligatoria.
- La realización de un aprendizaje general para estudiantes de más edad y que podrían basarse en su escolarización básica para llevar a cabo un trabajo más avanzado.

Quienes acudan a una universidad exclusivamente para hacer progresar su educación general deberían buscar programas de primer ciclo concebidos para cumplir ese objetivo. Se precisan programas en donde la rama principal de estudios sea común a todos con opciones escasas o nulas. Estos centros serían instituciones ideales para la preparación de los profesores de la escuela básica.

comprensiva y abierta permitirán una mayor adaptabilidad a situaciones transitorias de empleo.

Pérez Díaz (1981), Alonso (1984) y Gutiérrez (1984) han puesto de manifiesto el desajuste entre formación de titulados y demandas sociales (¿productivas?). Desde la perspectiva de estos y otros autores la calidad podría equipararse al criterio de valor (frente al de mérito) explicitado por Guba y otros; es decir, podríamos hablar de calidad de un sistema universitario en la medida que adecúa *su producción* a la demanda de *consumo*.

Para los defensores de este concepto de calidad (al que evidentemente añadirán el calificativo de *buena preparación profesional*) es obvio que el objetivo primordial de la universidad es el de la generación de una cualificada y no excedentaria mano de obra.

4.3. *Calidad y planificación*

Sin duda alguna el tema se complica cuando desde el punto de vista del actor social (gobierno, sociedad o sector productivo), se quieren anuar el objetivo de masas con el de profesionalización, manteniendo un sistema estructural y de planificación idéntico al de los años cincuenta. Toma aquí pleno sentido el concepto de calidad desde la perspectiva de la planificación de los medios del sistema universitario.

Si la *open university* aportó una significativa innovación en la estructura y funcionamiento de la enseñanza universitaria; en la última década el *sistema presencial*, sometido a múltiples presiones (una de ellas la económica), se ha visto en la necesidad de reestructurar o poner en marcha *nuevas formas* de satisfacer la demanda de la educación superior de unos nuevos tipos de población. Los trabajos discutidos en el 29.º Seminario del Centro para la Investigación e Innovación Educativa — O.C.D.E.— (París, 6-7 diciembre) apuntan claramente a un continuo incremento porcentual en el acceso a estudios superiores de grupos de población diferentes a los graduados en la escuela secundaria. El estudio del ministerio de educación holandés (Oijen, 1990) pone de manifiesto las consecuencias de los cambios demográficos y de nuevas demandas en su sistema de educación superior. En él se apunta la predicción de que en el plazo no superior a 40 años, el 30 % de la mano de obra del país tenga educación superior. El título del seminario (*Strate-*

gies for university planning meeting the needs of a new clientele), dentro del Programme on Institutional Management in Higher Education de la O.C.D.E., es suficientemente indicador de los nuevos retos que se le plantean a sistemas universitarios, como el español, que han permanecido con una anquilosada estructura pese al proceso de masificación sufrido.

El profesor Stern (1990), decano en Berkeley, enfatiza la necesidad de *reordenar* las estructuras de planificación y funcionamiento de la universidad para dar respuestas a: programas especiales para adultos, a programas de estudios presenciales pero a tiempo partido, o a programas de postgraduados ⁵. Un dato tal vez anecdótico para algunos: en el inicio del curso 87-88, el 41 % de los estudiantes de *colleges* americanos lo eran a tiempo parcial; el porcentaje llegaba al 64 % en los de estudios graduados. La Universidad de Harvard tenía 14.000 estudiantes en el programa especial de adultos (educación a distancia + tiempo presencial).

En resumen, la calidad de un sistema o de una institución de educación superior radica *también* en su capacidad de planificar y adecuar recursos a nuevas demandas o situaciones.

La innovación organizativa de nuestras universidades es una más de las asignaturas pendientes que deberían aprobarse en una muy próxima convocatoria, pese a lo difícil que será si se sigue dependiendo de *otros*

5. A simple título de curiosidad (Stern, 1990): el estado de California ha seguido durante 25 años un *Master Plan* de organización de su enseñanza universitaria. El Plan comprende tres sectores:

- University of California: 9 *campus*. Con énfasis en la investigación y en los estudios postgraduados. Con un presupuesto para el curso 88-89 de 6.000 millones de dólares, con aportación pública del 35 %.
- State College University: 20 *campus*. No tiene doctorado y la investigación es muy limitada. Su énfasis está en los estudios a tiempo parcial. Un presupuesto de 2.400 millones de dólares, con una aportación pública del 55 %.
- Community College: 107 *campus*. Con estudios a tiempo parcial y programas informales. Un presupuesto de 2.300 millones de dólares.

Una pregunta ingenua: ¿se podría pensar en *Planes* por Comunidades Autónomas? Evidentemente que no me limito a utilizar el término *Plan* como la simple proliferación de centros universitarios.

órganos de decisión. El mérito de la calidad, en relación a la planificación, no será de las universidades:

«Las Comisiones [del Consejo de Universidades] serán dos: una de Coordinación y Planificación y otra Académica.

a) La Comisión de Coordinación y Planificación, cuyo Presidente será el del Consejo de Universidades, estará constituida por los responsables de enseñanza universitaria en los Consejos de Gobierno de las Comunidades Autónomas que hayan asumido competencias en materia de enseñanza superior y por aquellos miembros del Consejo de Universidades que el Presidente designe...» (art. 24.4. LRU, 1983).

4.4. *El matrimonio docencia-investigación*

La relación entre docencia e investigación ha sido objeto de continuo debate y análisis (Kasten, 1984; Crawshaw, 1985; Elton, 1986). Dado que estos dos ámbitos constituyen el marco principal de actuación del profesor, así como los objetivos más importantes de todas las universidades, no es de extrañar la extensión que inevitablemente genera esta relación; máxime cuando ciertas investigaciones (Marsh, 1984) han puesto de manifiesto la existencia de una relación negativa. Está claro que esta relación negativa surge al operativizar ambas variables en la opinión de los alumnos sobre la calidad de la docencia impartida por el profesor y la productividad de publicaciones. De todas formas, la tesis de que docencia e investigación han de caminar unidas en la individualidad de cada profesor universitario (Medina, 1986) es objeto de controversia, y la realidad pone de manifiesto el deterioro de la primera en aras de la segunda.

De otra parte, está surgiendo en algunos sistemas europeos la práctica de acotar centros de investigación en el propio seno de las universidades, amén de la organización de la investigación en cuanto a su ubicación en determinados *campus* (vid. nota 5). El profesor Husen (1988) afirma:

«Hay que señalar a los que creen que investigación y enseñanza siempre estuvieron unidas en un matrimonio armonioso y que es imposible la enseñanza universitaria sin un estrecho contacto con la investigación activa, que el «matrimonio» es relativamente reciente en los países en que se celebró. Se produjo por primera vez en Alemania en el siglo XIX y esta alianza se copió en otros países, incluso Estados Unidos» (p. 145).

Si a esto se añade que por motivos financieros en casi todos los países que han vivido la explosión de matrícula la enseñanza basada en la investigación ha de limitarse a unas pocas instituciones, tal vez pueda parecer duro el planteamiento de Husen. Mas la realidad de la mediocridad reviste mayor dureza desde la perspectiva social.

4.5. *La perspectiva económica de la calidad*

Solmon (1989) señala que la perspectiva económica de la calidad tiene mucho de tautológica, ya que considera que la calidad de las instituciones o de los programas es mejor cuando sus impactos positivos, tanto en los individuos como en la sociedad en general, son mayores.

No es necesario investigar para llegar a esa conclusión: son precisamente los impactos los que han servido para determinar la calidad de la institución. La perspectiva económica de la calidad debería responder, afirma Solmon (p. 689), a otra cuestión diferente:

«...cuál es el coste de esta calidad superior y cuál el de los beneficios adicionales, para saber así si la tasa de rendimiento de los gastos realizados para desarrollar una enseñanza de calidad superior es comparable a las tasas de rendimiento de otros usos alternativos de estos fondos».

No es difícil argumentar la dificultad de prueba de esta propuesta. Es prácticamente imposible identificar y evaluar toda la gama de beneficios que se derivan de la asistencia a centros de distinta calidad. De aquí que, con demasiada frivolidad, los economistas hayan globalizado el *impacto positivo* en el indicador de los salarios de los egresados de una determinada universidad. A partir de este vulnerable indicador (desde una perspectiva del desarrollo global del estudiante), las investigaciones han arrojado cierta unanimidad en cuanto a las variables asociadas al efecto calidad-salario, siendo precisamente esas variables las utilizadas para establecer los conocidos y criticados *rankings* de las universidades norteamericanas: gastos por alumno, salarios de los profesores, puntuaciones en las pruebas de admisión, tamaño y tipo (*community college*, *college*, universidad) de la institución, y no muchas más (Foster y Rodgers, 1980; Solmon, 1981; Astin, 1985).

Es obvio advertir que la fundamentación última de la relación entre educación e ingresos se sustenta en una de las consecuencias de la

teoría del capital humano, según la cual la educación —y por tanto su calidad— aumenta la productividad (Chiswick, 1973). La linealidad y simplicidad de la perspectiva económica de la calidad queda reflejada en el trabajo de Foster y Rodgers (1980), quienes, al resumir las investigaciones al respecto, concluyen:

«...la calidad, sin duda, es importante. Ahora bien, esto no significa que su importancia se deba a los conocimientos y habilidades adquiridos por los alumnos. Significa, simplemente, que los alumnos de los buenos centros aprenden mejor a adaptarse a las normas sociales, o que los empresarios prefieren a los alumnos de este tipo de instituciones. Se precisan estudios que nos muestren por qué asistir a centros de calidad significa obtener ingresos superiores. Por ahora baste saber que sucede así» (p. 33).

5. La calidad como reputación o fama de las instituciones

Nadie se hubiera atrevido a pronosticar que sobre el indicador *tramos de investigación concedidos* se iba a construir la primera medida de *calidad* de las universidades españolas, a partir de la cual se daría cuenta a la sociedad del nacimiento histórico del *ranking de las universidades españolas*. El 12 de enero de 1991 (vid. *El País*) entrábamos en el *circo de las clasificaciones*, emulando así al sistema norteamericano.

Astin, en su trabajo-resumen de 1985, *Achieving Educational Excellence*, afirma que puede ser considerado un fenómeno psicológico la tradición de jerarquización de las instituciones americanas de educación superior:

«Existe en las mentes de los educadores y de muchas personas legas una serie de creencias compartidas (un tanto folclóricas, si se quiere) acerca de cuáles son las instituciones mejores o de mayor calidad. Este folclore, el cual constituye la base para la jerarquía institucional en la educación superior americana, es también la fuente de la calidad como reputación social. De acuerdo a este punto de vista, la calidad es todo aquello que la gente cree que es» (p. 25).

De los múltiples trabajos realizados por Astin y colaboradores sobre el tema unas claras conclusiones se obtienen:

a) Las quince instituciones (de un total de tres mil) situadas en la parte superior llevan décadas en dicha posición.

b) La fiabilidad de los diferentes *ratings* realizados por profesores de diferentes instituciones y áreas es muy alta; la creencia es firme.

c) El folclore de la jerarquización crea un efecto de halo, pues se mantiene la misma puntuación para aspectos o programas muy diferentes.

d) Las puntuaciones otorgadas por los *jueces* a cada institución están influidas por la familiaridad (conocimiento) que tenían de éstas.

e) La reputación de los programas graduados está fuertemente influenciada por el tamaño del programa y de la institución.

f) Las puntuaciones de calidad de los diferentes tipos de estudios dentro de una institución presentan escasa variabilidad.

g) Los *ratings* son muy estables a lo largo del tiempo.

h) Los análisis multivariados de los *ratings* ponen de manifiesto que las puntuaciones de selectividad en el acceso, el gasto por alumno y el número de departamentos ofreciendo doctorados son las tres variables que ofrecen una más que adecuada estimación de la puntuación global de calidad obtenida en los estudios a nivel nacional.

i) Los análisis ponen igualmente de manifiesto que sólo dos dimensiones de la calidad son consideradas por los *jueces*: la calidad científica de su profesorado —*scholarly excellence of faculty*— y el compromiso con la docencia del profesorado, así como su pedagogía e innovación curricular.

6. La dimensión *objetiva* de la calidad

Dentro de la *filosofía* o folclore del *ranking*, expertos y administraciones, afirma Astin (1985), han ido a la búsqueda de unos indicadores que alejarán el subjetivismo presente en las opiniones sobre las mejores universidades americanas. En el fondo, considero que explicitar tales indicadores no persigue otra cosa que perpetuar el estado de opinión actual, suministrando al opinante una información donde seguir sustentando su opinión. En consecuencia, el concepto de calidad asociado a los recursos o medios, a los *productos elaborados*, o a la oferta de

estudios a los clientes son matizaciones o precisiones al concepto globalizador de calidad proyectado por una institución y en donde el *marketing* juega también su papel.

6.1. *Calidad en relación a los recursos o inputs*

Cuatro tipos de recursos son tomados en consideración: profesorado, instalaciones, alumnos y dinero ⁶. Comenta Astin (1985, p. 38) que podríamos reducir los indicadores al exclusivo del dinero, pues con éste se pueden comprar los otros tres (salarios, edificios y equipos, becas). Si se eliminan las instalaciones (por su evidente relación con los recursos financieros) y se centra la atención en los otros tres, los análisis ponen de manifiesto la elevada relación que tienen entre sí, así como con los resultados de opinión sobre la calidad ⁷. Aún más duro: los niveles de selectividad constituyen la variable más significativa. En consecuencia, no es de extrañar que el famoso informe *A Nation at Risk* pidiera a la universidad elevar sus niveles de exigencia en las admisiones como medida directa e inmediata de elevar su calidad.

6. Los datos aportados por Astin (1985, p. 40) son reveladores: las 14 universidades que exigen la puntuación mayor en las pruebas de selectividad tienen unos gastos por alumno del 250 % (enseñanza) al 800 % (investigación) superiores a la media de las 2.690 instituciones analizadas.

7. Los datos aportados por Astin (1989, p. 9) ponen de manifiesto que en las instituciones que se exige mayor selectividad —mayor puntuación— (principal factor asociado a la posición en el *ranking* de calidad) se dan las siguientes características:

En los alumnos:

- mayor porcentaje con nota de sobresaliente en la escuela secundaria;
- mayores ingresos económicos en la familia;
- mayor porcentaje de padres con estudios universitarios.

En las instituciones:

- mayor gasto por alumno;
- mayores salarios de los profesores;
- mayor coste de la matrícula del alumno.

La conclusión es obvia: ¡todo es de la misma saca! En nuestras universidades, ¿cuántas de estas variables tendrían *variabilidad*? Creo que nos quedamos sin *ranking*.

Si miramos hacia dentro, los siguientes artículos de la LRU no ofrecen lugar a duda acerca de la adecuación de plantearnos este enfoque de la calidad de la enseñanza universitaria:

«Corresponde al Gobierno, oído el Consejo de Universidades, establecer los procedimientos de selección para el ingreso en los centros universitarios» (art. 26.1. *LRU*).

«El acceso a los centros universitarios y a sus diversos ciclos de enseñanza estará condicionado por la capacidad de aquéllos, que será determinada por las distintas Universidades, con arreglo a módulos objetivos establecidos por el Consejo de Universidades» (art. 26.2. *LRU*).

«El Gobierno establecerá el régimen retributivo del profesorado universitario, que tendrá carácter uniforme en todas las Universidades» (art. 46.1. *LRU*).

«Los créditos tendrán la consideración de ampliables, excepto en los siguientes casos:

a) El crédito correspondiente a la plantilla de funcionarios docentes de la Universidad, excluidos los conceptos retributivos a que alude el apartado dos del artículo cuarenta y seis de la presente Ley.

b) El crédito correspondiente a la plantilla de funcionarios no docentes» (art. 55.1. *LRU*).

6.2. *Calidad en relación a los productos*

Es fácil argumentar que en último término lo que importa es la calidad del producto elaborado (sus graduados); pero una vez más el problema estriba, por una parte, en determinar las especificaciones de ese producto; por otra, el de admitir si ese análisis del producto puede hacerse independientemente del análisis de la materia inicial y de los medios de producción. Variables como las ganancias profesionales, los puestos de reputación pública conseguidos, premios, resultados académicos (múltiples tasas o ratios), etc., son tomadas en consideración desde esta perspectiva objetivadora de la calidad. Los estudios ponen otra vez de manifiesto la asociación de estas variables con la reputación o fama, produciéndose un círculo de difícil análisis causal.

En ocasiones, el producto ha sido definido en términos del simple rendimiento interno (notas) de los estudiantes, o de los indicadores de abandono. En esta perspectiva hemos de citar la propuesta de modelo

de evaluación de la Enseñanza Superior española efectuada por Fernández, Lillo y Martínez (1989): el criterio de calidad —definido indirectamente en el concepto de rendimiento— es operativizado en tres indicadores: rendimiento académico, abandonos y productividad investigadora. El mérito de esta propuesta estriba en el intento de conceptualizar un esquema de relaciones de una multiplicidad de variables referidas a los profesores, alumnos e institución con dicho rendimiento institucional. Sin embargo, es difícil no valorar como reduccionista esta concepción de la calidad de una universidad. Por supuesto que al ser un modelo elaborado para una aplicación práctica ha tenido que optar. El debate se centrará, por tanto, en la valoración que hagamos de la opción.

6.3. *Calidad en relación a la oferta de estudios*

Permítaseme una burda comparación: si consideramos la cantidad de los estudios (títulos o programas) que se ofertan en una universidad como indicador de calidad, el *marketing* habrá que dirigirlo o enfatizar la oferta de un *surtido variado* donde pueda elegir el consumidor. Por el contrario, si queremos pasar del supermercado a la tienda especializada, el indicador de calidad será el de los *productos exclusivos* que se tienen en el escaparate. En estos momentos sería difícil pronosticar qué tipo de calidad elegirían las autoridades políticas centrales y autonómicas con respecto a las nuevas titulaciones para sus universidades (no olvidemos que un título se implanta si así lo decide el gobierno —central o autonómico—).

A título anecdótico y para información sobre uno de los sistemas con mayor influencia en la adopción de modelos de evaluación de la calidad (el estadounidense), diremos que en el curso 81-82 el 61 % de los graduados de *college* (4 años) en las 13 universidades de mayor selectividad (frente al 17,7 % de la media para 1.520 instituciones) lo fueron en: ingeniería, ciencias físicas, matemáticas y biología, e historia y ciencias políticas, y sólo el 1,6 % en *business* (22,2 % en la media) o el 0,2 % en educación (12,7 % en la media). Es claramente asumido por el sistema que el enfoque del contenido de calidad de los *colleges* se resume en las *liberal-arts*. La asociación de las ingenierías a este contenido puede ser explicado, afirma Austin (1985, p. 48), por la cantidad de financiación pública de su investigación.

Llegados a este punto cabe preguntarnos si hemos agotado o no la conceptualización de la calidad de la educación universitaria; o si lo

queremos formular de otra manera: ¿no debería tomarse en consideración el progreso (desarrollo) que la persona tiene a lo largo de sus años universitarios, en el análisis de la calidad de la educación universitaria?

Una respuesta afirmativa supondría asumir una visión comprensiva de la educación en la universidad, de nula tradición en nuestro contexto y que se pone de manifiesto en indicadores tan claros como la ausencia de servicios de ayuda y asistencia a los alumnos, amén del propio enfoque de la docencia. Remito a otro lugar (Echeverría y Rodríguez, 1989; Rodríguez, 1990), en donde analizamos y proponemos acciones específicas en esta dirección. Aquí nos limitaremos a un breve planteamiento conceptual que, como es lógico, se enmarca en la perspectiva del *valor añadido* que apuntábamos en las primeras páginas.

7. La calidad como desarrollo del potencial del alumno

La verdadera calidad de una institución universitaria, señala Astin (1985), radica en su capacidad para lograr el mayor desarrollo posible de sus miembros (alumnos y profesores). En términos económicos, una institución será de mayor calidad que otra en la medida que consiga el *máximo valor añadido* para los mismos *inputs*. Dicho valor añadido es entendido tanto como los diferenciales del grado de desarrollo intelectual, afectivo, personal y social de los estudiantes desde el inicio al final de sus estudios universitarios, como el de los profesores —tanto en su dimensión pedagógica como académica y de productividad científica—

Al asumir un concepto multidimensional del talento, podrá hablarse de niveles de calidad diferentes en el desarrollo de cada una de las dimensiones del talento dentro de una misma institución, de diferencias de calidad entre dimensiones e instituciones y, por el contrario, de niveles semejantes de calidad en el desarrollo de diferentes dimensiones del talento en diferentes instituciones.

El trabajo de Astin *et al.* (1984), una investigación realizada bajo los auspicios del Instituto Nacional de Educación de los EE.UU., ha constituido un sólido fundamento a este enfoque de la calidad de la educación universitaria. El problema radica en la dificultad de obtener datos más precisos sobre cada uno de los objetivos planteados por esta perspectiva. El listado que ofrecemos es una síntesis extraída de la obra de Astin (1985, pp. 63-67).

Una terminología imprecisa para definir el desarrollo del talento

- Adquirir conocimiento y competencias, así como las habilidades para su significativa aplicación.
- Desarrollar actitudes, valores, perspectivas y capacidad para un continuo, aprendizaje.
- Convertirse en una persona cultivada.
- Pasar de ser un aprendiz pasivo a un aprendiz activo.
- Provocar un aprendizaje significativo.
- Ser capaz de expresarse en uno de los diferentes lenguajes desde una perspectiva crítica.
- Mejorar las habilidades de pensamiento crítico, competencia interpersonal e implicación social. Saber definir objetivos, obtener recursos para su logro, usar la autoridad sabiamente,
- Saber utilizar, para cada problema, los métodos adecuados de resolución.
- Saber evaluar críticamente lo que se ha aprendido.
- Desarrollar las habilidades verbales, numéricas, el conocimiento sustantivo, racionalidad, tolerancia intelectual, sensibilidad estética, creatividad, integridad intelectual y sabiduría.
- Desarrollar la madurez emocional, tolerancia, empatía y habilidad de liderazgo.

Es evidente que la terminología utilizada para expresar los objetivos de la educación universitaria desde esta perspectiva está más próxima a una concepción humanista y holística del estudiante. En consecuencia, la dificultad de constatar el logro de los objetivos es un reto a superar.

Una cuestión previa a la posibilidad de evaluar esta concepción de la calidad sería la de preguntarnos acerca de la aceptación que tal concepción tiene en los círculos universitarios y gubernamentales. En el contexto europeo su aceptación es prácticamente nula. La influencia de la clásica concepción germana de la Universidad para élites, ocupada y preocupada por la exclusiva dimensión intelectual, ha motivado un olvido claro, pese al proceso de masificación acaecido, de las nuevas necesidades psicosociales de los alumnos.

En el contexto norteamericano, y debido a la peculiar estructura de su sistema de enseñanza superior, así como a la tradición institucional de los servicios de orientación y ayuda al estudiante ha merecido una mayor atención, especialmente en los años de estudios no graduados. No ha de extrañar la existencia de una larga tradición en la puesta en marcha de programas de ayuda especial para grupos de alto riesgo (p. ej., abandono de estudios), así como la preocupación de su calidad a

través de su sistemática evaluación (Levin y Levin, 1991). Incluso en los últimos años se observa una preocupación institucional por incluir métodos e instrumentos para evaluar el aprendizaje total del alumno y no sólo su dimensión académica⁸. Un dato que es necesario señalar es el que se refiere a la permanencia en el tiempo de ciertos logros en las dimensiones actitudinales y de desarrollo personal de los alumnos. Las escasas investigaciones realizadas en el contexto norteamericano ponen de manifiesto la mayor estabilidad de los logros académicos en términos económicos (posición, empleo, ganancias) frente a la relativamente pronta desaparición de los asociados al desarrollo personal. Por supuesto que la intensidad de las acciones para conseguir los primeros es siempre mucho mayor que la de los segundos.

Punto final

A lo largo de las páginas anteriores se han aportado ideas y argumentos sustentados en la investigación evaluativa sobre la calidad de la Universidad. La selección de variables e indicadores, la utilización de uno u otro enfoque de evaluación y, sobre todo, la finalidad de ésta no son ajenas al concepto de calidad que se sustenta; muy al contrario, abordar la evaluación de la calidad de la educación universitaria constituye el punto de partida de la definición que de la calidad hace el usuario de la evaluación. La toma de decisiones sobre lo que se acuerda entender por calidad marca el inicio de un proceso que debería de conducir a la mejora del sistema universitario en sus diferentes niveles de concreción. Evidentemente que ésta no es la única finalidad de la evaluación, pero sí la que merece la pena perseguir.

Los diferentes trabajos seleccionados para esta obra pretende ofrecer enfoques, modelos, experiencias; en definitiva, concreciones de la calidad institucional de la universidad a la luz de la evaluación de la misma.

8. The Association of American Colleges, así como *The National Governor's Association* y *The American Assambly of College Schools of Bussiness* han puesto en marcha acciones encaminadas a que las instituciones universitarias americanas adopten programas específicos para evaluar el impacto de su educación en el desarrollo global del alumnos. Los trabajos de Boyer et al. (1987), Cramer & Assc. (1990), Morantes (1987) y Potter (1986) recogen especificaciones de éstas y otras iniciativas.

LA EVALUACION INSTITUCIONAL: UNA PERSPECTIVA GENERAL

JOSE-GINES MORA RUIZ
Universidad de Valencia, España

Introducción

Tras unas décadas en las que las universidades han experimentado cambios radicales que las han transformado en instituciones de masas, en las que ha crecido el número de centros, así como el tamaño de cada uno, y en las que ha aumentado notablemente el presupuesto global del sistema universitario, las instituciones universitarias se enfrentan en estos momentos a un nuevo reto: mejorar la calidad del servicio que proporcionan a la sociedad. Cuatro razones justifican esta preocupación por la calidad:

1. Ante la creciente competencia económica internacional, en la que nuevos países industriales emergen, la única solución razonable que tienen los países más desarrollados para mantener su nivel de competitividad, sin deteriorar sus condiciones sociolaborales, es a través de la mejora de la productividad y del desarrollo tecnológico. La educación, especialmente la superior, está siendo considerada de nuevo como una de las piezas claves para afrontar el futuro económico de estos países (Levin y Rumberger, 1989).

2. La calidad de todos los procesos productivos, y en concreto la de los servicios, se está convirtiendo en la actualidad en una de las exigencias cotidianas de los consumidores en las sociedades desarrolladas.

3. En un momento de relativa estabilidad en cuanto a alumnos y presupuestos, es lógico que las universidades vuelvan la vista hacia la calidad de la educación que están suministrando y que de algún modo podía haberse deteriorado en los períodos anteriores de gran expansión de la demanda.

4. Los altos y crecientes costes totales de la educación superior demandan lógicamente un mayor control sobre las universidades por parte de la sociedad que las sustenta, de modo que todos están de acuerdo en que la calidad y la eficiencia deben de ser garantizadas y controladas (Bauer, 1988).

Como consecuencia de este interés por la calidad se ha generado en los países desarrollados una situación que ha sido denominada de «estado evaluativo» y que «intenta insertar una forma particular de ética competitiva definida externamente, como la primera fuerza directriz del desarrollo de la educación superior» (Neave, 1988; véase su traducción en este mismo libro). 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, además de una verificación del funcionamiento del sistema, un método para mejorarlo.

La preocupación por la calidad de las instituciones universitarias esta generando, especialmente en los países más desarrollados, todo tipo de mecanismos para su evaluación. Una revisión general de todos ellos puede encontrarse en Mora (1991). Tan (1986; con traducción en este mismo libro) revisa los métodos utilizados en Estados Unidos. Van Vught (1991; también traducido en este libro) hace un resumen de la situación general de la evaluación institucional en Europa y América.

La evaluación institucional se define como un intento de medir como 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 cuales son las metas que persigue la institución, lo que en algunos casos puede ser una primera dificultad para su puesta en práctica.

Se pueden encontrar dos enfoques básicos para las evaluaciones institucionales: el que centra su atención en las magnitudes contables y de control (in't Veld, 1990), y el que enfatiza los aspectos organizativos del proceso de aprendizaje (Premfors, 1986). Para algunos ha existido una tendencia general a desarrollar el primero, en conexión con valora-

ciones de la eficiencia y de los factores económicos, en detrimento de otros aspectos como la efectividad y el desarrollo profesional y educativo (Pollitt, 1987). Mientras que los partidarios del primer enfoque acusan a los 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. De un modo general se puede afirmar que todo proceso de evaluación institucional conlleva generalmente algún tipo de tensiones entre los individuos afectados por la evaluación, las instituciones sometidas a ella y los poderes públicos afectados (Una discusión detallada sobre esta problemática se encuentra en Westerheijden (1990b cuya traducción puede verse en este libro).

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. Aspectos básicos que deben ser objeto de revisión en toda evaluación institucional son los siguientes (Miller, 1979):

- Metas y objetivos de la institución.
- Aprendizaje de los estudiantes.
- Rendimiento del profesorado.
- Programas académicos.
- Servicios.
- Gerencia.
- Aspectos financieros.
- Consejo de gobierno.
- Relaciones externas.
- Grado de preocupación por la mejora.

Distintos tipos de evaluación institucional.

Las evaluaciones pueden ser clasificadas, según quien las realice, bajo dos tipologías: 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. El uso de revisiones externas tiene la ventaja de la imparcialidad y de la utilización de expertos que son competentes en técnicas de evaluación. 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.

Desde otro punto de vista, se pueden clasificar las evaluaciones institucionales según el fin perseguido por la evaluación: evaluaciones para el control de la institución por parte gubernamental; 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.

Unas evaluaciones están enfocadas hacia 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). Bajo otro punto de vista se puede clasificar una evaluación según sea formativa, concentrada en el proceso, o sumativa, centrada en los resultados.

Respecto a las distintas situaciones en las que se pueden desarrollar las evaluaciones, Frackmann y Maassen (1987) señalan la existencia de tres posibilidades: reguladas por el estado, reguladas a través de un organismo «amortiguador» y autorreguladas.

Según sea el mecanismo básico utilizado para llevar a cabo la evaluación, se pueden encontrar dos sistemas alternativos, que en la práctica se encuentran entremezclados con diferentes intensidades. Son los siguientes:

— *Revisiones por colegas (peer review)*. Este método se fundamenta en la hipótesis de que el sistema universitario 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. La puesta en práctica de estas evaluaciones puede hacerse usando equipos de evaluación externos a la institución, internos o mixtos.

— *Indicadores de rendimiento.* Se trata de utilizar indicadores, generalmente cuantitativos, que suministran información operativa sobre el funcionamiento de las instituciones y sobre su eficacia.

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 el 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. Como señala Sizer (1990a; véase traducción en este libro) ambos sistemas no deben confrontarse sino que los indicadores de rendimiento deben servir para facilitar la revisión por colegas. Una discusión detallada sobre las ventajas e inconvenientes de ambos mecanismos, con especial referencia a la situación en Holanda puede verse en Vroeijenstijn y Acherman (1990; con traducción en este libro).

Evaluaciones Autorreguladas

Se agrupan bajo la denominación general de evaluaciones autorreguladas las que son promovidas por la propia institución universitaria, realizadas por ella misma o por agencias externas, pero, en cualquier caso, con carácter eminentemente voluntario. Este tipo de evaluación surge históricamente en Estados Unidos, 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 calidad y efectividad de las instituciones. Una de las piezas claves del proceso de acreditación de las instituciones universitarias 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 una 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 Estados Unidos como en otros países,

para la evaluación de las instituciones universitarias. Una descripción detallada del proceso de autoevaluación en la Universidad de Alberta, Canadá, puede verse en Holdaway (1988; su versión traducida se encuentra en este libro). Una descripción más general del mecanismo de autoevaluación para una unidad universitaria se encuentra en McDonald y Roe (1984; traducido en este mismo libro) y en Roe *et al.* (1986).

Antes de pasar a revisar los mecanismos más habituales de la evaluación institucional autorregulada, veamos el marco general en el que estos procesos tienen sentido, es decir, cuales son las condiciones generales que debería cumplir todo sistema de evaluación autorregulado. Siguiendo a Kells y Vught (1988) esos 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.
- La 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.

Autoestudios

El autoestudio suele ser la pieza clave de las evaluaciones autorreguladas. En esencia un 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. 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.

Un rasgo esencial 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 autococimiento para saber exactamente que se esta haciendo, con que grado de perfección y a que 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 valen la pena (Chambers, 1984).

Analizar globalmente una institución significa considerar cada una de sus partes esenciales. Pueden citarse las siguientes como objeto central del autoestudio (Chambers, 1984):

- La misión, metas y objetivos de la institución.
- Los programa y *curricula*.
- 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.

Características del autoestudio

Un hecho esencial y diferenciador de los autoestudios es que tienden a fijarse fundamentalmente en la efectividad de las instituciones. Kells (1988) 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 ade-

más 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 y valorar las metas de la organización, 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.
- 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.

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.

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.

A partir de la segunda mitad de los años setenta se despertó en todos los países 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 fueron las restricciones económicas y una tendencia 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 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 eficien-

cia 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, 1982a, 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 en el Reino Unido; este estado de cosas es descrito así por Kogan (1986:125): «El giro hacia la adopción de mecanismo gerenciales esta 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 éste 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: un cierto 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 de las universidades 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 en los que suelen centrar preferentemente su atención las entidades que asignan los fondos a estas instituciones.

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 to-

davía: la 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 (Frackmann, 1987; véase su traducción en este libro).

La evaluación del rendimiento 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. 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 el cual es el objetivo primordial de las instituciones; desarrollar, en segundo lugar, una jerarquía de objetos 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).

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».

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, 1981; Sizer, 1990a).

Incluso superados los problemas conceptuales señalados, en el momento de evaluar el rendimiento institucional surgen algunos problemas más inmediatos como pueden ser los siguientes (Bogue, 1982):

- La evaluación de la efectividad exige el uso de una metodología que no esta 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.

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 controvertido, conviene dedicarles una atención especial. «Los indicadores de rendimiento han sido un tema central de la agenda de la educación superior desde el principio de los setenta, tanto por el interés de los gobiernos y de las agencias encargadas del reparto de los fondos como por el de las mismas instituciones de educación superior» (Kells, 1990a: 5). Los indicadores no sólo son un instrumento técnico para la evaluación, sino que deben ser considerados dentro de un complejo mundo de tensiones políticas entre la necesidad de control público sobre la educación superior y la de reparto por la autonomía de las instituciones universitarias.

El concepto de indicador de rendimiento

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). Como ya se ha señalado con anterioridad, la utilización de indicadores de rendimiento para evaluar el rendimiento de las instituciones universitarias se desarrolló 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, el uso de los indicadores de rendimiento puede ayudar considerablemente a establecer un cierto or-

den y a hacer previsiones en un sistema tan complejo como es en la actualidad el universitario. Cuando los indicadores están suficientemente desarrollados ayudan a entender la situación de los problemas con respecto a unas variables determinadas, a ver como es su evolución en el tiempo y a hacer previsiones para el futuro. 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, 1982a).

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, 1982a) 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 fijar 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).

Tipos de indicadores de rendimiento

Existen multitud de indicadores de rendimiento que se han ensayado para recoger alguna faceta importante de la vida de las instituciones

universitarias. Una clasificación amplia de todos ellos puede hacerse basándose en tres criterios diferentes (Dochy *et al.*, 1990; véase su traducción en este mismo libro):

- Según la función:
 - Indicadores docentes.
 - Indicadores de investigación.
 - Indicadores de servicios.
- Según el objeto:
 - Indicadores de *input*.
 - Indicadores de proceso.
 - Indicadores de *output*.
- Según el tipo de medida:
 - Indicadores cuantitativos.
 - Indicadores cualitativos.

Otra clasificación de carácter funcional 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, puede haber indicadores de rendimiento que midan la:

- Disponibilidad.
- Publicidad.
- Accesibilidad.
- Utilización.
- Adecuación.
- Eficiencia.
- Efectividad.

- Impacto externo.
- Aceptabilidad.

La utilización de los indicadores de rendimiento

Cabría preguntarse cuál es el grado real de utilización de los indicadores de rendimiento y qué tipos de estos indicadores son más utilizados por las universidades. Cuenin (1987; 1990, véase traducción en este mismo libro) realizó una encuesta sobre el uso de indicadores en un conjunto de universidades de países de la OCDE. Según ella, los indicadores son utilizados en la práctica para valorar toda una variada serie de cuestiones, aunque sólo unos pocos tienen un uso generalizado entre las instituciones.

En el Reino Unido es donde la utilización de los indicadores de rendimiento en la educación superior se desarrolló en primer lugar. En 1985 un informe gubernamental (*Jarrat Report*, 1985) recomendaba su uso para evaluar las universidades y para hacer comparaciones entre ellas. Como consecuencia un comité de las propias universidades estableció un conjunto de 54 indicadores especialmente de *input* y funcionamiento, ya que sólo se incluían dos de *output* (Sizer, 1990a).

En Holanda se está actualmente poniendo en marcha, de común acuerdo entre gobierno y universidades, un conjunto de indicadores de rendimiento que pretenden completar el sistema de revisión por colegas de las universidades que también recientemente fue implantado. La propuesta más reciente está constituida por un conjunto de 70 indicadores, gran parte de ellos de carácter cualitativo, que están siendo discutidos en la actualidad y cuyo uso puede convertirse en obligatorio próximamente (Spee, 1990; véase su versión castellana en este libro).

En un reciente informe promovido por el *Programme on Institutional Management in Higher Education* de la OCDE (Kells, 1990a) se recogían la situación respecto a los indicadores de rendimiento en once países. Además de Holanda y el Reino Unido, los indicadores de rendimiento están siendo utilizados en otros países como Australia y Finlandia, y están siendo objeto de estudio y preparación avanzada en Austria, Canadá y Suecia. En España cabe citar la propuesta de uso de indicado-

res para evaluar la docencia universitaria formulada por De Miguel (1991; una versión actualizada puede encontrarse en este libro).

Discusión sobre el uso de los indicadores

En este mismo libro se encontraran varios artículos expresamente dedicados a los indicadores de rendimiento. Especialmente en dos de ellos, Kells (1990b) y Frackman (1987), se discuten con todo detalle las ventajas e inconvenientes que representa el uso de estos indicadores. A ellos remitimos al lector. Sin embargo algunas ideas generales sobre esta cuestión se recogerán en este apartado.

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 la partes implicadas y sean utilizados para objetivo 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».

Otro de los problemas graves 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, 1982a).

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.

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.

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 real de los indicadores afirma Cuenin (1987: 126) que «las opiniones están divididas (...), pero prevalece la actitud de escepticismo». Esta misma actitud de escepticismo era encontrada por Romney (1978) en un estudio realizado en Estados Unidos 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. Findlay (1990) llega a afirmar que los indicadores de rendimiento son un síntoma de una cultura educativa «thacheriana» que ha creado una mentalidad obsesiva de autoanálisis y autojustificación.

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 mejorarla. 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 sobre todo 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 (Dochy *et al.*, 1990).

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 los indicadores más idóneos y el de llegar a un acuerdo entre todas las partes implicadas, tanto para su uso como para su interpretación.

EVALUACION DE LA CALIDAD BASADA EN EL CONTROL FRENTE A LA EVALUACION DE LA CALIDAD BASADA EN LA MEJORA *

TON I. VROEIENSTI

Asociación para la Cooperación de las Universidades Holandesas,
Países Bajos

HANS ACHERMAN

Universidad de Amsterdam, Países Bajos

Introducción

El sistema de evaluación externa de la calidad, tal y como se desarrolla bajo la tutela de la Asociación para la Cooperación de las Universidades Holandesas (VSNU), se basa primordialmente en la revisión por colegas. En su evaluación del primer año del sistema holandés, Kells (1989) afirma que uno de los mayores logros de las universidades holandesas ha sido:

«el no adoptar «medidas de rendimiento» como base del sistema. Aunque existía una inicial atracción hacia ese tipo de enfoque tan simplista de la evaluación, los holandeses se dieron perfecta cuenta del reduccionismo de este sistema y aprendieron de la disfunción, por no decir del desastre, de la experiencia británica al respecto. Mientras sigue vigente cierto interés por los indicadores de rendimiento, éstos no se eligieron inicialmente como base del sistema.»

El debate acerca de la revisión por colegas y los indicadores de rendimiento está estrechamente relacionado con los objetivos que sustentan el rol de la evaluación de la calidad. La cuestión más importante es la siguiente: ¿pretende la evaluación de la calidad controlar y medir

* Traducción íntegra del capítulo 6: *Control Oriented versus Improvement Oriented Quality Assessment*, perteneciente a L.C.J. GOEDEGEBUURE; P. A. M. MASSEN y D. F. WESTERHEIJDEN (EDS), *PEER REVIEW AND PERFORMANCE INDICATORS*. Utrecht: Lemma, pp. 81-101, 1990.

la calidad con el fin de tomar decisiones respecto a la planificación y distribución de los recursos basándose en los resultados, o bien, se emplea la evaluación de la calidad para influir en ella y mejorarla? ¿está la evaluación de la calidad orientada hacia el control o bien orientada hacia la mejora? Una cuestión relacionada con las anteriores es si la evaluación de la calidad queda incluida en la estructura de auto-regulación de los sistemas universitarios o en la estructura del control estatal y gubernamental.

En este capítulo se tratarán estas cuestiones, centrándose en primer lugar en la cambiante relación entre el gobierno y las universidades a la luz del concepto de autonomía. Después, se pasará a debatir el sistema de control externo de la calidad tal y como se ha desarrollado en los Países Bajos, especialmente en los que respecta a los aspectos de auto-evaluación y revisión por colegas, así como al rol de la VSNU como estamento coordinador. Este sistema queda pues contrastado con el concepto de indicadores de rendimiento, al anteponer diversas objeciones al uso de los indicadores de rendimiento —objeciones que se apoyan en el ejemplo del uso (abuso) de dicho enfoque en los Países Bajos. En base a ello se formulan algunas observaciones finales.

2. La universidad holandesa: ¿autónoma o marioneta?

La política de la educación superior holandesa ha experimentado notables cambios desde 1985, con un marcado énfasis en la autonomía institucional y el control remoto por parte del gobierno.

Respecto a la autonomía de las universidades, surge la cuestión de si dicha autonomía es real, ficticia o limitada. A las universidades se les prometió autonomía y autorregulación, pero al mismo tiempo el gobierno estableció unos límites: «No pueden permitirse ciertas formas de actuación por parte de las instituciones» (Maassen y Van Vught, 1989). El gobierno tiene sus propias metas respecto a la educación superior y pretende controlar su desarrollo. En lugar de conceder responsabilidades a las universidades y mantenerse al margen, el gobierno adopta una actitud de «control remoto». En lugar de ser autónoma, la universidad holandesa está en el peligro inminente de convertirse en una marioneta. Esta ambigua actitud queda también reflejada en la evaluación externa de la calidad de las universidades holandesas.

El debate acerca del papel y responsabilidades del gobierno y de las universidades no ha concluido por el momento, pero en lo que respecta a la educación, el resultado en cuanto a los roles de los actores principales parece ser:

- el rol principal del gobierno se basa en establecer, como guardián de la educación, las condiciones «límite» (la cantidad de dinero disponible, quien debe ser admitido en las universidades, la duración de los programas, etcétera).
- La función principal de las universidades estriba en decidir qué tipo de programas ofrece;
- la función principal del profesorado es la estructuración de sus programas, conservando una buena calidad educativa y orientando a sus alumnos a través de los mismos.

Desde el punto de vista de las universidades, la principal preocupación se dirige a poder ofrecer una educación de alta calidad dentro de las condiciones límite establecidas por el gobierno (especialmente tras severas restricciones de presupuesto y la imposición por el gobierno de una duración máxima de cuatro años para las titulaciones) así como a lograr convencer al público de que el profesorado, dadas las circunstancias, hace lo posible por asegurar al máximo la calidad. Así pues, el problema principal es si la mejora de la calidad es posible y la cuestión principal es cómo asegurar que los profesores estén dispuestos a adaptar sus métodos de enseñanza a las cambiantes condiciones límite. Mintzberg (1983) puede aportar una respuesta a estas cuestiones:

«El cambio de la burocracia profesional (como las universidades) no se produce cuando los nuevos administradores anuncian grandes reformas, ni tampoco cuando las tecno-estructuras gubernamentales intentan poner a los profesionales bajo su control. Por el contrario, el cambio se produce a través del lento proceso de cambiar a los profesionales...»

La cuestión está en cómo cambiar la actitud de los profesionales con respecto a sus contribuciones a un programa educativo dado. O, para decirlo de otro modo, ¿a quién están dispuestos a escuchar?. La respuesta obvia es que ellos están dispuestos a escuchar a colegas a los que consideran como sus superiores. Por consiguiente, la respuesta elemental al programa de mejora es la revisión por colegas y no el «control» por parte de los administradores, inspectores o similares.

Sin embargo, el gobierno no está tan interesado en la mejora del programa como lo está en demostrar al público en general y al parlamento en particular, que posee el control y que toma decisiones justificables respecto a la distribución del dinero, la duración de los programas, etc. Por consiguiente, el problema del gobierno está en cómo probar que se toman las decisiones correctas en todos aquellos casos en que los cambios de las condiciones límite puedan afectar la calidad. Si pueden medirse y compararse dichos efectos, queda evidente una «prueba» de ese tipo. Los indicadores de rendimiento se consideran muy útiles en cuanto a la medición de los resultados de la política gubernamental e institucional.

Sin embargo, la revisión por colegas no se basa en los indicadores de rendimiento. Los colegas no están siquiera interesados en este tipo de indicadores: confían en su juicio personal. En resumen, para los que toman las decisiones, la diferencia está entre si uno es capaz de basarse en juicios personales o bien en indicadores cuantitativos. La respuesta de las universidades holandesas es que prefieren basarse en los mejores expertos disponibles.

3. Evaluación externa de la calidad para el autocontrol o control estatal

En abril de 1986, después de la publicación del documento base sobre normativa en la enseñanza superior (HOAK —Educación Superior: Autonomía y Calidad—), en las consultas que se llevaron a cabo entre el Ministerio y las instituciones de educación superior, la conclusión que se alcanzó fue que la evaluación de la calidad es, en primera instancia, responsabilidad de las instituciones. Por consiguiente, éstas deben establecer un sistema coordinado y externo para evaluar la calidad. Sin embargo, tanto el gobierno como la inspección también están involucrados en el proceso de evaluación de la calidad. En el segundo Plan para la Educación Superior y la Investigación (HOOP), queda reflejado que la responsabilidad del gobierno sobre la evaluación de la calidad es complementaria a la de la evaluación realizada por las instituciones de educación superior. Por este motivo, un independiente Inspectorado de la Educación Superior se encarga por ley de este cometido. El Inspectorado también es responsable de los aspectos metodológicos de las evaluaciones realizadas por las instituciones (Ministerio de Educación y Ciencia, 1989).

El salvaguardar la calidad de la educación es una tarea del gobierno, establecida por la constitución. Si se acepta la nueva filosofía gubernamental, en la que la auto-regulación y la salvaguarda de la calidad por parte de las propias instituciones son una consecuencia, las actividades del gobierno pueden entenderse bajo una perspectiva distinta. En lugar de dirigir la atención hacia el control de la calidad en sí mismo, el gobierno debe prestar atención a la evaluación de la calidad por parte de las instituciones. ¿Cumplen las instituciones sus promesas con respecto a la garantía de calidad? ¿Se realiza de un modo correcto?

Si se acepta que éste es el rol del gobierno, puede cuestionarse la descripción siguiente de los comités ministeriales de reconocimiento en el HOOP (Ministerio de Educación y Ciencia, 1989):

«Finalmente, los comités de reconocimiento son un instrumento importante del gobierno para diagnosticar si la calidad y la eficiencia de la organización de la educación e investigación de un sector o una disciplina necesitan ser mejoradas.»

Bajo estos puntos de referencia, los comités ministeriales de reconocimiento entran en conflicto con el concepto de autonomía institucional. Ello refleja muy poca confianza en la habilidad de las instituciones de educación superior, por medio de los comités visitantes organizados bajo los auspicios de la VSNU, para clarificar al gobierno el estado y la calidad de la educación.

Además el papel de la inspección no queda del todo claro. Se ha aceptado que la inspección se hará cargo de la meta-evaluación, o sea de la evaluación externa. Solamente cuando fracasa la evaluación externa de la calidad por parte de las instituciones de educación superior es cuando actúa la inspección, estableciendo un comité visitante. No queda tampoco claro el método como se lleva a cabo la meta-evaluación y es objeto de debate entre la inspección y las instituciones de educación superior. Respecto el memorandum de la inspección sobre meta-evaluación (Inspectorado de la Educación Superior, 1990), da la impresión de que la inspección está más ocupada en la contra-evaluación que en la meta-evaluación. Además de la meta-evaluación, la inspección tiene sus propias responsabilidades sobre la calidad de la educación. Ello se realizará a través de una inspección sobre los aspectos de la educación superior que quedan al margen de los términos de referencia de los comités visitantes. Se hace necesaria una sincronización de todas las actividades relacionadas con el campo de la evaluación de la calidad.

La ambigua actitud mostrada por el gobierno, prometiendo autonomía por una parte y por otra tratando de controlarlo todo, también favorece la discusión acerca del uso de los indicadores de rendimiento.

4. Evaluación de la calidad y revisión por colegas

4.1. *Un esbozo*

Las universidades han aceptado seriamente el reto de la nueva filosofía y la prometida autonomía. Han mantenido su promesa de establecer un sistema externo de evaluación de la calidad (VSNU, 1988b; Vroeijenstijn, 1989). Finalmente, la evaluación externa de la calidad debe cubrir tanto la educación como la investigación. Por el momento, las universidades holandesas están trabajando en el esbozo de un sistema de evaluación de la calidad, que cubra todos los aspectos y actividades. Por el momento, pueden distinguirse los siguientes elementos.

*Evaluación de la calidad de la investigación: Sistema de financiación condicional*¹

Aunque el sistema de financiación condicional, debido a sus características básicas, puede entenderse como una especie de revisión por colegas, las diferencias con el sistema de revisión por colegas en el campo de los programas educativos son notables. En primer lugar, es mucho más burocrático, ya que se basa en «informes escritos» en lugar de observaciones personales y diálogo directo. En segundo lugar, difiere el nivel organizativo. En el campo de la educación, el énfasis primordial estriba en la responsabilidad (organizativa) respecto a la estructura y la coherencia de un programa a nivel de profesorado, así como en las medidas adoptadas para mejorar su calidad. En el campo de la investigación, el énfasis principal está en saber si los resultados esperados son lo suficientemente prometedores como para proseguir con la finan-

1. N.E. En 1983 el Ministerio introdujo un nuevo sistema de financiación de la investigación, dejando los niveles de matrícula y basado en una serie de normas y criterios como:

- Calidad científica y relevancia social de la investigación.
- Necesidad de aprobación, previa a la financiación, de los planes de investigación por comités externos.

ciación. Y en tercer lugar, las metas son distintas: mejora del programa frente a la separación entre programas de investigación buenos y malos, al igual que ocurre, por ejemplo, en la acreditación de programas.

Evaluación de la calidad en educación

Respecto a la evaluación de la educación, las universidades holandesas han desarrollado o están desarrollando tres instrumentos, que están estrechamente conectados entre sí y son complementarios: el comité visitante, el reconocimiento del programa y los comités de reconocimiento. Estos tres elementos se debatirán posteriormente, seguidos de una discusión acerca de las características del sistema de evaluación externa de la calidad.

El comité visitante.—El comité visitante es el instrumento más importante dentro del sistema de evaluación externa de la calidad educativa. Las visitas son estructuradas, cíclicas y cubren a nivel nacional todos los programas del nivel universitario. La tarea del comité visitante consiste en dar una opinión del rendimiento obtenido.

Reconocimiento del programa.—Hasta el momento, todos los programas para graduados de las universidades holandesas quedan garantizados por un documento gubernamental, el Estatuto Académico, que contiene, entre otras cosas, la lista de los programas aprobados, incluyendo una descripción de los requisitos necesarios para obtener el grado de doctor. Así pues, el Estatuto Académico se ha considerado como una especie de garantía de la calidad. En la estructura de la nueva filosofía sobre autonomía y flexibilidad, el Estatuto Académico será abolido. Ello lleva aparejado el problema de cómo asegurar los niveles de los programas. Un comité visitante no significa una respuesta a este problema dado que su visión es retrospectiva: dado un programa y establecidos sus objetivos, ¿cómo se puede evaluar el modo con el que el profesorado intenta alcanzar dichos objetivos? Para asegurar la calidad de los programas, el conjunto de las universidades holandesas establecerá un Comité para el Reconocimiento del Programa. La principal tarea del comité consistirá en revisar el programa propuesto y los objetivos formulados y evaluar posteriormente si pueden alcanzarse el nivel requerido para obtener el grado. Es un juicio «a priori» (Bluemink, 1990). El reconocimiento del programa también es importante a la luz de la futura competitividad en el mercado europeo a partir de 1992.

Comité de reconocimiento.— En algunos casos, el resultado de una visita puede mostrar que resultaría de gran utilidad el explorar nuevas direcciones con respecto a la innovación del programa de algunas disciplinas. Las universidades holandesas ya están familiarizadas con los «comités de reconocimiento» gubernamentales, especialmente en el campo de la investigación; lo cual supone un paso adelante para el establecimiento de sus propios comités de reconocimiento en el terreno educativo. Dentro del marco de la VSNU, ya se ha decidido un protocolo a este respecto (VSNU, 1988c).

Sistema de evaluación externa.—En general, el propósito de la evaluación externa de la calidad consiste en contribuir a) a la mejora de la calidad, b) a la rendición de cuentas ante el público respecto a los niveles de educación e investigación, y c) a la gestión de la calidad en el profesorado, en la universidad y en el sistema universitario en su conjunto. En los Países Bajos, el sistema de evaluación externa de la calidad tiene las siguientes características:

- Se basa en dos importantes pilares: a) el auto-estudio por parte del profesorado que es visitado, y b) el comité de expertos externos, o sea el comité visitante.
- Está orientado hacia la disciplina, o mejor dicho, hacia el programa. Se elige el máximo nivel posible de agregación. Por ejemplo, la revisión de Geografía incluye Geografía Social, Geografía Física, Planificación Urbanística, Demografía y Pre y Protohistoria.
- Durante el primer ciclo el énfasis se coloca principalmente en los programas educativos. En un futuro, tanto la investigación como los servicios comunitarios quedarán incluidos en una evaluación combinada.
- Es a nivel estatal: por ejemplo, todas las facultades con programas pertenecientes al campo determinado por el comité visitante son visitados por el mismo comité.
- Es cíclico: tras un ciclo de cinco años comienza de nuevo por la primera disciplina.
- Cubre toda la universidad: en un período de cinco años se evaluarán todos los programas universitarios graduados.

- Es público: por ejemplo, el comité visitante remitirá un informe a la VSNU, quien hará público dicho informe.

Observando estas características, se puede llegar a la conclusión de que el sistema holandés de evaluación de la calidad satisface la mayoría de requisitos del sistema de auto-regulación descrito por Kells y van Vught (1988) y Cook (1989); está basado en la institución más que en el gobierno; las evaluaciones interna y externa son complementarias; la auto-evaluación es la piedra angular del sistema (véase la siguiente sección); la medición de la calidad toma las metas y objetivos formulados por el profesorado como puntos de partida. Sin embargo, el sistema holandés difiere en dos puntos respecto a estos requisitos. En primer lugar, por el momento el sistema no es todavía comprensivo. La evaluación de la calidad de la educación y de la investigación se realizan separadamente. En el próximo ciclo, que se iniciará en 1994, deberá existir una conexión. En segundo lugar, los resultados de los estudios y las visitas no son confidenciales sino públicas. Dado que una de las funciones del sistema es la rendición de cuentas, es necesaria la publicación de los informes. El informe público también une el sistema de evaluación de la calidad basado en la institución con la función de la inspección, tal y como queda señalado por la ley: la meta-evaluación del modo cómo operan las universidades en las áreas de la calidad y de garantía de la evaluación de la calidad.

4.2. *Auto-evaluación y revisión por colegas.*

Auto-evaluación

El sistema de evaluación externa de la calidad, basado en la evaluación por parte de expertos externos, no es un fin en sí mismo, sino que complementa la gestión interna de la calidad. El eje entre la evaluación de la calidad externa e interna es el informe del auto-estudio, que se supone toda facultad debe entregar al comité visitante. Este informe de auto-estudio es el resultado de una evaluación interna llevada a cabo en la facultad con anterioridad a la visita del comité visitante. El informe de auto-estudio tiene un triple objetivo:

- estimular la gestión de la calidad interna;
- preparar internamente a la facultad para la visita del comité visitante;
- proporcionar información básica para el comité visitante.

La propia facultad decide como se efectuará la evaluación interna. Puede emplearse la guía propuesta por Kells (1983).

En la *Guía para la evaluación externa* de la VSNU se ofrece una checklist para la auto-evaluación (véase el apéndice de este capítulo). Los items de la lista pueden verse como piezas de un rompecabezas que, conjuntamente, pueden ofrecer una visión de la calidad. Dado que el comité visitante debe analizar diversos auto-estudios, aquel ofrece un formato preestablecido de informe. La presentación y el contenido del informe de auto-estudio están relacionados con la función del comité visitante: formarse una opinión del programa en términos del contenido de los programas, del proceso educativo, de la organización y gestión del programa y de los graduados. Los capítulos del informe de auto-estudio se corresponden ampliamente con estas áreas de atención. Debe tenerse en cuenta que cuando se comparan las áreas de atención del listado con la serie de indicadores de rendimiento para la educación en el estudio de Dochy y otros, se observa una asombrosa semejanza entre ambas listas. Sin embargo, la expresión «indicador de rendimiento» no se encuentra en la Guía. La información obtenida a partir de la auto-evaluación es la directriz del comité visitante para la entrevistas del profesorado. El comité discutirá y verificará los hechos. Pero ni los hechos ni los números hablan por sí mismos. Los colegas deben interpretarlos. Los hechos y los números sólo logran su significación en el diálogo entre los expertos y el personal implicado.

Al tratar los items separadamente, no es suficiente una *descripción*, también debe hacerse un *análisis* y debe hacerse constar la posición del profesorado respecto a la disciplina. Ello ofrece una oportunidad para adentrarse más profundamente en aspectos consistentes y en áreas de interés. Finalmente, después de la descripción y del análisis de la situación existente, deben indicarse los *medios previstos* para abordar las deficiencias encontradas.

Por el momento presente, las facultades visitadas en 1988 y 1989 han remitido cerca de 50 informes de auto-estudio. La calidad de los informes no es siempre la misma, ya que es una nueva experiencia para la mayor parte del profesorado. Pero en general, han demostrado ser de gran utilidad y haber influido en la gestión interna de la calidad.

Revisión por colegas

¿Quiénes son los colegas que visitan las distintas facultades y que dan su opinión acerca de la calidad de la educación? La evaluación de la calidad externa se realiza por eminentes expertos en el área y que al mismo tiempo son colegas. Es de gran importancia que el conjunto de la disciplina tenga confianza en el comité. Por consiguiente, los decanos de las facultades de cada disciplina, reunidos en el Consejo Disciplinario del VSNU, nombran a los expertos que integran el comité. El protocolo establece los criterios que deben tenerse en cuenta cuando se selecciona a los miembros de dicho comité, que consta aproximadamente de siete miembros. Ello guarda relación con la división del trabajo y la cobertura de cualquier ausencia temporal de uno de sus miembros. Al proponer los expertos deben realizarse todos los esfuerzos posibles para cubrir al máximo todas las especializaciones de una disciplina. Ello se hace particularmente necesario cuando se combinan diversas (sub) disciplinas. Los expertos deben ser elegidos tanto de dentro del sistema universitario, como de potenciales empleadores y de organizaciones profesionales, en la medida en que existan. Al buscar expertos dentro del sistema universitario es prudente una distribución equilibrada entre personas retiradas y en servicio activo en las universidades. También es importante la presencia en el comité de alguien con conocimientos del proceso educativo. El protocolo asume además que en el comité deberá existir al menos un experto externo. El secretario/a del comité visitante es aportado por la oficina de la VSNU.

La distribución entre los potenciales empleadores y el mundo académico no ofrece problemas, las grandes industrias como Philips y Shell están deseosas de formar parte de los comités. Hasta la fecha no ha resultado difícil el encontrar expertos extranjeros, aunque todos los miembros deben hablar holandés. En los comités de 1988, 1989 y 1990 habían miembros provenientes de Bélgica (10), Alemania (2), Gran Bretaña (1), Austria (1), Suiza (1) y Estados Unidos (3).

Un comité visitante tiene una función dual. Por una parte, el objetivo es la mejora y, por otra parte, la rendición de cuentas. La función de «mejora» está estrechamente relacionada con el *feedback* que el comité pueda aportar al profesorado. El profesorado puede intercambiar puntos de vista sobre problemas educativos en un diálogo con colegas. El comité visitante sostiene el espejo en el que debe mirarse el profesorado. Dado que el comité también visita otras facultades posee una

buena visión general de posibles soluciones a los problemas, y puede sugerir mejoras.

Respecto a la función de rendición de cuentas, el comité de expertos externos puede legitimar la evaluación y la gestión de la calidad interna. El comité observa lo que realiza la facultad con respecto a la gestión de la calidad. ¿Existe un sistema de evaluación de la calidad interna? ¿Funcional? El comité puede asegurar a la sociedad (parlamento, gobierno, inspección) que el profesorado se toma en serio el tema de la calidad. En su informe público, el comité da cuenta del estado real de la educación en la disciplina concerniente.

La función del comité visitante es la de emitir una opinión basada en la información aportada por el profesorado y por los debates sostenidos «in situ» acerca del nivel de la educación y de la calidad del proceso educativo, incluyendo la organización del programa y la calidad de los graduados, así como la de aportar sugerencias sobre la mejora de la calidad.

El término «emitir una opinión» no debe interpretarse como la emisión de un juicio y la clasificación de «bueno o malo». Tampoco implica la aprobación del programa. Al comité no le concierne la acreditación o reconocimiento del programa, ni tampoco el ranking de programas, profesorado o universidades. La meta de la evaluación externa de la calidad consiste en detectar, mediante un diálogo con el profesorado, puntos importantes y áreas de interés.

1.4.3. *El rol de la VSNU*

Después de que las universidades holandesas decidieran hacerse cargo de la evaluación externa de la calidad, la VSNU jugó un importante papel en la preparación y diseño de un protocolo para el sistema así como en la comprobación del mismo (Acherman, 1988; VSNU, 1988b; Vroeijenstijn, 1989). Tras una serie de fructíferas visitas, la VSNU decidió adoptar un determinado enfoque, siendo la coordinación y facilitación la función de la VSNU.

La dirección del comité visitante corresponde a la VSNU. Los miembros del comité son nombrados por la Junta Rectora de la VSNU. La oficina de la VSNU es responsable de la coordinación, en la persona del

coordinador de la evaluación externa de la calidad. Durante la primera reunión del comité visitante, el coordinador explica a los miembros de dicho comité lo que se espera de ellos y les da las instrucciones pertinentes. La oficina de la VSNU también organiza seminarios para el profesorado implicado en la preparación de la auto-evaluación y de la visita del comité. Mientras que los comités están trabajando, ni el coordinador ni la VSNU interfieren al comité visitante, a no ser que el propio comité solicite ayuda. Tras la finalización de las visitas y del informe, el comité lo entrega a la Junta Rectora y es publicado por la VSNU.

Durante el período de pruebas los informes eran confidenciales y todavía no se tiene ninguna experiencia del papel que la VSNU desempeñará después de la publicación, pero a la luz de la auto-regulación del sistema universitario, la VSNU puede desempeñar una función estimuladora en cuanto a la gestión de la calidad del sistema universitario. En primer lugar, puede mantenerse un debate en la Junta Disciplinaria basado en los informes de auto-estudio y en los informes aportados por el comité visitante. Además, la Junta Rectora de la VSNU (formada por miembros provenientes de las Juntas Ejecutivas de las universidades) pueden discutir los resultados.

Dichos resultados pueden constituir la base de acuerdos interdisciplinarios e interuniversitarios respecto a la (re)distribución y la especialización. El resultado de todo ello puede ser un medio natural de contracción y expansión en lugar de interferir con el gobierno; una auto-regulación en lugar de una regulación estatal.

En este momento (marzo de 1990) los comités de cuatro disciplinas han entregado los informes de 1989. Los comités de 1990 han iniciado sus actividades en otras cinco disciplinas. Si además tenemos en cuenta las visitas «en período de prueba» de los programas graduados, 15 disciplinas con 50 programas graduados pertenecientes a 86 facultades han estado o están sujetas a una evaluación externa de la calidad. La VSNU ha comenzado la preparación de las visitas de 1991 para otras cinco disciplinas.

1.4.4. *Observaciones finales*

Las universidades holandesas han escogido la evaluación externa de la calidad, dirigida hacia la mejora y basada en la revisión por colegas.

Así pues, las experiencias obtenidas demuestran que el profesorado está dispuesto a dialogar con sus colegas. Las entrevistas suelen ser por lo general abiertas y francas. Las personas están dispuestas a escuchar la opinión y el consejo del comité visitante. A pesar de que el comité nunca aporta nuevas cuestiones (y sin tener en cuenta si la auto-evaluación está bien hecha), el punto de vista externo puede aportar una dimensión distinta al problema.

En la revisión por colegas se pierde la posibilidad —si es que existe alguna posibilidad— de medir la calidad exacta en un momento dado. El utilizar un *ranking*, aún en el caso de que alguien quisiera hacerlo, queda al margen de la cuestión, pero hay que añadir que el procedimiento de la revisión por colegas y sus resultados tienen una gran influencia en la calidad. Los resultados de la visita proporcionan al profesorado un punto de partida para la mejora. En algunas ocasiones el comité, al romper moldes, actúa como catalizador de la innovación. En algunas ocasiones el profesorado emprende una acción directa, basada en las recomendaciones del comité, en otras no es posible actuar a corto plazo ya que se hace necesario un diálogo con la Junta Ejecutiva de la Universidad. Pero lo que resulta evidente es que numerosos grupos de profesorado utilizan los resultados de las visitas.

A pesar de que la evaluación externa de la calidad por parte de la VSNU se está desarrollando bien, existen algunos puntos que requieren especial atención.

- Como se ha señalado anteriormente, en un futuro próximo debe haber una conexión entre la evaluación de la educación y la evaluación de la investigación.
- Está muy bien el hecho de evaluar la calidad a nivel de facultad pero, ¿qué ocurre con la contribución de los servicios centrales y la institución en su conjunto? Será necesario complementar las revisiones de las facultades con una revisión de los servicios administrativos centrales y la dirección general.
- Un problema con el que se enfrentan todos los comités es la falta de una adecuada información cuantitativa por parte de la facultad. En un futuro próximo debe desarrollarse un buen sistema de datos.
- ¿Cómo supervisar el empleo que haga el profesorado de los resultados? ¿vale la pena esta inversión intensiva de tiempo?

- Debe prestarse una especial atención a la sincronización de todas las actividades, llevadas a cabo por distintos grupos, dentro de la estructura de la evaluación de la calidad. El profesorado corre el peligro de quedar «saturado» por causa de la atención prestada a la calidad.
- Muy importante para el éxito o el fracaso del sistema de auto-control es la actitud del gobierno y de la inspección. ¿Qué uso hará el gobierno de los resultados? ¿Será la función de la inspección meta-evaluadora o contra-evaluadora? Será necesario que el gobierno permanezca a la expectativa, observando si las universidades aseguran la calidad y en caso afirmativo de qué modo lo hacen. No deberá interferir directamente en base a los resultados, pero deben conceder a las universidades un plazo de entre cinco y diez años para establecer un adecuado sistema de funcionamiento. El memorandum explicativo surgido a la luz de la nueva ley sobre educación superior (WHW) refleja la actitud basada en la nueva filosofía del gobierno. El ministerio es de la opinión de que si el resultado de la evaluación de la calidad muestra que la misma es insuficiente en algunos aspectos, las instituciones deberán actuar para procurar una mejora. Si tal acción tiene éxito quedará reflejado en la siguiente evaluación. Si por el contrario, la calidad es todavía insuficiente después de la segunda visita, el ministerio podrá aplicar sanciones marcadas por la ley y, por ejemplo, clausurar el programa (Ministerio de Educación y Ciencia, 1989a).

5. Evaluación de la calidad e indicadores de rendimiento

5.1. *Un «lenguaje común» con respecto a los indicadores de rendimiento*

En la misma reunión (Abril de 1989) en la que se alcanzó un acuerdo sobre el sistema de evaluación externa de la calidad, también se llegó a un consenso con respecto a intensificar los debates acerca de los indicadores de rendimiento. El gobierno y las instituciones de educación superior deben colaborar para desarrollar un «lenguaje común» con respecto a los indicadores de rendimiento, que pueda emplearse en el diálogo entre el gobierno y las instituciones de educación superior. Se creó un comité técnico en el que participaban el Ministerio de Educación y Ciencia, el Inspectorado, la VSNU y el Consejo del HBO (de las institucio-

nes de educación superior no universitarias). Las actividades del grupo estarán respaldadas por la investigación de Wijnen y colaboradores. Tras la publicación del estudio (Segers *et al.*, 1989; véase también el trabajo de Dochy *et al.* incluido en esta obra) surge la cuestión de si se ha creado un «lenguaje común». Los resultados del proyecto de investigación pueden considerarse como un diccionario del «lenguaje de los indicadores». La conclusión a la que se ha llegado es que no existe consenso en cuanto a un determinado tipo y número de indicadores de rendimiento. El gobierno y las universidades no hablan el mismo lenguaje. Un factor sorprendente de la discusión relativa a los indicadores de rendimiento es que siempre hay dos partes encontradas: el gobierno pone un fuerte énfasis en la importancia de la utilización de los indicadores de rendimiento, se muestra optimista acerca de la posibilidad de determinar los indicadores correctos; Las instituciones de educación superior, por lo general, se muestran reservadas y escépticas acerca de la utilización de los indicadores de rendimiento. ¿Por qué no puede lograrse un lenguaje común? Las siguientes barreras para el diálogo pueden dar algunas pistas al respecto.

El término «indicador de rendimiento» es muy confuso

A pesar de los intentos de definir el significado y la función de los indicadores de rendimiento, las palabras inducen a confusión. El problema estriba en que los indicadores de rendimiento no tienen una relación directa con el rendimiento de una institución, pero deben siempre interpretarse a la luz de una meta u objetivo dados. El problema puede ilustrarse con un ejemplo aportado por los propios investigadores (Segers *et al.*, 1989). La ratio sexo dentro de la población estudiantil no es un indicador del aprendizaje o del rendimiento de la universidad, pero si es un indicador a tener en cuenta por parte del gobierno ya que muestra en qué medida se ha alcanzado el objetivo «emancipación». La conclusión a la que puede llegarse es que el gobierno ha de realizar mayores esfuerzos por alcanzar su objetivo emancipador, pero no dice nada acerca del rendimiento o la calidad de la institución.

Los indicadores de rendimiento tienen diversas funciones

El desarrollo de un lenguaje común queda también entorpecido por el hecho de que la gente atribuye diversas funciones a los indicadores de rendimiento. A pesar de que las universidades evitan el empleo del

término, reconocen la importancia del uso de dicho tipo de datos. A modo de cifras clave, estos datos desempeñan un importante papel en la supervisión y evaluación. Por otra parte, el gobierno contempla los indicadores de rendimiento como un instrumento de la planificación y distribución. En diversos países pueden observarse intentos de distribuir fondos basándose en los indicadores de rendimiento. (Brons, 1990).

Transformación de los indicadores en estándares

Observando las distintas funciones atribuidas a los indicadores de rendimiento, surge el temor de que los indicadores queden transformados en estándares. La ratio del éxito puede ser un indicador para constatar el objetivo «conseguir que se graduen el mayor número de alumnos». Una ratio de éxito del 80% indica una mayor consecución de objetivos que una ratio de éxito del 60%. Pero, ¿dice algo esa cifra acerca del rendimiento del profesorado o de la calidad de la educación? ¿Es el rendimiento de la universidad X con una ratio de éxito propedéutico del 80% superior al rendimiento de la universidad Y con una ratio de éxito propedéutico del 60%? ¿O bien ocurre que la universidad X ha bajado sus niveles? ¿O bien la universidad x es más selectiva durante el primer curso? Los indicadores de rendimiento provocan más cuestiones de las que responden.

Los usuarios de los indicadores de rendimiento tienen distintos objetivos y metas

Otra razón para no alcanzar un lenguaje común es la indistinción: ¿quién quiere qué? Los indicadores de rendimiento poseen un sólo significado a la luz de unos objetivos claramente formulados. Las metas y objetivos del gobierno y de las universidades se cubrirán mutuamente por lo general, pero también pueden entrar en conflicto. El gobierno busca instrumentos para comprobar en qué medida se alcanzan sus metas y objetivos. Así pues, necesita información, pero necesariamente bajo la forma de indicadores de rendimiento. Debe conseguirse un acuerdo entre el gobierno y las universidades acerca de qué datos son necesarios y de qué manera éstos deben conseguirse. El gobierno debe elegir sus propios indicadores, pero debe explicar porqué los ha elegido.

Las universidades deben penar además en la información necesaria para evaluar la realización de sus metas y objetivos. Los datos deberán

ser más cualitativos que cuantitativos. Por este motivo, la opinión de los colegas será indispensable, tal y como se viene empleando en las universidades holandesas dentro del sistema de la evaluación externa de la calidad. El empleo de una serie restringida de indicadores de rendimiento acarrea el peligro de «tener una vida propia» a costa del rendimiento real.

La conclusión alcanzada es que el estudio de Segers *et al.* no produjo un diccionario para un lenguaje común, sino más bien un catálogo de todos los posibles indicadores de rendimiento para evaluar la realización de las metas y los objetivos previstos. Tanto el gobierno como las instituciones de educación superior pueden usar el catálogo para elegir sus indicadores.

5.2. *La Teología en los Países Bajos: un caso de (ab)uso de los indicadores de rendimiento*

A mediados de 1989 se publicaron numerosos comentarios y reacciones en los periódicos holandeses acerca del informe del Comité Ministerial sobre Reconocimiento de la Teología (RCT, 1989). La calidad de las facultades holandesas de Teología era precaria. El comité declaró que había examinado la Teología desde siete perspectivas: productividad, cualificación personal, número de tesis, número de jóvenes investigadores, evaluación de los programas de investigación dentro de la estructura del sistema de financiación condicional, participación del personal en comités editoriales de revistas científicas y participación no especializada.

En cuanto al índice de productividad, el comité estableció una media de 1 libro cada 5 años y de 1 artículo anual para la dedicación a tiempo completo. Constatándose por parte del comité que la media internacional era de 1 libro cada 4 años y de dos artículos anuales. Basándose en un sistema ponderado de las publicaciones (1 libro en holandés, 10 puntos, en una lengua extranjera, 15 puntos; un artículo en holandés 1 punto y en una lengua extranjera 3 puntos), el comité estableció el *ranking* de las facultades y extrajo conclusiones acerca de su nivel científico. La conclusión a que se llegó fue que ninguna facultad alcanzaba un nivel internacional y que tan sólo dos de ellas alcanzaron un nivel nacional. Pero tanto los estándares como la ponderación se basaron en decisiones arbitrarias. ¿Por qué se le concedieron 10 puntos a un libro escrito en holandés y 15 a un libro en lengua extranjera, en lugar de 4 y 8 puntos

respectivamente, tal y como había estipulado el comité ministerial de Ciencias Educativas y Pedagogía? ¿Por qué un libro cada cinco años en lugar de 1 libro cada 4 ó 6 años? ¿Por qué motivo son equivalentes 10 artículos en holandés a un libro, mientras que basta con escribir cinco artículos en lengua extranjera para conseguir el mismo número de puntos que con un libro? Si se replantean los estándares y las ponderaciones, los resultados varían. La conclusión es que trabajando con ponderaciones y estándares arbitrarios los resultados deben manejarse con mucha precaución.

Lo atractivo de trabajar con indicadores de rendimiento es que se llega a conseguir una clara imagen de los puntos fuertes y débiles. Pero de todos modos esos análisis deben manejarse con cuidado y deben complementarse con otra información. Como queda de manifiesto en el informe, el comité no tuvo en cuenta la calidad de las publicaciones ni tampoco hizo indagaciones acerca de cómo y del porqué. El comité examinó el trabajo científico de las facultades de Teología a través de siete pares distintos de gafas. Según el par de «gafas» que se colocara, la imagen era más o menos oscura. La mayor objeción que puede achacarse a este modo de utilizar las cifras es el grado de absolutismo que el comité confiriera a dichas cifras. Una vez establecido un estándar, las cifras ya no son indicadores sino criterios. Al leer el informe uno tiene la impresión de que el comité ha examinado la Teología de los Países Bajos a través de diversas ventanas, pero que no entró en el edificio para hacer indagaciones.

6.5.3. *Observaciones finales*

Un lector malevolente de la sección 8.4 del proyecto del Plan de Educación Superior e Investigación de 1990 (Ministerio de Educación y Ciencia, 1989:319) puede sintetizar las palabras del ministro del modo siguiente: «definir en qué consiste la calidad, establecer niveles para ello, tomar algunos indicadores de rendimiento para medir en qué grado se han alcanzado los estándares. Sobre esta base se pueden tomar decisiones».

Evidentemente, los indicadores de rendimiento pueden emplearse para comprobar si las metas y objetivos propuestos se han alcanzado, pero los objetivos deben quedar muy claramente definidos y debe concretarse qué indicadores de rendimiento se emplean y por qué. También

resulta obvio que es muy peligroso el usar indicadores de rendimiento para la evaluación y el control de la calidad. Los indicadores de rendimiento en un sentido estricto (indicadores cuantitativos) nunca pueden ofrecer una buena imagen de la calidad. El caso de Teología nos demuestra cuan cuidadosos debemos ser con su tratamiento. El riesgo de que los indicadores se conviertan en criterios es muy real. Incluso puede llegar a ser más peligroso, pues en ocasiones las decisiones acerca de la formulación de los indicadores de rendimiento se fundamentan en bases arbitrarias. Las cifras debieran ser el inicio de una revisión por colegas y un diálogo con el profesorado. El saber el por qué de una realidad es más importante que el saber que la realidad es tal como es.

Los indicadores de rendimiento pueden desempeñar un papel, aunque inferior, en la evaluación de la calidad. Los indicadores de rendimiento en su sentido más amplio (cualitativo y cuantitativo) son el punto de partida para enfocar la calidad. Tomados en un sentido restringido, cuantitativo, desempeñan su función al apoyar la opinión de los colegas. La opinión de los colegas debe estar basada en hechos y cifras, pero nunca puede ser reemplazada por indicadores de rendimiento.

APENDICE

CHECKLIST PARA UN AUTO-ESTUDIO

1. El lugar dentro de la organización

- El enfoque de la evaluación externa de la calidad centrado en la disciplina no siempre se corresponde al modo en el que la unidad que ha de ser visitada se adapta, en términos organizativos, a la organización de la universidad. Así pues, sería deseable comenzar el auto-estudio con un capítulo que describa la particular estructura organizativa así como la posición del profesorado o de la disciplina, dentro de la estructura universitaria; también deben mencionarse los departamentos participantes en el diseño del programa.
- ¿Qué comités participan en la disciplina? ¿Cuáles son sus funciones y como están constituidos?

2. Los alumnos

- El tamaño de la matrícula de alumnos.
- Características de los alumnos matriculados (sexo, edad, origen geográfico, estudios previos o experiencia profesional).
- La prevención de las deficiencias y problemas derivados; problemas derivados de los niveles educativos anteriores.
- Actividades relacionadas con la información y la admisión.
- Tendencias de la matriculación en años anteriores; cambios significativos de la matriculación;
- Tasas de éxito de las etapas propedéutica (cursos comunes) y graduada.

3. Breve descripción de los programas

General

- ¿Se han formulado los objetivos?
- ¿Se ha especificado el nivel requerido de habilidad y de conocimiento de los graduados?
- ¿Difieren los objetivos de acuerdo con los programas?
- ¿En qué medida difieren los objetivos de programas similares en otras universidades?
- ¿En qué medida se alcanzan los objetivos?
- ¿Son los objetivos apropiados aún en situaciones cambiantes?
- ¿Existen planes para modificar los objetivos?
- ¿Existen planes para poner relieve los aspectos distintivos del programa o para reforzar su actual perfil?

Curso propedéutico (común y selectivo)

- ¿Cuál es la estructura general del nivel propedéutico? ¿Qué materias se imparten? ¿Cuáles son los pesos proporcionales de las materias?
- ¿Existen requisitos en términos de relaciones mútuas entre las materias?
- ¿Difiere este nivel propedéutico del de otras instituciones?
- ¿Qué materias desempeñan una función orientadora y selectiva en el nivel propedéutico? ¿Cuál es la evaluación de la función orientadora, selectiva y de referencia?
- ¿Es satisfactorio el nivel propedéutico a la luz de los cursos siguientes?
- ¿Cuáles fueron las premisas para elegir esta estructura para el nivel propedéutico?

- ¿Se han producido cambios estructurales en dicho curso durante los últimos años?

Etapas de estudios graduados

- ¿Cómo está organizada la etapa graduada? ¿Cuál es la relación entre los cursos optativos y las materias fundamentales?
- ¿Qué asignaturas fundamentales se ofrecen?
- ¿De qué modo se organizan, por lo general, las asignaturas fundamentales? ¿Cuáles son los componentes esenciales del programa y cuáles son los optativos?
- ¿Cuál es la dimensión del programa establecido? ¿Cuál es la variación de los elementos optativos?
- ¿El perfil de varias materias fundamentales en comparación con otras universidades. ¿Qué condujo a decidirse por un determinado perfil para una materia fundamental. ¿Cubre dicho perfil las expectativas?
- ¿Existen problemas de contenido en las materias fundamentales?
- Organización del trabajo práctico, proyectos y tesis de graduación.

Oferta de otros cursos

- Cursos de postgrado (formación profesional y del profesorado).
- Cursos y seminarios.
- Servicios profesionales y cursos contratados
- Becas de investigación;
- ¿Cuántos alumnos pasan a los estudios de postgrado? ¿El nivel de los postgraduados satisface las expectativas?
- Experiencia en servicios y cursos contratados.

Programas a tiempo parcial

- ¿En el programa, se ofrecen también cursos a tiempo parcial?
- En caso afirmativo, en qué medida difieren los programas a tiempo parcial, en términos de estructura y material?

4. El proceso educativo

- ¿Qué métodos docentes se emplean en la adquisición y transmisión del conocimiento (clases, organización de seminarios, tutorías, trabajo práctico, etcétera)?
- Cómo se emplea la informática en los cursos?
- En términos de teoría educativa y didáctica, ¿está diseñado el programa de tal modo que pueda completarse en el tiempo previsto sin demasiada dificultad?
- ¿Existen problemas específicos de contenido, por ejemplo, la transición de un nivel propedéutico estructurado (o curso básico) o la etapa graduada, formación de los alumnos, elaboración de la tesis de licenciatura y proyecto final?.
- ¿Cómo se evalúa a los alumnos? a) ¿De qué manera? (elección múltiple, respuestas libres); b) ¿Cuándo (parcialmente, evaluación única)?

5. Organización y dirección del programa

Política educativa

- ¿De qué manera se formula la política educativa? ¿Qué comités están implicados?
- ¿La política relativa a la disciplina está unida a la política institucional? ¿Existe conexión con los planes de desarrollo?

Política de personal

- ¿Cuál es la política respecto a la *ratio* docencia/investigación?
¿Cuál es la carga docente?
- ¿Cuál es la política respecto a la ampliación de profesorado en los cursos comunes o selectivos y en las asignaturas fundamentales?
- ¿Cuál es la política respecto a cobertura de personal en seminarios, supervisión de tesis, trabajo práctico o experiencia profesional?
- ¿Existen factores que dificultan la consecución de una adecuada política de personal?

Gestión de la calidad interna

- ¿Qué atención se presta a la gestión de la calidad interna? ¿Existe una evaluación sistemática? ¿Qué sistema de evaluación se emplea?
- ¿Cómo se maneja la innovación educativa? ¿De qué modo se organiza la ayuda educativa? ¿Existe colaboración con el Centro de Investigación y Desarrollo para la Educación Universitaria?
- ¿Qué atención se presta a la profesionalización del personal (cursos de dirección, capacitación didáctica o educativa)?

Progreso del estudio, supervisión del estudio y orientación del estudio

- ¿Se han evaluado durante los últimos años las horas de estudio del curso común y selectivo, las de la etapa graduada o ambas? En caso afirmativo, ¿cuáles fueron los resultados?
- Dichos resultados, ¿han producido cambios en el diseño del programa o en la supervisión del estudio?
- ¿Se registra el progreso del estudio? ¿De qué manera? ¿Conducen esos registros a un temprano reconocimiento de los proble-

mas y a acciones preventivas o de recuperación destinadas al alumno en particular o al diseño del programa?

- ¿Se ha llevado a cabo una investigación acerca de los motivos de sobrepasar el plazo de permanencia? En caso positivo, ¿cuáles fueron los resultados y cuáles fueron las medidas resultantes?
- ¿Cómo se organiza la orientación académica y cuál es su efecto?
- ¿Cómo se organiza la supervisión del estudio y cuál es su efecto?
- Información del programa *a)* para alumnos futuros, *b)* durante el programa y *c)* como preparación para introducirse en una profesión.

Servicios

- Servicios educativos (salas de lectura, áreas para trabajo práctico, etcétera);
- Situación respecto a laboratorios.
- Situación respecto a bibliotecas;
- Situación respecto a equipamiento informático;
- Tamaño del presupuesto para servicios.

Interacción entre los niveles central y de facultad

- ¿Cómo se organiza las relaciones entre la facultad o disciplina y el Consejo Ejecutivo de la Universidad con respecto a la educación y, en particular, a la promoción de la calidad?
- ¿Cómo se relaciona la política educativa de la facultad con la política educativa de la institución?
- ¿Se beneficia la facultad de los servicios aportados a nivel central para promocionar la calidad?

6. **Los graduados**

- ¿Establecen los futuros empleadores estándares que el graduado debe alcanzar? ¿Puede definirse con claridad el futuro de la profesión? ¿Han cambiado con los años dichas definiciones?
- ¿Dónde van a parar los graduados?
- ¿Cuál es la tasa de desempleo entre los graduados?
- ¿Existen contactos con los antiguos alumnos?
- ¿Cómo ven el programa los antiguos alumnos?
- ¿Los contactos con los futuros empleadores son estructurales o accidentales?
- ¿Cuál ha sido la política para introducir en lo posible el programa dentro del mercado laboral?

LA EVALUACION DE LA UNIVERSIDAD Y SU CONTEXTO POLITICO: GESTION DE LA CALIDAD Y TOMA DE DECISIONES EN LA EDUCACION SUPERIOR *

DON F. WESTERHEIJDEN
Universidad de Twente, Países Bajos

Introducción

Se discuten en numerosos países europeos nuevos sistemas de evaluación de la calidad en la educación superior y con frecuencia se han venido realizando los primeros pasos para su puesta en práctica. En lo que se refiere a la educación superior de los Países Bajos —el presente artículo se centrará fundamentalmente en dicho sistema— se vienen desarrollando, por separado, nuevos procedimientos para la evaluación de la calidad de la investigación (desde 1983) y de la educación (1988/89).

En todos los lugares, los objetivos del gobierno respecto al sistema nacional de educación superior constituyen el tema central del desarrollo de nuevas políticas. Se tienen mucho menos en cuenta los puntos de vista y los intereses de otros actores, lo cual tiene sus consecuencias especialmente para la fase de implementación de las normativas referentes a la evaluación de la calidad: aparecen efectos no previstos ni perseguidos que pueden dar origen a resultados contrarios a las expectativas mantenidas en el momento de la introducción de la normativa. Una mayor atención a las posturas y reacciones de estos otros actores, especialmente las instituciones de educación superior, podría abrir camino a una explicación de esas imprevistas consecuencias. Se examinará de este modo el sistema holandés de evaluación de la calidad de

* El presente trabajo es traducción íntegra de la versión revisada del autor de la ponencia presentada en el Coloquio Internacional sobre *La Pedagogía Universitaria*. Universidad de Barcelona, 17-20 de octubre de 1990.

la educación superior —que, por otra parte, queda descrito más detalladamente en otros trabajos de este volumen—, lo cual puede dar lugar a conclusiones aplicables a otros países europeos.

La estructura del presente artículo es la siguiente: una primera sección contiene las herramientas necesarias para el análisis (una *teoría de conducta política* y el entorno en el cual se sitúa dicho mecanismo, p. ej., la red de educación superior). El propio análisis constituye la segunda sección, mientras que en la tercera sección se exponen algunas observaciones finales.

1. *Conducta política en la educación superior*

El tema central de este artículo es el siguiente: ¿Cómo pueden explicarse las reacciones de las instituciones de educación superior, sus unidades básicas (facultades, grupos de investigación y cátedras), así como las de los miembros individuales del personal académico, ante los nuevos sistemas de evaluación de la calidad? La cuestión queda planteada en los niveles medio e inferior de los tres señalados por Kogan (1984: 62): intra-institucional, relaciones entre las instituciones y la macro política y la política del centro. El resultado de la política de centro se tomará aquí como un dato, expresado en las metas del Ministerio de Educación y Ciencia.

Como toda explicación, requiere una teoría para singularizar aquellos hechos que sirven como causas y efectos de la explicación, de entre la innumerable riqueza de hechos (Popper, 1965 [1963]: 23), en primer lugar deberá esbozar dicha teoría, la denominada *teoría de la conducta política*.

1.1. *Esbozo de una Teoría de conducta política*

La teoría elegida en el presente artículo se ha desarrollado para explicar la conducta de los *actores* (bien individuos humanos y organizaciones) en los *sistemas sociales* (sistemas de dos o más actores) y ha sido detallada posteriormente en lo que respecta a la situación de la toma individual de decisiones en las organizaciones del gobierno (De Vree, 1982, 1990; Lieshout, 1964; Lieshout y De Vree, 1985). Ha sido comprobada, con resultados mixtos pero no del todo insatisfactorios

(Westerheijden, 1986); también se ha empleado con cierto éxito en un re-análisis de los estudios sobre política de la educación superior (Van Vught y Westerheijden, 1989). Es una teoría política en el sentido de que los procesos de poder e interés se contemplan como los *mecanismos* que pueden explicar la conducta de los actores. Por consiguiente, este enfoque coincide con lo que Kogan (1984: 62) denominó como el *mayor dato* (sic-DFW) *descriptivo* a nivel institucional de la educación superior, especialmente el modelo político de Baldrige (1971).

1.1.1. *Conducta individual y poder*

Cada teoría debe iniciarse con uno o varios axiomas que el investigador simplemente formula, pero cuya autenticidad no puede probarse de antemano: «...ya que lo que se emplea para comprobar todo lo demás no puede comprobarse a sí mismo: una cadena de pruebas debe comenzar en algún lugar. Dar tal prueba es tan imposible como inútil» (Bentham, 1982: 12-13). En ese caso, tal axioma contiene el principio que rige la conducta de los actores básicos; por el momento asumiré que se trata de individuos humanos. El motivo de escoger una teoría sobre los *macro* fenómenos, partiendo de *micro* axiomas, es que, en el sistema educativo, como en cualquier otro sistema, «las influencias externas no fluyen en el sistema por un equivalente de ósmosis (ello significaría el abandono de la intervención humana por una metafísica holística)» (Archer, 1981: 31).

El axioma de la *teoría de la conducta política* comienza en, se encuentra siempre en los fundamentos de toda teoría de la *racionalidad subjetiva* del individuo, especialmente en la perspectiva de la nueva teoría de la organización (De Alessi, 1983): el individuo elige la mejor alternativa que conoce. Dicho de un modo algo más formal: *la conducta individual* consiste en la elección de una particular alternativa conductual de entre una serie de alternativas subjetivamente disponibles, con la probabilidad de que cualquier alternativa elegida sea proporcional a la utilidad marginal de dicha alternativa. Nuestro universo se rige por las leyes de la termodinámica, las cuales afirman —por lo que respecta al presente argumento—, que ninguna energía puede ser creada ni destruida (*primera ley*), solamente puede ser transformada en otros tipos de energía que no son igualmente utilizables, y que en cada proceso natural, algún tipo de energía se *pierde* de forma no utilizable (*segunda*

ley) ¹. Por consiguiente, toda criatura viviente necesita una entrada regular de energía para permanecer viva. En nuestro mundo, la energía establece *los límites de la conducta* (Simms, 1983). De acuerdo con ello, se le puede dar algún contenido empírico al axioma de la estimada utilidad subjetiva, introduciendo el siguiente supuesto: la probabilidad de que cierta alternativa conductual sea elegida está en función de la cantidad de *energía* producida por los resultados estimados de dicha alternativa, de la probabilidad estimada de que dichos resultados se alcancen y de la cantidad de energía que el individuo posea. Sin embargo, la mera posesión de las fuentes de energía no es suficiente, el individuo necesita ser capaz de utilizarlas. Además, también debe poseer *información* (teorías) ². Así pues, un concepto básico de esta teoría, *el poder*, puede definirse como la capacidad de un individuo de emplear la cantidad de energía a su disposición (Lieshout y De Vree, 1985: 134). Esta es una definición que sigue la tradición de Thomas Hobbes, quien describe el poder de un individuo «...(para hacer una generalización), como sus medios actuales de obtener algún bien futuro...» (Hobbes, 1981 [1651]: 150). Esta definición atañe no sólo a las situaciones sociales, sino también, y muy especialmente, a aquellas situaciones que pueden describirse como *juegos contra la naturaleza* y que incluyen a un solo individuo. Por supuesto, la definición debe tomarse en sentido universal, ya que para cualquier cambio que un individuo desee hacer en su entorno necesitará poder. De acuerdo con ello, el poder *puede* no ser en sí mismo la última meta que el individuo persiga, pero sin embargo *debe* ser su objetivo inmediato. De estas premisas parte la proposición clave de la teoría: *la conducta individual no se dirigirá exclusivamente a la adquisición de poder* (Lieshout y De Vree, 1985: 133-134).

1.1.2. *Sistemas sociales y reputación de poder*

En los sistemas sociales, p. ej., en aquellos sistemas que incluyen a dos o más individuos, ambos individuos intentan tomar el poder *para obtener algún bien futuro*. Pero si se influencia mutuamente (el supuesto de un sistema social implica que así sea) se siguen consecuencias de

1. Para una atractiva introducción del significado de la termodinámica, véase: Atkins, 1984.

2. No es necesario que sean conscientes. En muchas ocasiones, la mayoría de nuestras teorías son *background knowledge*, inconscientemente tomadas como fundamento o incluso encarnadas en nuestros sentidos. Ver también: Popper, 1965 (1963): 238.

largo alcance. Por *influencia* se entiende toda conducta de un individuo (*i*) que altere las probabilidades de otros individuos (*j* o *k*) al elegir una de sus alternativas de conducta, sin tener en cuenta si el individuo *j* pretende los efectos de esa conducta.

Dado que la probabilidad de que una alternativa de conducta sea elegida depende de la relativa utilidad subjetiva de la alternativa, la influencia presupone un cambio en la utilidad de los resultados estimados de dicha alternativa, en la probabilidad estimada de la consecución de los resultados o en ambos. Por consiguiente, la influencia supone que *j* sea consciente de la conducta de *i*. En otras palabras, la influencia de *i* depende de la información de *j*. Como consecuencia de ello, la influencia no se sitúa primariamente en el poder de *i* (aunque *i* necesite poder para ser capaz en efecto de imponer sanciones positivas o negativas a *j*), sino en la percepción subjetiva de *j* de que *i* posee tal poder. Esta expectación puede ser denominada como *reputación de poder*, credibilidad o quizá prestigio. En las situaciones sociales, la reputación de poder sustituye por tanto al concepto central de poder³. Siempre que los individuos no consideren necesario el comprobar los poderes mutuos (lo cual constituye una alternativa arriesgada, ya que puede conducir a una escalada de poder e incluso eventualmente al uso de la violencia, por lo cual no se elegirá con mucha frecuencia), deberán basarse en la reputación (en palabras de Stigler [1961: 224]: «la reputación economiza la búsqueda»), lo cual se considera a menudo como el aspecto más importante del poder social (entre otros: Barry, 1976: 88; Kaplowitz, 1978: 133; Morgenthau, 1978: 87; De Vree, 1982: 80).

En el contexto de la educación superior, y especialmente en la investigación —para anticipar las siguientes secciones—, la reputación de poder es predominantemente una reputación científica. Una elevada reputación científica se construye por medio de publicaciones; las publicaciones pueden conducir a la obtención de una (mejor) posición; las publicaciones también pueden dar pie a la apreciación por parte de otros científicos, quienes, a su vez, pueden disponer de ayudas u otros tipos de financiación; de ahí pueden seguirse las posibilidades de publicar nuevos artículos. En ese sentido, existe en el sistema social de la ciencia, corrientes de retroalimentación en la reputación de poder o *ciclos de credibilidad* (Latour y Woolgar, 1979; Rip, 1988; véase también la figura 1.

3. Esta teoría del poder social incorpora la *ley de reacciones anticipadas* de Friedrich.

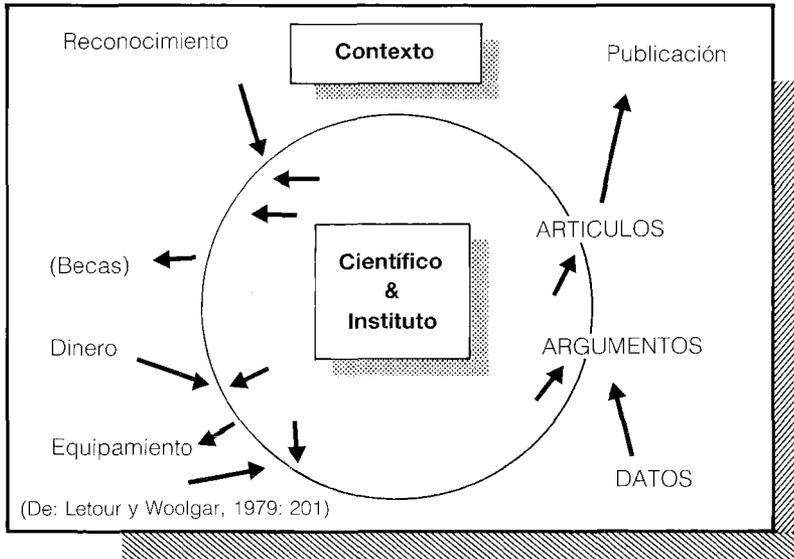


Figura 1. *Ciclo de credibilidad para un científico*

1.1.3. Organizaciones

Siempre que en un sistema social la conducta de los individuos esté equilibrada (que no sea escaladora ni desescaladora) durante cierto período de tiempo y posteriormente se caracterice por una división del trabajo, se la denomina conducta colectiva. La conducta colectiva convierte al entorno en más predecible para cada uno de los individuos, de tal modo que las consecuencias a largo plazo puedan pesar en los argumentos de utilidad de los individuos ⁴. Ello hace posible que empleen la energía a su alcance con más eficacia; en otras palabras, incrementa su poder. Por consiguiente y hablando en términos generales, los individuos tenderán a dejar que persista la conducta colectiva o a crearla en el caso de que no exista. Un esfuerzo consciente y exitoso de llevar a cabo la conducta colectiva se le denominará aquí una *organización*.

La predicibilidad de una organización implica la existencia de reglas que intentan limitar la conducta de los individuos. Ello implica que sea

4. Resulta de utilidad el resaltar el carácter de «suposición» de las teorías, ya que no se presupone que los individuos realicen cálculos de utilidad conscientemente.

necesaria alguna forma de *jerarquía* (bien formal o informal) que permita a los individuos imponer sanciones a otros individuos, con el fin de inducirles a una conducta ajustada a las reglas ⁵. Así pues, una organización puede definirse como: una forma jerárquicamente ordenada e institucionalizada de conducta colectiva para el beneficio de obtener ciertos objetivos (Lieshout y De Vree, 1985: 139).

1.1.4. *Redes inter-organizativas*

Las redes que conforman las organizaciones son sistemas sociales en los que las organizaciones son los *elementos* constituyentes. En principio, pueden considerarse por tanto como organizaciones en el entorno de los individuos. La diferencia estriba en que las redes inter-organizativas aportan menos predicibilidad (menos orden, menos jerarquía) que las organizaciones. En lo que respecta a la conducta de los elementos constituyentes, ello significa que tendrán menos incentivos para realizar inversiones a largo plazo que los individuos de las organizaciones. Sin embargo, en los ambientes gubernamentales existen redes relativamente estables, con un alto grado de especialización de las organizaciones constituyentes, concomitantemente, altos grados de dependencia mutua y relaciones jerárquicas. Esta última característica es una consecuencia de la habilidad del gobierno, por medio de su virtual monopolio de poder, de extraer impuestos de su entorno, que a su vez son distribuidos entre otras organizaciones gubernamentales en concepto de presupuesto anual. Asociado a ello, el gobierno también aporta legitimación a la conducta de las organizaciones gubernamentales pertenecientes a la red.

1.2. *Posiciones y poder en la Red de la Educación Superior*

Tras esta disgresión teórica, retomaré la cuestión central. La *raison d'être* del sistema de la educación superior son la educación y la investigación. Las universidades (y otras instituciones educación superior)

5. Es necesaria una jerarquía, ya que, aunque existan incentivos para que cada individuo mantengan una cierta conducta colectiva, le resultará aún más atractivo el recoger beneficios sin compartir las cargas. Conducta que suele denominarse como de «cabaigar en solitario» o «escurrir el bulto». (Entre otros: Olson, 1971 (1965); Buchanan, 1973; Alchian y Demsetz, 1972.)

pueden ser consideradas como las instituciones centrales del sistema. Para sus medios (económicos) de subsistencia dependen primariamente de las instituciones políticas, o sea, Parlamento y Ministerio de Educación y Ciencia. Desde la perspectiva de las universidades, sus relaciones con las instituciones políticas es lo más importante, simplemente porque dichas instituciones deciden los presupuestos; los presupuestos establecen *los límites de la conducta* de las universidades al igual que los de las demás organizaciones «...el rango de servicios que cualquier agencia puede otorgar queda determinada en último lugar por el dinero que está autorizada a gastar. El dinero tiene la palabra en la administración al igual que en todas partes» (Rourke, 1969: 25). Las relaciones con otras organizaciones son mucho menos importantes para las instituciones de educación superior. Casi todos los recursos económicos son aportados por el ministerio, mientras que los subsidios aportados por la Organización de los Países Bajos para la Investigación Científica (NWO) o los contratos con organizaciones sociales y comerciales (especialmente firmas comerciales) significan solamente una pequeña parte del presupuesto. Además, las instituciones políticas también son importantes en lo que respecta a la legitimización de las universidades en la sociedad (especialmente en cuanto a los efectos civiles de las titulaciones), aunque el campo (nacional e internacional) de la disciplina y en algunas ocasiones las organizaciones sociales y comerciales pueden desempeñar también un papel a este respecto⁶. No debe esperarse que la política de las universidades puedan diferir sistemáticamente, y a largo plazo, de las preferencias de los que establecen las normativas en el ministerio, dado que esas personas que establecen las normas en el ministerio tienen sus preferencias en lo que respecta a la educación superior y tienen la capacidad de usar las relaciones (financieras) para llevar a cabo sus preferencias. Es precisamente esta relación la que convierte al sector de la educación superior en una red inter-organizativa, en un *conglomerado semi-feudal* de organizaciones (Allison, 1971), en el que desde el punto de vista de las relaciones de poder la posición del Ministerio de Educación y Ciencia es superior a la de cada una de las instituciones de educación superior. Aunque hay que añadir que el gobierno no es omnipotente: depende de las instituciones de educación superior para unos muy necesarios resultados educativos y de investi-

6. Respecto a la legitimación, especialmente los estudios de teología en los Países Bajos, dependen más de organizaciones sociales (p. ej., las iglesias) que del gobierno. Algunos de ellos incluso están financiados básicamente por las iglesias.

7. Véase también: Clark, 1983.

gación. Las relaciones llevarán por tanto las características de una dependencia mutua aunque desigual, con la mayor parte del poder concentrado en el gobierno. Más adelante retomaré este punto.

De entre las organizaciones gubernamentales en el campo de la educación superior, el Ministerio de Educación y Ciencia constituye la organización central. El sistema de educación superior e investigación muestra algunas peculiaridades como consecuencia de la *naturaleza especial del conocimiento*⁷. Por expresarlo brevemente, la organización del *conocimiento* es como una gran estructura matriz, organizada de acuerdo con las dimensiones de las disciplinas científicas por una parte y las instituciones educativas y de investigación y las organizaciones del gobierno por la otra (véase la figura 2). Podemos denominarlas como las dimensiones académicas y administrativas, respectivamente.

La estructura matriz tiene consecuencias a nivel individual de universidad. En primer lugar, dicha estructura matriz se repite más o menos a sí misma en la doble estructura de las universidades, por una parte en la organización *burocrática* de la institución y por otra parte en la estructura académica y el campo disciplinario. Según Clark (1983: 17) el aspecto académico de las universidades es: «una estructura plana de elementos entrelazados... una federación, o, mejor dicho, una coalición

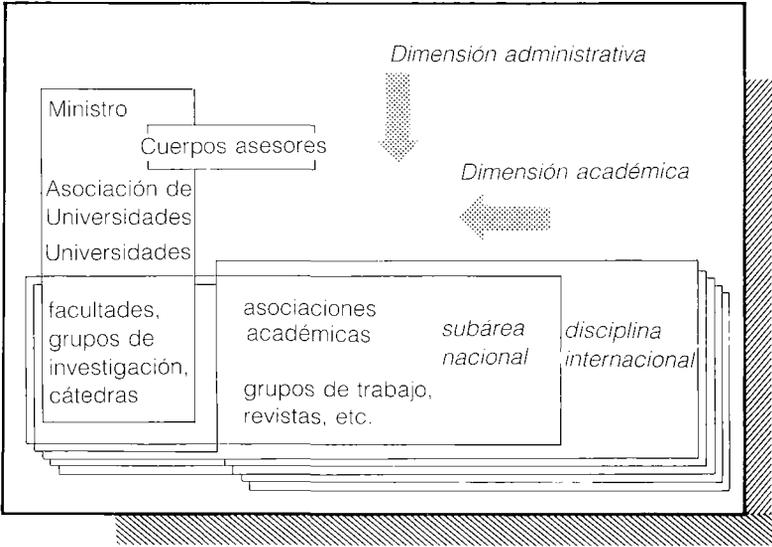


Figura 2. Matriz de la educación superior

más que... una burocracia». A pesar de que en la *teoría de la conducta política todas* las organizaciones son coaliciones (dado que todo individuo participa bajo el punto de vista de su utilidad personal), debe ponerse de relieve el carácter especial de la organización académica; tal y como Clark, Whitley y otros autores han señalado, se parece a una estructura artesanal o gremial en lugar de una estructura burocrática (Clark, 1983: 113; Kogan, 1984: 62; Whitley, 1984: 14-15). Lo cual en este caso significa que las legítimas relaciones de poder estén más basadas en la disciplina que en la institución. Al igual que en los talleres de los maestros gremiales del medievo, la principal unidad operativa o básica es la cátedra⁸. En tales unidades operativas existen relaciones jerárquicas, pero sólo en lo que se refiere a unas cuantas capas. Las relaciones entre las unidades son fundamentalmente horizontales, especialmente con otras unidades de la misma disciplina. Las unidades de una disciplina en su conjunto forman la mayor parte del *subcampo nacional*: la comunidad de científicos de una disciplina, quienes conjuntamente ejercen poder sobre sus miembros y confirman las reputaciones, como por ejemplo a través de la revisión por colegas (Whitley, 1984). Este control sobre los empleados desde fuera de la institución, con su énfasis en la experiencia y la libertad académica, le confiere un carácter bastante difícil al sector administrativo universitario.

Ni puede controlar detalladamente el proceso primario de la docencia y la investigación, ni tampoco dispone de derechos exclusivos o posibilidades para el desempeño de su función. En cambio, debe compartir con el subcampo nacional la influencia sobre los procesos primarios. Todo ello convierte a la cátedra o unidad básica no sólo en un punto central de una universidad, sino que le confiere un carácter inusualmente fuerte y central en los procesos de toma de decisiones» (Clark, 1983: 31-32).

En segundo lugar, la estructura matriz divide a su vez el aspecto académico de una universidad en líneas disciplinarias. Se sugiere y con razón que: «la unidad central de los sistemas académicos está orientada hacia la disciplina, dado que es más costoso abandonar el propio campo de experiencia que el abandonar la propia universidad o facultad...» (Clark, 1983: 33, 30). Por consiguiente, las líneas disciplinarias de las universidades son fisuras reales.

8. La descripción de las relaciones universitarias están centradas en la situación europea y más concretamente en la holandesa, más que en la situación norteamericana. Para una comparación véase: Clark, 1983.

Lo que se aplica a la dirección de una universidad también puede aplicarse a altos niveles de la dimensión administrativa: el ministerio carece de la experiencia necesaria para supervisar detalladamente el aspecto académico de las universidades, al menos del modo como lo realizan los administradores de cualquier universidad. *El poder monetario* que ostenta el ministerio significa primordialmente un poder sobre la dimensión administrativa de las universidades. En lo que respecta a la dimensión académica, el poder financiero del ministerio se limita a conceder o negar medios económicos, sin una posibilidad *directa* de comprobar que éstos sean empleados del modo que el ministerio pretende. Ello no implica por supuesto que tal poder no carezca de importancia. Sin embargo, y tal como ocurre en numerosos países, los actores del gobierno central parecen estar cada vez más convencidos por la idea de que necesitan más información acerca de las actividades de las instituciones de educación superior y de que deben diseñar un conjunto de estrategias e instrumentos efectivos para influir en la actuación de dichas instituciones (Van Vught, 1989).

Desde el punto de vista de estos actores del gobierno central, ello constituye un motivo para diseñar y llevar a cabo cambios de largo alcance en el sistema de educación superior. Los nuevos procedimientos empleados en Holanda para evaluar la calidad de la educación superior pueden considerarse desde esta perspectiva.

2. **Conducta política en los procedimientos de evaluación de la calidad**

2.1. *Nivel intermedio: Gobierno frente a las instituciones*

En el presente artículo, es mi deseo el poner de relieve que la gestión de la calidad no puede ser analizada desde el punto de vista de un solo actor, sino que se trata de una interrelación entre el gobierno y los restantes actores (universidades, facultades y personal académico), así como entre estos últimos actores entre sí. Esta observación podría carecer de importancia en el caso de que coincidieran los puntos de vista, pero distintas posturas dentro del entremado dan lugar a *prioridades y percepciones parroquiales* ante los distintos objetivos⁹. En primer lugar,

9. Además, las categorías de actores pueden ser a su vez subdivididas, lo cual aquí no se tendrá en cuenta.

se centrará la atención en las metas de nivel político medio, establecidas entre el gobierno y las instituciones. Posteriormente, pasará a referirme a las metas intra-institucionales.

2.1.1. *Metas del proceso de la calidad según el gobierno*

Una importante meta de toda la política del gobierno, desde la subida al poder del gobierno neo-conservador en 1982, es la *rentabilidad del dinero*, lo cual tiende a ser operativizado en forma de reducción del presupuesto. En lo que respecta a la educación superior, se trata más de una coacción que de una meta-coacción que siempre se cierne como *telón de fondo* y que produce importantes efectos.

Oficialmente, las metas de la educación superior apenas se prestan a discusión —no se encuentran fácilmente en las publicaciones del gobierno—. Parecen tener un carácter evidente y todo el mundo parece estar de acuerdo con ellas. Según la Constitución, el gobierno es responsable de la *salud* de la educación, lo cual no deja de ser una meta bastante abstracta. Por ejemplo, en los últimos años, se ha creado la Inspección de Educación Superior, pero hasta el momento presente no se ha clarificado cuales son las tareas concretas de dicha Inspección. Ni tan siquiera se reflejan con suficiente claridad las metas del gobierno con respecto a la calidad y evaluación de la calidad en el documento clave de la política educativa contemporánea en los Países Bajos, el informe HOAK. Las declaraciones apenas van más allá de la siguiente afirmación: «A los procesos y resultados en educación superior... se les ha prestado menos atención explícita, en lo que respecta al control de la calidad, de lo que sería deseable hoy en día» (Ministerio de Educación y Ciencia, 1985: 35). Dicho control de calidad debe ser *fiable*, lo cual significa que deben emplearse procedimientos tales como la revisión por colegas y los indicadores de rendimiento, siendo el ciclo de control de calidad responsabilidad de las instituciones en primer lugar.

En cuanto a su aplicabilidad, la definición de control de calidad debe incluir también las demandas de los empleadores (Ministerio de Educación y Ciencia, 1985: 36). Quizá la reticencia por parte del gobierno se ajuste a la filosofía de *mando a distancia*, pero, aparte de la orientación de mercado, manifiesta por la inclusión de las demandas de los empleadores, ello implica que la calidad puede ser cualquier cosa que las instituciones definan como tal. Sin embargo, queda claro que precisa-

mente la calidad es lo que deben garantizar las instituciones de educación superior con el fin de *merecer* la autonomía que el gobierno dice que poseen.

A pesar de que el gobierno desea *controlar a distancia* con el fin de no quedar sobrepasado por los esfuerzos requeridos para dirigir en detalle la educación superior, los costes de la información (o costes de transición) de la política de la educación superior se mantienen elevados: la calidad de la educación y la investigación debe ser evaluada y evidenciada ante el gobierno, ya que de otro modo dicha información no podría ser utilizada en la elaboración de normativas. Sin embargo, el gobierno ha conseguido con éxito el traspasar la mayor parte de la carga a las instituciones de educación superior y a las organizaciones intermedias ¹⁰ que han accedido a realizar por sí mismas las evaluaciones de la calidad (Ministerio de Educación y Ciencia, 1985: 36).

2.1.2. *Metas según las Instituciones de Educación Superior*

Cave *et al.* (1988: 29) citan en el *Informe Jarrat*: «los objetivos y metas de las universidades están definidos solamente en términos muy amplios». Por lo que respecta a las instituciones de educación superior la situación no presenta mejor aspecto que la de las instituciones británicas sobre las que escribe el Comité Jarrat; demasiadas metas de las universidades holandesas no están articuladas o lo están muy vagamente. De todos modos, parece que las metas de los gobiernos actuales no son rechazadas por las instituciones.

Sin embargo, no puede esperarse que las instituciones de educación superior suscriban la meta gubernamental del *valor del dinero*, al menos no en el sentido de que deseen menores presupuestos. Según la teoría *principal-agent*, se supone que los dirigentes institucionales maximizan su *residual* (Alchian y Demsetz, 1972), lo cual puede conseguir, entre otros medios, por medio de una potencialización del presupuesto universitario (Niskanen, 1971). Otra medida para maximizar el residual consiste en supervisar estrictamente las actividades del profesorado. Ello implica que exista una armonía de intereses entre el gobier-

10. Sin embargo, el Ministerio de Educación y Ciencia ha subvencionado de manera sustancial el inicio del ciclo de control de la calidad, especialmente en HBO.

no y la Junta Rectora de las instituciones: ambos desean información acerca de las actividades en las *unidades básicas*.

Además de estas metas nacionales y organizativas, las instituciones desarrollan metas adicionales (o eslogans) para subrayar su *identidad corporativa*, como, por ejemplo: llegar a estar entre las cinco primeras universidades o convertirse en una universidad *empresarial*.

2.1.3. *Organizaciones intermedias*

Las organizaciones intermedias (especialmente la VSNU —Asociación para la Cooperación de las Universidades Holandesas—) son, en muchos aspectos, organizaciones de las administraciones centrales de las instituciones de educación superior. Sus puntos de vista coinciden en gran medida con las de las Juntas de Gobierno. Actuando en el sistema a nivel nacional, algunos aspectos de los puntos de vista del Ministerio de Educación y Ciencia puede esperarse que aparezcan también en los objetivos de la VSNU. La labor desarrollada de estas organizaciones acerca de la evaluación de la calidad no constituye una parte considerable de sus actividades totales y, aunque iniciadas recientemente, estas tareas pueden ser o llegar a ser parte esencial de la organización (aquellas tareas que una organización considera como los objetivos primarios que deben alcanzarse).

La KNAW (Real Academia Holandesa de Ciencias), la organización central en el proceso VF (sistema de financiación condicional de la investigación), es un caso aparte a este respecto. Creada a principios de siglo, mantiene una posición independiente y consolidada. No está tan estrechamente asociada a las instituciones de educación superior como lo está la VSNU. Más bien puede considerarse como una organización de los sub-campos nacionales. La organización de los comités de colegas para evaluar la investigación universitaria no constituye una parte importante de su labor, ni tampoco es parte esencial (por el momento) de la KNAW. Lo mismo puede decirse *mutatis mutandi* de los restantes cuerpos organizativos incluidos en el proceso VF (NRLO —Comité Nacional para la Investigación Agrícola— y KIVI —Real Institución de Ingenieros—).

2.1.4. *Relaciones entre el gobierno y las instituciones*

La mayor parte del *juego* de la evaluación de la calidad en la educación superior se juega entre estos dos actores: el gobierno frente a las instituciones y sus representantes, las organizaciones intermedias.

He afirmado anteriormente que el gobierno y las instituciones sostienen una relación de *mútua dependencia*. Dado que todos los elementos de un sistema social son mutuamente dependientes, *por axioma*, esta afirmación no aporta ninguna información adicional. Algo más informativa es la afirmación de que esta *mútua dependencia* es desigual, estando concentrada la mayor parte del poder en el ministerio de Educación y Ciencia. El ministerio aporta los medios de subsistencia y parte de su legitimación a las instituciones de educación superior. El presupuesto es la base de gran parte de la relación burocrática entre la institución y sus facultades y, quizá aun más, entre la institución y el Ministerio de Educación y Ciencia. Por consiguiente, la evaluación de la calidad es potencialmente amenazadora para las instituciones. A su vez puede ser también prometedor, pero en momentos en que la ciencia esté *en estado de reposo* (Ziman, 1987) o en que se produzcan recortes presupuestarios, pesa más la amenaza que la probabilidad de conseguir altos presupuestos. En tales circunstancias es de esperar una conducta a la defensiva por parte de las instituciones.

La posición de poder que sustenta el gobierno se demuestra a partir del hecho de que sus iniciativas no acostumbran a poder ser vetadas por las instituciones, pero también es manifiesta su debilidad en lo que respecta a la información, por ejemplo en el procedimiento VF. Este se ha diseñado para diferenciar entre la mala, buena y excelente investigación, con el fin de que los presupuestos se distribuyan de menor a mayor entre las tres categorías. Pero, sin embargo, durante su realización la información que obtuvo el gobierno fue sobre si los programas eran suficientes o insuficientes (lo cual apareció rara vez). Por consiguiente, en el caso del procedimiento VF, puede llegarse a la conclusión de que las inversiones del gobierno para la adquisición de información obtuvieron muy poco rendimiento.

El poder de las instituciones se basa fundamentalmente en la información: solamente ellas pueden ofrecer educación e investigación; el gobierno no puede organizar, sin un gran esfuerzo, instituciones alternativas de educación superior (Bijleveld y Goedegebuure, 1989: 107), aun-

que en principio posea más que poder que cada institución aislada — razón por la que afirmo que el gobierno posee más poder que cada una de las instituciones de educación superior—. Sin embargo, permanece la fuerza contraria basada en la información: solamente las instituciones de educación superior (y sus organizaciones intermedias) pueden aportar al gobierno información acerca de sus actividades.

Las Juntas Rectoras de las instituciones están en desventaja a este respecto: con el fin de supervisar sus facultades y sus unidades básicas requieren la misma información que solicita el gobierno, pero para potenciar al máximo su residual no deben a su vez estar supervisadas demasiado estrechamente, por lo que no desean que el gobierno obtenga demasiada información. Ello podría explicar la cooperación de las instituciones y sus organizaciones intermedias con los procedimientos inducidos por el gobierno de evaluación de la calidad, en forma de revisión por colegas, mientras que de modo simultáneo se oponen a los esfuerzos del gobierno por introducir los indicadores de rendimiento en estos procesos de evaluación de la calidad, —lo cual facilitaría la comparación entre las diversas instituciones (Vroeijenstijn y Acherman, 1990: 93-94).

En la esfera de las reputaciones científicas, la evaluación de la calidad puede igualmente ser prometedora o amenazante. Si se contempla la distribución de reputaciones sobre un número fijo de elementos en la estructura de la educación superior, la situación resulta un juego de suma cero: el aumento de una posición por parte de una institución significa una pérdida para una o más instituciones. El hecho de que una institución invierta sus esfuerzos en potenciar la calidad de su investigación depende de la distancia que aquella perciba entre su posición en el ránking y las posiciones (y conducta) de sus competidores. Si la evaluación de la calidad de sistema de educación superior en su conjunto está en el candelero (por ejemplo en comparaciones internacionales) todas las instituciones tienen un gran interés por conseguir una alta calidad, aunque en ambos casos debe tenerse en cuenta que las instituciones pueden valorar distintamente su reputación científica: por ejemplo, una universidad que pretenda estar entre las cinco primeras universidades europeas en lo que respecta a investigación le dará una mayor estimación a su reputación científica que otra universidad más encaminada a la formación de *undergraduates*.

Ya sea desde una armonía de intereses o desde el poder ejercido por el ministro, las universidades (y sus organizaciones intermedias) in-

vierten en la información recopilada por la evaluación de la calidad. No debe subestimarse esta inversión: la VSNU calcula el gasto del procedimiento de visitas, incluyendo la preparación de los auto-estudios, en una media de 350.000 florines holandeses por disciplina (equivalente a unos 150.000 ecus). Cada institución debe contar con 31.000 florines o unos 13.000 ecus (VSNU, 1988 a: 28-34). Sin embargo, la recompensa de tal inversión puede significar tanto una mejora de posición a nivel institucional de la Junta Rectora como unos efectos positivos en el juego contra el gobierno.

2.2. *Nivel institucional: Juntas Rectoras frente a Facultades y Unidades Básicas*

Aunque la mayor parte de las implicaciones de la evaluación de la calidad de desarrollan en el nivel intermedio tanto de gobierno como de instituciones, las propias evaluaciones se llevan a cabo mediante interacciones entre los científicos a nivel individual o en pequeños comités y las instituciones de educación superior o facultades. Aunque, *en el juego de la evaluación*, puede esperarse que las consecuencias de más alto nivel conduzcan a una conducta previsor, de tal manera que en lugar de una relación *director-agente* se origine un triángulo con un *agente* y dos *directores*, lo cual implica el análisis. Anteriormente mencioné a los dos *directores*, así pues, me centraré a continuación en los dos *agentes*.

2.2.1. *Áreas científicas*

Los científicos que evalúan la calidad de los programas de investigación según el procedimiento VF, así como los comités visitantes que figuran en las evaluaciones educativas organizadas por la VSNU (les denominaré *asesores de la calidad* para abreviar) ocupan, temporalmente, posiciones en las encrucijadas entre las dimensiones administrativas y disciplinarias de la educación superior. Sus posiciones permanentes se encuentran generalmente en las *unidades básicas* de las instituciones de educación superior, en las que la dimensión disciplinaria tiene también una poderosa influencia (Spaapen y otros, 1988: 14, 18).

Las áreas científicas juegan un importante papel en los procedimientos de evaluación de la calidad de los Países Bajos, dado que dichos procedimientos están más o menos organizados según las líneas científicas. Las áreas científicas, como toda organización social, tiene normas implícitas (Kuhn, 1970 [1972]). Dichas normas pueden entrar en conflicto con las metas de las instituciones. El poder relativo de las demandas y objetivos de la dimensión científica varía según la dependencia mutua entre los investigadores de una disciplina; en el caso de que sean paradigmáticamente homogéneos el poder es elevado, siendo menos en las disciplinas menos cohesionadas (Whitley, 1984).

2.2.2. *Relaciones de poder dentro de las instituciones*

Tradicionalmente, la organización de la universidad holandesa se ajustaba fielmente al modelo del *sistema basado en la cátedra* (Clark, 1983). Ello implica que el equilibrio de poder dentro de las universidades se decantan hacia los poseedores de la cátedra, o sea, a los catedráticos. Los objetivos de las altas esferas de las instituciones contaban menos que los objetivos de los catedráticos. Desde aproximadamente 1970, el equilibrio de poder ha sufrido un cambio de dirección hacia los niveles medios de la institución, p. ej., los grupos de investigación y el profesorado (Spaapen y otros, 1988: 13-14). En los últimos años, el nivel central de la institución ha ganado mayor poder como consecuencia de las ajustadas condiciones financieras; la dependencia de las unidades de nivel inferior respecto al nivel central se ha clarificado con respecto a épocas anteriores¹¹. Así pues, los objetivos administrativos han adquirido mayor peso en la toma de decisiones de las instituciones de educación superior. Sin embargo, existen diferencias significativas en situaciones concretas: en algunas instituciones, el personal administrativo *siempre* ha disfrutado de un mayor poder respecto al personal académico que los administradores de otras instituciones.

2.2.3. *Las posiciones del profesorado y de las unidades básicas*

Puede suponerse que el profesorado y las unidades básicas consideren la evaluación como una estrecha supervisión de sus actividades.

11. Al igual que ha ocurrido en el Reino Unido (véase Kogan, 1984: 66).

Una estrecha supervisión implica menor nivel de autonomía académica de la que existía previamente, o, en los algo menos honoríficos términos de la teoría *director-agente*: menos posibilidad de esquivarse. La cooperación del profesorado en los procedimientos de evaluación de la calidad es, por tanto, peor para sus propios intereses en cuanto más información generen las evaluaciones para los *directores*. Sin embargo, la dependencia económica del profesorado respecto a estos actores de alto nivel impide su no-cooperación; para el profesorado, la cooperación en los procedimientos de evaluación de la calidad es considerada como un mal menor. Las posiciones de estos *agentes* han deteriorado sus relaciones con las autoridades académicas (a nivel central de la institución, del gobierno o de ambos).

Cuando se contempla este proceso más detalladamente, se descubre que a nivel de facultad la estructura matriz del personal administrativo y académico tiene también sus consecuencias: los primeros intentan conseguir alguna influencia sobre los segundos mediante la adquisición de la información aportada por los procedimientos de evaluación. Desde este punto de vista, la cooperación en la evaluación de la calidad puede significar una ventaja para los administradores. A este respecto, deseo hacer hincapié que desde el punto de vista de la calidad, incluso el personal académico puede percibir ciertas consecuencias positivas derivadas de la evaluación de la calidad, especialmente la obtención de cierto tipo de retroalimentación en su docencia e investigación a partir de los alumnos y de los colegas de la facultad (en los auto-estudios) o de los colegas externos (los asesores de la calidad).

A nivel de facultades pueden hacerse distinciones entre los diversos tipos de disciplinas. En las disciplinas paradigmáticamente homogéneas es posible alcanzar una unidad de criterio a nivel nacional e incluso internacional respecto a la calidad, mientras que en las disciplinas paradigmáticamente heterogéneas es prácticamente imposible (Whitley, 1984; Spaapen y otros, 1988: 15; Van der Meulen y Westerheijden, 1990: 15). Las disciplinas paradigmáticamente heterogéneas no sólo se encuentran en las humanidades y las ciencias sociales, sino también en las áreas científicas: en el campo de la ingeniería mecánica, por ejemplo, recientemente casi todos los profesores, a requerimiento del Consejo Asesor de Política Científica (RAWB 1989: 26), se abstuvieron de juzgar la investigación ajena a su propia especialidad dado que existía muy escasa comunicación entre las especialidades. Así pues, los profesores ocupan distintas posiciones en lo que respecta a los procedimien-

tos de evaluación de la calidad. Los profesores de las disciplinas heterogéneas pueden ejercer una mayor presión en favor de una evaluación de la calidad más flexible respecto a la disciplina, argumentando que en su modo de enseñar o de investigar existen distintas tradiciones o de que *tal o cual* indicador de rendimiento no es adecuado o de que el comité visitante es parcial. En la otra cara de la moneda aparece el hecho de que una disciplina homogénea tiene una posición más ventajosa frente a la institución y el gobierno (Spaapen y otros, 1988: 17-18). Por consiguiente, puede soportar mejor las presiones que una disciplina heterogénea. En ese caso, los procedimientos de evaluación de la calidad puede dirigirse primordialmente hacia los objetivos intrínsecos de la disciplina, en lugar de ser una mezcla de objetivos intrínsecos y objetivos políticos, predominantemente (Spaapen y otros, 1988: 18; De Veert, 1990: 68). Así pues, para el profesorado de disciplinas homogéneas la evaluación de la calidad puede resultar menos amenazante que para los de las disciplinas heterogéneas.

2.2.4. *La posición de los asesores de la calidad*

A nivel individual, la evaluación de la calidad requiere una considerable inversión de tiempo: tanto el presidente como los secretarios/as pueden emplear la mitad o incluso todo su tiempo disponible durante un año en el comité de reconocimiento¹². Los miembros ordinarios de los comités visitantes deben conformarse con algo más de una semana de trabajo (VSNU, 1988a: 33), aunque ésta parece ser una estimación bastante conservadora (información privada).

Como se ha reflejado anteriormente, los asesores de la calidad se encuentran entre dos *directores*. Además, su posición está en la encrucijada entre la institución como organización y la subárea nacional (la organización social de la disciplina en el conjunto del país). Pero, dado que en los últimos años la oposición entre los dos subsistemas y el gobierno ha aumentado, dentro de la política de la educación superior, respecto a la diferencia entre la subárea nacional y la institución, podemos mantener —en aras de una simplificación— que solamente hay dos *directores* con los que contar. Esta situación queda reflejada en la figura

12. Datos aportados por el proyecto de investigación «Evaluation of reconnaissance committees» de Van der Meulen y Westerheijden (1990).

3. ¿Qué representa esta situación para la conducta de los asesores de la calidad?

Una vez concluido su trabajo, los asesores de la calidad deben regresar a sus puestos permanentes de científicos en las universidades (o profesionales en su campo). Sus reputaciones para el poder están asociadas, en términos generales, con dichas posiciones permanentes. Por consiguiente, la presión hacia las subáreas nacionales y la institución será fuerte. Si los asesores de la calidad sucumben ante esta presión desarrollarán una *identificación con el área*. Por otra parte, una *identificación con el gobierno* significaría que los asesores de la calidad conceden prioridad a otras consideraciones a más corto plazo: p. ej., un informe acerca de la calidad con resultados de *bueno* o *excelente* ante los ojos del gobierno puede elevar la propia reputación para el poder ante los funcionarios del gobierno, los cuales pueden jugar un papel importante en el éxito del asesor, p. ej., aportando dinero para llevar a cabo propuestas, o bien en la utilidad del asesor o de su disciplina, como, p. ej., consiguiendo protección en épocas de recortes presupuestarios.

Unida a una identificación con el gobierno aparece la tendencia del asesor de la calidad a presentar un informe *bueno* o *excelente* a los ojos de la organización intermedia que ha realizado los procedimientos de evaluación. En ambos casos, el asesor de la calidad intentará encarecidamente entregar un informe que aporte al *director* la información que deseaba. Finalmente, existen razones estratégicas para que los asesores de la calidad no se aparten excesivamente del gobierno, aun en el caso de que se identifiquen con el área, ya que si los procedimientos de evaluación de la calidad no son considerados positivos por el gobierno, existe el peligro de que sea la Inspección la que asuma la tarea de evaluar la calidad de la educación superior. Para los científicos, los comités visitantes de colegas que informan a una organización intermedia más o menos controlada por la propia organización son menos amenazantes y, por tanto, preferibles a las visitas de los funcionarios del gobierno.

Ambas identificaciones son posibles; teóricamente no puede predecirse cual de ellas elegirán los asesores de la calidad. Empíricamente, los comités de reconocimiento (que pueden considerarse próximos a los asesores de la calidad) eligen ambas posturas. De entre doce comités de reconocimiento estudiados, seis se identificaban con la tarea asigna-

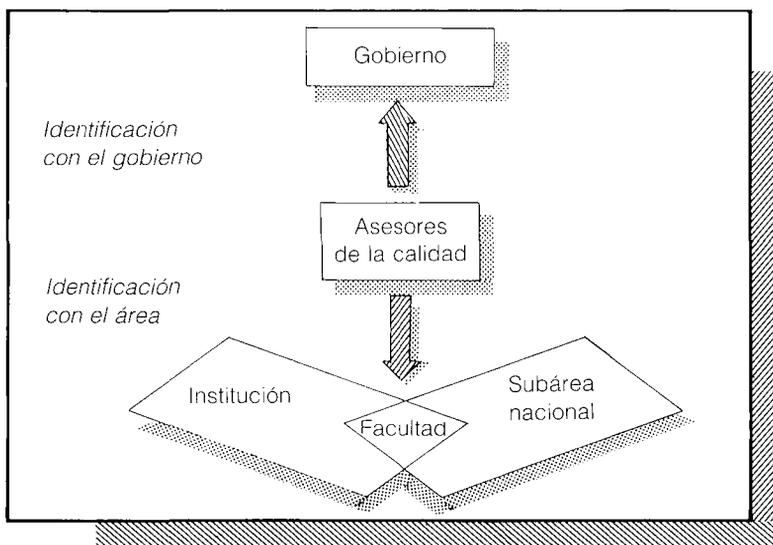


Figura 3. Posición de los asesores de la calidad.

da por el gobierno, dos ocupaban una posición intermedia y cuatro se identificaban con el área. De entre estos cuatro últimos, dos intentaban claramente defender su disciplina frente a las intenciones del gobierno (reducción); incluso evitaban el emitir juicios acerca de la calidad ¹³. Por el momento no existen resultados de investigación sobre la identificación de los comités visitantes según el procedimiento VSNU. Debido a las características de su tarea asignada, la cual, aparte del informe final para la VSNU, incluye sugerencias *in situ* a las facultades acerca de los contenidos y procesos de la educación, su posición está próxima a la subárea nacional y a las instituciones. El énfasis mostrado en los informes de los comités durante su primera ronda de visitas acerca de las restricciones impuestas por el gobierno a una educación de calidad (escasos presupuestos y currícula insuficientes) puede considerarse como una muestra de su identificación con las facultades frente al ministerio (VSNU 1988c: 13-14).

13. En uno de ambos casos, el comité de reconocimiento de Económicas, la estructura paradigmática de la disciplina, conjuntamente con varias sub-disciplinas dispares, no fue concluyente en cuanto a una evaluación global de la calidad.

3. Observaciones finales

3.1. *El dilema de los procesos de evaluación universitaria*

En las secciones precedentes se han analizado las posiciones, poder y objetivos de los principales actores participantes en los procedimientos de evaluación de la calidad de la educación superior holandesa. ¿Pueden estos datos, en combinación con el argumento teórico presentado anteriormente, llevar a conclusiones sobre los aspectos políticos de los procedimientos de evaluación de la calidad de la educación superior de Holanda?

Aunque pueda considerarse trivial, una primera observación es que, por razones de utilidad propia, la evaluación de la calidad no debe analizarse como una materia aislada, ya que está inextricablemente mezclada en los *círculos de credibilidad* de la educación superior, en los que se combina con escasos recursos y reputaciones científicas.

Dadas las diversas posiciones de los actores implicados, la evaluación de la calidad posee distintos significados para los distintos actores. Para el gobierno es un medio de conseguir información acerca de la educación y la investigación que puede emplear en su política. Para los líderes institucionales (Juntas de Gobierno) es a la vez un medio de supervisar las actividades del profesorado y de las unidades básicas inferiores y (en el caso de que no permita una supervisión demasiado estrecha de las Juntas de Gobierno por parte del Ministerio) un medio de intentar maximizar el presupuesto. En cuanto al profesorado, la elección de cooperar con la evaluación de la calidad es considerada como un mal menor; se trata de los *agentes* cuya posición queda más perjudicada.

Continuando con las observaciones previas, cabe esperar que los actores estén en su mayoría a favor de los métodos de evaluación de la calidad que les proporcionen un mayor poder; el poder lo proporciona la información que permite la supervisión de los actores del nivel inferior mientras que es impermeable a la supervisión de los actores del nivel superior.

Por consiguiente, la popularidad de los indicadores de rendimiento, que permiten visiones detalladas y comparaciones de las actividades de las instituciones puede ser explicada por estos motivos. Los aspectos

tos de poder quedan todavía más manifiestos si se observa que: «El Gobierno no está interesado en aquellos indicadores de rendimiento que expresan el rendimiento de los objetivos institucionales, sino solamente en aquellos indicadores que muestran el nivel de rendimiento de los objetivos del sistema» (Maassen y Weusthof, 1989). La preferencia de las organizaciones intermedias, instituciones y facultades por aquellos métodos que requieren una interpretación por parte de expertos (como, p. ej., ellos mismos o sus colegas más próximos), en otras palabras, por métodos en la línea de la revisión por colegas, se deriva del mismo argumento ¹⁴.

Dado el equilibrio de poder (dependencia mutua, pero desigual) entre el gobierno y las instituciones, no puede predecirse el resultado de los procesos políticos que envuelven la evaluación de la calidad. El poder del gobierno, basado en derechos formales y en la dependencia económica de las instituciones de educación superior, se muestra en la imposición de sus normativas en la educación superior, como en la estructura de doble línea con el procedimiento VF por un lado y la nueva filosofía directiva del documento HOAK. El concepto clave del HOAK de *gobernador a distancia*, en sí mismo un reconocimiento de los límites del poder del gobierno; sin embargo, el gobierno todavía impone su voluntad y los instrumentos de su nueva política a los demás actores de la educación superior, quienes deben aceptarlos sin la posibilidad de poder cambiarlos sustancialmente (Van Vught y Westerheijden, 1989) ¹⁵.

En el caso de que las instituciones de educación superior actúen como una colectividad, se produce un monopolio bilateral, pertrechándose frente al gobierno. Estando las posibilidades de una acción colectiva, *ceteris paribus*, negativamente relacionadas con el tamaño del grupo (Olson, 1971 [1965]) es de esperar la existencia de un monopolio bilateral en el sector universitario —13 universidades— en mayor grado que en el sector HBO —alrededor de unas 60 instituciones—. Los efec-

14. Compárense, p. ej., los puntos de vista adoptados por los funcionarios del gobierno Paardekooper y Spee (1990: 56-57) con los de los miembros del VSNU, Vroeijenstijn y Acherman (1990: 93-97).

15. Frente a este razonamiento puede argumentarse que las instituciones de educación superior no desean cambiar los nuevos instrumentos de control, dado que perciben que pueden serles de utilidad. La historia de la introducción de nuevas medidas políticas, como, p. ej., la estructura de dos líneas, muestra que las instituciones de educación superior son incapaces de cambiar aquellas normativas a las que se oponen sistemáticamente (Bijleveld, 1989).

tos de una acción colectiva pueden observarse en el procedimiento VF, en el que, como he mencionado anteriormente, las universidades se las arreglaron para ocultar al gobierno la mayor parte de la información, de tal modo que la política resultó poco desfavorable (prácticamente no se produjo ninguna redistribución de los fondos). Las universidades emplearon su monopolio colectivo de la información que el gobierno necesitaba para su política. El desarrollo de la evaluación de la calidad educativa no se encuentra todavía en un estadio que permita conclusiones comparativas.

En el presente artículo se reflejan numerosos cambios acaecidos en el desarrollo de las relaciones entre el gobierno holandés y las instituciones de educación superior durante los últimos años. Respecto a la nueva filosofía de *mando a distancia* se hace especial hincapié en el carácter *mandatario*, que es primordialmente una estrategia política para reemplazar los anteriores instrumentos de control (instrumentos detallados significan costes elevados de información) por medios más globales de supervisar a los *agentes* de la educación superior. Sin embargo, las evaluaciones de la calidad que constituyen una parte importante de estos medios más globales también tienen sus propios costes de información, costes que, como he cuestionado en numerosas ocasiones, no son precisamente bajos.

Finalmente, deseo retomar la primera observación que hice en esta sección, ya que el trillado argumento de que la evaluación de la calidad está unida a los procesos políticos de la educación superior tiene consecuencias de largo alcance. Todos los actores cooperan en los procedimientos de la evaluación de la calidad por la razón de que esperan conseguir de ella efectos positivos en cuanto a su reputación por el poder (o al menos, como en el caso del profesorado, efectos menos negativos de lo que les acarrearía la no cooperación. Pero cada uno de los actores espera que los efectos positivos se produzcan a partir de otros aspectos o de otros usos de la evaluación de la calidad, distintos a los propuestos por los demás actores, por lo cual se produce un conflicto inevitablemente. ¿No es posible tener una evaluación externa de la calidad sin esos conflictos? Me temo que no. Dentro de las unidades reducidas, o sea, las facultades, pueden operar ciertos sistemas un tanto informales de velar por la calidad, ya que todos los integrantes de la unidad están sensibilizados ante una *buena educación* o una *buena investigación*. En ese caso, el control social puede aportar las necesarias sanciones, tanto positivas como negativas. En los sistemas sociales

más amplios, los aspectos políticos pasan a ser dominantes. En ese caso, si los actores no esperan que se produzcan consecuencias sustanciales (tales como reducción de presupuestos o credibilidad científica adicional) a partir de la evaluación de la calidad, nadie se tomará la molestia de emprender un procedimiento que significa un alto coste de tiempo, energía y dinero. En el momento en que se introduzcan esos intereses concretos, la evaluación de la calidad se convierte en una arena en la que se disputa el juego de poder de las instituciones de educación superior, así como entre las instituciones de educación superior y el gobierno; por consiguiente, la conducta política dominará a la conducta orientada hacia la calidad. Esto es, supongo, el dilema de la evaluación de la calidad no solamente en los Países Bajos, sino en todos los sistemas de educación superior en los que el gobierno central desempeña un rol dominante.

Sin la expectativa de consecuencias reales, los incentivos para organizar la evaluación de la calidad serán inexistentes; con la expectativa de consecuencias reales, la evaluación de la calidad se convierte en un juego de poder.

La cuestión es, por supuesto, ¿qué hacer ante este dilema? Deben los gobernantes evitar el evaluar la calidad, ya que ésta es incapaz de ofrecerles una imagen no distorsionada de la situación? O, por el contrario, ¿deben aceptar el *mundo maquiavélico* y llevar a cabo procedimientos de evaluación de la calidad sabiendo que estarán distorsionados por la conducta política? Pero, aun en ese caso, ¿por qué preocuparse por entrar en un juego de poder de un modo tan rocambolesco?

Una postura intermedia entre ambas posibilidades del dilema puede parecer precaria, aunque no estrictamente imposible. Puede ciertamente encontrarse una vía intermedia si los incentivos son lo suficientemente generosos como para animar a los actores no centrales a cooperar en la evaluación de la calidad, pero al mismo tiempo lo suficientemente recortados como para que se considere no oportuna la opción de una escalada de *combate político* (siendo los métodos de uso del poder caros y arriesgados, en términos de credibilidad o de poder que deben ser invertidas). Dado el fenómeno psicológico de los niveles de aspiración¹⁶ —que generalmente se adaptan mejor en ascenso que en des-

16. Introducido por vez primera en las teorías de toma de decisiones por Simon, 1957: 247, 253.

censo—, esta precaria vía intermedia se encontrará con mayor facilidad cuando se produzca un *pequeño aumento de presupuesto* que cuando la situación esté estancada o incluso con recortes de presupuesto. En el contexto holandés, en el que el presupuesto estatal para educación sólo es sobrepasado por la deuda nacional, no debe esperarse una situación en la que se incrementen los presupuestos: las perspectivas de que prospere una fructífera vía intermedia para la evaluación de la calidad de la educación superior holandesa son, por consiguiente, bastante débiles.

3.2. *El dilema y el procedimiento de evaluación del VSNU*

Por razones analíticas, sería interesante confrontar el dilema con los esfuerzos realizados hasta el momento en los Países Bajos en favor de la evaluación de la calidad, especialmente el procedimiento VSNU. Se pretende que sea un medio para la mejora de la calidad, como se ha indicado anteriormente, no un medio (gubernamental) de control de la calidad. En palabras de Hans Acherman (1990: 182), uno de los participantes en su diseño y experimentación piloto:

«Por consiguiente, no se contemplaban recompensas económicas o sanciones, ni tampoco ninguna forma de ranking. Por el contrario, se consideraba contraproducente el enfoque empleado en ocasiones por el gobierno, consistente en emplear todo tipo de incentivos económicos y amenazas sobre aquellos que cumplieran o no cumplieran las normas.»

Esto parece contradecir lo expuesto en el dilema, principalmente, que un procedimiento de evaluación de la calidad a nivel nacional no puede funcionar sin sanciones. Estoy de acuerdo con Acherman en que las sanciones en previas normativas educativas hubieran producido efectos contraproducentes en la evaluación de la calidad, al mantener el profesorado una actitud reticente al máximo en sus interacciones con las *agencias de control* del gobierno¹⁷. El hecho de que en el procedimiento VSNU no se *contemplaran* sanciones directas de tipo económico o de reputación, no aporta pruebas suficientes de que no existiera algún tipo de sanciones. En primer lugar, uno puede imaginarse que el control social existente dentro de las facultades posee efectos semejantes a los mencionados anteriormente. En segundo lugar, y probablemente más

17. Véanse también Archer, 1981: 45: «...el dominio negociador... (del gobierno) generó gran descontento y oposición».

importante, ciertamente existen sanciones controladas por el gobierno relacionadas con el procedimiento de evaluación de la calidad. El VSNU inició su procedimiento de evaluación de la calidad como consecuencia de unas negociaciones entabladas en 1986 con el gobierno, quien originalmente planeaba confiar la evaluación de la calidad al recientemente creado Inspectorado de la Educación Superior. A las universidades no les quedaba otra elección que, o bien ejercer algún tipo de control sobre este instrumento, o bien dejarlo todo en manos de las *agencias de control* del gobierno. Además, el gobierno no se conformó con renunciar a su control sobre la evaluación de la calidad; a cambio, las universidades tenían que aceptar realmente el seguir los consejos de los comités visitantes. Están obligadas a reaccionar ante los informes (que están a disposición del público) en sus informes y planificaciones anuales o bienales, que a su vez también están disponibles al público. Aparte de la amenaza no demasiado fuerte de la denuncia pública, el gobierno se ha reservado el derecho de intervenir y tomar medidas cuando los consecutivos comités visitantes informan sobre el mismo tipo de problemas en una facultad determinada, sin que la institución tome medidas ante el problema. En ese caso, con una demora de unos 10 años, aparecen las sanciones económicas en este procedimiento de evaluación de la calidad. Ciertamente, la conexión de las sanciones con la conducta es mucho menos directa y, por tanto, menos pronunciada que en anteriores iniciativas gubernamentales, pero, de todos modos, no puede considerarse al procedimiento VSNU una completa falsificación del dilema.

3.3. *Algunas observaciones acerca de las evaluaciones internacionales de la calidad universitaria*

Cualquiera que sea el contenido de la política sobre educación superior a nivel europeo —si es que es posible hablar de política de educación superior europea—, 1992 es una fecha clave para la política de educación superior de los gobiernos y de las instituciones. En este sentido, 1992 se está convirtiendo en una auténtica *profecía*. En los albores de este movimiento aparecen ideas acerca de las evaluaciones internacionales de la calidad de la educación universitaria, así como de comparaciones de niveles nacionales de calidad. Si no ya por razones económicas y políticas (¿pueden competir nuestros graduados en el mercado europeo?), este tema se está convirtiendo en un asunto interesante, con un ojo puesto en el intercambio de programas de estudiantes y de personal (especialmente el ERASMUS) y con el otro en una política de

personal para todos los empleadores (privados y públicos) respecto a los graduados de la educación superior. Tómese como ejemplo el interés despertado por *Liberation* en diciembre de 1989 con su lista de las mejores universidades de Europa.

En esta escena internacional emergen las mismas complicaciones que las observadas en el caso de una sola nación. Es especialmente complicado en el caso europeo el hecho de que la evaluación de la calidad no puede estar aislada de su contexto político, dado que cada país elabora distintos sistemas políticos con sus propias peculiaridades, que a su vez dan pie a distintos intereses, objetivos y puntos de vista de las instituciones implicadas. Además en estos distintos contextos, los objetivos educativos pueden diferir ampliamente (aunque quizá en menor grado en las universidades que en la educación superior no universitaria); así pues, la calidad deberá definirse a escala nacional o institucional, lo cual convierte a la comparación de la calidad en un asunto precario, por no decir otra cosa peor. El número de complicaciones se multiplica por el número de países.

La toma de conciencia de que 1992 se aproxima es en sí misma un incentivo para las instituciones, el gobierno y los demás actores —aunque alguna de sus consecuencias no se vean con claridad—. El dilema queda sin resolverse, más aun, la estructura de los incentivos adquiere una menor transparencia para los actores (incentivos y *contra-incentivos* nacionales y/o europeos) y crece la incertidumbre acerca de lo que deparará el futuro (¿se materializarán los esperados futuros incentivos?) Por estar razones parece que no vale la pena el organizar un *fuerte* y voluminoso procedimiento para la evaluación internacional de las universidades que requiera la recogida de numerosos indicadores de rendimiento; ello significaría una inversión demasiado costosa en términos de trabajo adicional, tiempo y energía frente a unos beneficios escasos e inciertos. Un viable procedimiento de evaluación internacional deberá ser *ligera* y *simple*.

El requisito de tener en cuenta las distintas metas educativas y los distintos contextos de las universidades europeas se contradice con el requisito de mantener un procedimiento simple. No veo otro camino para este *dilema europeo* distinto al dilema presentado anteriormente. Sugeriría que, si bien fuera mejor comenzar por un aspecto del dilema (un procedimiento sencillo, preferentemente una participación voluntaria) y avanzar luego hacia un procedimiento metodológicamente más satisfac-

torio, una vez que las instituciones implicadas y el gobierno hayan *aprendido mediante la práctica* que el balance entre inversión y beneficios es positivo. Ello, sin embargo, no evita la posibilidad de que el balance resulte negativo para las universidades, lo cual no es imposible dado que la mayor parte de los procedimientos de evaluación realizados en Europa son iniciativas del gobierno, pero no parece improbable que las universidades no valoren el hecho de ser incluidas en las florecientes redes europeas de la educación superior y su evaluación.

¿COMO EVALUAR LA CALIDAD DE LA EDUCACION SUPERIOR EN EUROPA OCCIDENTAL? *

FRANS A. VAN VUGHT
Universidad de Twente, Países Bajos

1. Introducción

Ultimamente se están discutiendo nuevos sistemas de evaluación de la calidad de la educación superior en muchos países de Europa Occidental y, a menudo, se están dando los primeros pasos para iniciar la puesta en marcha de tales sistemas. En varios países parece difícil combinar, en dichos sistemas, los objetivos del gobierno en cuanto a los sistemas de educación superior, por un lado, con las opiniones y características de las instituciones de educación superior, por otro.

En este artículo se discuten las experiencias en la evaluación de la calidad de la educación superior en varios países de Europa Occidental, así como en Canadá y los Estados Unidos. De esta discusión deduciré un modelo general de evaluación de la educación superior, en el que se combinan elementos que han sido útiles en varios entornos. Sugeriré que el modelo general que se presentará ofrece un buen punto de partida para el desarrollo de un sistema de evaluación de la calidad en algunos países de Europa Occidental.

2. Experiencias en los Estados Unidos y el Canadá

En países tales como los Estados Unidos y el Canadá, la idea del mercado y la competencia está muy arraigada en la educación superior. Hay un poder considerable en la cumbre de las instituciones de ense-

* El presente artículo es traducción de la ponencia no publicada que el autor presentó en el *I Congreso Internacional sobre Calidad de la Educación Universitaria* (ICE Universidad de Cádiz, Puerto de Santa María, marzo 1991).

ñanza, con una junta directiva o su equivalente y un auténtico director ejecutivo. Esas condiciones existen en estos países porque sus constituciones han delegado considerables poderes en esas instituciones. La influencia gubernamental es limitada. Se supone que las instituciones de educación superior quieren autorregularse para no perder recursos, estudiantes y becarios a favor de sus competidores.

En los Estados Unidos la creciente diversificación de las instituciones y la falta inicial de criterios definidos centralmente, junto a la ausencia de iniciativas de los estados, produjo, a finales del siglo XIX, un cierto nivel de caos en el sistema de la educación superior estadounidense. Si las instituciones no hubieran respondido a este caos, probablemente se habría producido la intervención del gobierno estatal, lo que no era deseado por las instituciones. Por consiguiente, las instituciones desarrollaron en los 50 años siguientes, y han perfeccionado en los últimos 25 años, dos procesos de evaluación de la calidad.

El primero, muchas veces mal entendido en Europa, es la acreditación. La acreditación de una institución de educación superior, o de un programa de estudio en concreto dentro de una institución, consiste en un procedimiento de autoevaluación por la organización que pide la acreditación, seguida por la visita de un equipo de asesores externos y de una discusión final por una junta de colegas que decide sobre la acreditación, utilizando criterios de acreditación preexistentes. En EE.UU. la acreditación tiene dos formas: la de las instituciones, llevada a cabo por agencias regionales controladas por las propias instituciones; y la de las especialidades (o programática), que se lleva a cabo a nivel nacional por organismos controlados por los profesionales.

El segundo procedimiento de evaluación de la calidad en las instituciones de educación superior en los EE.UU., y que de ninguna manera está asociado a la acreditación, es el proceso de examen sistemático de los programas de estudio, que ha sido utilizado en algunas universidades durante más o menos una década para evaluar la calidad de los programas, con el fin de facilitar la selección y, en algunos casos, para establecer una base para la redistribución de recursos marginales (Barak, 1982; Kells & Van Vught, 1988).

En el Canadá, los esfuerzos realizados para evaluar la calidad han tomado un camino algo diferente. Aunque algunos de los sistemas de educación superior, organización a nivel provincial, han discutido la po-

sibilidad de instrumentar un sistema de acreditación, no se ha implantado esa filosofía tal y como se había hecho en EE.UU. En la Columbia Británica, por ejemplo, las escuelas técnicas y los colegios universitarios han decidido por consenso utilizar dos aspectos del sistema de acreditación: la autoevaluación y las visitas por colegas a las instituciones (British Columbia, 1984). De un modo similar, la forma de la evaluación de la calidad utilizada en universidades tales como la universidad de la provincia de Alberta es un enfoque en el cual se utiliza un proceso de autoestudio diseñado por el profesorado y en el que se organizan visitas externas de forma crítica. De este modo, en sistemas de educación superior donde las instituciones tienen un alto grado de autonomía, y donde el poder en la cumbre de la institución es relativamente alto, se han desarrollado dos formas de evaluación de la calidad que han sido utilizadas durante un período de tiempo relativamente largo. La primera es un método colaborativo, consensual, interinstitucional, instituido por los colegas que, o bien utiliza criterios preexistentes e incluso decisiones ministeriales sobre la suficiencia de un programa o institución (acreditación), o bien autoestudios y visitas por colegas, pero que no utilizan criterios y decisiones (como es el caso de la Columbia Británica). La segunda forma de evaluación de la calidad proviene de la propia institución y suele estar bien integrada en el sistema de gestión de la misma.

3. Evolución en Europa Occidental

Contrariamente a los EE.UU. y el Canadá, los sistemas de educación superior en Europa Occidental han sido controlados fuertemente por los gobiernos. En esos sistemas de educación superior, el número de instituciones es relativamente pequeño y la financiación generalmente proviene del Estado.

Durante los años setenta y ochenta los sistemas de educación superior en Europa Occidental tuvieron que hacer frente a un número de cambios de gran envergadura. La mayoría de esos cambios se puede relacionar de una forma u otra con un cambio en las estrategias gubernamentales hacia la educación superior. Un factor de fundamental importancia fue la subida al poder, en muchos de esos países, de gobiernos conservadores. La llamada filosofía del «máximo rendimiento a tenor del dinero invertido» (*value for money*) condujo al fin de la más o menos ilimitada financiación pública de la educación superior por parte de los gobiernos. Se utilizó una nueva ideología para legitimar nuevas

políticas gubernamentales, con el fin de «abolir la anarquía y las pérdidas» (Archer, 1981; 47). En la práctica eso implicaba, entre otras cosas, que la financiación pública de la educación superior iba a estar vinculada cada vez más a los resultados de esas instituciones. Como consecuencia, la cuestión de cómo evaluar los resultados o la calidad de la enseñanza superior llegó a ser uno de los temas centrales de discusión de los últimos diez años en la educación superior de Europa Occidental.

Un segundo e importante cambio relacionado con la toma de decisiones políticas en la educación superior es el auge de la estrategia gubernamental de la llamada «autorregulación» (Van Vught, 1989; Neave & Van Vught, 1991). Durante la segunda mitad de los ochenta, los ministerios de educación y las instituciones de educación superior, especialmente en los países del noroeste de Europa, han dialogado sobre lo deseable que sería aumentar la autorregulación por las propias instituciones de educación superior. Varios gobiernos de países occidentales han recomendado, durante este período, la desregulación por parte de los ministerios, así como una mayor competitividad e innovación por y entre las instituciones de educación superior. Al mismo tiempo, los gobiernos querían más rendimiento por el dinero invertido, una sociedad más consciente de la calidad de los programas educativos, la eliminación de la duplicación en los programas, junto con una organización y una planificación más creativas, y todo esto durante un período de considerable austeridad financiera. Este complejo proyecto ha sido analizado por Van Vught (1989) como una forma diferente de dirección centralizada, tanto indirectamente (con ministerios «dirigiendo a distancia») como *post hoc* (con un énfasis más fuerte en la evaluación *ex-post* de los resultados de la institución).

El desarrollo de estrategias gubernamentales, dirigidas hacia una mayor autonomía de las instituciones de educación superior, fue generado por los agentes gubernamentales con el deseo de crear una mayor innovación, así como un mayor grado de adaptación a las necesidades de la sociedad en cuanto a los sistemas de educación superior. A su vez, se quería una sociedad y una comunidad cada vez más conscientes de la calidad de programas de estudios, lo que significa que se debía desarrollar una evaluación de la calidad creíble, algo que era prácticamente inexistente en los sistemas de educación superior de la Europa Occidental antes de los años ochenta.

En algunos países de Europa Occidental surgieron nuevos intentos de establecer sistemas de evaluación de la calidad a partir de los he-

chos mencionados anteriormente. El sistema de evaluación de la calidad fue iniciado bien por las autoridades gubernamentales centrales (como fue el caso de Francia y el Reino Unido) (Staropoli, 1987; Cave *et al.*, 1988), o bien fue negociado entre los representantes gubernamentales y los dirigentes de las instituciones de educación superior (como fue el caso de los Países Bajos) (Kells & Van Vught, 1988). En otros países se suscitaron discusiones sobre la evaluación de la calidad, pero todavía no se ha introducido ningún sistema (por ejemplo, en Bélgica, Finlandia, Noruega, Suecia y España).

Los informes sobre las experiencias de los países de Europa Occidental con sistemas de evaluación de la calidad han sido relativamente pocos hasta ahora y parecen ser intentos más bien parciales en relación con lo que se esperaba teóricamente de ellos. La experiencia francesa muestra que, en este país, un cuerpo constituido a nivel nacional tiene el poder de inspeccionar las instituciones de educación superior con la prerrogativa de establecer su propia agenda. El enfoque británico de la evaluación de la calidad parece ser el resultado de un proceso altamente político, en el que la campaña para «más rendimientos por menos inversiones» ha enfatizado fuertemente la importancia de los indicadores de rendimiento, «*rankings*» y tendencias competitivas (para un desglose véase: Goedgebuure, Maassen & Westerheyden, 1990).

En los Países Bajos, el esfuerzo más sistemático para una evaluación cíclica de la calidad, realizada por la misma institución, ha sido iniciado por el sector universitario. Durante un año de éxito experimental se llevaron a cabo autoevaluaciones de cuatro disciplinas en todos los centros de enseñanza. Hubo 26 autoevaluaciones, ayudadas por talleres y materiales desarrollados por la propia Asociación para la Cooperación de las Universidades Holandesas (VSNU, 1987) y hubo visitas *in situ* en todos los programas (realizadas por un equipo para cada disciplina, que actuó en todo el país durante un período de tres meses). Posteriormente, el ciclo de las visitas se ha fijado en cinco años y todavía se están debatiendo las características del sistema (Kells & Van Vught, 1988; Vroeijenstein & Acherman, 1990).

En los politécnicos holandeses se está discutiendo un sistema similar con revisiones, tanto de los programas como de toda la institución, que empezará su fase operativa dentro de poco tiempo (Rullman, 1990).

Observando las diversas experiencias sobre la evaluación de la calidad de la educación superior en Europa Occidental, se puede decir,

hasta ahora, que se han suscitado importantes discusiones y que se ha dado un enfoque global a la evaluación de la calidad. En concreto, el sistema holandés parece poder ofrecer un número de elementos útiles que se pueden integrar en un modelo general de evaluación de la calidad. Sin embargo, las experiencias realizadas hasta ahora en Europa Occidental demuestran también que un desarrollo más amplio de la base teórica de dicha evaluación podría influir positivamente en la introducción de un sistema de evaluación de la calidad en uno o más países determinados. Es por eso que se introduce en los siguientes párrafos un modelo general, integrado y autorregulador de la evaluación de la calidad.

4. Un modelo integrado y autorregulado de la evaluación de la calidad

La introducción de una estrategia gubernamental de autorregulación en varios países de Europa Occidental, ofreciendo una mayor autonomía a las instituciones de educación superior, puede conducir a que los ministerios de educación interesados en una menor regulación, en una mayor innovación y en la mejor adaptación a las necesidades de la sociedad, saquen la conclusión de que un sistema de evaluación de la calidad en el que las instituciones de educación superior jueguen un papel importante podría ser un mecanismo adecuado para pedir responsabilidades a las mismas. Además, suponiendo que la mayoría de las instituciones de educación superior prefiriese ejercer más control sobre su propia suerte, tanto en los tiempos buenos como en los malos, parece razonable esperar que dichas instituciones prefieran sistemas de evaluación de la calidad en los que ellas mismas estén presentes en el control de los procesos de revisión.

El siguiente modelo autorregulador de la evaluación de la calidad se basa en la idea de que dicha evaluación debería estar, en gran medida, organizada y controlada por las instituciones de educación superior, y ser producto del esfuerzo colaborativo entre los profesores. Se presenta aquí un modelo como punto de partida general para aquéllos a quienes, en un contexto específico, se les pueda pedir la propuesta de un sistema de evaluación de la calidad.

El modelo se basa en tres cuerpos de pensamiento. El primero consta de las lecciones que se pueden sacar de los sistemas de evaluación

de la calidad en los EE.UU. y el Canadá. El segundo se deriva de la base conceptual utilizada por los planificadores de los proyectos de revisión de programas e instituciones de los Países Bajos y que fue resumido anteriormente por Kells & Van Vught (1988). El tercero es lo que se podría denominar la idea de un sistema integrado que fuese compatible con la supervisión ministerial característica del contexto europeo y con los sistemas presupuestarios y de toma de decisiones de la universidad europea, de modo que se pueda poner en práctica un sistema de revisión de programas.

Las experiencias de los Países Bajos desde 1984, y las obtenidas en el Canadá y los EE.UU. durante un considerable período, indican claramente que el sistema de evaluación de la calidad debe satisfacer algunas características básicas si se quieren conseguir resultados positivos. No todas esas características se dan en este momento en los ejemplos existentes en Europa o en muchos de los sistemas de revisión de programas de los EE.UU.

El primer requisito es que debe haber un acuerdo y un mecanismo adecuado que defina claramente el papel del agente que va a llevar a cabo la revisión, estableciéndose con precisión los objetivos de la misma para, a través de ellos, poner los medios y los profesionales competentes que hagan que el sistema funcione. Para una determinada universidad que quiere organizar sus propias revisiones surge la necesidad de llegar a un acuerdo consensual entre los dirigentes de los programas que se quieran revisar, respecto de las reglas y procedimientos del sistema, para delimitar con claridad las responsabilidades dentro del equipo coordinador y para generar el interés y el entendimiento de los miembros de las unidades. En la mayoría de los sistemas universitarios europeos se requiere un acuerdo negociado con el ministerio de educación (como es el caso de los Países Bajos) para asegurar, por ejemplo, el control de las universidades sobre el sistema de revisión a cambio del compromiso de evaluar todos los programas en un período determinado. En los Países Bajos este acuerdo se basa en un *quid pro quo*, en el que el derecho a desarrollar libremente programas en las instituciones (en respuesta a un llamamiento por parte del ministerio en favor de la innovación y de un espíritu más competitivo) se podía conseguir con el compromiso de desarrollar y llevar a cabo un sistema de evaluación de la calidad. Otras situaciones exigirían otros tipos de negociación. En Holanda se utilizó inicialmente la VSHU para coordinar el proyecto de revisión. Otros abogaban por una organización completamente neutral,

un organismo nuevo, una agencia no-gubernamental o un inspectorado. Para que funcione bien el sistema se necesita una agencia imparcial, que sea respetada, con un personal adecuado, con objetivos claramente definidos y dotada con los fondos necesarios.

El segundo lugar, el sistema se debe basar en la autoevaluación, de modo que pueda ser evolutivo y orientado a fomentar las mejoras. Dicho simplemente, para que los miembros de la comunidad académica puedan hacer los cambios necesarios, deben participar en un proceso que les permita descubrir cuáles son los problemas, proponer soluciones y estar dispuestos finalmente a poner en práctica los resultados. Esto exige un proceso de consideración y de análisis, y no sólo la preparación de un informe descriptivo (Kells, 1988). Sólo mediante este proceso puede cumplirse completamente la obligación que tiene un sistema de evaluación de la calidad de rendir cuentas públicas. En efecto, la experiencia de los Países Bajos en su primer año ha confirmado este hecho dentro del contexto europeo. Las facultades informaron que el proceso, aunque nuevo para algunos e introducido de un modo algo apresurado, había sido útil. De un modo igual, la autoevaluación debería llevarse a cabo teniendo en cuenta dos factores: los objetivos del programa que está siendo revisado, junto con algunas normas del campo en cuestión, y en qué medida dicho programa parece estar funcionando correctamente. Además, para un resultado idóneo, los talleres de autoevaluación y los manuales deberían estar disponibles para poder ayudar a los directores de los programas cuando empiece el proceso. Esto ha funcionado bien en alguna parte de los Países Bajos y su utilización está ampliamente extendida en EE.UU.

También se debe tener acceso, durante el proceso, a los datos adecuados. Esto implica tiempo, energías intelectuales y recursos para recoger la información necesaria y las opiniones sobre el logro de los objetivos y la marcha del proceso de la autoevaluación. Las instituciones de educación superior difieren mucho en su capacidad para obtener información útil y ponerla a disposición de las personas interesadas en los procesos de análisis y planificación, sobre todo a nivel de programas. Los planificadores de un sistema de evaluación de la calidad deben tener en cuenta este tema. Los procesos más eficientes de autoevaluación se basan en encuestas a los estudiantes y en las opiniones de los directores de los departamentos y de los responsables de la facultad. También debe existir un acceso especial a los datos universitarios que se recogen regularmente.

La autoevaluación interna debe ser revalidada y puesta en otra perspectiva mediante una visita externa llevada a cabo por expertos imparciales, que son invitados a pasar al menos dos días en el centro discutiendo el informe de la autoevaluación, revisando otros documentos, entrevistando al personal y a los estudiantes y preparando un informe, tanto escrito como oral, para los participantes en el programa y para los responsables de la universidad. Si los visitantes externos son invitados por una universidad en particular, deben recibir instrucciones sobre la naturaleza de los asuntos que esa universidad quiere abordar y la clase de informes que desea. Los visitantes que formen parte de un sistema a nivel nacional deberían recibir material de orientación sobre la naturaleza del programa de revisión e, idealmente, deberían asistir a un programa de formación para visitantes externos antes de ejercer como tales. Las experiencias del sistema de acreditación en los EE.UU. y del sistema de revisión de programas de las universidades holandesas, donde un equipo visita varios lugares, confirman la utilidad de esa formación.

Los resultados del programa institucional de autoevaluación y las reacciones detalladas del equipo visitante deberían ser confidenciales, para permitir la discusión, la reacción y los cambios sin interferencia externa, para potenciar la confianza en el sistema y en sus dirigentes, admitiendo que algunas cosas no pueden cambiarse rápidamente a causa de la naturaleza de la administración de la institución y la complejidad de sus objetivos y estructuras. Se puede informar públicamente sobre la naturaleza del proceso de revisión (por ejemplo, a través del Ministerio de Educación), dando una información general sobre los resultados de la revisión. El carácter cíclico del sistema de evaluación asegura los cambios y el mantenimiento de los criterios establecidos previamente, ya que los equipos posteriores buscarán información sobre las labores realizadas en el ciclo anterior. El compromiso de los dirigentes de la institución con el sistema y con la puesta en práctica de los resultados asegura también estos objetivos.

Lo ideal sería que el sistema fuese propulsado a través de un conjunto de incentivos, o que existiese la posibilidad, al menos, de que las revisiones tuviesen alguna repercusión. En los procesos de revisión de programas de algunas universidades se ha acordado desde el principio que, con la ayuda de una comisión de profesores y otros miembros del personal, algunos de los recursos marginales de la institución puedan ser redistribuidos, bien competitivamente, bien negociándolos para ca-

da programa en función de los resultados de las revisiones. Tales sistemas han conseguido el respeto en las instituciones en las que han sido empleados. De todas formas, debería haber un sistema continuo para revisar los resultados de la autoevaluación y de las revisiones externas, que se planteara medidas de procedimiento, cambios de política y atención a los problemas y oportunidades que se hayan descubierto o se haya agudizado por las revisiones. Uno de los pasos más importantes que todavía no se han dado en el sistema holandés es la consideración de los resultados en un contexto más amplio, más allá de las facultades revisadas aisladamente. Este punto será discutido más adelante. Finalmente, la institución individual que está desarrollando un sistema de evaluación de la calidad para sus propias necesidades, o bien los planificadores de un sistema a nivel regional o nacional, debe darse cuenta de que su puesta en marcha llevará mucho tiempo, probablemente dos o tres ciclos, si se desea que el sistema funcione bien y logre sus objetivos plenamente. Especialmente en Europa Occidental, donde la idea de la revisión de programas no es tan corriente, se necesitarán como mínimo dos ciclos para que las experiencias positivas del primer ciclo se filtren a la comunidad universitaria, generándose confianza y obteniendo las consecuencias de las mismas. Los ministerios que aprueban un sistema que va a ser dirigido por las instituciones superiores deberían proteger (no interferir) el sistema mientras se está desarrollando la experiencia y se llevan a cabo los primeros cambios como resultado de las revisiones.

Los elementos básicos del sistema de evaluación de la calidad que se han mencionado más arriba todavía no están uniformemente presentes en los sistemas que han sido desarrollados hasta ahora en Europa. Algunos confían en el autoestudio y otros no. Algunos han sido dirigidos por las universidades o por agencias neutrales, mientras que otros no lo han hecho. Algunos parecen tener personal suficiente, otros tienen muy pocos recursos. Algunos han contado con visitantes cuidadosamente seleccionados y preparados, otros parecen no darles tanta importancia. La confidencialidad no se valora igual en todos sitios. La disponibilidad de datos a nivel de programa o de facultad y la existencia de procesos para considerar sistemáticamente los resultados de las revisiones (más allá del programa en cuestión) son puntos que deberían tenerse en cuenta en los sistemas de educación superior en Europa Occidental. Las estructuras y la forma en que el poder está distribuido en las instituciones de la educación superior obstaculizan el flujo de información útil y la construcción de sistemas de datos para todas las facultades. El

nivel de prerrogativas en las funciones de cada facultad individualmente obstaculiza, cuando no hay incentivos, la creación de sistemas que permitirán la evaluación de la calidad en todas las instituciones de educación superior.

5. Aspectos adicionales

Esto nos lleva a la noción de integración. La integración de intereses nacionales o regionales con los de la institución de educación superior y los de sus programas se debe tener en cuenta en cualquier sistema europeo de evaluación de la calidad. Asimismo, poco beneficio se podrá obtener de un sistema que permita a los programas individuales o a las facultades usar el sistema de evaluación de la calidad de forma aislada de sus colegas en otras facultades, de los responsables de la universidad o de otros funcionarios. Además, si los resultados no están integrados adecuadamente en el sistema presupuestario y de planificación de la institución, los efectos del uso de dicho sistema de evaluación serán limitados. Igualmente, es poco lógico concentrarse solamente en las revisiones de programas académicos. Seguramente los sistemas de mantenimiento, las bibliotecas, oficinas, etc., necesitan ser revisados, ya que tienen un impacto importante en los programas académicos, docentes y de investigación, y pueden beneficiarse de las autorrevisiones orientadas hacia la mejora. Finalmente, en algunas situaciones, debe ser posible fomentar el planteamiento de problemas globales de la institución y estrategias de planificación a partir de un ciclo completo, o incluso parcial, de revisiones de programas en una institución determinada. También debe ser posible que un organismo estatal averigüe, manteniendo la confidencialidad, la esencia de los problemas que se reiteren o los asuntos de política educativa que vayan apareciendo a lo largo de las revisiones de programas. Esto se está intentando ahora en los Países Bajos y Francia.

Se debería añadir una serie de aspectos adicionales al modelo básico de evaluación de la calidad que se acaba de presentar para que sea un auténtico sistema integrado.

En primer lugar, debería existir una supervisión ministerial o gubernamental cuidadosamente negociada, junto con un sistema de evaluación de la calidad regional o nacional, para asegurar que el sistema funciona y que las instituciones de educación superior no rechazan los resultados

o son interesados por las facultades u otras unidades. La amenaza de una intervención, en caso de que el sistema no funcione después de uno o dos ciclos de visitas, va a ser tomada muy en serio por los responsables de las instituciones de educación superior. Aunque no se deben entregar al ministerio informes individuales sobre el sistema, los responsables del mismo sí deben presentar informes generales de la situación y una lista de los problemas de política educativa que hayan surgido.

En segundo lugar, debería haber participación de todas las facultades en los procesos de autoevaluación y en la revisión de los resultados de las visitas e informes de la evaluación dentro de una determinada institución, para mantener el objetivo de conseguir altos niveles de calidad, mejorándolos y evaluándolos continuamente. Asimismo, la existencia de visitantes cuidadosamente seleccionados y de alto nivel, nacionales y extranjeros, tanto académicos como profesionales de la industria, favorece el afán de superación y el mantenimiento de un sistema abierto y orientado a la mejora.

En tercer lugar, se debe desarrollar un sistema claro y bien conocido que incorpore los resultados relevantes del proceso de revisión al sistema presupuestario anual de la institución de educación superior y a su plan de desarrollo o a su plan estratégico. Si esto no es así, el resultado no será muy relevante. Naturalmente, este será el caso en cualquiera de los sistemas de incentivos, competitivos o no, que redistribuyan recursos marginales en la institución en cuestión.

Finalmente, cada institución de educación superior debería plantearse un ciclo integrado de evaluación y planificación que incluya revisiones de programas (académicos y administrativos), realizando algunos cada año siguiendo un ciclo de cinco o siete años, y estudiando al final de cada ciclo la política global de la institución y otros asuntos surgidos en las revisiones anuales de programas o en cualquier planificación o estrategia que sean consecuencia de la revisión. Esta es la mejor forma de aprovechar el esfuerzo que suponen las revisiones de programas, beneficiando a todos los niveles de la universidad.

6. Conclusión

Este trabajo pretende situar las revisiones autorreguladas de programas universitarios en su perspectiva histórica y comparativa, y sugerir

los elementos, tanto básicos como adicionales, que se deberían tener en cuenta cuando se quieren establecer y usar tales sistemas. Se han analizado las experiencias obtenidas en entornos diferentes, con el fin de dar forma a esta perspectiva. Naturalmente, se parte de la base de que un sistema bien integrado en su entorno será mucho más efectivo que aquél que ha sido incorporado de forma imprecisa en los procesos fundamentales de una institución de enseñanza superior.

EVALUACION DE LA CALIDAD DE LA ENSEÑANZA: UNA REVISION CRITICA DE LA BIBLIOGRAFIA Y LA INVESTIGACION *

DAVID L. TAN

Universidad de Oklahoma, Estados Unidos

Calidad... se sabe lo que és, pero no se sabe expresar lo qué es. Sin embargo esto no es una contradicción. Algunas cosas son mejores que otras, es decir, tienen más calidad. Pero cuando se intenta decir lo qué es la calidad, separada de las cosas que la poseen, ¡no se tiene ni idea! No hay nada de qué hablar. Si no se puede determinar qué es la Calidad, ¿cómo se sabe qué es, o siquiera que existe? Si nadie sabe lo qué es, entonces en la práctica no existe. Pero en la práctica sí existe. ¿En qué se basan si no las calificaciones? ¿Por qué si no la gente paga fortunas por unas cosas y arroja otras a la basura? Obviamente algunas cosas son mejores que otras... pero ¿qué es lo «mejor»?... Así pues, nos movemos en círculo, rodando mentalmente sin encontrar algo en que agarrarnos. ¿Qué demonios es la calidad?

(Pirsig, 1974, p. 184).

«¿Qué demonios es la calidad?» preguntaba Pirsig. En efecto, esta es una pregunta que se ha discutido y debatido desde el principio del género humano. Sin embargo no existe una respuesta definitiva. El problema básico que subyace a esta dificultad es la carencia de una teoría de la calidad y la falta de acuerdo en cuanto al concepto mismo. Dado que la calidad es polifacética y varía con la percepción subjetiva, no permite llegar a un consenso universal. Al mismo tiempo, si embargo, sabemos intuitivamente que las cosas que regularmente usamos o manejamos varían en calidad, desde aparatos y automóviles hasta alimentos y libros.

* Trabajo publicado originalmente en *Research in Higher Education*, 24, 3, pp. 223-265, 1986.

¿Por qué es la calidad un concepto tan esquivo? La principal razón se basa en el hecho de que diferimos en nuestro concepto de calidad. La calidad, como la belleza, se encuentra en el ojo del observador; posee un significado diferente para gente distinta. Por ejemplo, algunos pueden considerar que un automóvil Toyota es de gran calidad a causa de la fiabilidad de su motor y su excelente récord de mantenimiento, mientras que un Ferrari puede ser aclamado por otros por su línea deportiva y su alta velocidad. Un equipo de fútbol americano puede ser considerado de calidad por su número de victorias, mientras que otro puede serlo por el elevado número de goles por partido. Algunas personas piensan que las películas y programas de televisión son de calidad por sus escenas de acción, mientras que otros los juzgan por su contenido o por su excelente representación. Tal como sugieren estos ejemplos, no es posible establecer un criterio o medida absoluta de calidad.

Sabemos que no podemos tener un consenso universal del concepto de calidad. Si el concepto es tan esquivo ¿por qué estamos tan obsesionados en categorizar las cosas basándonos en la calidad? ¿Por qué es tan importante saber qué programa de televisión o película es la mejor en una semana concreta, qué equipo universitario de fútbol americano es el primero de la nación, qué libro obtiene el récord de venta mensual, y a qué restaurante deberíamos acudir para una cena de calidad?. Se especula que tal obsesión es inherente a nuestra cultura democrática, en la cual son valores centrales la competitividad y la excelencia. Cualquiera que sea el motivo, el punto fundamental es que cuando categorizamos las cosas, esencialmente queremos decir que existen algún tipo de consistencia en nuestra percepción de calidad. Aunque no es posible una consistencia absoluta, puede haber algunos atributos de la calidad sobre los que todos estemos de acuerdo. Si es así, hay cierta esperanza de que podamos entender mejor el concepto de calidad.

En la enseñanza superior, la mayoría de los profesores esperan cierto consistencia en la percepción de la calidad, aunque el entorno de este nivel de enseñanza se caracteriza por una gran diversidad de personas e instituciones. Cuando se considera a Harvard como una de las mejores instituciones del país, la mayoría de los profesores se dan cuenta de la variedad de criterios usados para determinar ese estatus. Algunos pueden aducir como razones de su elección su excelente profesorado o a los currícula que se imparten de esta universidad, mientras que otros pueden hacerlo por sus magníficos estudiantes o por la abundan-

cia de sus recursos financieros. Estas opiniones difieren bastante. Sin embargo, a pesar de la diversidad, parece lógico que si se pueden categorizar algunas de estas opiniones comunes por medio de una serie de parámetros, existe cierta esperanza de que se pueda alcanzar un consenso general en el asunto de la calidad. Suponiendo de una serie de parámetros o atributos de calidad se pudieran identificar consistentemente en Harvard, ¿significa necesariamente que la calidad de todo tipo de instituciones de enseñanza superior se puede juzgar basándose en los mismos parámetros? Si es así, ¿deberían todas las instituciones considerar a Harvard como modelo de excelencia?

Evidentemente, no todas las instituciones tienen que asumir la tradición de Harvard ni considerar que los parámetros de calidad de esta universidad se les puede aplicar necesariamente a ellas. Los *colleges* locales, por ejemplo, difieren de Harvard por sus fines, objetivos, énfasis curricular, y estudiantes, pero también cumplen una función vital en nuestra sociedad. Ofrecen, al contrario que Harvard, un currículum más fuerte en el área técnica y profesional. Más importante aún, proporcionan oportunidades educativas a adultos que trabajan, personas mayores y otra gente que de otro modo no hubieran tenido la oportunidad de asistir a ninguna institución. Alterar estos objetivos para adaptarlos a los de Harvard no significa necesariamente que la calidad de esos *colleges* fuese a aumentar. A largo plazo, esta "acomodación" puede dañar su importante finalidad de proporcionar oportunidades de educación superior a aquellos que de otro modo no las hubiesen tenido.

Aunque no se pueden aplicar los mismos atributos de la calidad a todas las instituciones, si pueden utilizarse cuando nos referimos a un mismo tipo de universidad, centro ó *colleges*. Si cabe esta posibilidad, el conocimiento acerca de los parámetros de la calidad puede proporcionar información valiosa a las instituciones que pretenden mejorarla. Si el profesorado de una universidad ha alcanzado la excelencia en investigación, deberían proporcionarse entonces los medios para que este profesorado logre la máxima productividad; y si los recursos de la biblioteca están asociados con la calidad de esta productividad, entonces la dotación y el aumento presupuestario de la biblioteca necesariamente aumentarán la efectividad del profesorado y la calidad general de la institución.

La necesidad de obtener este tipo de conocimientos ha originado miles de publicaciones. Como era de esperar, las opiniones, metodolo-

gías de investigación y los resultados que aportan estos trabajos son tan variados y complejos como el concepto mismo de calidad. Con todo, estas publicaciones se pueden clasificar en tres grandes tipos: 1) estudios de reputación; 2) estudios en base a indicadores objetivos y 3) estudios basados en correlaciones. Los estudios de reputación fueron los primeros en investigar la calidad de la educación superior. Estos estudios emplean evaluaciones subjetivas del profesorado, directores de departamentos o decanos como medio para la evaluación de los programas. Al contrario, los estudios a partir de indicadores objetivos no utilizan evaluaciones subjetivas. Valoran los programas por medio de variables objetivas como la productividad en investigación del profesorado, los recursos financieros o resultados de los estudiantes. Por último, los estudios basados en correlaciones tienen como finalidad principal identificar las variables que se correlacionen con la calidad del profesorado o del programa. Estas variables se llaman "correlatos de calidad. Ejemplos de correlatos son el tamaño del departamento, el presupuesto asignado, los fondos de la biblioteca, los salarios del profesorado, y la productividad en investigación del profesorado. Los tres tipos de estudios han contribuido a una mayor comprensión de la calidad de la enseñanza superior, pero aún se necesita más conocimiento, ya que la mayoría de los trabajos al respecto presentan alguna limitación metodológica. En este artículo, junto con los principales resultados, se discuten los puntos fuertes y débiles de estos tres tipos de estudios.

Estudios de reputación

Estudios de reputación son aquellos que utilizan la opinión de expertos para clasificar los programas. La mayoría se centran en programas de postgrado, aunque algunos también se ocupan de programas profesionales y de licenciatura.

Estudios sobre Enseñanza de Postgrado

Los mejores estudios de reputación conocidos sobre la enseñanza de postgrado son los dos informes subvencionados por el Consejo Americano de Educación (ACE): el Informe de Allan Cartter, *Una Evaluación de la Calidad de la Enseñanza de Postgrado (1966)*, y su réplica, *Una Clasificación de la Enseñanza de Postgrado (1970)*, dirigida por Kenneth Rose y Charles Andersen. Antes de estos dos estudios del ACE,

otros estudios sobre reputación fueron dirigidos por Hughes (1925, 1934) y Keniston (1959). El primer estudio de Hughes fue uno de los primeros estudios publicados acerca de la enseñanza de postgrado. Lo que impulsó a Hughes, rector de la Universidad Miami en Ohio, a realizar su estudio fue su preocupación por la falta de información acerca de los centros de postgrado para facilitar a los licenciados de su universidad que deseaban continuar sus estudios. Por ello, Hughes evaluó la calidad de la enseñanza de postgrado en treinta y ocho de las sesenta y siete instituciones que entonces impartían el doctorado. El primer paso de su estudio fue recopilar una lista de profesores distinguidos expertos en veinte campos o áreas del saber elegidos por el profesorado de su universidad. Posteriormente remitió unos cuestionarios a estos profesores solicitando la clasificación de sus programas de doctorado. Un segundo informe de Hughes en 1934 empleó una metodología similar, pero ampliando su muestra para incluir treinta y cinco disciplinas en cincuenta y nueve universidades, clasificando a los departamentos en dos categorías adecuados o distinguidos utilizando las evaluaciones de cuarenta profesores. Aunque los dos informes de Hughes no eran ideales en términos de la metodología de investigación utilizada, Hughes merece ser llamado el pionero en la investigación de la calidad. Como Lawrence y Green (1980) señalaron, Hughes fue el primero en defender el uso de profesores como evaluadores, el primero en enfatizar la importancia de revisar programas a nivel de postgrado y el primero en utilizar la calidad del profesorado como un criterio importante para evaluar la calidad de un programa.

A pesar de las alabanzas, los dos estudios de Hughes también han despertado las críticas. Dos de las críticas más frecuentes se han centrado sobre el bajo número de expertos (entre veinte y sesenta) y la desproporcionada representación geográfica de estos, que tendían a proceder del Noreste y el Medio Oeste. A pesar de estas críticas, los resultados de los trabajos de Hughes continúan citándose a menudo, normalmente para establecer comparaciones.

Aunque los informes de Hughes hayan sido muy criticados, no se intentó mejorar o actualizar esta metodología de trabajo hasta el estudio realizado por Keniston a finales de los 50. La principal meta de Keniston (1959) era determinar el rango académico de la Universidad de Pennsylvania en relación a un grupo de veinticuatro universidades similares pertenecientes a la Asociación Americana de Universidades (AAU). Este trabajo, como los de Hughes, utilizó opiniones mediante un equipo de

jueces. El estudio se basaba específicamente en sesiones de consulta con directores de departamento, a los que se pedía que nombrasen los cinco primeros departamentos de postgrado en sus campos, combinando la calidad del programa y del profesorado. Juntando los datos de las sesiones, Keniston generó una lista ordenada y categorizada de las veinticinco instituciones y, a efectos comparativos, comparó la lista con la de Hughes (1925) para detectar cambios en los rankings de las mismas instituciones. En efecto, se encontraron algunos cambios. La Universidad de Harvard, por ejemplo, había pasado del segundo lugar en el informe de Hughes al primero; la Universidad de California (Berkeley) pasó del noveno al segundo lugar; la Universidad de Yale pasó del quinto al cuarto; y la Universidad de Michigan pasó del octavo al quinto (Keniston 1959).

Como la mayoría de estudios de reputación, el estudio de Keniston ha sido criticado. Al igual que en los informes de Hughes, se ha considerado que el equipo de expertos era demasiado pequeño, y que la utilización de directores de departamento como jueces es la menos idóneo para reunir información sobre la calidad de una institución. Las razones en las que se apoyan los críticos para justificar esta última opinión son las siguientes:

- 1) La mayoría de los directores de los departamentos eran catedráticos, generalmente de edad avanzada, y frecuentemente no eran los profesores más distinguidos en los departamentos, ni tampoco los más representativos de sus pares en cuanto a su especialización en el área o disciplina académica respectiva (Cartter, 1966);

- 2) Igualmente los jefes de departamento eran generalmente de mayor edad que el resto del profesorado, tendían a ser más tradicionalistas y de ahí que podían haber calificado más alto los departamentos más prestigiosos y menos a los menos prestigiosos (Cartter, 1966; Lawrence and Green, 1980);

- 3) El hecho de haber sido antiguos alumnos podía haber sido un sesgo importancia ya que la mayoría de los jefes de departamento seleccionados se habían graduado en las instituciones evaluadas (Lawrence and Green, 1980).

Otra crítica frecuente al estudio de Keniston se centra sobre las dificultades del investigador para distinguir entre las medidas de prestigio

del profesorado y las medidas de prestigio sobre la enseñanza impartida (Cartter, 1966).

Cuando se les presta atención, las críticas constructivas pueden utilizarse para mejorar diseños metodológicos subsiguientes, como fue el caso en los dos estudios de ACE mencionados. En el informe Cartter de 1966, se tuvo gran cuidado en considerar varias inconsistencias metodológicas que había en los informes de Hughes y Keniston. Por ejemplo, se consiguió una representación geográfica más equitativa. Cartter tomó una muestra de 106 universidades ubicadas en diversas regiones del país. No obstante, al margen de la mejora se mantuvieron aspectos muy criticados de los estudios previos. Uno de ellos fue el uso de un equipo de expertos como jueces. A pesar de las fuertes críticas, la validez de esta estrategia metodológica fue defendida por el investigador en estos términos:

«Este estudio es una encuesta de opinión informada. Las opiniones que hemos buscado es lo que en un tribunal de justicia se llamaría «el testimonio de peritos» —personas de cada disciplina cualificadas para juzgar, quienes por su preparación poseen conocimiento y son objetivas, y que por su participación en la vida académica en sus respectivos ámbitos se han ganado el respeto de sus colegas.» (Cartter, 1966, p. 8).

Consciente de otras críticas contra Hughes y Keniston, Cartter estableció tres propósitos en sus estudios. Primero, pretendía actualizar los dos estudios previos para ver qué cambios habían ocurrido en los mismos programas. Segundo, quería «ampliar la evaluación para incluir las principales universidades de los Estados Unidos basándose en la suposición de que el mayor desarrollo se produciría en las 10-15 universidades tradicionalmente más destacadas» (1966, p. 3). Tercero, quería examinar y comparar técnicas de medida objetivas y de reputación.

Cartter trabajó con un total de cuatro mil evaluadores: directores de departamentos, profesores distinguidos con años de experiencia y profesores jóvenes destacados. Se les pidió a todos que clasificaran cada programa de doctorado de su propia área de conocimiento teniendo en cuenta dos variables:

- 1) La calidad del profesorado que impartía el postgrado.
- 2) La eficacia del programa.

Las categorías de respuesta para la primera variable iban desde *destacada a insuficiente*, mientras que las categorías para la segunda iban de *muy atractiva a nada atractiva*. Para ambas variables existía la opción de *información insuficiente*. Es importante señalar que en la segunda variable Cartter confundió la eficacia del programa con su atractivo. Sin embargo, a partir de esta escala, Cartter recopiló puntuaciones para cada departamento y generó calificaciones nacionales de 1.663 programas de postgrado en 29 disciplinas académicas de 106 universidades. Ocho de los programas clasificados entre los diez primeros (a saber, los programas en la Universidad de California [Berkeley], la Universidad de Chicago, la Universidad de Columbia, la Universidad de Harvard, la Universidad de Michigan, la Universidad de Princeton, la Universidad de Wisconsin y la Universidad de Yale) eran similares a los clasificados por Hughes y Keniston. No obstante, Cartter mejoró significativamente la metodología de investigación utilizada en los estudios previos. Prestó una atención especial a la fiabilidad y a la validez de los métodos de investigación utilizados en el estudio. Primero, comprobó y analizó la representatividad de la muestra de evaluadores empleando información biográfica sobre ellos. Segundo, realizó pruebas detalladas de la fiabilidad de los *rankings* obtenidos en tres campos de estudio: economía, inglés y física. Cartter observó que sus clasificaciones sobre la calidad eran fiables ya que habían permanecido inalterables independientemente del rango académico de los evaluadores, el tipo de institución del que proceden, el nivel de titulación y el tipo de disciplina. Cartter también investigó la idea extendida de que un número reducido de expertos puede sesgar los resultados. Sus hallazgos revelaron que se requiere un mínimo de cincuenta personas competentes para asegurar la fiabilidad de los resultados.

Por último, Cartter comparó sus clasificaciones con las obtenidas mediante mediciones objetivas, como tamaño del departamento, recursos de la biblioteca, prestigio del profesorado, salario, y publicaciones efectuadas. Cartter encontró coincidencias entre los dos tipos de clasificaciones y concluyó:

«parece probable que si se incluyesen suficientes para construir indicadores objetivos —permitiendo variaciones según el tamaño de la institución y el énfasis de la universidad en determinados campos de estudio— los resultados de nuestra evaluación subjetiva casi se replicarían exactamente» (1966, p. 18).

El primer informe ACE recibió una fuerte respuesta de la comunidad académica, tanto en críticas severas como en alabanzas desmedidas. En 1970, se efectuó una réplica del primer estudio de ACE, esta vez realizado por Roose y Andersen. El propósito original de este segundo informe era cumplir el compromiso de Cartter de llevar a cabo un seguimiento a los cinco años del primer informe. Como era de esperar, el segundo informe consistió en una nueva evaluación de los mismos departamentos del primer informe. Roose y Andersen, sin embargo, señalaron que no llevarían a cabo ningún estudio que fomentara o inflara «egos institucionales». Más bien, su estudio pretendía proporcionar información. Estos investigadores esperaban que los lectores de su trabajo «pensaran en términos de escalas de calidad en lugar de pensar en un orden jerárquico específico». A fin de animar a los lectores a pensar de esta manera, presentaron escalas de puntuaciones en lugar de puntuaciones absolutas de los departamentos (como sucedió en el informe de Cartter).

Roose y Andersen también vieron la necesidad de «proteger al consumidor potencial de la enseñanza de postgrado contra programas inadecuados». Por ello su estudio se amplió para incluir instituciones menos conocidas. Específicamente, su muestra abarcó a 131 instituciones frente a 106 en el primer estudio y 36 campos del conocimiento, mientras que en el primero se emplearon sólo 29. Al mismo tiempo, se examinaron 2.626 programas de postgrado, con un incremento del 55% sobre el primer estudio subvencionado por la ACE. Como Cartter, Roose y Andersen no añadieron las puntuaciones de los departamentos a los *rankings* institucionales. Las tres razones que dieron para no hacerlo, fueron originalmente formuladas por Cartter. Primero, dado que no todas las instituciones encuestadas tenían programas doctorales en todos los campos, no podrían obtenerse *rankings* institucionales adecuados. Segundo, aunque todas las instituciones tuviesen programas doctorales en todos los campos, asignar pesos a campos diversos sería difícil, si no imposible. Tercero, dado que la especialización departamental es el principio organizativo básico en la vida académica, serviría de poco dar información sobre una institución de tipo medio. A pesar de estas razones, algunos investigadores no han estado de acuerdo con Cartter y han desarrollado *rankings* institucionales agregando las puntuaciones de los departamentos. Entre ellos se encuentran Magoun (1966), el National Science Board (1969), Petrowski, Brown y Duffy (1973), y Webster (1983).

Aunque los dos informes del ACE mostraban una mejora sustancial en metodología de investigación respecto a los estudios de reputación

previos, sin embargo han ocasionado muchas críticas. Quizás la crítica más frecuente haya sido que los estudios del ACE ignoraron un elevado número de instituciones de educación superior (Lawrence y Green, 1980; Conrad y Blackburn, 1985a; Petrowski *et al.*, 1973). También se han criticado las encuestas realizadas por no considerar los programas menos populares en instituciones menos conocidas y los que tienen un profesorado de inferior renombre... Además, estos estudios no tuvieron en cuenta los departamentos en el amplio campo de la educación, y en otros más pequeños, pero importantes, campos profesionales como la agricultura, la administración de empresas, el derecho, la odontología y la medicina (Petrowski *et al.*, 1973).

Otra crítica frecuente a los informes del ACE ha sido que medían más el prestigio que la calidad (Lawrence y Green, 1980; Dolan, 1976). Dolan indicó que aunque los informes midiesen el prestigio, la metodología empleada en estos informes no era apropiada para esta tarea. La reputación, señaló Dolan, es a menudo tan difícil de medir como la calidad y los informes del ACE no medían ni la una ni la otra adecuadamente. Knudsen y Vaughan (1969) también señalaron esta dificultad, sugiriendo que esto se debe a que el prestigio está demasiado influenciado por el tamaño del departamento y la familiaridad con las publicaciones de sus profesores. Gallant y Prothero (1972), aunque estaban de acuerdo con ello, examinaron la situación desde otro punto de vista. No consideraron que la reputación del profesorado se viese afectada por el número absoluto de profesores, sino por la existencia o no de lo que ellos denominaron «masa crítica». La masa crítica, según estos autores, es el tamaño mínimo que requiere a un departamento para impartir un programa de alta calidad, aunque no lo garantiza. Siguiendo este razonamiento, el prestigio puede estar tan influenciado por la presencia de esta masa crítica como por el tamaño del departamento.

Otra crítica a las clasificaciones de ACE ha sido que pueden causar un efecto “halo”, el también llamado efecto Hertz-Avis, al establecer un orden jerárquico entre las instituciones. Esta crítica se basa en las muchas limitaciones metodológicas de las clasificaciones del ACE, que podrían usarse incorrectamente como bases para establecer un orden jerárquico engañoso (Webster, 1961; Dolan, 1976; Lawrence y Green, 1980). Además, algunos críticos han indicado que el efecto halo originado por estas clasificaciones puede reforzar el “status quo” institucional e impedir innovaciones y mejoras. Dolan (1976), por ejemplo, indicó que sólo los departamentos académicos grandes, clásicos, son favorecidas

por estas clasificaciones. Lo que Dolan quería decir era que los departamentos clasificados muy bien no estarían tan motivados ni serían tan entusiastas en cuanto a implementar otras ideas creativas o demasiado innovadoras en sus programas que pusieron en peligro su estatus. Los programas más afectados por esta actitud serían aquellos que exigen un trabajo más innovador como los postgrados de carácter interdisciplinar o experimental, dado que no serían debidamente reconocidos o no se les concedería el tiempo necesario para su desarrollo y, eventualmente, podrían no tener éxito.

Los estudios del ACE también se han criticado por no tener en cuenta todas las medidas válidas sobre la calidad de los programas de doctorado (Dolan, 1976; Webster, 1981; Lawrence y Green, 1980). Las clasificaciones del ACE se basaban únicamente en dos variables (prestigio del profesorado y del programa) y por tanto, según los críticos, no pueden considerarse como si midiesen la calidad global de los programas doctorales. Sin embargo, Petrowski y sus colegas (1973), encontraron esta observación sin fundamento al descubrir que existe muy poco o ninguna relación entre el prestigio de programas de doctorado analizados y los de licenciatura u otros programas de postgrado no encuestados. No obstante, la falta de adecuadas medidas de la calidad constituye la crítica más permanente.

Otra crítica a los estudios del ACE se ha centrado en la polémica utilización de un equipo de expertos como evaluadores. Además del problema del sesgo ya considerado, los críticos han señalado que dado que un número muy elevado de evaluadores eran graduados o profesores de los programas más prestigiosos, sus evaluaciones presentarían también sesgos característicos de antiguos alumnos (Lawrence y Green, 1980; Cartter, 1966). Ambos tipos de sesgos pueden afectar las clasificaciones de forma engañosa. Algunos autores han ofrecido sugerencias para mejorar estos aspectos. Lawrence y Green (1980), por ejemplo, recomendaron que se incluyesen como evaluadores personas ajenas al mundo académico y otros individuos que puedan tener más conocimiento acerca de calidad de programas.

Por último, se han criticado los informes del ACE por el tiempo empleado entre la fecha en que fueron recogidos los datos y la fecha en que se informó de las clasificaciones. La duración de los estudios, aseguran los críticos, hace que las clasificaciones no reflejen las condiciones reales de una universidad determinada en la fecha en la que son emitidos los informes (Webster, 1981; Petrowski et al., 1973).

El único grupo de estudios sobre reputación que no se han discutido son ciertos trabajos que categorizaron a las instituciones añadiendo las puntuaciones de los departamentos a la puntuación institucional, a pesar del ruego de Cartter de que no se hiciera. Entre estos estudios se encuentran los efectuados por Magoun (1966), el National Science Board (1969), Petrowski et al, (1973), y Webster (1983).

Magoun (1966) recopiló datos del informe de Cartter y agregó las puntuaciones del departamento a dos tipos de categorías institucionales global y por divisiones que cubren los campos de las humanidades, las ciencias sociales, las ciencias biológicas, las ciencias físicas, y la ingeniería. Como ya hemos dicho, dado que Cartter utilizó dos tipos de puntuaciones, una basada en la calidad del profesorado del postgrado y otra en la eficacia del programa, Magoun generó en total cuatro rangos institucionales. Las categorías institucionales por divisiones se obtuvieron dividiendo la suma de todas las puntuaciones de los departamentos de cada división entre el número total de programas clasificados en la división. La ordenación de las instituciones con carácter global se obtuvo sumando todas las puntuaciones de las divisiones divididas entre cuatro (cuatro porque se excluyó la ingeniería).

Magoun comparó ambos tipos de clasificaciones, hallando grandes similitudes. Las instituciones que aparecían en los grupos superiores de la lista en la clasificación por divisiones también aparecían en la mitad superior de la lista en la clasificación por divisiones también aparecían en la mitad superior de la lista en la general. Por ejemplo la Universidad de Iowa, que logró un tercer lugar en el *ranking* en dos campos (humanidades y ciencias sociales) y un segundo lugar en ciencias biológicas, se ubicó en la clasificación institucional global en la primera categoría del *ranking*. Esta y otras similitudes llevaron a Magoun a la conclusión de que agregar puntuaciones del departamento o una puntuación institucional no sólo era posible, sino que refleja mejor el prestigio global de la institución (matemáticamente, se esperaban similitudes porque las puntuaciones institucionales globales no eran más que cálculos matemáticos utilizando las puntuaciones divisionales).

Otro intento de obtener *rankings* institucionales añadiendo las puntuaciones de los departamentos se utilizó en un informe publicado por el Consejo Nacional de la Ciencia (National Science Board, 1969) titulado «Educación de Postgrado: Parámetros para Programas Públicos». Este estudio se basaba en las clasificaciones de Cartter sobre el presti-

gio del profesorado de los cursos de postgrado. El Consejo Nacional de la Ciencia añadió las puntuaciones de los departamentos (las escalas de Cartter sobre prestigio del profesorado) a las puntuaciones institucionales y ordenó jerárquicamente las instituciones. La lista mostraba cincuenta instituciones en las tres categorías superiores de un sistema de siete categorías. Aunque no se identificaran las instituciones que ocupaban los primeros puestos, se dividieron en nueve regiones geográficas. Su distribución no fue tan uniforme como esperaba el Consejo Nacional. La mayor concentración de instituciones se ubicaba en la región Atlántica Media (con trece instituciones), seguida de la Noreste Central (con diez instituciones), mientras que la menor concentración se hallaba en las regiones Montañosa y Sureste Central, con solamente una.

Petrowski *et al.* (1973) realizaron el tercer y más reciente intento de obtener *rankings* institucionales. Los investigadores emplearon en este estudio tres métodos para identificar lo que denominaron «universidades nacionales». Estas se consideraron instituciones de categoría superior que deberían ser subvencionadas por el gobierno federal para llevar a cabo la enseñanza de postgrado e investigación. El primer método incluyó la obtención de una puntuación institucional haciendo el promedio de las puntuaciones de todos los departamentos (categorizados o no) en una institución determinada (los datos provenían de los estudios ACE). El segundo método categorizó las instituciones de acuerdo con el número de programas que habían alcanzado los *rankings* de ACE. El tercer método utilizó un esquema de pesos, considerando tanto el número de departamentos clasificados como los *rankings* alcanzados por estos departamentos. Con este método se obtuvo una puntuación media ponderada de la institución multiplicando las puntuaciones medias obtenidas mediante el primer método con la proporción de departamentos clasificados por el número total de disciplinas encuestadas. Estas puntuaciones ponderadas se usaron posteriormente como base para elaborar *rankings* institucionales.

Para comprobar la plausibilidad de estos *rankings*, Petrowski *et al.* los comparaban con los del Consejo Nacional de la Ciencia. El Consejo categorizó las instituciones en nueve regiones geográficas; por lo que Petrowski *et al.* agruparon las instituciones en las mismas regiones. Se hallaron similitudes entre los *rankings* en 7 de las 9 regiones. Las dos regiones en que las listas no coincidieron fueron la Atlántica Media y Atlántica Sur. Además, los investigadores encontraron ubicadas en la categoría superior. Por ejemplo, las treinta primeras universidades eran

instituciones grandes y se habían clasificado muy bien más de 16 de sus departamentos. Por último, los investigadores fueron precavidos en cuanto a sus resultados. Señalaron que su lista jerarquizada de instituciones o cualquier otra que haya intentado identificar «universidades nacionales» solamente para obtener apoyo financiero produciría inevitablemente un grupo de instituciones selectivas en las cuales la admisión sería muy difícil. Los investigadores insistieron en desaconsejar la mediocridad de las instituciones, pero tampoco no pertenecer al grupo de élite.

Webster (1983) realizó el cuarto intento para obtener *rankings* institucionales anadiendo las puntuaciones de los departamentos. Webster llegó a su puntuación institucional añadiendo las puntuaciones sobre calidad del profesorado recopiladas por el *Conference Board of Associated Councils* (Jones, Lindzey, and Coggeshall, 1982). Los centros de postgrado en ciencias físicas y matemáticas, humanidades, ingeniería, ciencias biológicas y ciencias sociales fueron clasificados en base a su puntuación institucional. Se confeccionó una lista con los 30 centros clasificados en primer lugar. Las primeras 10 instituciones eran las universidades de California (Berkeley), Stanford, Harvard, Yale, Princeton, Chicago, California (Los Angeles) Michigan y Wisconsin (Webster, 1983).

Además de los estudios de ACE y aquellos que agregaron las puntuaciones de los departamentos a los *rankings* institucionales, se realizaron otros trabajos de reputación sobre cursos de postgrado por Ladd y Liset; Clark, Harnett y Baird (1976) y, más recientemente, por la *Conference Board of Associated Research Councils* (Jones *et al.*, 1982). Ladd y Lipset, utilizando básicamente la misma metodología que Hughes y Keniston, pidieron a 4.000 profesores de *colleges* y universidades que identificaran 5 departamentos conocidos porque sus profesores eran destacados en sus disciplinas respectivas. Se tuvieron en cuenta solamente los departamentos que recibieron al menos un 10% de votos de los encuestados en cada uno de los 19 campos estudiados. Las categorías de muchas instituciones fueron similares a las obtenidas en las investigaciones anteriores (cf. Scully, 1979). Por ejemplo, la Universidad de Chicago se clasificó en sexto lugar en este estudio, en quinto por Hughes (1934), en sexto por Keniston (1959), en noveno por Cartter (1966), y en séptimo por Roose y Anderson (1970). Otro ejemplo es la Universidad de Harvard, que fue categorizada la segunda en este estudio y por Cartter (1966), y primera por Hughes (1934) y Keniston (1959).

El estudio de Clark y sus colegas (1976), subvencionado por el *Council of Graduate Schools and the Educational Testing Service*, utilizó variables objetivas y de reputación, para clasificar los programas. Profesores y estudiantes proporcionaron evaluaciones sobre la calidad de 73 departamentos en los campos de química, historia y psicología. Estas evaluaciones se combinaron con 30 variables objetivas para llegar a dos grupos de indicadores:

- 1) Indicadores orientados hacia la investigación (incluyendo variables como tamaño del departamento, prestigio del profesorado, habilidad académica del estudiante, publicaciones del profesorado, y recursos físicos y financieros).

- 2) Indicadores sobre experiencia docente (incluyendo relaciones interpersonales de los profesores, clasificaciones de antiguos alumnos sobre experiencias de disertación, y clima académico).

Se halló una elevada correlación entre las variables dentro de cada grupo, pero no entre grupos. Evidentemente, los entrevistados creían firmemente que las variables incluidas en cada grupo eran de naturaleza distinta. Como resultado, Clark *et al.* señalaron rápidamente que sus clasificaciones «no eran clasificaciones de la calidad global del programa doctoral, sino más bien clasificaciones del profesorado que imparte estos programas, reflejando esencialmente esto en los informes de sus investigaciones». Este hallazgo apoya el punto de vista compartido por varios críticos que consideran incorrecto, o al menos precipitado, el que algunos investigadores hayan identificado el prestigio del profesorado con la calidad del programa.

El estudio más extenso y reciente sobre acreditación de los cursos de postgrado es el informe en 5 volúmenes del Conference Board of *Associated Councils* (Jones *et al.*, 1982). Este Consejo está compuesto por el *American Council of Learned Societies*, el *American Council of Education*, the *National Research Council*, y *Social Science Research Council*. El estudio del Consejo valoró la calidad de programas doctorales en 32 disciplinas en más de 200 universidades con programas de doctorado (un total de 2.699 programas). El estudio se realizó basándose en 4 principios claramente definidos:

1. La importancia de los resultados del estudio para organismos nacionales y estatales;

2. El deseo de estimular el énfasis constante en la calidad de la educación de postgrado;

3. La necesidad de que las evaluaciones actuales tengan en cuenta las innovaciones introducidas en los programas desde el estudio de Roose Anderson; y

4. El propósito de ampliar el número de medios utilizados en estudios evaluativos de programas de postgrado (Jones *et al.*, 1982a).

Bajo este contexto, el Consejo tuvo extremo cuidado de analizar las críticas contra los estudios de reputación previos e intentó resolver algunas de sus deficiencias metodológicas. Sin embargo, reconoció que, no obstante, independientemente de las medidas que se utilizasen, no todos quedarían satisfechos, especialmente aquellos que insistían en que la calidad de la enseñanza de postgrado no se puede cuantificar. Además, según el Consejo, la utilidad de cualquier evaluación depende de la validez y fiabilidad de las variables usadas. El Consejo consideró 3 criterios para la inclusión de variables. El primer criterio era que las variables deben estar relacionadas con la calidad de los programas de doctorado e investigación. El segundo criterio tenía que ver con la viabilidad para recoger y recopilar datos fiables sobre estas variables a efectos de hacer comparaciones a nivel nacional de programas en disciplinas concretas. Tercer, el Consejo solamente consideró variables que se pudieran aplicar a la mayoría de las disciplinas estudiadas.

Usando este método, el Consejo recopiló, analizó, e informó sobre datos y resultados en 5 estudios que abarcaban programas de (1) humanidades, (2) ciencias sociales y de la conducta, (3) ciencias biológicas, (4) ciencias matemáticas y físicas; y (5) ingeniería. Dado que estos estudios contienen variables y datos, los examinaremos a continuación en detalle.

Humanidades

En este estudio, se examinaron un total de 522 programas de doctorado en historia de arte, lenguas clásicas, literatura y lengua inglesa, francesa, alemana y española, lingüística, música y filosofía. Se utilizaron 12 variables agrupados en 4 áreas generales: 1) tamaño del programa, 2) características de los graduados, 3) resultados de los estudios de reputación previos, y 4) tamaño de la biblioteca de la universidad.

Dentro del grupo relativo al tamaño del programa, se incluyeron tres variables:

a) El número de profesores en cada programa en diciembre 1980 (esta información se obtuvo directamente de las universidades participantes).

b) El número de graduados en los programas durante 5 años, desde 1975 a 1980 (esta cifra figura en los datos recopilados en la *Encuesta sobre Doctorados Obtenidos del National Research Council*).

c) El número total de estudiantes de postgrado a tiempo completo y a tiempo parcial —en diciembre de 1980— que estaban matriculados en los programas y que pretendían obtener doctorados. Las variables utilizadas por el Consejo en los 5 informes se resumen en la Tabla 1.

En relación con las características de los graduados se recopilaron datos sobre 4 variables.

a) La proporción de graduados que habían recibido alguna beca estatal o ayuda durante la enseñanza de postgrado desde 1975 hasta 1979.

b) La media de años desde la matriculación en el centro de postgrado hasta recibir el doctorado.

c) La proporción de graduados en cada programa durante los años 1975-1979 que al cumplir los requisitos para el doctorado declararon que tenían ofertas firmes de trabajo dentro o fuera del mundo académico.

d) La proporción de graduados del programa entre los años 1975-1979 que al cumplir los requisitos del doctorado declararon que tenían ofertas firmes de trabajo en universidades que impartían cursos de doctorado (esta variable pretendía medir el éxito de los graduados para alcanzar posiciones académicas).

El tercer grupo de variables en este trabajo resultados de estudios de acreditación, se basó en respuestas a la encuesta del Consejo en 1981, e incluían:

a) la clasificación media sobre la calidad del profesorado del programa,

Tabla 1. Variables usadas por el CBAC

Variables	Humanidades	Sociales	Biol.	Mat.	Ingeniería
N.º de profesores	X	X	X	X	X
N.º de graduados	X	X	X	X	X
N.º de estudiantes	X	X	X	X	X
N.º de estudiantes con beca	X	X	X	X	X
Media de años para alcanzar doctorado	X	X	X	X	X
Proporción de graduados con empleo	X	X	X	X	X
Proporción de graduados con empleo académico	X	X	X	X	X
Prestigio del profesorado	X	X	X	X	X
Eficacia del programa	X	X	X	X	X
Mejora del programa	X	X	X	X	X
Familiaridad de los evaluadores	X	X	X	X	X
Indicador global sobre biblioteca	X	X	X	X	X
Proporción de profesores con subvenciones		X	X	X	X
Gastos de investigación		X	X	X	X
Publicaciones de los profesores		X	X	X	X

(X) significa que esta variable se utilizó en el estudio.

b) la clasificación media de la eficacia del programa para formar futuros investigadores/científicos,

c) la clasificación media sobre las mejoras en la calidad de los programas durante los últimos 5 años,

d) la clasificación media de la familiaridad de los evaluadores con el trabajo del profesorado del programa.

La última variable se utilizó como control para comprobar la fiabilidad de las 3 anteriores.

Por último, sobre los recursos de la biblioteca de la universidad se utilizó un índice compuesto (estimados con datos de 1979 y 1980). Este índice, basado en datos recopilados por la *Association of Research Libraries*, describía el tamaño de la biblioteca en la universidad en que se ubicaba cada programa.

Ciencias Sociales y de la Conducta

En este estudio se presentaron datos sobre 639 programas en antropología, economía, geografía, historia, ciencias políticas, psicología y sociología. Además de los 4 grupos de variables utilizados en el estudio humanidades, se añadieron dos más:

1) Datos sobre apoyos a la investigación, estimados por (a) la proporción del profesorado del programa con ayudas de investigación de la *Alcohol, Drug Abuse, y Mental Health Administration*, los *National Institutes of Health*, o la *National Science Foundation* durante el período de 1979-80, y (b) los gastos totales destinados por cada universidad a las actividades de investigación y desarrollo en un campo específico durante el año 1979.

2) Relación de publicaciones, evaluadas a través de las siguientes variables: (a) el número de artículos publicados por el profesorado del programa en los años 1978-80, y (b) la proporción de profesores del programa con uno o más artículos publicados durante estos mismo años.

Ciencias Biológicas

Se revisaron un total de 616 programas en bioquímica, botánica, biología celular/molecular, microbiología, fisiología y zoología. Se utilizaron las 6 áreas de variables utilizadas en el estudio anterior sobre ciencias sociales y de la conducta. La única diferencia fue que, en lugar de utilizar la proporción de profesores con uno o más artículos publicados como una variable, se utilizó una variable sustituto que estimaba la influencia global de artículos publicados. El Consejo declaró que esta

nueva variable no estaba diseñada para medir el impacto de artículos del profesorado, sino para medir el impacto de las revistas que incluyen trabajos relacionados con un programa de doctorado concreto.

Ciencias Matemáticas y Físicas

Se recopilaron y presentaron datos sobre 596 programas en química, informática, geología, matemáticas, física, y estadística/bioestadística. Las variables fueron similares a las utilizadas en el estudio de ciencias biológicas.

Ingeniería

Se revisaron en este informe exactamente 326 programas en ingeniería química, civil, eléctrica y mecánica. Las variables fueron similares a las utilizadas en el estudio de ciencias biológicas.

Algunos resultados de los estudios del CBAC

A partir de los resultados entre los cinco estudios anteriores el Consejo encontró diferencias entre disciplinas. Por ejemplo, en el estudio de humanidades, los programas de inglés, como promedio, tenían el número mayor de profesores (31), mientras el de música ocupaba el segundo lugar con 20 profesores. En los programas de inglés también se graduaron más estudiantes (una media de 44 doctorandos por programa en el período de 1975-79) y tenían una matrícula más elevada (una media de 62 doctorandos en diciembre de 1980). En cambio, los programas de alemán presentaban una media de sólo 9 profesores, 13 graduados y 15 doctorandos.

El Consejo también realizó el análisis de correlación (Pearson) con todas las combinaciones posibles de las variables utilizadas en cada estudio. Por ejemplo, en el estudio de humanidades, se encontró que 2 variables el número de graduados en el programa y la calificación del profesorado mediante escalas estaban altamente correlacionadas. Los resultados de correlación en otras disciplinas fueron similares.

Quizás la parte más significativa de este estudio ha sido probar la fiabilidad de tres ítems de la encuesta utilizada para evaluar al profes-

rado. Utilizando el análisis de correlación en dos mitades, se estableció la fiabilidad global de los resultados de la encuesta en cada disciplina en los cinco tipos de estudios. Se obtuvo como resultado un alto grado de fiabilidad de los juicios de los evaluadores en estos tres puntos de la encuesta.

Como paso final, el Consejo se defendió frente a algunas de las críticas contra estudios de reputación. En respuesta a la crítica de que las clasificaciones de prestigio pueden estar afectadas por juicios inexactos de evaluadores que no están familiarizados con el programa que han de evaluar, el Consejo realizó varias observaciones. Primero encontró que sus clasificaciones variaban en función del nivel de familiaridad de los evaluadores. Específicamente, los evaluadores calificaban más alto al profesorado con el que tenían una familiaridad considerable ésta era una categoría utilizada que aquellos con los que tenían poca familiaridad. Este hallazgo, sin embargo, no sorprendió. Sin embargo, en análisis posteriores, se encontró que las clasificaciones de evaluadores no familiarizados tenían efectos insignificantes sobre las clasificaciones globales.

Al considerar la crítica de que las evaluaciones puedan estar afectados por la proximidad geográfica de los evaluadores, en este estudio se reagruparon los programas en dos categorías, cerca y lejos, dependiendo de la proximidad de los evaluadores al lugar donde se imparten los programas. Analizados de nuevo las clasificaciones del profesorado, se encontró que, aunque los evaluadores les deben a los programas cercanos clasificaciones más altas que a los de fuera de la región donde residían, las diferencias eran insignificantes. El Consejo consideró a estas diferencias como un «efecto secundario», ya que los evaluadores tendían a clasificar más alto aquellos programas con los que estaban más familiares, bien cercanos o no. Asimismo, las elevadas correlaciones entre las clasificaciones medias de los dos grupos de programas, los cercanos y los lejanos, mostraron que el rango ocupado por los programas no estaba influenciado de forma significativa por la proximidad geográfica de los evaluadores.,.

Al considerar otra crítica, a saber, que las encuestas de evaluación pueden estar afectadas por sesgos característicos de antiguos alumnos, el Consejo analizó sus propios datos y encontró evidencias que apoyan esta afirmación: existían diferencias en las evaluaciones del profesorado entre evaluadores que fueron antiguos alumnos y los que no lo

fueron. Afortunadamente, la proporción de evaluadores que eran antiguos alumnos de los programas que se les pedía evaluar era pequeña y por tanto tenía «muy poco» efecto sobre los resultados globales de la encuesta. Se llegó a esta conclusión calculando de nuevo las clasificaciones medias del profesorado en algunos programas seleccionados excluyendo la evaluación de los antiguos alumnos y comparando estas clasificaciones con las medias originales. En el estudio de humanidades, por ejemplo, las clasificaciones medias de 485 de los 522 programas habían permanecido inalterables (aproximadamente variaban en una décima). Se observaron diferencias similares insignificantes para las ciencias sociales y de la conducta, las ciencias biológicas, las ciencias físicas y matemáticas e ingeniería. Convincentemente, en este estudio se llegó a la conclusión de que el sesgo determinado por el hecho de que el evaluador hubiera sido antiguo alumno no jugaba un papel significativo en las calificaciones y por tanto, las puntuaciones de los evaluadores con estas características no se excluyeron.

El estudio también consideró la crítica de que estas evaluaciones pueden estar afectadas por la interacción entre los intereses de los evaluadores por la investigación y el énfasis que los programas de doctorado hacen sobre la investigación. Algunos críticos ha indicado que dado que algunos programas pueden ser muy especializados en áreas restringidas de investigación, este campo de especialización quizás no se conozca y no sea evaluado adecuadamente. El Consejo reconoció la dificultad de identificar intereses por la investigación dentro de los programas; por ello sólo intentó hacerlo en el campo de la física y la estadística /bioestadística. Se dividieron los evaluadores de cada disciplina en dos grupos, según su área de especialización. Por ejemplo, en física, el primer grupo incluía especialistas en partículas elementales y estructura nuclear, mientras que el segundo grupo estaba compuesto por especialistas de las demás áreas de la física. Se volvieron a calcular las clasificaciones medias del profesorado para cada uno de estos dos grupos y los resultados globales indicaron que la especialidad de los evaluadores tenía efecto sobre sus clasificaciones, pero no el suficiente para modificar la ubicación o clasificación de los programas. Se llegó a una conclusión similar en el caso de los programas de estadística/bioestadística.

En resumen, el Consejo ha desarrollado un estudio sobre acreditación de programas doctorales muy mejorado. Las metodologías utilizadas en la recolección de los datos de los programas e instituciones y

las evaluaciones realizadas han sido claramente mejores que las efectuadas anteriormente. Sin embargo, tenían limitaciones. La primera fue que el Consejo no intentó definir el concepto de calidad. Se acercó a una definición cuando especificó que sus encuestas eran medidas de la calidad percibida en relación o un programa no calidad en un sentido ideal o absoluto. Por ello, el Consejo no jerarquizó los programas. Argumentaron que de intentarlo supondría no captar la calidad real de los programas. Otra limitación del estudio fue que el Consejo supuso simplemente que el prestigio del profesorado era la principal variable indicadora de calidad. Aunque esto no es una crítica fuerte, se podría haber combinado los resultados de la encuesta sobre el profesorado con datos obtenidos sobre variables objetivas para llegar a una definición o medida más adecuada de la calidad. Los estudios futuros deberían considerar estas lagunas.

Estudios sobre programas profesionales y de pregrado

Una de las críticas principales a los estudios de reputación es que la mayoría se centran en programas de postgrado. Teniendo en cuenta esta crítica, algunos investigadores se han centrado en programas de pregrado y profesionales.

Un intento de valorar programas profesionales fue un estudio que clasificó los programas de 17 campos profesionales: arquitectura, negocios, educación, ingeniería, derecho, biblioteconomía, enfermería, farmacia, trabajo social, odontología, agricultura, periodismo, medicina, optometría, salud pública, teología y veterinaria. Utilizando un enfoque similar al de Roose y Andersen, 1.180 decanos de escuelas profesionales identificaron en un cuestionario los mejores programas de sus respectivas especialidades. Posteriormente se juntaron las respuestas y se clasificaron los programas por institución. Como ejemplos los primeros programas de arquitectura fueron los de la Universidad de California (Berkeley), Massachusetts Institute of Technology, Universidad de Harvard, Universidad de Pennsylvania y la Universidad de Princeton; mientras que los primeros en programas de asistencia social fueron los de la Universidad de Columbia, la Universidad de Chicago, la Universidad de Michigan, la Universidad de Brandeis y la Universidad de Case Western Reserve. En un estudio de seguimiento, Blau y Margulies añadieron la música como el campo décimo octavo y respondieron a las críticas a su primer informe. Primero rebatieron la crítica de que sus descubrimientos

sólo tenían por objeto crear «efecto halo» institucional. Según ellos, las escuelas profesionales realmente difieren en calidad, y proporcionar esa información a los consumidores es un servicio público. Segundo, defendieron la baja tasa de respuesta de los decanos (sólo el 36%) en su primer informe (los críticos indicaban que una baja tasa podía sesgar los resultados) proporcionando evidencia de que sea alta o baja la tasa de respuesta no afectaba a sus *rankings*. En su estudio de seguimiento Blau y Margulies lograron una tasa más alta de respuesta (79%) y descubrieron que los *rankings* eran los mismos que en el estudio original.

Cartter y Solmon en 1977 realizaron otro estudio de programas profesionales. Este estudio clasificaba programas en derecho, educación y negocios. Básicamente utilizaban la misma metodología de Blau y Margulies, enviando cuestionarios a profesores y decanos solicitando calificaciones sobre la calidad del profesorado y programas en un escala del tipo Likert. Las cinco primeras escuelas de derecho del país eran las Universidades de Harvard, Yale, Stanford, Michigan y Chicago; las cinco primeras en educación eran Stanford, Harvard, Chicago, Universidad de California en Los Angeles y Universidad de California (Berkeley); y las cinco primeros puestos en negocios lo ocupaban Stanford, Harvard, Massachusetts Institute of Technology, Chicago, y Universidad de Carnegie Mellon. Cartter y Solmon encontraron grandes similitudes entre sus hallazgos y los de Blau y Margulies. También existían diferencias que, según Munson y Nelson, podían deberse simplemente a las diferencias en la muestra e instrumentos utilizados.

En un estudio de Carpenter y Carpenter (1970) que clasificaba 44 programas de bibliotecas y documentación, se incluían, además, los siguientes estudios sobre reputación de cursos profesionales: un estudio de Cole y Lipton (1977) sobre facultades de medicina; dos estudios de la *MBA Magazine* (1974, 1975) sobre los 15 primeros programas de negocios y un estudio de la *Juris Doctor Magazine* (1976), que clasificaba las 20 primeras escuelas de derecho del país. El estudio sobre las escuelas de derecho era original dado que usaron como evaluadores a 1.300 lectores de la revista, además de 58 decanos. El estudio halló consistencia entre ambos grupos.

Por últimos, otro grupo de estudios sobre reputación valoraba programas de pregrado. Un estudio de Johnson (1978) identificaba una lista de 10 instituciones líderes por su influencia a nivel nacional. Las diez instituciones eran las universidades de Columbia, Harvard, Massa-

chusetts Institute Technology, Princeton, Stanford, California (Berkeley), California (Los Angeles), Chicago, Michigan y Yale. Esta lista resultó ser similar a la lista generada por Petrosky *et al.* (1973). Astin y Solmon (1981) en otra valoración de educación a nivel de pregrado, utilizando un enfoque bastante diferente, identificaron los principales departamentos de pregrado en economía, historia, inglés, química, sociología y biología. Los investigadores pidieron a los evaluadores que categorizaran los departamentos basándose en 6 criterios: 1) La calidad global de la educación a nivel de pregrado. 2) El nivel de preparación de los estudiantes para escuelas de postgrado o profesionales. 3) El nivel de preparación de los estudiantes para el mundo laboral después de la universidad. 4) El compromiso de los profesores con la enseñanza a nivel de pregrado. 5) Los éxitos profesionales o académicos de los profesores. 6) innovación del curriculum y de la metodología didáctica (Astin y Solmon, 1981, p. 17). Se hizo una lista de las diez primeras universidades cuyos programas reunieran al menos uno de los seis criterios y que no estuvieran incluidos en la selección efectuada por Roose y Andersen (1970). Como ejemplo, los principales centros de economía, que satisfacían estas dos condiciones eran los de Swarthmore, Dartmouth, Haverford, Amherst, Carleton y Claremont's Men, mientras que en biología estaban entre los primeros Reed, Amherst, Carleton, Pomona y Bowdoin, Más importante aún, Astin y Solmon descubrieron que el prestigio en una disciplina o campo en una institución proporcionaba una estimación razonable de prestigio en otros programas de pregrado en la misma institución. Astin y Solmon justificaron estos resultados diciendo que determinados sesgos y efectos halo habían afectado sus evaluaciones. Además, descubrieron que la calidad podía ser estimada con bastante exactitud a partir de hechos conocidos, como gastos para bibliotecas, edificaciones y objetivos educativos, y los títulos concedidos en ciencias naturales e ingeniería. Dado el alto coste para llevar a cabo estudios de reputación y las limitaciones inherentes a las escalas de estimación, Astin y Solmon concluyen afirmando la necesidad de que se realicen más estudios sobre reputación, especialmente teniendo en cuenta más campos y más criterios de calidad.

Crítica a los estudios sobre reputación.

Los estudios sobre reputación han sido criticados por sus limitaciones metodológicas. Se critica normalmente la falta de una definición clara o una medición objetiva de la calidad. En primer lugar casi todos

insisten en que lo que realmente se mide en estos estudios es el renombre o el prestigio de programa, institución o profesorado, pero no la calidad de los resultados (Dolan, 1976; Lawrence y Green, 1980). Segundo, existen, además los problemas de sesgo, bien porque los evaluadores hayan sido antiguos alumnos de los programas o centros que están evaluando (Webster, 1981; Cartter, 1966; Lawrence y Green, 1980) o bien porque están familiarizados con los programas que se les pide que evalúen (Blackburn y Lingerfelter, 1973; Lawrence y Green, 1980). Tercero, se puede considerar que un programa tiene más prestigio que otro sólo cuando se comparan en una escala relativa. Cuarto, los estudios de reputación han sido criticados por haberse centrado generalmente en programas de postgrado y en las primeras instituciones del país. Las críticas señalan que no existe información acerca del estatus de escuelas normales, universidades y *colleges* regionales estatales, *colleges* femeninos y otras instituciones que también efectúan una función educativa superior vital (Petrowski et al. 1973; Conrad y Blackburn, 1985a). Quinto, se han criticado igualmente estos estudios porque han contribuido a producir el llamado efecto Hertz-Avis estableciendo un orden jerárquico entre las instituciones de educación superior (Lawrence y Green, 1980). Sexto, algunos estudios de reputación han sido criticados por no considerar el entorno institucional, como el tamaño institucional y la mezcla cultural de los estudiantes, que pueden afectar a las evaluaciones (Astin y Solmon, 1981; Dolan, 1976; Lawrence y Green, 1980). Por último, algunas críticas señalan que algunos *rankings* no reflejan la reputación y práctica actual de los departamentos debido a que muchas estimaciones se basan en datos pasados e impresiones (Lawrence y Green, 1980; Cox y Catt, 1977).

A pesar de tales críticas los estudios sobre reputación tienen muchos méritos. En particular, han sido muy informativos sobre la excelencia de programas académicos a nivel de doctorado. Por ello, muchos investigadores propugnan la continuación de estos estudios, particularmente aquellos que han utilizado metodologías más adecuadas. Este sentimiento es expresado por Conrad y Blackburn (1985) de este modo:

«La controversia sobre estudios de reputación no debería detener a los investigadores a llevar a cabo dichos estudios en el futuro. Si estos estudios están diseñados para responder a las críticas recibidas, estamos convencidos de que pueden efectuar una contribución importante para evaluar la calidad en la educación superior. Especialmente si se realizan refinamientos metodológicos, si la calidad se evalúa mediante criterios múltiples, si se utilizan normas de calidad, y si la calidad no se evalúa simplemente en

escuelas de «élite», los estudios sobre reputación podrán ayudar a clarificar la calidad de los programas a nivel de licenciatura, cursos profesionales y de postgrado.»

Estudios a partir de indicadores objetivos

En lugar de utilizar las calificaciones otorgadas por los profesores, jefes de departamento y decanos, los estudios en base a indicadores tienen como propósito valorar la calidad mediante medidas objetivas. Dado que existe poca teoría para guiar a los investigadores acerca de las variables que son apropiadas para medir la calidad, los investigadores las seleccionan según su propio criterio. Como resultado, se han utilizado diversas variables. Sin embargo, estas variables se pueden categorizar en cinco tipos generales según tengan relación con: 1) el profesorado, 2) los estudiantes, 3) los recursos del departamento o de la institución, 4) los resultados, y 5) otros criterios.

Estudios basados en el profesorado

Quizás los estudios más conocidos sobre indicadores son aquellos basados en el profesorado. Los investigadores que realizan este tipo de estudios mantienen que la calidad de un departamento depende de la calidad de su profesorado. Sugieren que si el profesorado tiene alta calidad el departamento tendrá necesariamente también una calidad elevada. El problema, sin embargo, estriba en la falta de acuerdo en cómo medir la calidad del profesorado. La mayoría de investigadores han utilizado los premios recibidos o su origen académico.

Un ejemplo de estudio que no empleó la productividad de investigación del profesorado para clasificar programas fue el informe de Bowker (1965). En este informe, el investigador clasificó las escuelas de postgrado en dos grupos de disciplinas, ciencias sociales y humanidades y ciencias, y utilizó un índice de calidad basado en cuatro criterios: 1). El número de antiguos becarios (Woodrow Wilson) entre el profesorado 2). El número de estos becarios que asisten a la institución. 3). El número de profesores premiados por el *American Council of Learned Societies Award* y 4). El número de becarios (Guggenheim Memorial) entre el profesorado. Las 10 primeras instituciones en los dos grupos de disciplinas fueron las universidades de Columbia, Harvard, Yale, California

(Berkeley), Chicago, Stanford, Michigan, Princeton, Indiana y Pennsylvania (Bowker, 1965).

A excepción del estudio de Bowker, la mayoría de los estudios basados en el profesorado habían utilizado la productividad de la investigación como variable para mediar la calidad. Por ejemplo, los estudios de Somit y Tonenhaus (1964), Lewis (1968), Knudsen y Vaughn (1969), Glenn y Villemez (1970), Clemente y Sturgis (1974), Wispe (1969), Cox y Catt (1977), Liu (1978), Gill (1975) y House y Yeager (1978).

Somit y Tanenhaus (1964), en su intento de clasificar los departamentos de ciencias políticas, usaron publicaciones del profesorado como la medida objetiva de la calidad del departamento: a mayor número de publicaciones, mayor calidad. Si embargo, los autores tuvieron cuidado al hacer inferencias sobre instituciones de menor categoría. Declararon que los escasos registros de publicaciones en instituciones de menor categoría no significaban necesariamente que los profesores en esas instituciones fueran de baja calidad o deficientes en formar a sus estudiantes en productividad de investigación. Los autores sugirieron que más horas de docencia y no poder acceder a bibliotecas adecuadas podían haber causado escasez en el número de publicaciones del profesorado en estas instituciones.

En otro estudio, Lewis (1968) utilizó también la relación de las publicaciones del profesorado como medida de calidad, esta vez para valorar los 17 primeros departamentos de sociología identificados en el informe Cartter de 1966. También se consideró el volumen de las publicaciones de los doctorandos (como indicador para medir la efectividad de la formación en el postgrado). Las publicaciones de profesores y estudiantes se midieron por el número de artículos, informes y notas de investigación, así como comentarios adicionales (excluyendo cartas al editor), publicadas en la *American Sociological Review* desde 1955 a 1965. Estos datos constituyeron la base para categorizar los departamentos. Lewis comparó las clasificaciones obtenidas con las de Cartter y encontró que, con excepción de unos pocos programas (como los de las universidades de Minnesota, Northwestern y Washington), había una relación consistente entre el prestigio (estimado en el trabajo sobre acreditación de Cartter) y la productividad de investigación (estimada de forma objetiva).

Knudsen y Vaughan (1969), utilizando una metodología similar, examinaron también la relación entre medidas de prestigio y objetivos en

los departamentos de sociología. Las medidas objetivas empleadas aquí para categorizar programas eran registros de publicaciones del profesorado y graduados en los programas. Los datos sobre las publicaciones se midieron de dos maneras: 1) El número total de publicaciones. 2) El número de publicaciones por persona (para controlar el tamaño del departamento o centro). Se consideraron las contribuciones académicas de las siguientes revistas: la *American Sociological Review*, el *American Journal of Sociology* y *Social Forces*. Se diseñó un esquema de pesos para medir el significado de cada tipo de contribución. Se asignó un peso de 48 a una monografía teórica o de investigación, 24 a un libro de texto, 16 a un libro editado y a un artículo en la *American Sociological Review*, 12 a un artículo en el *American Journal of Sociology* y ocho a un artículo en *Social Forces* o unos apuntes sobre investigación en la *American Sociological Review*. Este esquema de pesos permitió que medir y categorizar la productividad en investigación de cada departamento. Las clasificaciones obtenidas de esta forma se compararon con las estimaciones del estudio de Cartter. Globalmente, Knudsen y Vaughan encontraron similitudes entre las dos clasificaciones, pero solamente en las instituciones destacadas o claramente superiores. Por ejemplo, de estos se clasificaron 13 instituciones dentro las 15 primeras de ambas listas las universidades de California (Berkeley), Harvard, Chicago, Michigan, Columbia, Princeton, Wisconsin, California (Los Angeles), Stanford, Northwestern, Yale, Cornell y Minnesota. Por debajo de estas 15 instituciones las clasificaciones no fueron tan consistentes (Knudsen y Vaughan, 1969).

Otro estudio basado en indicadores objetivos sobre los departamentos de sociología fue un informe de Glenn y Villimez (1970) que examinó 45 programas de este campo. Este estudio era una réplica del estudio de Knudsen y Vaughan salvo que cubría un período diferente e incluía más publicaciones. Recordando el estudio de Knudsen y Vaughan, se consideraron datos sobre publicaciones de la *American Sociological Review*, *American Journal of Sociology* y *Social Forces*. En el estudio de Glenn-Villimez, sin embargo, se consideraron, además, las contribuciones de otras 19 revistas, incluyendo *Sociology of Education*, *American Sociologist*, *Social Problems* y *Social Science Quarterly* (para una lista completa, ver Glenn y Villimez, 1970). Otra diferencia respecto al estudio de Knudsen y Vaughan fue que Glenn y Villimez consideraron los apuntes sobre investigación y los comentarios extensos además de libros y artículos. Los investigadores también revisaron el sistema de pesos diseñado por sus predecesores y lo llamaron Índice Global Glenn-Villimez.

Por ejemplo, en lugar de asignar un peso de 48 a las monografías teóricas o de investigación, asignaron un 30 (los pesos se basaban en respuestas a un cuestionario a partir de una muestra de 109 sociólogos a los que se pidió que asignaran pesos a cada tipo de publicación). Utilizando el Índice Global Glenn-Villimez, que se consideraba como una medida más adecuada de la cantidad y calidad de la productividad de investigación de los sociólogos que el índice Knudsen-Vaughan, se clasificaron los programas doctorales y las clasificaciones se compararon a las generadas por Knudsen y Vaughan. Glenn y Villimez hallaron consistencias entre ambas valoraciones, pero sólo para las instituciones de élite. Los programas clasificados muy alto en las universidades de Columbia, Chicago, Michigan, Harvard, Wisconsin y California (Berkeley) habían recibido una categoría similar en ambas listas. Por debajo de estas instituciones, las diferencias eran bastante significativas (Glenn y Villimez, 1970).

Clemente y Sturgis (1974), en un intento de repetir el estudio de Glenn-Villimez, utilizaron el Índice Global Glenn-Villimez para convertir los datos sobre publicaciones de 2.205 doctores en sociología entre los años 1940-70 en puntuaciones de productividad en investigación del profesorado. El propósito principal del estudio era examinar si existía una correlación entre la calidad del entrenamiento en el postgrado y la productividad de investigación. La calidad de entrenamiento del postgrado se operativizó mediante 4 escalas de evaluación. Las primeras tres escalas fueron recopiladas directamente de los estudios de Keniston (1959), Cartter (1966), y Roose Andersen (1970). La cuarta escala, creada por Clemente y Sturgis, era esencialmente una composición obtenida combinando las tres anteriores. Cuando se utilizó el análisis de correlación para probar la relación entre entrenamiento en el postgrado (medido a través de las cuatro escalas) y la productividad de investigación, los resultados revelaron sólo una débil relación entre ambas estimaciones (Clemente y Sturgis, 1974).

En un estudio sobre los departamentos de psicología, Wispe (1969) utilizó registros de publicaciones del profesorado para clasificar los departamentos. Pero el propósito principal de su estudio era examinar las razones de la productividad de los departamentos de psicología: ¿un departamento es más productivo (medido mediante datos sobre publicaciones) que otro porque es de mejor calidad o por que es mayor? Wispe trabajó con una muestra de 28 departamentos de psicología que entre 1895 y 1965 habían concedido al menos un doctorado en psicolo-

gía y descubrió que el tamaño del departamento estaba relacionado con la productividad de investigación. Wispe encontró una correlación bastante alta entre las dos variables: 0,82 (p. 0,1) en 1950 y 0,86 (p. 0,1) en 1960. Otros investigadores como Drew (1975), Guba y Clark (1978), Elton y Rodgers (1971), Elton y Rose (1972), Hagstrom (1971), y Knudsen y Vaughan (1969) habían encontrado resultados similares. Sin embargo, a nivel de pregrado no se ha descubierto que el tamaño del departamento esté relacionado con la calidad. Este resultado fue confirmado por Astin y Solom (1981), que encontraron una correlación negativo entre las dos variables.

Otro estudio sobre departamentos de psicología fue un informe de Cox y Catt (1977) que clasificaba a los departamentos de psicología según las contribuciones académicas del profesorado a 13 revistas subvencionadas por la *American Psychological Association* entre los años 1970 y 1975. Cuando se compararon estas clasificaciones con las evaluaciones de los mismos departamentos efectuadas por Roose y Andersen (1970), Cox y Catt hallaron que las clasificaciones Roose Andersen no reflejaban ya la productividad académica actual de estos departamentos (cf. Lawrence y Green, 1980, p. 10).

No todos los estudios sobre indicadores objetivos que utilizan la productividad de investigación del profesorado como medida de calidad se han centrado en programas de psicología y sociología; algunos se han centrado en otras disciplinas. El estudio de Liu (1978), por ejemplo, categorizó los departamentos de ingeniería mecánica según el número de citas de los profesores. En este trabajo se utilizó una muestra de 232 profesores y un total de 3.264 citas de edición de 1975 del *Science Citation Index*. Liu empleó la media de citas para determinar los rankings de departamentos porque creía que era una medida mejor que el número total de citas. Cuando Liu comparó sus clasificaciones a las elaboradas por Gill (1975) encontró grandes diferencias entre las dos excepto en un aspecto interesante. Los cinco primeros departamentos (los de las universidades de Stanford, Minnesota, California, Massachusetts Institute of Technology y California Institute of Technology) eran los mismos en ambos estudios excepto que el orden era diferente. Otro estudio similar sobre una disciplina diferente fue realizado por House y Veager (1978), que valoraron los departamentos de economía basándose en el número total de páginas publicadas por profesores titulares en 45 importantes revistas de la disciplina.

Se debe señalar que no todos los estudios han utilizado la productividad de investigación del profesorado como indicador de calidad. Además de las características del profesorado utilizados por Bowker (1965), algunos investigadores han utilizado los años de experiencia docente, las calificaciones académicas del profesorado, los premios recibidos por el profesorado, y otros datos similares, pero como Blackburn y Lingenfelter (1973) explicaron, la mayoría de estos indicadores no eran tan válidos como la productividad en la investigación para indicar la calidad del departamento.

El uso de las características del profesorado como indicadores de calidad no se acepto tanto como cabría esperar. El primer problema que causa diversidad de opiniones es suponer que la productividad en la investigación determina la calidad del profesorado y que por tanto ambas variables son iguales. Algunos investigadores han indicado que no son necesariamente lo mismo. El segundo problema en el que igualmente no hay acuerdo es suponer que la mejor forma de medir la calidad de un departamento es a través de la calidad del profesorado (medida por la productividad en la investigación). La habilidad y resultados de los estudiantes, y los recursos financieros se consideran frecuentemente como factores igualmente importantes para clasificar los departamentos de alta calidad. Por tanto, se necesita más investigación para explorar hasta qué punto el profesorado y otras características juegan un papel en el aumento de la calidad.

Estudios que utilizan variables sobre los estudiantes

Aunque lo más común es utilizar las características del profesorado para medir la calidad, algunos investigadores han utilizado las características de los estudiantes. Entre los investigadores que han empleado variables relativas a los estudiantes para medir la calidad de los programas están Knapp y Greenbaum (1953), Krause y Krause (1970), Dubé (1974), Tidball y Kristiokowski (1976), Astin y Henson (1977), Astin y Solmon (1979) y Glower (1980).

En el estudio de Knapp-Greenbaum, se valoraron las instituciones basándose en la proporción de sus alumnos que obtuvieron título de doctor. En el estudio de Krause y Krause, los *colleges* de pregrado fueron valorados de acuerdo con el número de sus graduados que habían escrito artículos para *Scientific American* entre 1962 y 1967; en el

estudio de Dubé, se clasificaron 100 *colleges* de pregrado según el número de graduados que fueron admitidos en las escuelas de medicina en 1973-74; en el estudio de Tidball-Kristiakowski, se clasificaron los *colleges* según el número de sus alumnos que obtuvieron título de doctor; en el de Astin-Henson, se utilizó como medida de la calidad de la institución el índice de selectividad de los estudiantes (medido a través de su rendimiento académico en el primer curso); en el estudio de Astin-Solmon, se usaron las preferencias de *colleges* de los buenos estudiantes como base para medir la calidad; y por último, en el estudio de Glower, se utilizaron el éxito en las carreras de los antiguos alumnos según datos biográficos en *Who's Who In Engineering* para medir la calidad de la escuelas de ingeniería.

El uso de características de estudiantes como indicadores de la calidad de programas ha recibido críticas. Aunque «la calidad de los estudiantes es un criterio de excelencia por propio derecho» porque «los estudiantes bien cualificados son un elemento esencial de un excelente programa» (Blackburn y Lingenfelter, 1973, p. 8), existe pocas evidencias empíricas que apoyen la relación directa entre excelencia de los estudiantes y calidad de los programas. Como mucho, la mayoría de los estudios han podido demostrar que ciertas características como rendimiento y número de estudiantes están correlacionadas positivamente con la calidad de los programas, especialmente a nivel de postgrado, pero no han explicado la gran varianza de los programas en relación con la calidad. Además, no se llega a un acuerdo sobre el tipo de variables que podrían usarse para medir la calidad de los estudiantes. Kuh (1981) sugirió que se utilizase la experiencia como un indicador de la excelencia de los estudiantes, pero esto también es difícil de medir. La mayoría de los investigadores, según Kuh, han confiado en índices sustitutos de la experiencia de los estudiantes en lugar de mediciones directas. Evidentemente, se necesita desarrollar índices más adecuados para medir la experiencia y excelencia de los estudiantes.

Estudios que utilizan variables sobre recursos

Además de la características de los profesores y los estudiantes, algunos investigadores han utilizado los recursos para medir la calidad. Estos investigadores sugieren que la calidad se puede medir por los recursos departamentales, institucionales, y humanos. Por ejemplo, el número de profesores, personal auxiliar y estudiantes; el valor de las

edificaciones (bibliotecas, laboratorios, despachos, salas de informática); las dotaciones; los gastos por estudiante y por profesor; los sueldos de los profesores; los fondos de investigación; las ofertas de enseñanzas; y la diversidad de programas. Aunque se ha discutido mucho que los recursos son un importante factor para mejorar la calidad, pocos estudios han podido establecer la relación directa que existe entre ambos tipos de variables.

Ejemplos de estudios que han utilizado los recursos para medir la calidad son los informes de Kelso (1975), Glower (1980), y el Carnegie Council on Policy Studies in Higher Education (1976). Kelso categorizó las escuelas de derecho de acuerdo con un índice de recursos. Glower (1980) utilizó el número total de graduados según *Who's Who In Engineering* de 1977 y la cantidad total invertida en investigación para clasificar las escuelas de ingeniería. Glower especula sobre el supuesto de que la utilización de recursos como medida de la calidad de los departamentos es más fiable que las medidas de reputación porque este método puede clasificar a todos los programas, no sólo los visibles.

Quizás el estudio más conocido que utilizó los recursos para clasificar las instituciones fue un informe del *Carnegie Council on Policy Studies in Higher Education* (1976). En este informe el *Carnegie Council* clasificó las instituciones en 6 grandes categorías basándose en la cantidad de apoyo financiero federal recibida por cada institución, el número de doctorados concedidos, y el índice de selectividad de estudiantes de la institución. Las instituciones clasificadas en primer lugar por el volumen de fondos federales recibidos y el número de doctorados concedidos en 1973-74 (incluyendo doctorados en medicina si la institución tenía una escuela de medicina) eran las instituciones que conceden doctorados. En el siguiente grupo estaban las universidades y *colleges* de tipo general, seguidos por *colleges* de artes liberales, *colleges* e institutos de dos años, escuelas profesionales y otras instituciones especializadas, y por último institutos de educación permanente.,

No se comprende plenamente hasta qué punto se pueden utilizar los recursos como indicadores de calidad. Aunque por una lado se ha especulado que los recursos financieros son un factor importante para mejorar la calidad, existen pocas evidencias empíricas que prueben la relación directa entre los dos. Abbott y Barlow (1972), por ejemplo, en su estudio sobre programas de sociología, no pudieron encontrar suficientes evidencias que apoyasen esta relación. En su lugar, encontraron

variables como la productividad de investigación del profesorado y el número y tipo de títulos concedidos como variables intervinientes entre los recursos y el prestigio de una institución. Los recursos pueden ser un factor para mejorar la calidad, pero no de forma directa, evidentemente.

Estudios basados en los resultados

Los investigadores que realizan ese tipo de estudios sugieren que la calidad de los departamentos o programas no dependen tanto de los inputs (recursos institucionales y de departamentos) como de los resultados. Las variables de resultados son los productos de los estudiantes y de los antiguos alumnos.

Un estudio que utilizó los resultados para medir la calidad fue el de Knopp y Goodrich (1952). Estos autores consideraron la proporción de estudiantes que habían obtenido título de doctor como un indicador para medir la calidad de la institución. En un estudio posterior de seguimiento, esa variable fue sustituida por la proporción de estudiantes que habían conseguido becas de postgrado (Knapp y Greenbaum, 1953).

Otros estudios dentro de este grupo son los de Tidball y Kristiakowski (1976), el estudio del *Carnegie Council on Policy Studies in Higher Education* (1976), y el del *National Science Board* (1969). Tidball y Kristiakowski usaron el número de antiguos alumnos que obtuvieron el doctorado como medida de calidad; sin embargo, tanto el *Carnegie Council* como el *National Science Board* emplearon el número de doctorados concedidos.

No existe un acuerdo claro sobre la utilidad de emplear variables sobre los resultados como indicadores de calidad. Algunos investigadores afirman que los estudiantes muy productivos y con éxito determinan que un departamento sea de alta calidad, pero hay pocas evidencias empíricas que lo apoyen. Por lo contrario, existen investigadores y profesores que no están completamente convencidos de que los resultados sean indicadores precisos de la gran calidad de los programas. Astin (1962), por ejemplo, proporciona evidencias de que la mayoría de los resultados del aprendizaje de los estudiantes dependían más de la calidad de propios estudiantes que de la calidad de los programas. La relación entre inputs (como las habilidades de los estudiantes), resulta-

dos, y calidad de los programas, necesita ser considerada en futuras investigaciones.

Estudios basados en criterios múltiples

Los estudios analizados hasta ahora en esta sección han utilizado generalmente un enfoque con una sola variable para medir la calidad. Sin embargo algunos estudios han empleado el enfoque multivariado. Quizás los más controvertidos de los que han utilizado este enfoque son los trabajos de Gourman (1967, 1977a, 1977b, 1982, 1983). Estos informes han sido criticados severamente y justificadamente porque la metodología utilizado para obtener las calificaciones no se explicaba adecuadamente. Dado que simplemente se facilitaban los datos sin una adecuada explicación, Lawrence y Green (1980) llamaron a las calificaciones de Gourman «idiosincráticas e irreplicables». Otro crítico, Webster (1984) cuestionó la credibilidad de Jack Gourman como un evaluador de instituciones y proporcionó evidencias de que los informes de Gourman tenían serios errores.

En un trabajo más riguroso realizado por Brown (1967) se utilizaron 8 variables objetivas para valorar los *colleges*: 1) La relación de profesores con doctorado. 2) El sueldo medio y los ingresos adicionales de los profesores. 3) La proporción de estudiantes que continuaban estudios de postgrado. 4) El porcentaje de estudiantes que se graduaban. 5) El número de libros en la biblioteca para cada estudiante a tiempo completo. 6) El número total de profesores a tiempo completo. 7) La ratio profesor-alumno. 8) El total de ingresos por estudiante. Se clasificaron un total 1.121 instituciones anónimas según los 8 criterios anteriores; asignándose una puntuación de uno a 1.121 a cada una de las instituciones en cada criterio. Promediando las ocho puntuaciones para cada institución, Brown llegó a obtener un rango global por cada institución, que él llamó clasificación compuesta.

Dado que la calidad es un concepto multidimensional, el uso de variables múltiples puede ser el enfoque ideal para obtener un indicador comprensivo de la calidad. Desafortunadamente, pocos investigadores han utilizado este enfoque. El problema básico ha sido la dificultad en seleccionar una combinación «correcta» de variables que midan la calidad. Esto continuará siendo un problema si no tenemos una teoría de la calidad. Ahora bien, la mejor oportunidad de que se desarrolle una teoría de la calidad es mediante el uso de variables múltiples. Es el proble-

ma de la gallina y el huevo que tendrá que ser resuelto en estudios futuros.

Crítica a los estudios sobre indicadores objetivos

Los estudios sobre indicadores objetivos han contribuido claramente a nuestra comprensión de la calidad de la educación superior. Significativamente, han sacado a la luz un enfoque completamente diferente para medir la calidad: el uso de variables objetivas. Así, la principal ventaja de este enfoque es que hay una medida objetiva de la calidad.

Al mismo tiempo, los estudios sobre indicadores objetivos parecen tener también limitaciones metodológicas. Una limitación común en la mayoría de los estudios es que, a menudo, han utilizado la productividad del profesorado en el campo de la investigación como la única medida de la calidad. En muchos de estos estudios, los investigadores han sido criticados por igualar la calidad del profesorado con la calidad de los programas. Algunas críticas han indicado que la calidad del profesorado y de los programas no es lo mismo y que por tanto no pueden considerarse como iguales (Smith y Fiedler, 1971; Liu, 1978). Esto es la limitación más seria de la mayoría de los estudios que utilizan indicadores objetivos. Otras de las limitaciones frecuentemente citadas son: a) El fracaso de muchos estudios por incluir en la muestra solamente instituciones que son muy conocidas (Conrad y Blackburn, 1985a), b) La incapacidad de estos estudios para llegar a una serie de medidas objetivas que puedan usarse para estimar la calidad de instituciones muy diferentes, no sólo de las que ocupan los primeros rangos (Conrad y Blackburn, 1985a) y c) las dificultades derivadas por incluir variables múltiples. No obstante Webster (1981) argumentó que dado que la calidad de los programas es multidimensional necesariamente debe medirse utilizando variables múltiples. Por ello, algunas de estas limitaciones necesitan ser tenidas en cuenta y, quizás, resolverse en estudios futuros.

Estudios sobre correlaciones

Al contrario que los estudios de reputación e indicadores objetivos, los estudios basados sobre correlaciones no se han diseñado para obtener una medición de la calidad. Más bien, este tipo de estudios han

intentado identificar variables que se correlacionaban con programas de mucho prestigio (utilizando las escalas que usualmente utiliza ACE, para identificar programas prestigiosos). Las variables relacionadas con las clasificaciones obtenidas se denominan «correlatos de calidad». No todos los estudios, sin embargo, han usado las escalas sobre reputación como variable dependiente. Dos excepciones a la regla son los estudios realizados por Axelson (1960) y Conrad y Blackburn (1986).

Unos de los primeros investigadores que identificó los correlatos de programas de alta calidad fue Cartter (1966). Aunque su estudio constituye esencialmente un trabajo de reputación, Cartter identificó dos correlatos de calidad: las compensaciones al profesorado y los recursos de la biblioteca. Encontró que puntuaciones altas en ambas variables estaban relacionadas con programas muy prestigiosos. Sin embargo sólo había una correlación entre las dos variables, no una relación de causa-efecto. Esto quería decir que más recursos para la biblioteca o mayores compensaciones al profesorado no causaban necesariamente un aumento del prestigio del profesorado o viceversa.

Oromaner (1970), utilizando un enfoque similar al de Cartter, descubrió que el tamaño del departamento (estimado a través del número de profesores), la edad de los profesores, y sus antecedentes académicos se correlacionaban con programas de sociología muy prestigiosos (tal como habían sido evaluados por Cartter). Específicamente, Oromaner descubrió que los departamentos más prestigiosos de sociología se caracterizaban por un determinado tamaño del departamento, profesores relativamente jóvenes y profesores graduados en los departamentos más prestigiosos. Además, Oromaner halló que los departamentos destacados y especializados eran los más productivos en la concesión de doctorados a los estudiantes. Estos departamentos otorgaron el 45% de los 1.560 concedidos entre los años 1955 y 1964.

En otro estudio sobre departamentos de sociología, Abbott (1972) investigó variables de las instituciones y los departamentos que pudieran ser correlatos potenciales de la calidad (según las clasificaciones de ACE). Las variables institucionales examinadas fueron: a) la productividad total de investigación de la universidad, b) el número total de doctorados otorgados, y c) el número total de profesores contratados a tiempo completo. Las variables de los departamentos fueron las mismas. Los resultados del estudio indicaron que las variables institucionales estaban más correlacionadas con las clasificaciones de ACE que

con las variables de los departamentos. De acuerdo con esto Abbott (1972) llegó a la conclusión de que la clasificación de los departamentos de sociología no era tanto función de lo que pasaba en un departamento aislado como de la universidad en general.

Otro estudio sobre departamentos de sociología de Abbott y Barlow (1972) examinó inicialmente los ingresos anuales de los profesores, los volúmenes de la biblioteca, el número de profesores a tiempo completo y el valor de las edificaciones de una institución como correlatos potenciales de calidad. No obstante, Abbott y Barlow no encontraron ninguna evidencia sustancial que mostrase que estas variables estuvieran altamente correlacionadas con la calidad. En su lugar, encontraron variables sobre resultados (o funciones, como son identificadas por los investigadores), a saber, la productividad en investigación del profesorado y el número y tipo de títulos concedidos por una institución, presentaban correlacionadas más altas. De acuerdo con esto, estas funciones fueron denominadas variables intervinientes entre los recursos y el rango o prestigio de una institución.

En otro estudio sobre programas de sociología, Axelson (1960) intentó identificar los correlatos de 65 programas doctorales. Una notable diferencia de este estudio con los anteriores fue que el investigador no confió en las clasificaciones de Cartter para identificar programas muy prestigiosos. En su lugar, clasificó los programas según el número medio de publicaciones en la *American Sociological Review*, *American Journal of Sociology* y *Social Forces*. Utilizando sus propias clasificaciones como variable dependiente, Axelson encontró que los programas más productivos se caracterizaban por profesores que fueron en su mayor parte antiguos alumnos de grandes instituciones mayores (especialmente de las universidades del Medio Oeste) mientras que los programas con menor producción se caracterizaban por profesores que habían sido graduados en instituciones más pequeñas (por ejemplo, en universidades católicas privadas). Evidentemente, el origen de la preparación de los graduados en sociología era un buen predictor de la productividad en investigación.

Los estudios discutidos hasta ahora se han centrado en programas de sociología, también hubo algunos centrados en otras disciplinas. Un estudio de Elton y Rodgers (1971), por ejemplo, examinó los correlatos de los departamentos de física. Las variables examinadas como correlatos potenciales de las evaluaciones de Cartter de los departamentos de física fueron: 1) el número de áreas de especialización en un departa-

mento; 2) el número de profesores; 3) el número de doctorados concedidos entre 1960 y 1964, 4) el número de estudiantes a tiempo completo; 5) el número de estudiantes de primer año y 6) el ratio de estudiantes de tiempo parcial/estudiantes de tiempo completo. Un análisis discriminante múltiple utilizando estas variables reveló que colectivamente estas eran correlatos muy buenos de las clasificaciones de ACE (r^2 era del 75%). Específicamente, estas variables eran más eficientes prediciendo los departamentos clasificados por Cartter como muy aceptables y menos eficientes en la predicción de aquellos que habían sido clasificados como menos aceptables. Globalmente, los investigadores señalaron que estas «variables objetivas proporcionan una buena aproximación a las clasificaciones numéricas de los programas de postgrado establecidas en el estudio de Cartter» (Elton y Rodgers, 1971, p. 568).

Elton y Rose (1972) realizaron un seguimiento de ese estudio un año más tarde. En su trabajo, estos investigadores examinaron los departamentos de psicología con las mismas variables predictoras utilizadas en su anterior estudio. Sus resultados revelaron que los departamentos clasificados por Cartter como muy atractivos se caracterizaban por un número relativamente grande de profesores, un número elevado de doctorados concedidos, y una matrícula alta de estudiantes a tiempo completo; los departamentos atractivos (en segundo lugar) se caracterizaban por departamentos con rendimientos relativamente altos en las seis variables; los departamentos muy aceptables se distinguían por un número elevado de áreas de especialización; y por último los departamentos que eran menos aceptables (esta categoría de departamentos fue creada por Elton y Rose para incluir los departamentos no categorizados por Cartter) se caracterizaban por una elevada ratio entre estudiantes a tiempo parcial y a tiempo completo. Los investigadores llegaron a la conclusión a partir de estos resultados de que «las clasificaciones de Carter de 1966 se pueden haber basado más en el efecto halo derivado de factores de tamaño que en los criterios establecidos en las instrucciones de Cartter, como la accesibilidad de los profesores...». Añadieron que Cartter había igualado demasiado pronto la cantidad a la calidad. Elton y Rose señalaron que «aunque la cantidad y la calidad en educación no tienen por qué ser mutuamente excluyentes, puede ser imprudente... concluir que donde haya calidad si no hay un gran tamaño». (1972; p. 200).

Quizás el mejor ejemplo de un estudio sobre correlatos sea el informe de Hagstrom (1971). En este informe, Hagstrom intentó identificar las

variables relacionadas con la calidad a través de correlaciones estadísticas y análisis de regresión lineal múltiple con las estimaciones de Carter, usando una muestra de 125 departamentos de matemáticas, física, química y biología. Las variables independientes utilizadas en el estudio se agruparon en seis áreas: (1) tamaño de departamento, (2) producción en investigación, (3) facilidades para investigar, (4) características del profesorado, (5) características de los alumnos y (6) premios y cargos del profesorado. El tamaño del departamento fue medido por el número de doctores entre el profesorado. Para medir la producción en la investigación se utilizaron cinco variables: a) la media de artículos de investigación de los profesores durante 1961-66; b) la media de las de la investigación publicada en 1966, c) el número de artículos en revistas, d) el número de libros de texto en su carrera y e) el número de libros en su carrera. Las facilidades para investigar se estimaron a través de estas tres variables: a) la media de tiempo que se emplea en la investigación, b) la facilidad en obtener ayuda externa para la investigación y c) el porcentaje de profesores con ayudas a la investigación. Las características del profesorado, la cuarta área, se estimó a partir de: a) la calidad de los departamentos en los que se doctoraron los profesores y b) la selectividad del *college* donde se había graduado el profesorado. Para estimar las características de los estudiantes se emplearon cuatro variables que incluían a) la media de becarios en el departamento; b) la selectividad a nivel de pregrado de toda la institución; c) la media de ayudantes de investigación en el postgrado y d) la media de técnicos trabajando para el profesorado. Por último, el área de premios y cargos fue medida por a) la puntuación máxima por premios recibidos por los profesores en un departamento; b) la puntuación máxima por el número de participantes en comisiones de gobierno; y c) el porcentaje de profesores con cargos en sociedades o editoriales.

Hagstrom por medio de regresión lineal múltiple, halló que los mejores correlatos de calidad eran variables relacionados con la producción de investigación ($r^2 = 0,542$), especialmente dos variables, la media de artículos de investigación y la media de citas de las publicaciones. Las variables que presentaban a continuación mejores correlaciones pertenecían al bloque de características de los estudiantes ($r^2 = 0,542$), seguido por el bloque de premios y cargos ($r^2 = 0,436$), el bloque de facilidades para la investigación ($r^2 = 0,356$), y por último el tamaño del departamento ($r^2 = 0,323$). Aunque el tamaño del departamento era lo último, explicaba casi un tercio de la varianza del prestigio alcanzado por el departamento. Esto preocupó a Hagstrom. En un análisis poste-

rior, controló el tamaño del departamento y encontró que no afectaba a los efectos de las otras variables. Por ello, Hagstrom sugirió que «el tamaño es una condición casi necesaria para la excelencia de establecimientos científicos modernos. El tamaño grande permite que se combine la amplitud con la especialización esencial para trabajar en frentes, de investigación de rápido desarrollo». Hagstrom controló también las demás variables, señalando que los resultados de todas las variables, a excepción del bloque de oportunidades de investigación (especialmente una variable), estaban altamente correlacionadas con el prestigio del departamento. La única variable con una correlación parcial negativa era el tiempo medio utilizado en investigación. Hagstrom especuló que una razón posible de ello era que los profesores de departamentos menos prestigiosos no habían tenido que abordar una sobrecarga tan grande de trabajo administrativo como los profesores de departamentos prestigiosos.

Otro estudio que identificó correlatos de calidad fue el del *National Science Board* (1969). En este trabajo se identificaron las siguientes variables correlacionadas con las clasificaciones de Cartter.

1. Magnitud del programa doctoral (número de títulos concedidos).
2. Recursos economías estatales para la investigación y el desarrollo.
3. Ingresos no estatales destinados a la educación superior.
4. Centro donde se licenciaron los becarios del postgrado (becas NSF).
5. Centro donde se graduaron los doctorandos.
6. Selectividad para entrar a la universidad.
7. Selección de instituciones por los becarios de postgrado (becas NSF).
8. Estudiantes posdoctorales en ciencia e ingeniería.
9. Doctorados concedidos por profesor.
10. Número de doctorados concedidos por estudiante.

11. Ratio de doctorados/graduados.
12. Compensaciones a profesores titulares.
13. La proporción de profesores titulares entre el profesorado.
14. Ratio estudiante/profesor en el postgrado.
15. Departamento con siete o más profesores.

Significativamente, cuando el National Science Board (1969) examinó el número de doctorados concedidos por profesor y por estudiantes matriculados, halló que las instituciones muy prestigiosas habían producido el mayor número de doctorados por profesor y por estudiante.

Otro estudio que identificó los correlatos con las clasificaciones de Cartter fue el informe de Morgan, Kearney y Regens (1976). Este estudio investigó departamentos de postgrado en humanidades, ciencias sociales, ciencias naturales e ingeniería. Las variables independientes utilizados en el trabajo fueron 1) ingresos por estudiante durante 1966-67; 2) sueldo de los profesores; 3) ratio estudiante profesor (1969-70) y 4) volúmenes en la biblioteca. Utilizando el análisis de regresión múltiple, los investigadores hallaron que las 4 variables explicaban el 80% de la varianza de la calidad del profesorado de postgrado. El número de volúmenes en la biblioteca fue el predictor más fuerte de los 4, seguido por los ingresos por estudiante, el ratio estudiante-profesor y el sueldo de los profesores.

Otros investigadores que realizaron trabajos sobre variables correlacionadas con la calidad de las instituciones fueron los siguientes:

Jones (1969) y Lavender, Mathers y Pease (1971), hallaron que el ratio estudiante profesor se correlacionaba con las estimaciones de prestigio de Cartter;

Crane (1970) y Shichor (1970), encontraron igualmente que determinados aspectos de la movilidad y empleo del profesorado también estaban correlacionados;

Knuden y Vaughan (1969), Drew (1975) y Guba y Clark (1978); reafirmaron que la productividad de investigación era un buen correlato;

Drew y Karpf (1981), que hallaron que la ratio de publicaciones del departamento era un buen indicador de la calidad;

Solmon y Astin (1981), comprobaron que el tamaño, el prestigio, la selectividad, los gastos por estudiante y la concentración curricular de la institución estaban correlacionadas con la calidad de los programas de pregrado.

Perkins y Snell (1962), Jordan (1963), Cartter (1966), hallaron que la dotación de la biblioteca (medida por los gastos y sueldos de los bibliotecarios) era un importante correlato de la calidad.

Greg y Sims (1972), averiguaron que algunas características de los estudiantes, especialmente la calidad de los estudiantes y graduados, eran variables importantes de la calidad de programas profesionales en administración educativa.

El estudio sobre correlaciones más reciente es un estudio de Conrad y Blackburn (1986) que examinó las variables relacionadas con la calidad de un departamento a nivel de cursos master y doctorado en *colleges* y universidades regionales. Se analizaron 45 departamentos en 14 universidades estatales en 5 disciplinas (matemáticas, biología, educación, historia y química). La variable dependiente de este estudio no se basaba en el sistema de reputación, sino en revisiones de los programas obtenidos por varios equipos de evaluadores externos. Basándose en dichas revisiones, los programas fueron clasificados en una escala de 5 puntos por otros 3 expertos. Esta clasificación se usó como variable dependiente. Los 73 variables independientes utilizados en el estudio se agruparon en 5 áreas: 1) el profesorado; 2) los estudiantes; 3) el programa; 4) las instalaciones y 5) los medios de apoyo. Los hallazgos globales que indican que las medidas individuales y combinadas del profesorado (la productividad en la investigación, las ayudas recibidas, la edad y el estatus, el lugar donde obtuvo el título más alto, y horas de docencia), de los estudiantes (número y habilidad), del programa (proporción de programas a nivel de postgrado y concentración curricular) y las instalaciones (tamaño de biblioteca) eran las variables que más correlacionaban con la calidad del departamento. Algunas de estas variables no se habían identificado como correlatos en los estudios previos. Por tanto, Conrad y Blackburn indicaron que era prematuro que algunos investigadores (como Hagstrom, 1971) supusieran que todos los correlatos de la calidad habían sido identificados o examinados ade-

cuadramente, sobre todo porque las instituciones menos conocidas no se habían investigado. Por último, estos investigadores llegaron a la conclusión de que «los factores relacionados con la calidad de los departamentos son más multidimensionales en *colleges* y universidades regionales que en los departamentos de las universidades de élite».

Crítica a los estudios sobre correlatos

Los estudios sobre correlatos han tenido éxito al identificar las principales variables relacionadas con el prestigio, especialmente de los programas de postgrado en instituciones muy conocidas. Este éxito, según Conrad y Blackburn (1985), no significa necesariamente que investigaciones futuras que pretendan identificar correlatos de calidad a nivel de postgrado no puedan añadir nada nuevo. Por el contrario, si se consideran las limitaciones metodológicas de este tipo de estudios, se puede hallar nuevos conocimientos.

La primera limitación de la mayoría de los estudios sobre correlatos es que dependen demasiado de las clasificaciones del ACE. Por esta razón, muchos de estos trabajos están sujetos a todas las limitaciones relacionadas con las evaluaciones del ACE. Se puede mejorar la situación usando clasificaciones más recientes y mejores, como las desarrolladas por la *Conference Board* (1982), Conrad y Blackburn (1986) o estimaciones recientes obtenidas mediante variables objetivas.

La segunda limitación es que la mayoría de los investigadores sobre correlatos no han usado un enfoque teórico para identificar las variables relacionadas con la calidad. «Parecen haber revuelto en sus datos en busca de factores que puedan estar ligados empíricamente a la calidad, en lugar de identificar variables potenciales de acuerdo con una teoría de la calidad» (Conrad y Blackburn, 1985). Por tanto, se sugiere que se desarrolle una teoría de la calidad para proporcionar un marco de interpretación de las correlaciones obtenidas en muchos de estos estudios.

Otra limitación de este tipo de trabajos es su énfasis exclusivo sobre programas de postgrado. Los estudios futuros deberían centrarse sobre correlatos de calidad en programas profesionales y de pregrado, dado que apenas existe información sobre los mismos (Conrad y Blackburn, 1985).

A pesar de estas críticas, estos estudios han contribuido a nuestra comprensión de la calidad de programas. Como señalaron Conrad y Blackburn, aunque se ha criticado a los investigadores por no relacionar sus resultados con un andamiaje teórico, se han encontrado correlaciones que tienen fuerte consistencia para identificar factores relacionados con la calidad, y que tienen importantes implicaciones para la planificación de las instituciones. Además, si se consideran las limitaciones metodológicas asociadas a los trabajos realizados, este tipo de estudio puede generar más conocimientos en el futuro.

Conclusión

Los tres tipos de estudio sobre evaluación de la calidad han contribuido claramente a nuestra comprensión de este concepto, pero aún quedan por resolver cuatro asuntos controvertidos.

Primero, nadie sabe lo que están midiendo los estudios sobre reputación ¿prestigio o calidad? Algunos investigadores han afirmado que aunque sus estudios pretendían medir el prestigio de los programas, la calidad se medía también porque ambas variables están relacionadas. Los críticos han argumentado que las estimaciones del prestigio no eran más que una competición entre la popularidad basada en la percepción de los evaluadores de la fama del departamento y su familiaridad con los programas que tenían que evaluar. Muchos han indicado que los estudios de reputación no valoran la calidad. Además, se ha dicho que padecen serios problemas metodológicos, tales como sesgos relacionados con los evaluadores y los antiguos alumnos. Si lo que se valora en los estudios de reputación es el prestigio y no la calidad, deberíamos buscar un enfoque alternativo, como el uso de variables objetivos.

La utilización de variables objetivas para evaluar la calidad es un camino más prometedor, pero no ha tenido mucho éxito. El mayor problema ha sido suponer que la productividad de la investigación de los profesores es el principal indicador de calidad. Dado que existen pocas evidencias empíricas que apoyen la relación directa entre productividad en la investigación y calidad, algunos profesores han indicado que no debería ser el principal indicador de la calidad de un departamento. Otras variables igualmente importante, como la habilidad del estudiante, los recursos financieros y la programación curricular, deberían utilizarse como criterios. Suponer que la productividad en la investigación es el principal criterio de calidad, ¿no significa que las instituciones que no

ponen la investigación por encima de la docencia como sucede en los *colleges* de artes liberales y *colleges* locales no son instituciones de calidad? Este no es el caso, evidentemente. Claramente, la mejor forma de evaluar la calidad es el uso de variables múltiples. Pero se ha obtenido poco éxito. Parte del problema es que hay poca teoría que oriente a los investigadores en la selección de una combinación «correcta» de variables para medir la calidad. Esto continuará siendo un problema mientras no tengamos una teoría adecuada de la calidad. Esta situación de la gallina y el huevo no debería, sin embargo, desanimarnos en nuestros intentos de desarrollar una teoría de la calidad.

Tercero, nadie sabe cuál debería ser la definición de calidad. ¿Debería definirse por características de la calidad relativas a las instituciones de élite o a las que no son de élite? Si las características de la calidad e instituciones de élite pueden ser diferentes a las de las instituciones «normales» y viceversa, ¿cuáles deberíamos escoger como modelos en nuestros esfuerzos para mejorar la calidad? Quizás necesitemos establecer una serie de determinantes de la calidad, aplicables a tipos concretos de instituciones.

Por último, aunque la mayoría de los estudios han tenido éxito en identificar una variedad de correlatos relacionados con la calidad, ninguno ha examinado adecuadamente la relación entre calidad y el desarrollo educativo de los estudiantes. Astin y Solmon (1981), lo expresaron claramente: «¿Se desarrollan de forma diferente los estudiantes de instituciones evaluadas como excelentes que los de instituciones clasificadas en posiciones medias o bajas?». Esta pregunta sobre el «valor añadido», usando la terminología de Astin y Solmon, podría responderse mediante estudios longitudinales destinados a examinar el grado en el que varía el desarrollo del estudiante según los diferentes tipos de instituciones. La hipótesis es que las «instituciones excelentes serían aquellas que pueden influir significativamente en el desarrollo intelectual y académico de los estudiantes (y de los profesores también). Este punto de vista sobre el «valor añadido» o del ahora llamado «desarrollo del talento» ha sido defendido por Astin (1985) en su nuevo libro, *Achieving Educational Excellence*.

Finalmente, si se pueden resolver estas cuatro cuestiones y la mayoría de los problemas metodológicas de estos estudios, puede que lleguemos a una definición operativa de calidad, una definición que pueda guiar a aquellos de nosotros comprometidos en mejorar la calidad de nuestras instituciones.

SOBRE LA COMUNICACION DE LA CALIDAD PARA LA MEJORA DE LA INVESTIGACION Y EDUCACION SUPERIOR EN LOS PAISES BAJOS *

ARNOLD. A. J. SPEE
Ministerio de Educación y Ciencia, Países Bajos

1. Introducción

En Holanda se da mucha importancia a la calidad de la enseñanza superior. A principios de año, los examinadores de la OCDE (1990) exponían en el *Análisis sobre la política educativa en los Países Bajos* su conclusión de que la calidad de la educación holandesa, incluyendo la enseñanza superior, es excelente: «de hecho, con la convicción de que los holandeses son quizá demasiado críticos con su propio sistema, ha dado muy buenos resultados, tanto a corto como a largo plazo».

También hay, sin embargo, opiniones negativas sobre la calidad de la enseñanza superior. El Consejo Asesor de Enseñanza Superior (ARHO), en su informe publicado en diciembre de 1989, dudaba que Holanda pudiera proporcionar una educación e investigación óptima y aplicar las opciones que se crean en el proceso de unificación europea. El Consejo concluyó, tras realizar un sondeo entre expertos, que la imagen de los licenciados y diplomados universitarios no es totalmente positiva.

La Organización de los Países Bajos para la Cooperación Internacional en la Enseñanza Superior (NUFFIC, 1990) también expresa su inquietud sobre la evaluación de los diplomas holandeses de enseñanza superior en otros países. Según NUFFIC, el problema obedece más a una infravaloración que a una calidad deficiente.

* El presente trabajo es traducción íntegra de la versión revisada del autor del trabajo presentado por el autor en la 12th Annual European A.I.R. Forum on *Quality and Communication for Improvement*. Lyon, France, September, 9-12, 1990.

Desde esas autorizadas conclusiones se infiere que la calidad de la enseñanza e investigación holandesa está en entredicho. En este informe trataremos de responder a tres cuestiones principales:

1. Si en el marco de las modificaciones realizadas en el sistema de control del gobierno las instituciones educativas pueden adoptar las medidas necesarias para proporcionar una enseñanza y una investigación de calidad.

2. Si el proceso de evaluación de la enseñanza y la investigación es adecuado para elaborar juicios de calidad comparativos.

3. Si se informa con la suficiente claridad sobre la calidad de la enseñanza y la investigación al gobierno, a las instituciones, a los estudiantes y al público en general para poder mejorar la enseñanza y la investigación.

Estas cuestiones se examinan en los tres apartados que figuran a continuación. Primero se trata sobre la política de la enseñanza superior holandesa, y los aspectos a los que se da más importancia. A continuación se describe el proceso de evaluación de la calidad, y en tercer lugar se analiza la comunicación del resultado, según el modelo SMCRE (Fuente, mensaje, canal, receptor y efecto).

1. *Nuevos avances en la política de la Enseñanza Superior*

Desde 1985, con la publicación del informe Verde *Enseñanza Superior: autonomía y calidad*, el gobierno busca un sistema de control que optimice la autorregulación en la enseñanza superior y en la investigación. A continuación se exponen algunos puntos de esa tendencia:

- la abolición de regulaciones *ex ante* no específicas;
- el gobierno intervendrá para contrarrestar evoluciones indeseables sólo si los efectos de la autonomía de las instituciones son inaceptables;
- la intervención del gobierno consistirá principalmente en corregir las deficiencias que se hayan producido;

- los incentivos del gobierno tienen la finalidad de favorecer la creación de procedimientos de evaluación y asignación en el sistema para mejorar la rendición de cuentas.

Esos factores se especifican en el Proyecto de Ley sobre la Enseñanza Superior y la Investigación, presentado al parlamento en marzo de 1989 y que se espera entre en vigor en agosto de 1992.

En este párrafo se discutirán los principales puntos del Proyecto de Ley. El sistema se regula mediante un ciclo de planificación que consiste en un plan de desarrollo gubernamental de dos años (El Plan de Enseñanza Superior e Investigación, HERP) e informes anuales de las instituciones de enseñanza superior e investigación. Para que las instituciones tengan más autonomía tendrán más capacidad en la toma de decisiones sobre la enseñanza y la investigación, la admisión de estudiantes y la asignación interna de recursos financieros. Este incremento de autonomía requiere un detallado proceso de evaluación de la calidad, para posibilitar la rendición de cuentas del desarrollo del sistema de enseñanza superior e investigación.

1.1. Planificación de la enseñanza superior y la investigación académica

Los principales objetivos de los sistemas de enseñanza superior e investigación son:

- satisfacer las necesidades individuales en la enseñanza superior;
- responder a la demanda social de titulados, tanto en el aspecto cualitativo como en el cuantitativo;
- proporcionar el suficiente conocimiento científico, sobre todo en los casos en que de modo general la sociedad carece de poder adquisitivo.

La función del gobierno consiste en crear el marco que permita a las instituciones ofrecer la enseñanza y la investigación necesarias. Aparte de eso, el poder del gobierno se limita a:

- determinar la cuantía de los recursos financieros por institución;
- restringir el número de plazas para estudiantes cuando haya un

sustancial desequilibrio entre la demanda de los estudiantes y las necesidades de la sociedad;

- establecer las principales características del proceso de evaluación de la enseñanza y la investigación, que pondrán en práctica las instituciones.

El gobierno establecerá su política referente a la enseñanza superior y la investigación académica en un plan de desarrollo bienal, el Plan de Enseñanza Superior e Investigación (HERP), que se debatirá en el Parlamento y con las instituciones. Las instituciones educativas responden al plan en un párrafo de su informe anual.

1.2. *Enseñanza y admisión*

El Proyecto de Ley articula la formación de un conjunto variado y flexible de estudios. No se restringirá el desarrollo de los cursos. Los requerimientos sólo los puede establecer la ley (por ejemplo, abogados), regulaciones europeas (por ejemplo, médicos) u organizaciones profesionales (por ejemplo, contables).

Cada institución de enseñanza superior puede determinar el número de estudiantes que admite anualmente, que no puede ser inferior al 75 % del promedio de admisiones de los últimos años. Sólo en los estudios en que hay escasez de plazas, el número mínimo de admitidos debe ser el 125 % del promedio de admisiones en los últimos tres años.

En situaciones extremas, como por ejemplo si se registra un significativo desequilibrio entre la demanda estudiantil y las necesidades de la sociedad, se podrá imponer una limitación gubernamental de admisión de estudiantes, un *numerus clausus*. En ese caso se realizará un sorteo entre los estudiantes que aspiren a ingresar y que, en función de sus resultados académicos en la enseñanza secundaria, puedan optar a la obtención de una plaza por sorteo. El incremento de la participación de grupos minoritarios se asegura al dar a los estudiantes aspirantes mayores posibilidades en el sorteo.

Para facilitar la diferenciación entre las instituciones, se podrá exigir tener aprobadas un máximo de dos asignaturas específicas de la enseñanza secundaria para entrar en un curso. Los centros también podrán

determinar las tasas académicas, con un máximo establecido por el gobierno. El máximo es de 1.750 florines para el curso 1990/91, que se incrementará en 100 fl. anuales, hasta 2.250 fl. en 1995/96.

1.3. *Financiamiento*

El financiamiento de las instituciones tendrá un carácter más global que ahora. El número de *costes* por estudiante (la financiación basada en un promedio de personal docente por estudiante y un salario promedio por equipo docente), que ahora son cinco, probablemente se reducirá a dos. Se eliminará el depósito de fondos por departamento, independientemente del número de estudiantes.

Dado que las instituciones podrán establecer el importe de las matrículas (iguales o inferiores al máximo establecido), éstas se considerarán ingresos de los centros. El *coste* por estudiante se reducirá con la cuota máxima. Esto significa un cambio, pues ahora las instituciones actúan como meros recaudadores de fondos para el gobierno, y las cuotas no se tienen en cuenta en la financiación estatal.

Las subvenciones por estudiante se pagarán según su «rendimiento», en el que de momento sólo se hacen dos distinciones: los que obtienen el título y los que abandonan. Para posibilitar los cursos cortos y los diplomas parciales, la subvención se diferenciará en «años de estudiante», y se determinará por la cantidad de créditos que un estudiante obtiene al año.

1.4. *Investigación*

El objetivo del Proyecto de Ley es mejorar la calidad de la investigación académica. Mediante el sistema de Subvención Condicional, puesto en marcha en 1982, se ha evaluado una parte considerable de los programas de investigación. Este sistema se extenderá a todos los programas financiados con la *primera partida de fondos* (recursos económicos que el gobierno destina directamente a la universidad).

Para incrementar la calidad y la adecuación social de la investigación académica, parte de la primera partida de fondos se canalizará a través de los Consejos de Investigación (segunda partida de sub-

venciones). Esto permitirá establecer prioridades en la investigación académica nacional y, en algunos casos, internacional.

Para propiciar una excelente investigación académica se crearán escuelas graduadas. En la actualidad una comisión consultiva analiza las posibilidades de establecer esos institutos para ciertos campos de la investigación. Esos institutos estarán probablemente regidos por las autoridades universitarias.

Esas escuelas también se convertirán en institutos de formación para jóvenes investigadores con talento. Esos investigadores, una vez doctorados, trabajarán primero como asistentes de investigación y luego tendrán la oportunidad de acceder a un puesto permanente en la investigación.

1.5. *Evaluación de la calidad*

La clave del sistema de control son los procedimientos de valoración de la calidad. En el Proyecto de Ley se regulan diversos aspectos de la calidad. Cada institución designará un comité de exámenes, responsable de los estándares en el examen final. El Proyecto de Ley también establece los requisitos de entrada, las normas sobre los programas, el año propedéutico (primero) y la publicación de los cursos que se ofrecen. Asimismo, traza las directrices de las cualificaciones que se necesitan en ciertas profesiones.

La evaluación de la calidad se garantiza mediante la obligación de establecer comités visitantes para la enseñanza y equipos externos de evaluación para la investigación. Cada institución debe articular unas normas para la valoración de la calidad. La Oficina de Inspección de la Enseñanza Superior tiene la función de recomendar las técnicas y los puntos de vista aplicados por los comités de visita, y también puede ejercer un control sobre esos aspectos.

El registro central de cursos, introducido por la Ley de Educación Superior e Investigación, será de primordial importancia. El registro es una base de datos que incluye los cursos por universidad y por escuela técnica financiados por el gobierno, así como los programas o diplomas reconocidos y en los que los estudiantes puede recibir ayudas. Se indicará el coste por estudiante.

En principio se registrarán los cursos subvencionados hasta el momento. En caso de una repetida evaluación negativa de un curso, se podrá retirar del registro. En ese caso la subvención cesará y los estudiantes dejarán de recibir ayuda financiera.

2. El sistema de control de la calidad

A lo largo de los últimos años el Ministerio de Educación y Ciencia ha jugado un importante papel en la elaboración de una política para crear un sistema de evaluación de la calidad en la enseñanza superior (Paardekooper y Spee, 1990). Pero no ha actuado sólo: los administradores de las universidades y escuelas técnicas, sean independientemente, sea mediante sus asociaciones (la Asociación para la Cooperación de las Universidades Holandesas, VSNU, y el Consejo para las Escuelas Técnicas, HBO-Raad), también han participado en lo que se conoce como «desarrollo del proceso». Los consejos de investigación, la Organización Holandesa para la Investigación Científica (NWO), y la Real Academia de Artes y Ciencias (KNAW) han tomado igualmente parte en los debates, así como las organizaciones estudiantiles, profesionales y los grupos de interés. Finalmente hay que señalar asimismo el papel del Inspectorado de la Educación Superior.

El proceso el control de calidad partió de unos acuerdos entre el gobierno y las instituciones. Se alcanzó un acuerdo, por ejemplo, sobre la división de las responsabilidades entre los diferentes protagonistas, sobre quién hace el qué cómo y cuándo. También se llegó a un acuerdo sobre la evaluación del control de calidad, área en la que el Inspectorado de la Educación Superior y el Ministerio deben definir sus responsabilidades, por ser de su competencia.

Uno de los resultados más interesantes obtenidos al tratar de elaborar métodos de supervisión y mejora de la calidad de la enseñanza superior consiste en que el gobierno, casi por definición, queda relegado a una función secundaria, que consiste en controlar si se realiza la evaluación, cómo se lleva a cabo y cómo se aprovechan sus resultados. Es fácil comprender por qué Neave (1988), en la tipología de funciones del gobierno, describió la actual situación como la del *estado evaluador*. A medida que los mecanismos de evaluación y retroalimentación comienzan a funcionar con eficacia en los sistemas, habrá cada vez menos necesidad de que el gobierno intervenga directamente.

Según el principio de que las propias instituciones son las principales responsables del control de calidad, la política se basa siempre en la premisa de que las instituciones deben organizar y designar su propio sistema de control de calidad. Esto se considera como un modo de incrementar las posibilidades de éxito desde el principio. VSNU y HBO-Raad han asumido el papel de coordinadores de este proceso. El Ministerio y el Inspectorado de la Educación Superior participan en las discusiones, hacen sugerencias, establecen condiciones, formulan requerimientos y juzgan propuestas. Debido a la falta de experiencia en control de calidad que hay en Holanda, el sistema se puso en marcha con la idea de aprender con la práctica más que aplicando un sistema concreto. A medida que adquirimos experiencia y aprendemos de ella, el sistema se perfeccionará e institucionalizará gradualmente.

2.1. *Evaluación de la calidad de la investigación académica*

Desde 1982, el sistema denominado Financiación Condicional evalúa la calidad de los programas de investigación universitarios. Este sistema no limita la libertad de cada universidad de establecer sus prioridades y seleccionar proyectos de investigación, sino que, por el contrario, les concede la máxima responsabilidad, pues, al fijar sus prioridades y elegir, pueden establecer su propia identidad en términos de política de investigación. Pero, por otro lado, las universidades deben rendir las debidas cuentas de la política que siguen en sus planes y en su rendición de cuentas anual. Las conclusiones que extraen de la evaluación de su investigación son de particular importancia.

El Ministerio no evalúa el contenido de la investigación, sino que deja esto en manos de los equipos de evaluación independientes y externos, que basan sus juicios en la calidad de la investigación y del equipo de investigación, en la significación académica y, si procede, social de los descubrimientos científicos. Esos equipos dan cuenta de los métodos de evaluación utilizados y sus conclusiones se publican en informes.

Entre 1987 y 1989 se han evaluado más de 1.000 programas *ex post*. La evaluación se realiza por (grupos de) disciplina(s), lo cual significa que todos los programas (los proyectos de investigación en conjunto) de una disciplina se evalúan el mismo año. La evaluación la organizan los consejos de investigación NWO, KNAW, el Real Institutos de Ingenieros, KIVI, y el Consejo Nacional para la Investigación Agrícola, NRLO. El

92% de los programas se estimó satisfactorio, el 6 % insatisfactorio y sobre un 2 % no hubo consenso (véase Apéndice A).

Como se ha dicho anteriormente, en su rendición de cuentas anual las universidades explican las medidas tomadas respecto a los programas estimados negativamente. De 343 programas, 15 fueron considerados insatisfactorios en 1987, y sobre 3 no se pronunciaron. De esos 18 programas, 11 estaban acabados, 1 integrado con otro, y el personal y las inversiones de 5 programas habían ascendido. Dos de los 5 programas que habían mejorado de calidad luego fueron evaluados positivamente.

VSNU establece algunas conclusiones:

- la evaluación *ex post* se debe seguir realizando por equipos externos;
- la evaluación se debe hacer extensiva a todas las investigaciones financiadas por la primera partida;
- las combinaciones de disciplinas en relación con los equipos de evaluación externos se deben mantener igual;
- el plazo de tiempo debe ser de cinco años, en vez de ser de tres años más dos de *descanso*;
- aparte de algunos proyectos piloto, no hay punto de vista integrado sobre la evaluación de la investigación y la enseñanza;
- se pueden fijar tres categorías de evaluación añadiendo una categoría muy buena (*dobles plus*) a la positiva y la negativa;
- se puede reconsiderar si la evaluación debe realizarse sobre el proyecto o sobre el programa;
- en la evaluación de la financiación condicional se puede tratar de incluir la investigación en los institutos;
- la conveniencia de prestar más atención a los componentes internacionales de los equipos de evaluación externos;
- es conveniente reconsiderar la división de las disciplinas en el consejo de investigación;

En general, el procedimiento de evaluación parece ser satisfactorio, aunque quizá proceda realizar pequeños cambios, como por ejemplo el agrupamiento de la evaluación para pequeños campos de investigación y la identificación de investigaciones punteras por comités de evaluación externa.

2.2. *Evaluación de la calidad de la enseñanza*

El VSNU organiza las evaluaciones de la calidad de las universidades, en las que destacan los siguientes puntos:

- el autoanálisis es el pilar del proceso, que complementa las evaluaciones de calidad internas;
- el comité de visitas, compuesto por expertos, es el elemento más importante;
- cada disciplina se considera por separado la evaluación no se realiza por institución, sino por campo de estudio y por curso;
- es un proceso a escala nacional, lo cual significa que visitarán todas las facultades del mismo campo académico y se dispondrá de estudios comparativos de varios años;
- el proceso es periódico, pues cada área de estudio se evaluará una vez cada seis años;
- el proceso cubre los 25 campos de estudio y los 150 cursos existentes;
- el proceso se limita a la enseñanza, aunque más tarde se puede integrar en la investigación;
- el público tendrá acceso a los informes de los comités.

Las tareas y métodos de los comités de visita, así como los criterios de autoevaluación por parte de las instituciones, se establecen en un protocolo. El VSNU establece el programa de visitas.

En 1988 se realizó una evaluación piloto de historia, ciencias naturales y astronomía, psicología e ingeniería civil.

En 1989 se contaba con una detallada evaluación de geografía, matemáticas y tecnología de la información, lenguas no occidentales, diseño industrial e ingeniería aeroespacial. En 1990 han sido evaluables: biología, derecho, económicas, filosofía e ingeniería eléctrica. En 1991 se investigarán: medicina, lenguas occidentales, ciencias políticas y administración pública, teología y geología. A finales de 1991 habrá sido evaluado el 50 % de todos los programas graduados de las universidades holandesas.

Los comités estaban compuestos por eméritos (30 %), profesores (15 %), especialistas en educación (5 %), académicos extranjeros (25 %), y representantes de organizaciones profesionales y patronos (25 %). En cuatro casos un académico extranjero actuó de presidente del comité.

Los comités basan su evaluación, en primer lugar, en las autoevaluaciones realizadas por las universidades. Los comités consideraron que, en general, estas autoevaluaciones eran útiles, aunque algunas no fueron muy analíticas y algunas facultades veían el mundo, según el comité, con *lentes de color de rosa*. Por otro lado, el *libro de notas*, que recoge datos sobre los estudiantes, el personal, el porcentaje estudiantes/personal docente, los porcentajes de titulados, etc., es bastante incompleto y no facilita la comparación. En consecuencia, en el futuro estos comités también evaluarán las autoevaluaciones.

En general, las visitas de los comités estuvieron bien organizadas y no presentaron ningún problema. Se prestó mucha importancia a la recogida de información de los estudiantes: los comités organizaron reuniones con estudiantes, hablaron con ellos como miembros de comités de facultad y en algunos casos los estudiantes dieron sus opiniones por escrito.

En el informe de los comités suele figurar una descripción del método de evaluación, de los procedimientos de evaluación interna, los porcentajes de titulación, la enseñanza, el contexto internacional y las necesidades del mercado laboral. Los comités emiten su opinión sobre las deficiencias y los puntos fuertes de las facultades y dan las recomendaciones que juzgan pertinentes. En un comité las conclusiones no fueron unánimes y en el informe se incluye el punto de vista minoritario.

Los comités visitantes envían sus informes a las universidades participantes; pero transcurre demasiado tiempo desde que se efectúa la

visita hasta que los reciben, y su publicación también se demora en demasía. Ambos defectos deben subsanarse en el futuro.

De acuerdo con Vroeijensteijn (1991), existen algunos puntos que requieren una atención especial:

- De cara al futuro, la educación es el centro de atención. En un futuro próximo deberá existir una conexión entre la evaluación de la educación y la de la investigación.
- Está muy bien el evaluar la calidad a nivel de departamento, pero ¿qué ocurre con la contribución de los servicios centrales y de la institución en su conjunto? Se hará necesario el complementar las revisiones departamentales con una revisión de los servicios centrales y de la gestión general.
- Un problema con el que se enfrentan todos los comités es la falta de una buena información cuantitativa por parte de las facultades. En un futuro próximo debe desarrollarse un buen sistema de datos.
- ¿Cómo supervisar lo que hacen las facultades o lo que pueden hacer con los resultados? ¿Merece la pena esta intensiva inversión de tiempo?
- Debe prestarse una especial atención a la sintonización de todas las actividades en el marco de la evaluación de la calidad llevada a cabo por distintos grupos. Las facultades corren el riesgo de quedar estranguladas por causa de la atención prestada a la calidad.

3. *La comunicación de la calidad*

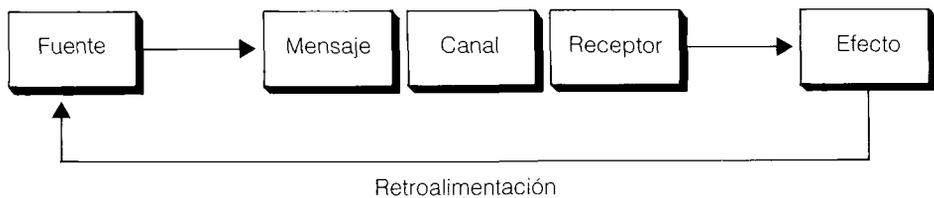
El principal objetivo del proceso de evaluación de la calidad es la autorregulación de las instituciones educativas. La autoevaluación es un medio de descubrir los aspectos positivos y negativos, las oportunidades y los riesgos. Además, tras la evaluación de profesionales y expertos ajenos al centro, las instituciones toman las medidas oportunas.

Por otra parte, se puede ejercer cierta presión externa sobre las instituciones, lo cual también induciría a mejorar la calidad. Ello se pue-

de hacer mediante la comunicación de la calidad, que implica asimismo una mejora de la rendición de cuentas del sistema de enseñanza superior e investigación.

Podemos describir el proceso de comunicación de la calidad de las universidades y escuelas técnicas holandesas mediante el modelo SMCRE (Fuente, Mensaje, Canal, Receptor, Efecto), desarrollado por Berlo (1962).

Puesto que durante el proceso de evaluación de la calidad apenas hay comunicación interpersonal, este modelo simplificado es suficiente para nuestro propósito (Oomkes, 1986):



Las fuentes de comunicación de la calidad son los equipos de evaluación externos para la investigación académica y los comités de visita para la enseñanza. El mensaje es el juicio de calidad de los equipos de evaluación externos y los comités de visita. El canal por el que se da a conocer el dictamen es el informe, del que se publicará un resumen en el registro. Los receptores del mensaje son las universidades y las escuelas universitarias (los productores de la calidad), la administración del gobierno y los inspectores (los consejeros y los responsables de la calidad) y los estudiantes y la sociedad en general (los consumidores de la calidad). Los efectos son los cambios que se realizan en la enseñanza y la investigación.

Para optimizar la comunicación se deben dar algunas condiciones relacionadas con cada elemento del SMCRE, como veremos a continuación.

3.1. *La fuente: equipos de evaluación y comités de visita*

La pertenencia y la composición de los equipos de evaluación externos y los comités de visita son importantes para evitar la selectividad en la comunicación de la calidad de la enseñanza y la investigación.

La pertenencia a un equipo de evaluación externo para la investigación se restringe, en la práctica, a académicos y colegas. Eso garantiza por un lado un alto grado de fiabilidad, que, por otro lado, es menor en la revisión de los compañeros (Goedegebuure, *et al.*, 1990). La revisión por los compañeros es un método flexible que se puede ajustar a las tradiciones en todas las disciplinas y es la única manera de evaluar directamente la calidad. Sin embargo, ocurre al igual que cuando queda algo en la caja negra, pues los juicios son resultado del proceso mental inverificable de los jueces. Esos dictámenes se deben complementar con información verificable, como, por ejemplo, con indicadores de rendimiento.

Los miembros de los comités de visita, que analizan el proceso de enseñanza y sus resultados, no son sólo especialistas en las áreas más importantes, sino también representantes de organizaciones profesionales y empresarios.

No parece haber ningún problema en designar representantes autorizados de organizaciones profesionales y empresarios. Es más difícil encontrar especialistas aceptables, pues debido al número limitado de académicos de habla holandesa, el número de expertos independientes es limitado, y, debido a la ocupación de los profesores en ejercicio, se recurre a los eméritos. Los educadores deben desempeñar una función más importante en esos comités, que en el futuro deben centrarse más en los aspectos didácticos, como el currículo, los objetivos pedagógicos, la supervisión, etc.

Los sindicatos estudiantiles han presionado para que se incluyan estudiantes en los comités de visita, argumentando que, como consumidores de la enseñanza superior, su opinión es importante. Sin embargo, teniendo en cuenta que el valor que se da al veredicto de un comité depende en gran parte de la autoridad, los estudiantes no deben ser miembros del comité puesto que su participación no incrementa su autoridad.

3.2. *El mensaje: el juicio de calidad*

Los mensajes deben ser claros y comparables. Por ello es importante el criterio con el que los equipos de evaluación externos y los comités de visita juzgan la investigación y la enseñanza.

Los equipos de evaluación externa analizan la investigación universitaria considerando las publicaciones en función de los siguientes aspectos:

- la calidad de las investigaciones, medida por las calificaciones de los investigadores y por sus publicaciones;
- la adecuación académica, la contribución a la ciencia;
- la coherencia de los programas, también en relación con la investigación realizada en otras instituciones;
- la significación social de los programas en relación con las necesidades de la sociedad.

La observación de estos aspectos contribuye en gran manera a la realización de unos dictámenes de calidad comparables, aunque hay que hacer algunas observaciones.

La evaluación es de naturaleza global, dado que se aplica a un programa (los proyectos de investigación en conjunto). Incluso si el programa se estima satisfactorio, es posible que algunas partes del programa estén por debajo del nivel estándar. Eso puede llevar a una conducta estratégica de los administradores de la universidad: integrar la investigación más defectuosa en un programa (que se espera) que vaya a ser mejor, o bien suprimir las partes más deficientes, para que todo el programa pase el control de calidad.

Otro punto es la dicotomía del juicio, positivo o negativo. De las conclusiones del VSNU se desprende que hay razones convincentes para introducir una tercera categoría de estándar muy elevado (de *double plus*). Esto satisfará a los académicos cuando consideran que está mucho mejor que los que están justos para aprobar. Y también proporcionará un criterio para redistribuir los fondos de investigación en función de la calidad.

Para la enseñanza no se establecen criterios de juicio. En el libro de las directrices del comité de visita y en el protocolo se habla profusamente sobre procedimientos, métodos, tipos de recogida de información, escuchar a ambas partes, etc.

Llama la atención un estudio de Segers, Dochy y Wijnen (1989) sobre indicadores de rendimiento. Estos investigadores entrevistaron a un

elevado número de partes interesadas: rectores universitarios, personal académico, responsables de toma de decisiones políticas, así como empresarios, organizaciones profesionales, estudiantes y padres. A continuación elaboraron una lista de 20 indicadores cuantitativos, y más de 50 indicadores cualitativos en los que había un considerable consenso entre los diferentes grupos (véase Apéndice B). Esos indicadores ofrecen una valiosa información para evaluar la calidad de la enseñanza y la investigación. Vroeijensteijn (1991) sugiere, sin embargo, que aunque estos indicadores de rendimiento no están señalados en la guía, se espera de las facultades que describan dichos indicadores en su autoevaluación, a la vez que sean objeto de diálogo con el comité.

3.3. *El canal: los informes y el registro*

El mensaje se debe enviar a los receptores por un canal *libre de interferencias*.

Los informes de los equipos de evaluación externos siempre se han dado a conocer, así como los dictámenes de la primera evaluación completa sobre la calidad universitaria realizada por los comités de visita (1989). Aunque los informes están a disposición del público en general, de hecho esas publicaciones sólo las leen los que están directamente interesados: los rectores de las instituciones de enseñanza superior, los funcionarios, etc.

Por consiguiente, el gobierno intenta facilitar la publicación de una guía del consumidor para el alumno.

Ello consistirá en una publicación regular que aporte informe sobre la calidad de los cursos, servicios para el estudiante y oportunidades profesionales. La guía del consumidor aportará información proveniente de los informes presentados por los comités visitantes, colegas internacionales, paneles de expertos y entrevistas con los estudiantes, con los licenciados, con organizaciones profesionales y empleadores.

Además del registro, se considera la posibilidad de publicar una guía de consumo para el estudiante. Será una publicación periódica con información sobre la calidad de los cursos, las posibilidades de los estudiantes y las oportunidades profesionales. La guía del consumidor incluirá información sobre los informes de comités de visita, comparacio-

nes internacionales, grupos de expertos y entrevistas con estudiantes, titulados, organizaciones profesionales y empresarios.

3.4. *Los receptores: instituciones, gobierno y consumidores*

Los comités de visita entregan sus informes a las instituciones educativas antes de su publicación, lo que permite corregir posibles errores. En los planes de desarrollo o en los informes anuales las universidades describen las acciones emprendidas siguiendo las recomendaciones de los informes.

El gobierno envía la revisión de los informes (elaborados generalmente por VSNU y HBO-Raad) al Parlamento, que puede convocar un debate sobre los resultados de la evaluación de calidad. En las consultas bilaterales y multilaterales el ministro discute los resultados con las instituciones de enseñanza superior. Puede darse el caso de que el ministro solicite a Inspección que se pronuncie sobre la calidad de un área que no se haya considerado, o que recabe otra opinión sobre la evaluación de calidad.

3.5. *El efecto: posibles cambios en la investigación y la enseñanza*

El principal criterio para evaluar el proceso de comunicación gira en torno al análisis de los cambios que produce el mensaje. Del estudio realizado por Bij Spaapen y otros se desprende que el proceso de evaluación de calidad de la investigación no ha aumentado significativamente las perspectivas de contar con una avanzada investigación y con profesores más preparados. La razón estriba en que los administradores están influidos en sus decisiones no sólo por consideraciones de política de investigación, sino también por otro tipo de preocupaciones, como el mantenimiento de sus presupuestos y de los miembros de equipo (Westerheijden, 1990a).

Esto se debe sobre todo a la naturaleza *ex post* de la evaluación, que no se pronuncia sobre la conveniencia de proseguir con los programas que se estiman satisfactorios.

Hofstee (1988) aboga por una relación directa entre el nivel del proyecto y la evaluación. Se ha tratado, aunque someramente, de solventar

este problema potenciando el papel de NWO, para obtener al mismo tiempo una evaluación del programa y una reasignación del presupuesto. Esto significaría un cambio de la situación actual, en la que la evaluación *ex post* lleva a determinar la capacidad de la investigación sin relacionarla directamente con los programas y los temas de investigación específicos.

De momento no se ha dado a conocer ningún estudio sobre los efectos de la evaluación cualitativa en la enseñanza. Un reciente trabajo de Weusthof (1989) sobre la autoevaluación, como parte del proceso de calidad, indicaba que su repercusión es muy limitada y que los rectorados de las facultades apenas utilizan los resultados de la evaluación. Sin embargo, mientras más institucionalizada está la autoevaluación, más atención se presta a los resultados. Es evidente que no se pueden extraer conclusiones generales de una sola evaluación de una parte del proceso de control de calidad, pero probablemente sea conveniente tomar algunas medidas para que los informes sobre la evaluación sean más efectivos. Como ya se ha mencionado, una de esas medidas consiste en incluir resúmenes de los trabajos en el registro de los cursos. Otra medida podría ser la de dar a los estudiantes más oportunidades de votar.

Conclusión

En la política educativa de la enseñanza superior destacan, a lo largo de los últimos años, dos características principales: autonomía y calidad.

Al decir autonomía nos referimos a que las instituciones de enseñanza superior y de investigación académica cuentan con más responsabilidad y libertad para conseguir sus objetivos educativos y de investigación.

La autonomía es una estrategia adoptada al considerar que si las instituciones pueden elegir por sí mismas y actuar en consecuencia, se realizarán mejor las diferentes funciones que hay que desempeñar en la educación y la investigación en los próximos años.

La evolución de la política educativa de la enseñanza superior —en referencia a la primera cuestión planteada en la introducción de este

trabajo— ha reportado más poder de decisión a las instituciones de enseñanza superior y de investigación, que no tendrán restricciones para realizar sus cursos. Las instituciones pueden determinar, dentro de unos amplios límites, el número de estudiantes que admiten y, en algunos casos, las calificaciones que se requieren para ingresar. Pueden fijar asimismo el importe de la matrícula por debajo del máximo establecido. La financiación por estudiante se pagará en función de los resultados académicos de los estudiantes. Al reducir a dos el número de estudiantes a los que se paga según su rendimiento, la financiación de las instituciones será aún más global.

El complemento inevitable de la autonomía es la rendición de cuentas. La clave del sistema administrativo reside en el proceso de evaluación de calidad, que proporciona, hasta cierto punto, juicios de calidad comparativos para responder a la segunda cuestión. La evaluación de la investigación se realiza según un conjunto de criterios claramente definidos, cuya aplicación posibilita que los dictámenes sean razonablemente comparables. En cualquier caso, la dicotomía de los juicios (positivo o negativo) se debe completar con una categoría muy buena o *doble plus*, que puede ser el criterio fundamental para una reasignación de fondos. También es conveniente dividir la evaluación, los programas, en varios proyectos, para que el dictamen sea más preciso. La evaluación de la enseñanza es difícilmente comparable dado que son los propios comités los que establecen los criterios. Los comités de visita deben aplicar los indicadores cualitativos —enumerados por Segers, Dochy y Wijnen (1989)— para que sus juicios sean más comparables.

La última cuestión hace referencia al efecto que tiene dar a conocer el dictamen sobre la calidad. Para que tenga más repercusión se puede incluir el resumen de los informes en el registro de los cursos y publicar los resultados en la guía del consumidor. Entonces los estudiantes podrán obtener valor (calidad) por dinero (tasas).

Se puede aumentar la repercusión de la evaluación de la investigación con la propuesta mencionada, canalizando más fondos de investigación a través de los consejos de investigación. De esta forma se relacionará la evaluación del programa con la reasignación del presupuesto. Al haber más poder de decisión fuera de las universidades, éstas tratarán de ofrecer la mejor educación e investigación posibles para aprovechar las opciones creadas con el actual proceso de unificación europea.

APENDICE A: FINANCIACION CONDICIONAL 1987-1989

Evaluación 1987

Disciplina	EAB	Número	Evaluación		
			posit.	negat.	indet.
Lengua	NWO	40	40	0	0
Literatura	NWO	55	54	0	1
Historia	NWO	31	30	1	0
Historia del arte	NWO	8	8	0	0
Biología	NWO	39	37	2	0
Química	KNAW	65	58	6	1
Bioquímica	KNAW	12	11	1	0
Biofísica	KNAW	11	11	0	0
Farmacía	KNAW	5	4	1	0
Psicología	KNAW	53	48	4	1
Ingeniería mecánica	KIVI	24	24	0	0
TOTAL		343	325	15	3

Evaluación 1988

Disciplina	EAB	Número	Evaluación		
			posit.	negat.	indet.
Teología	NWO	29	24	5	0
Medicina y odontología	NWO	162	150	3	9
Arqueología	NWO	9	9	0	0
Geología	NWO	12	12	0	0
Geografía	NWO	20	20	0	0
Estudios ambientales	NWO	5	5	0	0
Pedagogía	KNAW	27	23	4	0
Educación	KNAW	16	13	2	1
Sociología oriental	KNAW	38	31	7	0
Admón. Pública	KNAW	17	13	4	0
Sociología	KNAW	14	10	1	3
Arquitectura	KIVI	16	14	1	1
Ingeniería civil	KIVI	6	5	1	0
Geodesia	KIVI	4	4	0	0
Diseño industrial	KIVI	3	3	0	0
Ingeniería aeroespacial	KIVI	4	4	0	0
Ciencia agrícola	NRLO	53	50	3	0
TOTAL		435	390	31	14

Evaluación 1989

Disciplina	EAB	Número	Evaluación		
			posit.	negat.	indet.
Psicología	KNAW	27	25	1	1
Derecho	KNAW	75	62	8	1
Economía	KNAW	45	42	2	1
Matemáticas	NWO	31	30	1	0
Tecnol. de la información	NWO	24	24	0	0
Ciencias naturales	NWO	54	54	0	0
Astronomía	NWO	5	5	0	0
Administración empresas	KIVI	10	7	3	0
Ingeniería social	KIVI	7	7	0	0
Minería	KIVI	4	4	0	0
Ingeniería electrónica	KIVI	18	18	0	0
TOTAL		300	278	15	7

APENDICE B: INDICADORES Y VARIABLES PROPUESTOS

ENSEÑANZA

Indicadores	Variables	
	cuantitativas	no cuantitativas
plantilla	<ul style="list-style-type: none"> - estructura por edad - % de trabajadores con un n.º dado de años de experiencia - movilidad laboral del profesorado - promedio de publicaciones por departamento - n.º de profesores con tesis 	<ul style="list-style-type: none"> - cualificaciones académicas del profesorado - motivación de los profesores para enseñar - política de contratación de personal
ambiente pedagógico	<ul style="list-style-type: none"> - tiempo, medios y resultados de actividades institucionalizadas para la renovación educativa 	<ul style="list-style-type: none"> - filosofía educativa - orientaciones innovadoras
infraestructura y material disponible	<ul style="list-style-type: none"> - equipos informáticos (p. ej. ordenadores), por facultad o instituto - programas informáticos por profesor, facultad o instituto 	<ul style="list-style-type: none"> - calidad de las instalaciones docentes
oferta educativa		<ul style="list-style-type: none"> - amplitud de elección de currículum en relación con el volumen de currículum obligatorio - rango de elección de las principales asignaturas - oferta de las áreas de estudio: regular, libre y experimental - posibilidades de seguir una formación a media jornada, prácticas, una educación con garantía de contrato posterior, formación de postgrado

Indicadores	Variables	
	cuantitativas	no cuantitativas
currículum		<ul style="list-style-type: none"> - eficacia de los métodos de revisión de los programas existentes o de nuevos cursos - claridad en las descripciones de los objetivos del curso - adecuación de los objetivos del curso a la realidad profesional - concordancia entre el contenido y los objetivos del currículum - cumplimiento de las funciones del currículum - corrección, aceptabilidad, límites y nivel de los objetivos curriculares
política educativa		<ul style="list-style-type: none"> - sistema institucional de garantizar la calidad - acuerdos con otras instituciones sobre la distribución o concentración de tareas - aplicación democrática de la política educativa
ingreso de estudiante	<ul style="list-style-type: none"> - % de estudiantes que superan por área el examen proedéutico - n.º de doctorados por área - tiempo invertido en cada área de estudio - % de abandonos 	
salida de estudiantes	<ul style="list-style-type: none"> - promedio del tiempo transcurrido entre la obtención del título y el primer trabajo - % de paro/titulados en cada área de estudio - % de titulados con tesis después de x años, por facultad (universidades) 	<ul style="list-style-type: none"> - relación de la enseñanza con las actividades profesionales x años después de la graduación - demanda/oferta de plazas de formación, asignadas a titulados en el mercado laboral (enseñanza técnica superior) - adecuación del número de titulados al mercado laboral y a la sociedad

Indicadores	Variables	
	cuantitativas	no cuantitativas
INVESTIGACION		
personal	<ul style="list-style-type: none"> - movilidad interna y externa de los investigadores 	<ul style="list-style-type: none"> - disponibilidad de orientación competente para la investigación - política de contratación laboral
medios materiales	<ul style="list-style-type: none"> - número de becas de investigación recibidas; fondos de los organismos del Ministerio y del sector privado - % de gastos de investigación por institución 	
organización y ejecución de la investigación		<ul style="list-style-type: none"> - objetivos del programa práctico y sólido - alcance del programa, teniendo en cuenta sus objetivos - valor científico del programa (universidades)
política		<ul style="list-style-type: none"> - evaluaciones de investigación por el instituto - relación entre los objetivos de la investigación y la reciente planificación sobre el área correspondiente - estímulos de la institución para investigar
colaboraciones		<ul style="list-style-type: none"> - investigación interdisciplinar - colaboración con institutos y centros de investigación - relaciones internacionales con investigaciones realizadas - contactos con proyectos de calidad (nacional, internacional)
datos sobre publicaciones y referencias	<ul style="list-style-type: none"> - n.º anual de publicaciones científicas a nivel nacional e internacional 	<ul style="list-style-type: none"> - calidad de las publicaciones

Indicadores	Variables	
	cuantitativas	no cuantitativas
méritos	<ul style="list-style-type: none"> - n.º anual de artículos en publicaciones (internacionales con sistemas de referencia) - n.º tesis/año 	<ul style="list-style-type: none"> - utilidad de los resultados de la investigación en el comercio y la industria - utilidad de los resultados de la investigación en la propia enseñanza - méritos científicos: los resultados generan nuevos acuerdos, resuelven problemas fundamentales del conocimiento científico - reconocimiento nacional e internacional

SERVICIOS

servicios externos	<ul style="list-style-type: none"> - investigación por encargo del gobierno, compañías, sociedades... - relaciones con organizaciones exteriores - contratos educativos - contratos de investigación - actividades de educación permanente: n.º, duración y participación - difusión de los resultados científicos - medios que utilizan los departamentos para otros cursos y las instituciones HVE en el proceso de unificación - asesoramiento académico e información - tutoría
--------------------	--

EVALUACION DE LOS DEPARTAMENTOS

ROD MCDONALD

Universidad de Murdoch, Australia

ERNEST ROE

Universidad de Queensland, Australia

I. Introducción

¿Qué es una evaluación de departamento?

Las universidades siempre han estado implicadas en diversos tipos de evaluaciones. Las formas más visibles son la acreditación de las enseñanzas en *colleges* de educación avanzada, la evaluación de los títulos académicos por sociedades profesionales relevantes y la evaluación de individuos durante toda la educación superior a efectos de su promoción.

Más recientemente cierto escepticismo sobre el valor de la educación superior, unido a la escasa participación y a las críticas sobre la calidad de la educación de este nivel de enseñanza, ha llevado a los gobiernos a imponer restricciones presupuestarias. Esto a su vez ha producido un nuevo tipo de evaluación: por ejemplo, si deberían cerrarse o fusionarse ciertas instituciones, si deberían cerrarse ciertos departamentos, si debería transferirse personal de un departamento o institución a otro, etc.

Hay muchos puntos de vista en relación a qué es una evaluación, pero en este artículo no abordamos este tema. Nos preocupan las evaluaciones de departamentos y centros (en este trabajo se utiliza el término «departamento» como sinónimo de una unidad académica centro,

* Trabajo publicado originalmente en *Higher Education Research and Development Society of Australasia* (HERDSA), n.º 1, pp. 1-31, 1984.

departamento, instituto ó cualquier otro tipo de unidad) que implican una investigación sistemática de su estructura y funcionamiento académico con el propósito de evaluar y mejorar su rendimiento. La palabra sistemática es significativa. Implica una operación bien planeada y dirigida. Una evaluación de departamento podría consistir en un examen autocrítico del propio departamento y en una evaluación externa a través de un panel de expertos, realizadas ambas actividades de forma sistemática. Cualquier cambio recomendado que se diseñe para mejorar la calidad del trabajo de un departamento debe señalar el camino para obtener un mayor rendimiento. Cuando el departamento esta funcionando muy bien, los cambios a proponer serán pequeños. Ocasionalmente pueden ser sustanciales. No es normal encontrar un departamento perfecto.

Para definir una evaluación de un departamento puede ser útil decir lo que no es. No es una evaluación retrospectiva que se limita a juicios sobre el rendimiento pasado: los departamentos son entidades vivas y los juicios sobre ellos deben mirar hacia adelante. Una evaluación departamental no es el momento para evaluar detalladamente el rendimiento individual del personal. Inevitablemente tendrá implicaciones para el trabajo de sus miembros, dado que su trabajo en conjunto es el trabajo del departamento. Al poner de relieve las implicaciones de la evaluación para el profesorado, el grupo de expertos deberá igualmente mirar hacia adelante, constructivamente, considerando en todo momento los derechos de los individuos.

Razones que justifican este trabajo

Las evaluaciones de los departamentos bien efectuadas son ejercicios constructivos que contribuyen a la planificación a largo plazo. Pueden constituir una respuesta inteligente a las demandas de cambios, a las situaciones problemáticas y otras eventualidades propias de la educación superior.

Pero esto sólo puede ocurrir si se efectúan apropiadamente. Hemos visto ejemplos de evaluaciones que consisten en juicios basados en los prejuicios de los miembros del equipo de evaluación; de evaluaciones que han ofendido a todo el personal del departamento; y de evaluaciones que no han diagnosticado las dificultades importantes de un departamento con problemas.

Este artículo ha sido escrito para ayudar a las personas implicadas en estas evaluaciones grupos de expertos, miembros del departamento y administradores a fin de que puedan llevarlas a cabo de forma profesional. Se ha preparado únicamente como guía e introducción y no como prescripción universal. Muchos detalles en cuanto a cómo se efectúa una evaluación dependen del tamaño de la institución y del departamento a evaluar. Un análisis más completo se puede encontrar en Roe, McDonald y Moses (1985).

¿Son una buena idea las evaluaciones de los departamentos?

Algunos sugieren que no debería efectuarse una evaluación sistemática de los departamentos. Las razones aducidas nos son familiares. Señalan el temor a que los evaluadores no sean competentes, o considerados, o que utilicen un criterio equivocado, o que las revisiones cuesten dinero y no sirvan para ningún propósito útil e, incluso, que puedan hacer daño. Las asociaciones del profesorado también están preocupadas porque podrían utilizarse las evaluaciones para empeorar la situación de los docentes de la enseñanza superior restando tiempo y energía a empresas académicas más importantes.

Es cierto que una mala evaluación puede hacer realidad alguno de estos temores. Pero los beneficios de evaluaciones bien realizadas han sido importantes. Nuestros colegas y nosotros mismos hemos presenciado ejemplos de departamentos a los que les ha dado un nuevo impulso y oportunidades para iniciar nuevos rumbos, de individuos que se dieron cuenta por vez primera de que eran importantes en el departamento, de mejoras en organización, en docencia y en otros aspectos producidos por una seria revisión introspectiva de las actividades que no se cuestionaban antes. Las evaluaciones también pueden proporcionar información para tomar decisiones en el ámbito de la planificación a nivel institucional.

II. Planificación de una evaluación

1. ¿Cuál es el propósito de la evaluación?

Existen muchas razones para evaluar un departamento, desde el deseo de sus miembros de obtener información que les permita mejorar el

rendimiento hasta el intento de una institución de justificar una decisión realizada en perjuicio de ese departamento. Esta última razón no se puede defender moralmente. En efecto, una evaluación que «simplemente» juzga un departamento es difícil de justificar. Si se emiten juicios sobre la calidad de un departamento, la institución tiene la obligación de ofrecer ayuda y oportunidades para remediar los defectos: las evaluaciones deben resaltar la relación entre evaluación y desarrollo y, siempre que sea posible, deben señalar el camino para mejorar el rendimiento.

Los autores sobre el tema enfatizan la necesidad de que las evaluaciones tengan un propósito honesto y constructivo. El propósito central debe ser la mejora, no el juicio: identificar prioridades de desarrollo, planificar a largo plazo y revisar de forma sistemática los cometidos y actividades actuales.

Los beneficios de la evaluación estriban tanto en el proceso preparación, recogida y análisis de los datos, interpretación y discusión de resultados como en el informe final.

2. *¿Cuáles son las cuestiones a revisar?*

El ámbito de las evaluaciones de departamento puede variar considerablemente. Sin embargo, es común que se preocupen de los cursos que ofrece el departamento, la docencia, las actividades profesionales y de investigación, el personal académico, la administración y las necesidades futuras.

En términos generales las cuestiones a revisar incluyen normalmente los siguientes aspectos, aunque lógicamente deben ser adaptados a cada institución concreta:

1. La adecuación, calidad y equilibrio de los currícula específicos del departamento así como la calidad de la enseñanza que imparte.

2. La adecuación, calidad y equilibrio entre el desarrollo de la investigación y las actividades académicas de carácter profesional efectuadas en el departamento.

3. Analizar si las instalaciones, la distribución de fondos entre docencia/investigación/equipamiento, la administración interna de los re-

cursos, las normas y disposiciones institucionales, las estructuras de apoyo, y la gestión de los cargos directivos, cooperan eficazmente al desarrollo del departamento y el progreso de los estudiantes.

4. Analizar si existe una relación adecuada entre el departamento y otros departamentos similares, el resto de la institución y organismos externos relevantes en el campo científico específico.

Como consecuencia de sus trabajos, normalmente se solicita al equipo evaluador que informe sobre:

5. Si el departamento debe continuar en su forma actual ó si debería cambiar su estructura.

6. Direcciones futuras deseables para el departamento.

Normalmente se pide al equipo de expertos que informen sobre estos asuntos, teniendo en cuenta las normas nacionales e internacionales propias de la disciplina y/o profesión. También se espera que no sólo juzgen sino que sugieran mejoras cuando lo consideren apropiado.

Además de las cuestiones anteriores, habrá otras menos comunes que parecen convenientes en un momento concreto pero que, con el tiempo, dejan de tener importancia y son reemplazadas por otras. Por ejemplo, solicitar a los expertos cómo se pueden subvencionar nuevos desarrollos en el departamento por medio de contratos.

También puede haber cuestiones específicas de una institución o departamento: por ejemplo, interacciones entre diversas disciplinas y áreas, analizar si el departamento prepara a sus graduados para el mundo profesional o el papel que desempeñan los seminarios dentro del departamento.

Es inevitable que en el curso de una evaluación, los expertos formen opiniones sobre los miembros que componen el departamento. En algunos casos estas opiniones condicionarán sus recomendaciones finales. Sin embargo, no debe confundirse la evaluación de individuos como miembros de un departamento cuando se efectúa dentro de una evaluación departamental con la evaluación en profundidad de los individuos. Normalmente no es apropiado que la evaluación departamental suponga una duplicación de los procedimientos que una institución tiene establecidas para evaluar el rendimiento de su profesorado.

3. *¿Cómo se debe organizar?*

Bajo este epígrafe hay varias cuestiones clave: qué procedimientos existen o deben establecerse para efectuar la evaluación; quién debe llevarla a cabo; qué información se debe buscar, de quién y cómo. Estas cuestiones son las que se consideran a continuación.

Se debe prestar especial atención al apoyo administrativo; las necesidades varían para cada evaluación. Recoger y resumir información, preparar las sesiones de trabajo, asistir a reuniones y elaborar el borrador y el informe definitivo pueden llevar un mes de trabajo a un administrativo competente. Esto debe tenerse muy en cuenta, porque si no se recibe el apoyo necesario, si el comité de evaluación no dispone de la información precisa y en la forma que desea, sus trabajos se verán afectados.

Algunos de los principios generales que deben tenerse en cuenta en la organización de una evaluación se establecen en la sección V de este artículo.

4. *¿Quién debe efectuar la evaluación?*

Existe un amplio abanico de posibilidades, desde un evaluador o un equipo totalmente externo a la institución a un equipo formado exclusivamente por personal del mismo departamento. La composición del equipo evaluador depende del contexto institucional y de la finalidad de la evaluación. No hay fórmulas fijas, pero pensamos que los miembros del equipo de expertos deberán tener:

- experiencia adecuada sobre la disciplina(s) y/o campo profesional,
- conocimiento del funcionamiento de la institución, y
- competencia sobre los principios y métodos de evaluación.

Hay diversidad de opiniones en relación a cómo se pueden garantizar mejor estos requisitos dentro del grupo que, a su vez, necesita mantener un tamaño razonable. Estamos a favor de formar los equipos de evaluación a partir de las cuatro primeras categorías que exponemos a

continuación, pero lógicamente existen argumentos a favor de cualquiera de las que se relacionan.

a) Expertos en la disciplina(s) externos al departamento y probablemente externos a la institución.

b) Representantes del campo profesional (si son apropiados).

c) Un oficial administrativo veterano de la institución, como secretario del grupo.

d) Un especialista de evaluación, como miembro fijo o asesor del grupo. (Tales especialistas pueden ayudar al equipo o planificar la evaluación para obtener el máximo beneficio de ésta, aconsejar procedimientos, metodología, y otros aspectos prácticos, e incluso pueden reunir, analizar y presentar los datos).

e) Miembros del profesorado de la institución (externos al departamento).

f) Miembros del departamento.

g) Delegado o alto cargo del ministerio.

h) Estudiantes y/o graduados.

Se considera que una buena parte de los miembros del equipo evaluador deben ser externos para asegurar la objetividad, imparcialidad y una nueva perspectiva en el enfoque del análisis. Los argumentos en contra de esto enfatizan el valor de autoexamen. Algunos afirman que lo ideal es lograr un ethos institucional donde la evaluación se inicie internamente y se lleva a cabo de forma responsable, en un clima de autonomía y honestidad. Pero incluso cuando un departamento es evaluado internamente deberá implicar en este trabajo uno o dos evaluadores externos «por las razones anteriores».

La mayoría de los equipos de evaluación se establecen por acuerdo, con miembros externos e internos de diversas proporciones. El tema de la representatividad también es complejo. Si la tradición local pide que todas las partes interesadas estén representadas directamente, el grupo puede resultar demasiado grande para ser eficiente. Si las evaluaciones

se realizan en dos etapas (el equipo puede actuar en la segunda fase, una vez realizado el autoexamen por el departamento) pueden necesitarse menos miembros. No obstante, más importante que la representatividad de las personas que forman el equipo de evaluación es que haya suficiente interacción entre estos y el profesorado del departamento.

5. *¿Qué información se precisa, de quién y cómo?*

Un asunto clave para llevar a cabo una evaluación departamental es decidir qué información debe obtenerse. Lógicamente, esto está condicionado a las cuestiones a evaluar. Cierta información será común o la mayoría de las evaluaciones; por ejemplo, memorias anuales recientes, información sobre el profesorado, datos sobre cursos ofrecidos, presupuesto del departamento, datos acerca de la matrícula de estudiantes y porcentajes de graduados, publicaciones, etc. No obstante los equipos de evaluación o los evaluadores piden a menudo mucha más información de la que suelen utilizar. Para que esto no suceda hay que clarificar el propósito de la evaluación.

Por otro lado, el extremo opuesto, la evaluación sin datos reviste enormes peligros. Existe muchos ejemplos de evaluaciones basadas en conversaciones, opiniones e impresiones que apenas aportan información y no tienen ningún tipo de utilidad. Este tipo de trabajos cada vez son menos frecuentes.

Resumiendo, la más eficaz es utilizar la cantidad suficiente de información. En caso de una evaluación global, se pueden seleccionar los aspectos a recoger a partir de la siguiente lista:

A. Filosofía, objetivos e imagen del departamento

... objetivos del departamento, como se relacionan las actividades con estos objetivos, peso de la enseñanza orientada hacia lo «profesional» en relación con la de carácter «general», imagen del departamento dentro y fuera de la institución, etc.

B. Profesorado y cargos directivos

... número, cualificación y experiencia docente, investigaciones, intereses profesionales, promoción y desarrollo del profesorado, normas y procedimientos de evaluación del personal, etc.

C. Alumnos

...normas de admisión, número de estudiantes solicitantes/admitidos, datos sobre progreso, rendimiento, traslado a otras universidades, abandonos, causas de los abandonos, etc... (Esta información se recogerá tanto sobre los estudiantes de pregrado como de los postgraduados, cuando proceda).

D. Currículum, docencia, evaluación

... objetivos generales, estructura (incluyendo la integración de diferentes temas y entre diferentes áreas), nivel de los contenidos, horas de docencia, balance entre teoría y práctica, textos y otros materiales didácticos, calidad de la enseñanza, adecuación del sistema de evaluación, calidad de la supervisión de los doctorandos, procedimientos de selección de los sistemas de orientación y evaluación, etc...

E. Investigación

... proyectos de investigación individual; investigaciones en equipo; integración de estudiantes de postgrado en proyectos de equipo; distribución de fondos para investigación, becas para investigar, áreas prioritarias, etc.

F. Recursos, instalaciones y organización

... presupuestos y gastos, criterios para distribuir los fondos, costos por estudiante, instalaciones, equipamiento, servicios de apoyo a la enseñanza, estructura organizacional y de grupo, planificación, etc.

G. *Perspectivas en el futuro*

...restricciones del personal, flexibilidad de la plantilla, desarrollos probables de los cometidos del departamento, cambios posibles de las demandas de estudiantes, etc.

La respuesta a la pregunta «de quién» se debe obtener la información es sencilla: de todas las partes que tienen un interés legítimo en esta evaluación. En una evaluación departamental pueden ser fuente de información el director, el profesorado y los estudiantes del departamento, los graduados, los decanos, los administradores centrales y las organizaciones profesionales. Los métodos utilizados deben adaptarse a la naturaleza de las preguntas formuladas y al número de personas implicadas. En la práctica, se suele utilizar entrevista con los directores de departamento, decanos y responsables de la administración, mientras que los cuestionarios son más útiles cuando hay que recabar información de grandes muestras, por ejemplo los estudiantes.

Las decisiones sobre estas tres preguntas qué información, de quién y cuándo será más oportunas para la efectividad de una evaluación si se tiene siempre presente el propósito, el énfasis, el ámbito y las cuestiones a revisar en una evaluación. Esto es tan obvio que resulta casi ridículo recordarlo. Sin embargo, es muy común que la búsqueda de información, una vez que ha comenzado, se desarrolle sin control. Los evaluadores actúan a veces entusiasmados por la carrera, y necesitan detenerse y reflexionar por qué y hacia dónde corren.

6. *¿Cómo se debe procesar la información?*

A veces mucha información disponible o recogida en una evaluación no se analiza en realidad. Ante la complejidad y variedad del conjunto de aspectos a revisar, a veces parece que los expertos llegan a conclusiones mediante impresiones en lugar de realizar un análisis exhaustivo. Este procedimiento que descansa en la suposición de que las impresiones subjetivas son un juicio objetivo es peligroso. Las interpretaciones personales de los datos y la concentración indebida en aspectos que interesan a un individuo concreto son inevitables. Sería poco realista suponer lo contrario. Sin embargo, se pueden reducir al mínimo estos sesgos si hay una decisión clara y recursos suficientes para analizar adecuadamente los datos más importantes.

El análisis de datos requiere pericia y experiencia si se quiere llevar a cabo con competencia. Aunque su estudio excede el propósito de este artículo, se ofrecen algunas «pautas» para que sean tenidas en cuenta por los equipos de evaluación.

a) No merece la pena realizar un análisis exhaustivo de datos que no aportan información relevante o que están considerando el problema de forma equivocada.

b) Los que efectúan la evaluación y los que toman las decisiones basándose en ellos se beneficiarán si planifican el análisis de los datos juntos.

c) Existen muchas formas de analizar datos recogidos en una evaluación.

d) Los aspectos éticos siempre aparecen de alguna forma en el análisis de datos.

e) Se debe tener mucho cuidado al pasar de la descripción a la relación y a la predicción.

f) Hay que esperar al análisis de datos para que nos indique el camino a seguir, no planificar demasiado.

g) El análisis de datos lleva más tiempo de lo que se espera.

Un estudio completo sobre el tema se puede encontrar en Winston (1984), del que se extrajo la lista anterior.

7. *¿Qué tipo de informe se debe efectuar y para quién?*

El informe sobre la evaluación constituye la consecuencia lógica del proceso de recogida y análisis de datos y de los juicios profesionales de los expertos. La posibilidad de que los datos puedan recogerse y analizarse para reforzar un informe ya escrito (al menos en la mente de alguien) ya se ha comentado y deplorado. Se espera que un informe presente los datos completa y fielmente, y que interprete los resultados con ingenio y objetividad. Los evaluadores deben seguir las convencio-

nes que tienen en cuenta los investigadores al escribir un informe de investigación.

Sin embargo, existen algunas diferencias. La más acusada es que en la mayoría de los casos el informe sobre una evaluación debe efectuarse para que tenga consecuencias prácticas inmediatas. No obstante, los que llevan a cabo una evaluación departamental pueden considerar su responsabilidad muy limitada; de ser así, al igual que en los informes de investigación se pueden presentar los datos, interpretados con cautela si se desea, y dejar que los lectores lleguen a sus propias conclusiones. En el otro extremo, el informe puede ofrecer conclusiones justificadas y recomendaciones firmes, cuando quienes lo efectúan asumen bastante responsabilidad para realizar cambios en el departamento.

Al elaborar un informe sobre una evaluación también hay que considerar la audiencia (o audiencias) a quienes está dirigido. Es relevante saber cómo se originó, así como el contexto donde tienen lugar. Un informe de evaluación puede escribirse para el rector, para una facultad o escuela, para el director del departamento o todos los miembros de éste. El evaluador debe tener en cuenta que aunque frecuentemente se dirija al rector, también leerán su informe los miembros del departamento y deberán beneficiarse de ello. Constituyen una segunda audiencia. Cualquier persona con un interés legítimo en los resultados de la evaluación es una audiencia potencial, siendo normal que esas personas y/o grupos puedan ver el informe definitivo.

8. *¿Cómo asegurar que las propuestas se lleven a cabo?*

Los primeros pasos para su puesta en práctica deben darse mucho antes de escribir el informe. Al planificar la evaluación departamental, debemos considerar las posibilidades de realización de cualquier propuesta que pueda surgir y tomar medidas oportunas para que estas puedan llevarse a cabo. Estas medidas incluyen la creación de un clima favorable, informando y consultando a los miembros del departamento cuando sea posible, y asegurando su compromiso con la evaluación, para que, al menos, muchos de ellos esperen que las propuestas de cambio se lleven a cabo y, en el mejor de los casos, se involucren personalmente en ello.

Otra medida que favorece la puesta en práctica es la implicación de las diversas audiencias en la redacción de aspectos concretos del informe, invitando a las partes interesadas a comentar un borrador previo, haciendo todo lo posible para que los resultados se reciban con una actitud favorable.

El conservadurismo de la mayoría de las instituciones educativas obliga a emplear procedimientos explícitos para poner en práctica las propuestas: las posibilidades de adoptar tácticas dilatorias y de hacer sólo cambios superficiales son demasiado grandes para dejar la puesta en práctica al azar. Debe identificarse la persona u órgano que en cada caso asume la responsabilidad de llevar a cabo recomendaciones, estableciendo límites de tiempo para efectuar la acción; por ejemplo, se puede requerir un informe de control o progreso sobre implementación después de dos años. A este informe se debe acompañar de una copia del informe original para que todas las partes puedan ver lo que se ha hecho y lo que queda por hacer.

Esto no quiere decir que cada una de las recomendaciones de los expertos deba siempre llevarse a cabo. Se puede decidir no poner en marcha determinadas propuestas de un informe porque costarían demasiado en tiempo, dinero u otros recursos, o porque las ventajas no compensan el costo, o porque hacer un cambio en un lugar puede traer consecuencias desfavorables en otro, o porque no se puede realizar con el personal existente, o porque el cambio es incompatible con la política o la filosofía de la institución. Estas y otras razones pueden hacer fracasar la puesta en práctica. Sin embargo, cuando se aducen razones para no llevar a cabo las recomendaciones, deberán examinarse con rigor, y no usar estos argumentos como simples excusas para no actuar y no llevar a cabo los cambios propuestos como resultados de una evaluación.

III. Fases en la planificación de una evaluación

Al llevar a cabo la evaluación de un departamento, ¿qué debe *hacerse* realmente? En este apartado se sugieren, no se prescriben, las etapas o pasos para efectuarla. Algunas de estas etapas se pueden omitir o combinar dependiendo de las circunstancias particulares, pero sólo si existen buenas razones para ello.

Un asunto importante en la planificación y ejecución de una evaluación departamental es saber quién va a ser responsable de cada cosa. Esto debe quedar muy claro, aunque el mismo individuo o grupo no efectúe todos los pasos. Muchas evaluaciones han fracasado por falta de claridad en este aspecto.

1. *Elegir el grupo de expertos*

Esta decisión la deberán realizar aquellas personas que tienen mayor responsabilidad en la evaluación: el rector, el consejo académico o el comité encargado de la supervisión y control de las evaluaciones en la institución. En caso de una evaluación decidida internamente, lo hará el mismo departamento o su director. Si existe un comité de supervisión, sus miembros tendrán que efectuar esta elección. (En el apartado II.4 de este artículo se incluyen algunas propuestas sobre la constitución de equipos de expertos).

2. *Tomar las decisiones mencionadas en la sección II*

Constituye todo el proceso de planificación y es obviamente la etapa más extensa, ya que necesariamente exige muchas horas de trabajo en equipo del grupo de expertos. Aunque se recomienda tomar pronto decisiones, puede haber razones para posponer algunas de ellas o comenzar efectuando acuerdos provisionales.

Por otro lado, no todas las decisiones son responsabilidad del equipo de evaluación. Si existe, por ejemplo, un comité de supervisión éste puede tomar las decisiones relativas a la finalidad, el ámbito, las cuestiones a revisar e incluso acerca de la puesta en práctica, antes de designar a los expertos. La naturaleza de una evaluación y a quienes va destinada pueden ser asuntos que competen a la institución, sobre los que toma las decisiones el consejo académico y/o el rector.

Por tanto, dependiendo de cómo se estructure la autoridad en las evaluaciones, las dos etapas anteriores pueden dividirse en subetapas, implicando en ellas a diferentes personas. Lo importante es que, en este momento, se haya establecido un equipo para efectuar la evaluación, cuyas responsabilidades estén definidas con claridad, y todas las decisiones sobre la planificación hayan sido tomadas y comprendidas por el

equipo y por los que van a ser evaluados. Si esto es así, entonces se puede pasar al paso 3.

3. *Operativizar las cuestiones o evaluar*

Si las cuestiones a evaluar se especifican detalladamente, entonces se puede saltar este paso. Sin embargo, normalmente, las cuestiones o aspectos a evaluar se formulan de manera muy general, especialmente si constituyen normas para todo tipo de evaluaciones en la institución. (Observar ejemplos del punto II.2. Cuestiones como: «Valorar la calidad académica del departamento» o «Decidir si el departamento debe continuar con su estructura actual o si necesita cambiar para ser más efectivo»). Cuando esto sucede es responsabilidad del equipo de evaluación convertir estas cuestiones en preguntas específicas, susceptibles de ser investigadas. Si se pregunta si el departamento debe conservar su estructura actual, el equipo de evaluación tiene que determinar qué necesita conocer para tomar la decisión, qué aspectos del funcionamiento debe estudiar, traduciendo las cuestiones a analizar a términos breves y prácticos, que faciliten los pasos sucesivos.

La forma de realizar esto depende de cómo el equipo considera que debe realizar su tarea. Una manera es identificando, tan pronto como sea posible, los aspectos principales que intervienen en la dinámica de un departamento y concentrarse en ellos. Pero primero el equipo de expertos tiene que poseer un conocimiento suficiente sobre el funcionamiento de los departamentos para poder decidir cuáles son estos asuntos.

4. *Localizar la documentación relevante sobre cada cuestión a evaluar*

Los cuestiones o aspectos que se han identificado como el objeto de la evaluación constituyen los puntos clave sobre los que se centra todo el proceso. Mucha información necesaria para revisar estas cuestiones ya consta en algún lugar, pero necesita ser localizado y recogida adecuadamente. Una de estas cuestiones podría ser «Tendencias en desarrollo, matrícula y rendimiento de los estudiantes en los cursos del departamento durante los últimos 10 años». Toda la información relevante sobre esta cuestión, en principio, puede encontrarse en los archivos del departamento o del centro.

5. *Decidir qué información adicional se necesita*

Cuando se haya localizado la documentación existente, habrá que reunirla (ver paso 6). Al mismo tiempo, el equipo deberá decidir qué información adicional se necesita para poder evaluar correctamente cada cuestión. Esta decisión normalmente supone para el equipo su primer conflicto serio entre lo óptimo y lo real. Las decisiones provisionales tomadas en el paso 3 van a parecer optimistas. Ante la inminente tarea de recoger la información, y quizás un poco impresionados por la enorme cantidad de ella, las decisiones del equipo respecto a lo que realmente se necesita saberse suelen ser más restrictivas.

6. *Reunir documentación existente*

Siendo realistas, el equipo puede optar por ser selectivo en lugar de exhaustivo. Al menos esperar a que los que proporcionan el material lo hagan de forma resumida cuando sea posible. Esta etapa puede efectuarse, a su vez en varios pasos, según el contexto. Por ejemplo, el grupo de expertos, sobre todo si ellos mismo recogen los datos, puede verse en la necesidad de dar instrucciones específicas a los que proporcionan la información señalando qué deben proporcionar y cómo hacerlo; o de mirar las muestras de los materiales disponibles antes de dar esas instrucciones. Si el equipo evaluador cuenta con un personal auxiliar, se le puede encomendar esta tarea facilitándole un esquema previo y/o dejándole tomar sus propias decisiones en relación a qué materiales se deben recoger y cómo. Aunque esta fase. «Reunir documentación existente» se expone aquí como un único paso, el equipo puede distribuir esta actividad en varias operaciones.

7. *Establecer procedimientos, diseñar instrumentos, etc. para recoger información adicional*

La secuencia de esta fase es similar a como se indicó en la introducción. El paso 5 decidir que información es necesaria se puede tomar parcialmente después del paso 4 y del 6. También es posible repetir los pasos 5 y 7, incluso más de una vez. En otras palabras, así como el estudio de la documentación existente puede sugerir la necesidad de obtener más información, el estudio de esa información nos puede llevar

a la identificación y recogida de más información. La necesidad de una etapa adicional de recogida de datos puede producirse antes o después de haber procesado una información, o tanto antes como después.

Sin embargo, cada vez que el equipo de evaluación decide buscar más información, tendrá que establecer los procedimientos para conseguir esa información, o si tiene la fortuna de disponer de auxiliares experimentados, se podrá pedir a esas personas que planifiquen la manera de obtenerla.

INFORMACION A EVALUAR POR LOS EXPERTOS

A continuación se enumeran algunos items relativos al tipo de información que podría utilizar un equipo de expertos para evaluar un departamento, aunque la lista no pretende ser prescriptiva. Parte de esta información ya consta en los registros y archivos oficiales y otra tendría que ser preparada o recogida.

1. Reglamentos de la institución y procedimientos establecidos para la evaluación de los departamentos, incluyendo las razones que justifican esta evaluación.
2. Un resumen sobre la estructura del departamento, su organización situación actual, y desarrollo futuro previsible.
3. Copias de los libros de actas e informes anuales del departamento.
4. Presupuesto del departamento.
5. Profesorado dotaciones, plantilla actual (nombres, edades, nivel académico, áreas de conocimiento, curriculum vitae, etc...) y personal de apoyo.
6. Información sobre cursos ofrecidos por el departamento estructura de los cursos, requisitos, esquema de los cursos, didáctica y métodos de evaluación.
7. Información sobre la carga docente del profesorado.

8. Datos sobre la matrícula durante los últimos 5 años: ratio estudiante/profesor, ratios de graduados, datos sobre promoción y abandonos (comparando estos datos, si es posible, con los de otras instituciones).

9. Estadísticas sobre calificaciones de los cursos impartidos por el departamento.

10. Información sobre los trabajos de investigación en el departamento: proyecto, publicaciones, otros outputs, becas y ayudas recibidas, etc...

11. Resultados de evaluaciones en cursos selectivos o cursos clave.

12. Salidas laborales de los graduados del departamento y, en el caso de cursos profesionales, cómo se relacionan con los objetivos del departamento.

13. Informes específicos sobre aspectos concretos: carga de trabajo del alumno, efectividad de los prerrequisitos para el acceso, opiniones de los graduados, causas del abandono de los estudiantes, etc...

8. *Recoger esta información*

La duración de esta fase varía mucho de acuerdo con el número de cuestiones a evaluar. Cuando la evaluación se limita a temas sobre los que existe documentación suficiente, las búsquedas adicionales son poco necesarias. En el extremo opuesto, «recoger la información» puede incluir largos cuestionarios a estudiantes, graduados, empleados, miembros de una o varias profesiones, además de entrevistas extensas, realizadas por los expertos u otras personas en su nombre, a muchos individuos incluyendo también todo el personal del departamento a evaluar.

9. *Analizar o procesar la información*

Esta fase, que se ocupa de elaborar la información pertinente, es similar a que se efectúa en el paso 6 sobre documentación existente. En ambas el propósito es sintetizar datos de forma que puedan ser escritos en un informe. En el caso de que la información se pueda cuantificar

como la recogida por cuestionarios se harán tablas. Los datos obtenidos a través de entrevistas se pueden registrar, transcribir y resumir.

10. *Interpretar los datos*

La importancia de este paso puede dar lugar a varias reuniones y discusiones del equipo de evaluación. Constituye el proceso de pasar de la descripción a la evaluación. El equipo tiene toda la información organizada, que esencialmente es una descripción detallada del funcionamiento del departamento. Se incluirán además las opiniones individuales acerca del funcionamiento del departamento, que también son evaluativas. El equipo debe hacer ahora *su* evaluación: preguntarse ¿qué significa todo esto? ¿a qué conclusiones se puede llegar? ¿qué importancia concedemos a estos hechos y a estas opiniones? ... y otras cuestiones similares. Como resultado el equipo de evaluación debe emitir su juicio profesional acerca del departamento.

Al interpretar datos y formular recomendaciones, los expertos deberán resistir la tentación de poner su sello personal en el departamento. A ellos se les pide que evalúen un departamento, no si personalmente prefieren hacer las cosas de otra manera.

Este es todo el consejo que podemos dar en un pequeño manual sobre el proceso de formación de juicio. Suponemos que un grupo de expertos constituido inteligentemente, por ejemplo, con varios bioquímicos, será capaz siguiendo los pasos que hemos explicado y utilizando sus conocimientos de llegar a juicios profesionales sobre un departamento de bioquímica.

En muchos de las evaluaciones es necesario afrontar cuestiones difíciles de forma clara. El profesor Richard Johnson, que ha presidido muchas de ellas en la Universidad Nacional Australiana, menciona un ejemplo de estas situaciones críticas: mientras que los miembros del equipo no confían el uno en el otro no se abordan las cuestiones problemáticas, hasta que alguien se impacienta y decide hablar abiertamente y entonces, con un sentimiento de alivio, el grupo comienza a encarar el verdadero problema del departamento y las medidas necesarias para mejorar su situación. Un presidente experimentado puede anticipar y acelerar este proceso.

11. *Elaborar un borrador del informe*

Un equipo de evaluación puede emprender esta tarea de formas diferentes, dependiendo en parte de cómo pretende estructurar su informe y de quién va a escribirlo. Algunas secciones del mismo pueden hacerse en borrador muy pronto (ya en los pasos 5 y 6); por ejemplo, datos sobre la historia del departamento u otra información descriptiva ya existente y que no necesita más investigación. Por otro lado, el equipo puede decidir no escribir nada excepto los apuntes de los miembros hasta que se disponga de todo el material, y se pueda dividir la tarea entre los miembros del equipo, responsabilizándose cada uno de una sección, o puede ser realizado por el presidente tras largas discusiones sobre qué debe incluirse en el informe. Incluso se puede hacer y considerar uno, dos o más borradores antes de que haya un borrador definitivo para difusión pública (ver paso 12).

12. *Distribuir el borrador del informe para que lo comenten las partes interesadas*

Todas aquellas personas afectadas por la evaluación deberán leer el borrador, invitándoles a aportar sus comentarios por escrito antes de una fecha determinada. Este paso no solo debe realizarse para cumplir el principio de efectuar las máximas consultas posibles en todas las etapas, sino también para garantizar que no se ha omitido o malinterpretado nada importante, dado que cualquier pecado de comisión o de omisión llamaría su atención.

13. *Editar y distribuir el informe definitivo*

Una vez que se han recogido todas las observaciones al borrador, el equipo de evaluación decidirá los cambios que deben introducirse. Entonces emitirá su informe definitivo. Normalmente todos los miembros del equipo deben estar de acuerdo con el contenido del mismo. En circunstancias excepcionales, uno o más de sus miembros pueden realizar un informe distinto, en desacuerdo con uno o más puntos de las observaciones o recomendaciones propuestas. Esos desacuerdos van a dañar, seguramente, la credibilidad del informe y van a limitar su efectividad.

14. *Iniciar la puesta en práctica de las recomendaciones*

Normalmente el paso 13 es el último del que es responsable el equipo de evaluación. Por supuesto, las consecuencias del informe son cruciales. De hecho, en términos de «costos», no se justificaría de otro modo la inversión. Sin embargo las medidas para poner en práctica las propuestas del informe dependen de otras personas. De ahí que muchas de estas observaciones y propuestas comiencen con términos como «normalmente» o «usualmente» dado que en definitiva dependen de quien tiene la autoridad en la institución en relación con la evaluación departamental. (ver pasos 1 y 2). Por supuesto, las evaluaciones no sustituyen a los procedimientos habituales establecidos para la toma de decisiones en las instituciones académicas, aunque pueden formar parte de ellos. Sin embargo, en determinadas circunstancias, el equipo de evaluación puede tener cierta responsabilidad en la supervisión o ser el receptor de informes concernientes a las observaciones obtenidas durante la puesta en práctica.

La puesta en práctica de las recomendaciones de los expertos se puede iniciar de muchas formas. El informe puede ir al rector y/o al consejo académico, que demanda al departamento la realización de algunas o todas las recomendaciones propuestas. Puede ir al rector y/o al consejo académico simplemente a título informativo, pero la responsabilidad de iniciar las acciones oportunas recae sobre el director del departamento, el decano u otro órgano académico. Puede ser una norma de la institución que la innovación se efectúe a partir de los informes de la evaluación o, por lo contrario, que «alguien» tenga que decidir en cada evaluación que recomendaciones se van a poner en práctica.

15. *Establecer procedimientos para controlar el progreso en la puesta en práctica*

La acción que se ha iniciado puede detenerse. Las propuestas del equipo de evaluación pueden perderse ante la enorme actividad que desarrollan los departamentos, o menos que haya algún mecanismo que asegure la *continuación* de la acción durante su realización. Se pueden elaborar informes de progreso bien individualmente o en grupo, pero lo verdaderamente importante es que estos informes se hagan y se entreguen a un receptor específico. Cualesquiera que sean los procedimientos que se establezcan para controlar el progreso de la puesta en

práctica, deben incluir las fechas propuestas para finalizarla, es decir las fechas en que todas las recomendaciones deban haberse puesto en práctica.

IV. Creación y mantenimiento de un clima adecuado para la evaluación

Gran parte de los consejos expuestos en este artículo sobre la planificación de una evaluación y los pasos de la misma, tienen como objetivo crear y mantener un clima apropiado, o sea, un clima favorable.

Esta constituye una difícil tarea que los promotores, planificadores y realizadores de las evaluaciones deberán prever y tener en cuenta. En este artículo se ha enfatizado un enfoque positivo y constructivo de la evaluación, pero es ingenuo suponer que esto es suficiente para crear un buen clima. Una afirmación como «sólo quiero ayudarle a mejorar su rendimiento» puede interpretarse como «hay algo equivocado en su rendimiento. Necesita mejorar». Puede pensarse que se desea hacer pública alguna deficiencia en el rendimiento del departamento.

Las afirmaciones con finalidad constructiva deben traducirse en acciones positivas o, al menos, en firmes compromisos de acciones positivas. Los compromisos negativos que los expertos no deben hacer son igualmente importantes: por ejemplo, que no existan órdenes encubiertas, que no se oculte información significativa sobre los procedimientos de la evaluación, que no se identifique a los que hayan facilitado información en privado, que los expertos no pasen por alto los comentarios efectuados sobre el borrador del informe, etc...

Los aspectos clave para crear y mantener un clima favorable incluyen los siguientes: explicitar los propósitos; escoger un equipo de evaluación con credibilidad (por ejemplo, no sólo con miembros externos, sino con miembros y/o un secretario completamente familiarizado con las normas y procedimientos administrativos de la institución); mantener informadas a todas las partes interesadas, con consultas amplias y reales; asegurar la confidencialidad cuando sea apropiada, proporcionar oportunidades para que el departamento haga comentarios antes de terminar el informe y crear y cumplir expectativas en cuanto a la implementación del informe. A menos que el personal del departamento crea que va a ocurrir algo como resultado de la evaluación, el clima de la

misma estará contaminado por el cinismo, que afectará a las futuras evaluaciones en la institución. Por otro lado, la puesta en práctica de cualquier evaluación puede ayudar al clima de futuras evaluaciones si se enfatizan más los puntos fuertes del departamento. Es importante realizar acciones constructivas en los puntos débiles, pero no menos importante es aprovechar los puntos fuertes y, en particular, reconocer a aquellos en los que hay un buen rendimiento y que, como se dá por sentado, obtienen poco reconocimiento y aprecio por sus méritos. Resumiendo, se debe reconocer y enfocar continuamente la relación entre evaluaciones y desarrollo, cambio y mejora.

Las evaluaciones se aceptarán mejor si son un aspecto normal de la vida de una institución, por ejemplo, si es una norma de la institución la evaluación regular de todos los departamentos y servicios. No es malo ser evaluado, nadie tiene por qué sentirse discriminado, todos somos tratados de igual modo.

V. Implicaciones para los programas institucionales

En este último apartado, se consideran puntos clave para crear y mantener un buen clima en las instituciones a favor de las evaluaciones. Como sirven para recordar las cuestiones importantes de las secciones II y III, pueden considerarse como pautas de acción para los programas institucionales en lo relativo a las evaluaciones de departamentos y centros.

1. Las evaluaciones deberán formar parte de las prácticas normales de las instituciones educativas, ayudando a los departamentos a dirigir su progreso y haciendo realidad la autoevaluación que la comunidad espera.

2. Para establecer un sistema de evaluaciones, una institución tiene que proporcionar la iniciativa en los siguientes aspectos:

- a) «vendiendo» el concepto de evaluación periódica del profesorado como un modo de contribuir a la salud académica de la institución y ayudar a realizar una planificación racional, actividad que es normal en todas las instituciones grandes;
- b) formando al profesorado sobre los objetivos, actuaciones y resultados de estas evaluaciones;

- c) proporcionando los medios necesarios para remediar la deficiencias detectadas en cada evaluación.

3. La institución necesita proteger los derechos individuales del profesorado mediante normas como las que a continuación se especifican de acceso a la información y confidencialidad de los resultados:

- a) Los propósitos de la evaluación, las áreas a examinar y las cuestiones a revisar deberán hacerse públicas. Deberá darse la oportunidad al personal del departamento de comentarlos antes de establecerlos definitivamente;
- b) todo el profesorado del departamento deberá tener acceso a los miembros del equipo de evaluación;
- c) todo el personal del departamento deberá conocer la relación de documentos e informaciones que van a ser revisados por los expertos;
- d) las personas que faciliten datos e informaciones durante el proceso de evaluación deben ser confidenciales exclusivamente para los miembros del equipo de evaluación.
- e) las discusiones del equipo de evaluación deberán ser secretas;
- g) aunque es imposible evaluar el rendimiento de un departamento sin evaluar al mismo tiempo el rendimiento de los individuos que lo forman, no deberán hacerse comentarios evaluativos sobre los individuos que se citan en el informe de evaluación;
- h) los miembros del departamento deberán tener la oportunidad de comentar el informe y facilitar sus sugerencias mientras sea todavía un borrador;
- i) el informe deberá estar a disposición de todos aquellos interesados, a menos que haya razones poderosas contra ello.

4. Los miembros del equipo de evaluación deberán:

- a) ser competentes e imparciales y ser vistos como tales por el departamento a evaluar;

- b) disponer al menos de un experto reconocido sobre la materia a evaluar externo a la institución, y cuando sea apropiado un profesional destacado en ese mismo campo;
- c) cuando se evalúen varios departamentos, si es posible, mantener a los mismos miembros; por ejemplo, al mismo personal administrativo para ayudar a los paneles y/o al mismo presidente del equipo de evaluación.

5. La institución ha de establecer pautas para realizar, informar e implementar las evaluaciones, enfatizando especialmente su naturaleza constructiva. Estos son algunos de los principios que deberán tener en cuenta los expertos, además de los ya establecidos anteriormente (apartado 3). Deberán incluir:

- a) el equipo de evaluación debe poseer datos relevantes antes de expresar sus juicios, no se puede dar recomendaciones sin datos; particularmente hay que ser muy cautos con los comentarios y opiniones que se efectúen sobre un sujeto cualquiera, independientemente de su experiencia, hasta confirmarla con evidencias;
- b) el equipo de evaluación deberá solicitar consejos profesionales (si es necesario) para asegurar que los procedimientos a utilizar son apropiados para el tiempo que hay disponible y para las circunstancias concretas de la evaluación;
- c) deberá consultarse también con el profesorado del departamento en cuanto a los procedimientos de evaluación;
- d) se deberá dar un amplio margen al departamento para recoger y facilitar datos que son necesarios, incluyendo la posibilidad de que el mismo departamento realice una evaluación previa, si esa es la norma de la institución;
- e) no obstante, se deberá terminar toda la evaluación en el menor tiempo posible para no perder la ocasión y la credibilidad.

6. El equipo evaluador deberá realizar un informe:

- a) que contenga recomendaciones claras que pueden implementarse;

- b) que ponga de manifiesto las limitaciones de los datos en que se basan las recomendaciones, y las limitaciones de la propia evaluación;
- c) que incluya las razones de todos los comentarios y recomendaciones citando datos relevantes cuando estén disponibles, para que la base sobre la que se argumenta esté clara.

7. La institución y/o el equipo de evaluación puede acelerar la puesta en práctica de las recomendaciones a través de las siguientes formas:

- a) sugiriendo medios para efectuar cambios;
- b) dando responsabilidad sobre cambios específicos o individuos concretos;
- c) delimitando un período para su puesta en práctica (por ej. dos años) con un seguimiento al final del período y un informe al organismo apropiado.

AUTOESTUDIO INSTITUCIONAL Y MEJORA DE LA CALIDAD *

E. A. HOLDAWAY
Universidad de Alberta, Canadá

Introducción

En este artículo se considera un autoestudio como sinónimo de una revisión de programa. Las actividades de un autoestudio van más allá de un autoexamen, ya que también incluyen la revisión que hace una universidad del funcionamiento de aquellas unidades que han realizado un autoexamen. Desde mediados de los setenta el interés en las revisiones de los programas universitarios ha aumentado rápidamente, especialmente en Estados Unidos, pero también en Gran Bretaña, Australia y Canadá. Harper (1986) ha observado que a las revisiones de programas se les está dedicando cada vez más tiempo, esfuerzo y personal. En Canadá varias universidades han desarrollado procedimientos detallados para llevar a cabo revisiones de programas, mientras que otras, o bien están poniendo en marcha procedimientos semejantes, o bien han manifestado interés en tal posibilidad. Las experiencias de la Universidad de Alberta, la segunda mayor universidad del Canadá, han tenido un efecto esencial sobre el desarrollo del interés en las revisiones de programas; estas experiencias están descritas en profundidad en este artículo.

Este artículo se apoya esencialmente en otros artículos de revisiones de programas universitarios, especialmente en el de Enns, Harris, Holdaway, Jopling y Kreisel (1986), e incluye material del informe anual del presidente del «Comité Asesor de Revisiones del Campus» de la Universidad de Alberta, algunas de cuyas frases son repetidas aquí literalmente.

* Este artículo apareció originalmente en Kells, H. R., y Vught, F. A. van (Eds.), *Self-regulation, Self-study and Program Review in Higher Education*, Utrecht: Lemma, 1988.

Veamos algunas definiciones que son esenciales. La *evaluación* puede ser considerada como una valoración limitada, basada normalmente en un criterio preestablecido. La *revisión* es tomada como una descripción y valoración más global de los antecedentes, del rendimiento actual y de los planes futuros y perspectivas del programa. Un *programa* puede ser un programa académico, por ejemplo, un programa de licenciatura, pero también puede ser un departamento o grupo de departamentos como proponía Gilmour (1980: 1):

«Ultimamente los departamentos académicos son una unidad de análisis más adecuada que los programas que confieren un grado, debido a que son a los primeros a los que se les asignan los recursos institucionales y forman el bloque básico del funcionamiento y de la estructura de gestión en la mayor parte de las instituciones de educación superior.»

Esto nos lleva a aceptar la definición de Arns y Poland (1980: 269) de *revisión de programa* como una «búsqueda de la evaluación global de un grupo coherente de actividades académicas en funcionamiento».

Resumen de literatura seleccionada

Propósitos

Varios autores han formulado los siguientes propósitos para la revisión de programas:

1. Informar a los responsables de la toma de decisiones sobre los puntos fuertes y débiles de los programas (v. gr. Peterson, 1980; Barak y Berdahl, 1978).
2. Determinar el nivel de los programas en relación con criterios preestablecidos o con el rendimiento de otros programas (v. gr. Council of Graduate Schools in the Unites States, 1979).
3. Proporcionar información para la planificación (v. gr. Cochran y Hengstler, 1984; Nicols, 1982; Barak, 1986; Benoist, 1986).
4. Ayudar a una institución en la toma de decisiones sobre los programas, su instalación, continuación, modificación, expansión o cierre (v. gr. Jany y Keller, 1981; Peterson, 1980).

5. Ayudar a una institución en la toma de decisiones sobre reducciones de gastos e incrementos en la eficiencia (v. gr. Barak y Berdahl, 1978; Belager y Tremblay, 1982).

6. Rendir cuentas del funcionamiento del programa (v. gr. Barak y Berdahl, 1978).

Más sucintamente Olscamp (1978) clasificó los propósitos de la revisión de programas en: justificar públicamente el programa, rendir cuentas cuantitativamente, y evaluar y mantener la calidad. Mientras que Feasley (1980) señaló como propósitos la planificación, la mejora y la justificación del programa. Jany y Keller (1981) observaron que las revisiones internas tienden a enfatizar la efectividad, mientras que las externas favorecen la eficiencia y la responsabilidad. Harpel (1986: 1) ha observado que, a pesar de que la demanda de revisiones de programa a escala estatal se ha incrementado en algunos estados norteamericanos, son más frecuentes los procesos internos; éstos pueden tomar la forma de una «respuesta instantánea a un problema observado o a una situación crítica... o también puede ser un proceso continuo que evalúa todos los programas durante un período de tiempo». Este último enfoque es el que será descrito en este artículo. Harpel (1986: 2) también daba este importante consejo: «El propósito de una revisión de programas debe ser definido claramente en un principio y reafirmado a lo largo del proceso si los resultados, tanto directos como indirectos, se quiere que tengan un impacto positivo en la institución.» Además, algunos autores han subrayado que las revisiones de programas no deberían ser utilizados para ocuparse de problemas que pueden ser resueltos a través de procedimientos administrativos ya establecidos.

Procedimientos

McCorkle y Achibald (1982: 143) han llegado a la conclusión de que las revisiones de programa son «ahora más rigurosas, más frecuentes y más globales en amplitud y en profundidad». Sin embargo, como Harpel (1986: 2) nos recuerda, «No existe ciertamente un único método correcto para llevar a cabo una revisión de programa». Los procedimientos de revisión de programa incluyen varios elementos importantes; la lista siguiente ha sido hecha por Craven (1980: 444-449), Dressel y Cammack (1972: 277-283), y por otros:

1. Identificar los propósitos de la revisión.
2. Identificar las cuestiones que se van a revisar.

3. Identificar los procesos que van a usarse.
4. Designar los métodos que se van a utilizar para reunir y analizar los datos numéricos, otras informaciones y opiniones.
5. Decidir cómo ha de ser la composición de los comités de revisión y cómo deben ser seleccionados sus miembros.
6. Decidir los criterios que se van a usar en la valoración de los programas.
7. Decidir cuál será la distribución final del autoestudio y de los informes de la revisión.
8. Establecer mecanismos para: (a) determinar el grado de acuerdo de las instituciones con las recomendaciones, y (b) supervisar la realización de esas recomendaciones a lo largo del tiempo.

Algunos escritores han subrayado la necesidad de evaluar los procesos de revisión usados en las universidades. Así, por ejemplo, Craven (1980) sugirió con este fin los siguientes criterios:

1. ¿Son útiles los resultados de la revisión para fundamentar distintas opiniones sobre decisiones alternativas?
2. ¿Mejora lo acertado de esas decisiones gracias a que se ha tenido en cuenta la información obtenida en las revisiones?
3. ¿Son realmente puestas en práctica las diferentes decisiones que han sido seleccionadas?

Diferentes universidades utilizan diversos procedimientos para la revisión de programas, que han sido descritos en la bibliografía (v. gr. Harris y Holdaway, 1983; Sloper, 1982) y en informes institucionales. Desde mediados de los setenta, la Universidad de Illinois en Urbana-Champaign ha estimulado la puesta en marcha de revisiones de programa a través de la organización de congresos y proporcionando información a todas aquellas instituciones interesadas en las revisiones. Los procedimientos utilizados en las diferentes instituciones varían en aspectos tales como la preparación de los informes del autoestudio, la composición del comité organizador, la participación de los administradores, del personal y de los estudiantes, la claridad en la comunicación, el nivel de publicidad del proceso, la frecuencia de las revisiones en

cada unidad, la amplitud de las revisiones, el efecto de retroalimentación sobre la propia unidad, la inclusión de programas de pregraduados y/o postgraduados, la inclusión de revisiones, de unidades de apoyo, la relación entre los comités revisores y el comité organizador, y la manera en que los revisores externos se relacionan con los internos.

La experiencia estadounidense ha tenido algunos efectos en Canadá. Los procedimientos utilizados en la Universidad de Alberta, descritos en este artículo, estuvieron fuertemente influenciados por los procedimientos usados en Illinois. Otras universidades canadienses han emprendido también amplias actividades de revisión. Por ejemplo, la McGill University (1982) tiene uno de los sistemas de revisión de programas más activos: comenzó sus revisiones de unidades académicas con ciclos de siete años en 1982, y en 1987 había completado más de 100 revisiones. Los objetivos de la McGill eran la mejora de los programas, el establecer prioridades y la distribución de los recursos. Se utiliza un comité interno formado por personal de la universidad y por estudiantes, presidido por algún decano prestigioso. Los procedimientos usados en la McGill incluyen la preparación de un informe de autoestudio, la preparación de otro informe por uno o varios revisores externos, y la preparación de un tercer informe por un comité interno que tiene en cuenta el informe realizado por el comité externo.

En otras universidades canadienses se han usado también revisiones para modificar la distribución de los recursos. La práctica más generalizada es el uso de revisiones *ad hoc* para examinar áreas con un problema en particular; sin embargo, en este artículo no se hace referencia a este enfoque.

Costes y beneficios

Cuando las universidades deciden llevar a cabo una revisión de programas siempre se preguntan: «¿Cuáles son los posibles beneficios y costes, no sólo los más obvios, sino también los que pueden ser más difíciles de percibir?»

Dressel y Cammack (1972), Russo, Brown y Rothweiler (1977), Arns y Poland (1980), Hoyt (1980) y otros han identificado estos beneficios: las revisiones de programas nos obligan a examinar aquellas materias que frecuentemente son ignoradas, a tener en cuenta aquellas otras que

están afectadas por posibles cambios, y de ese modo crear una atmósfera más receptiva; proporcionan un equilibrio cualitativo a los enfoques que hacen énfasis a los números; y pueden identificar cambios constructivos, creando una imagen institucional positiva.

Los principales costes involucrados son los gastos administrativos (salarios, honorarios, viajes, fotocopias, etc.) y el aumento de la carga de trabajo de los miembros de la unidad que va a ser revisada. Esta carga adicional de trabajo parece recaer principalmente sobre el director de la unidad que está siendo revisada. Pueden generarse consecuencias negativas, tales como un cierto sentimiento inicial de paranoia, pero actuando con sensatez el comité organizador puede minimizar la posibilidad de que se produzcan tales hechos.

Para mejorar la relación entre beneficios y costes son recomendables seguir estos criterios: concentrarse sobre los principales objetivos; reunir solamente aquellos datos más importantes; utilizar un sistema de revisión continuo creíble y aceptable, con preferencia a un sistema *ad hoc*; y utilizar un sistema de intercomunicación efectivo y oportuno.

Jacobson (1985) ha señalado que una preocupación común, manifestada en la reunión anual de 1985 del *National Center for Higher Education Management System* (NCHEMS), es la de que en los procesos de revisión de programas que habitualmente son puestos en práctica se da demasiada importancia a la acumulación de datos y muy poca a la evaluación. La Universidad de Alberta tuvo en cuenta todas las consideraciones anteriormente mencionadas para el desarrollo de sus métodos de revisión de programas. Ahora podemos considerar estos procedimientos a la luz de las discusiones anteriores.

Los procedimientos utilizados en la Universidad de Alberta

Antecedentes

La Universidad de Alberta se fundó en Edmonton en 1908. En diciembre de 1986 la matrícula en 16 facultades era de 21.527 estudiantes pregraduados a dedicación completa, de 2.775 estudiantes postgraduados a dedicación completa en 75 departamentos, y de 4.710 estudiantes a dedicación parcial. Aunque la Universidad de Alberta es desde hace bastantes años una institución grande y compleja, no se

emprendió un programa de revisión de un modo global hasta 1980. El actual proceso de revisión de programas tuvo su origen en un Comité de Prioridades Universitarias que en 1978 recomendó que se iniciaran las revisiones de todos los programas y unidades.

Estas recomendaciones se aceptaron. Un comité establecido para elaborar los criterios de referencia para las revisiones dedicó un tiempo considerable a la valoración de los procedimientos utilizados en otras universidades y recomendó que se adoptaran procedimientos simples que no interfirieran con la estructura ya establecida para la toma de decisiones. La universidad aprobó los principios y procedimientos siguientes:

1. La meta de las revisiones deberá ser la producción de información que sirva como base segura para la planificación y la mejora de los programas.

2. Los métodos que van a utilizarse en las revisiones deberán ser fácilmente comprendidos, efectivos, eficientes, responsables, flexibles, creíbles y realistas.

3. El Comité Asesor del Presidente para las Revisiones Académicas debe estar formado por un director y por otros dos miembros seleccionados por el presidente y el vicepresidente académico.

4. Los cometidos principales de este comité serán los siguientes:

a) asesorar al presidente sobre el orden en el que deben ser revisadas las unidades y sobre la composición de los equipos de revisión;

b) asesorar a los equipos de revisión sobre los requisitos de la revisión;

c) transmitir los informes de las revisiones al presidente; y

d) mantener a la comunidad universitaria informada sobre la marcha del trabajo.

5. Cada unidad que va a ser revisada deberá preparar un informe de autoestudio.

6. Se deberá nombrar un Comité de Revisión de la Unidad (CRU) por cada unidad, constituido por dos miembros de fuera de la universi-

dad dedicados a la misma disciplina, otro de dentro de la universidad dedicado a una disciplina afín y otro de la propia universidad sin especificar disciplina.

7. El programa de revisiones deberá realizarse a lo largo de varios años.

8. Cada CRU deberá consultar activamente con los estudiantes pregraduados y postgraduados, con el personal académico y no académico, con los graduados recientes y, cuando se considere oportuno, con los empleadores de los graduados.

9. Deberá recaer sobre el vicepresidente académico la responsabilidad de la supervisión de las medidas tomadas siguiendo las recomendaciones dadas por las CRUs.

En 1980 se nombró un Comité Asesor del Presidente para las Revisiones Académicas (CAPRA). Estaba constituido inicialmente por un director de departamento, el vicepresidente académico, el decano de estudios postgraduados e investigación y por un profesor. Cuando se formó el CAPRA la universidad decidió también revisar todas las unidades administrativas y de apoyo. El uso de un único comité tuvo la ventaja de unos costes mínimos y de unos procedimientos uniformes. Al CAPRA se le agregó un director de servicios técnicos y se le cambió el nombre por el de Comité Asesor del Presidente para las Revisiones en el Campus (CAPRC). Pasado el primer año, el vicepresidente académico se retiró del comité para poder estar más dedicado a la puesta en práctica de las recomendaciones y a la supervisión a largo plazo. En 1980 el presidente y CAPRC establecieron los siguientes principios:

1. La postura de la universidad frente a las revisiones será la de promover un cambio progresivo y fomentar la calidad de la vida académica.

2. Se esperan múltiples beneficios, tales como una mejora de la planificación y un mejor conocimiento del funcionamiento de los programas.

3. Se hará énfasis en la flexibilidad de las diferentes fases, proporcionando un número limitado de normas.

4. Mientras que estén siendo desarrollados los procedimientos de la revisión, se revisarán un número limitado de unidades.

5. Para evitar duplicaciones, la revisión del CAPRC y el proceso de acreditación deberán coincidir siempre que sea posible.

6. Se deberá evitar todo sensacionalismo y se respetará la confidencialidad. El CAPRC decidió que era fundamental un equilibrio entre los datos numéricos, la información descriptiva y la evaluación. Se pidió al Gabinete de Investigación y Planificación Institucional que proporcionara datos a todas las unidades para usarlos en sus actividades de autoestudio.

Procesos globales

En 1980-81 se iniciaron las revisiones de sólo cinco unidades de un modo deliberadamente prudente, de modo que se pudieran considerar cuidadosamente los pormenores del proceso. Actualmente, se revisan 15 unidades por año. En junio de 1987 se habían completado visitas *in situ* para 75 unidades, mientras que otras 50 no habían sido revisadas todavía.

El CAPRC proporciona las instrucciones a seguir y asesora a la unidad que va a ser revisada, la cual realiza por su parte un informe de autoestudio. El autoestudio se lleva a cabo generalmente en un período de seis a doce meses. Seagren y Bean (1981) observaron que éste era el caso más frecuente. Se espera que los informes estudien estos aspectos: historia, estado actual, resultados claves, una consideración clara y equilibrada de efectivos y áreas mejorables, estrategias para cambios futuros junto con los requerimientos para tales cambios, y las misiones a realizar tanto para los individuos como por los grupos. Las instrucciones que el CAPRC da a la unidad son de tipo genérico y no específicas. Como ejemplos se suministran copias confidenciales de informes de autoestudio realizados por otras unidades. Se espera que cada unidad realice un autoestudio maduro y serio, que involucre a todos o a la gran mayoría de los miembros de la unidad, que se forme un pequeño comité directivo para la preparación de los borradores, que se incluyan los *curricula vitarum* (los cuales pueden ser entregados en un informe aparte), y que se haga uso de un enfoque apoyado en los datos preferentemente a un enfoque dirigido por los datos.

El CRU está formado generalmente por dos revisores externos y por dos revisores internos (es decir, de la Universidad de Alberta), aunque puede incluirse también un miembro relevante de la profesión. Se espera que uno de los miembros de la universidad, denominado «internamente relacionado», conozca bien la unidad y que sea capaz de entender la composición y el funcionamiento de la unidad revisada. Del otro miembro de la universidad, denominado «miembro libre», se espera que represente los intereses de la comunidad universitaria. La unidad hará propuestas para elegir a los miembros externos del CRU y al «internamente relacionado». Después será el CAPRC quien discuta con el director de la unidad sobre estas propuestas y sobre la suya propia respecto del «miembro libre». La lista de propuestas se presenta al presidente para su aprobación. Los miembros del CRU deben ser personas de confianza tanto dentro de la unidad como fuera de ella. Se espera que el director de la unidad proporcione evidencias al CAPRC de que los cuatro a seis revisores externos propuestos son excelentes en sus especialidades y que son personas objetivas y honradas en sus juicios.

Una vez que el CRU se ha formado se enviará a cada miembro un documento de antecedentes sobre la revisión y el informe de autoestudio, aproximadamente un mes antes de la visita *in situ*. Ya en el campus, las relaciones del CRU con los miembros de la unidad serán abiertas y amplias:

«Sin intención de restringir la amplitud de la revisión, deberíamos prestar especial atención sobre cuales son, en opinión de los miembros del CRU, los puntos fuertes y débiles en la enseñanza de los departamentos y los programas de investigación, una evaluación del personal y de la cantidad de trabajo que soportan, de si son adecuados los recursos disponibles, de la eficiencia de la organización del departamento, y de las relaciones de este departamento con el resto. En particular, deberíamos poner más atención sobre cada oportunidad que tiene el CRU para mejorar los programas del departamento, las relaciones con organizaciones tanto internas como externas, y para una mejor utilización de los recursos disponibles. Pedimos también vuestros juicios sobre los planes del departamento para el futuro.»

El CRU envía sus informes al CAPRC normalmente dentro de los dos meses posteriores a la visita. Una vez que el informe ha sido aprobado por el presidente, se envía a la unidad, la cual prepara una respuesta propia. Tras haberla recibido, el CAPRC hace un resumen de los hechos principales de los tres informes y se reúne con el jefe de la unidad para discutir los hechos y conclusiones que resultan de la revisión. El presidente discute también por separado el sumario del CAPRC con los

miembros del CAPRC. Después tienen lugar discusiones por separado sobre el informe de autoestudio, el informe del CRU y la respuesta de la unidad, participando en ellas el CAPRC, el presidente, el jefe de unidad, el director o decano de quien depende el jefe de la unidad, y el Comité de Planificación y Prioridades. En cualquier momento los administradores pueden empezar a seguir las recomendaciones formuladas por la revisión.

En 1985 se hicieron dos modificaciones. La primera es que tras el encuentro con el presidente para la discusión de los tres informes y el sumario del CAPRC, el CAPRC escribe ahora a continuación una carta confidencial a los miembros del CRU para comunicarles sus puntos de vista sobre los resultados de la revisión a fin de que sean tenidos en cuenta. Los revisores valoran cada una de las críticas. La segunda modificación es que pasados unos 18 meses después del encuentro con el presidente, el CAPRC llama la atención del vicepresidente académico sobre algunos de los temas importantes cuya solución puede ser necesario revisar.

Se han celebrado alrededor de 125 reuniones del CAPRC durante cada año. Su duración es de una hora y participan en ellas sólo miembros del CAPRC, aunque puede incluir también otras personas tales como el jefe de la unidad revisada, los dos revisores internos, el CRU, el presidente, y los principales comités del campus. Cada miembro del CAPRC prepara unos cuatro informes para el presidente. El presidente del CAPRC, quien realiza las dos terceras partes de los nombramientos, contacta inicialmente con las unidades, vigila los procesos y dirige globalmente la oficina del CAPRC. Le ayuda un coordinador a tiempo completo.

La totalidad de los procesos de revisión, particularmente la fase de autoestudio, han tendido a prolongarse más de lo que se pensó originalmente, y esto ha retrasado la finalización del ciclo de revisiones. El CAPRC y el presidente decidieron revisar primero departamentos y facultades no departamentalizadas, al mismo tiempo que las unidades administrativas y de apoyo. Aún no se ha tomado ninguna decisión sobre el camino que se debe seguir una vez que hayan sido revisadas todas las unidades.

Para la revisión de las unidades seleccionadas se han utilizado diferentes enfoques. Este es un ejemplo tomado de la revisión del CAPRC

de 1986-88 para la Facultad de Educación. Tras la finalización de la revisión de cinco de los seis departamentos de educación, el decano de educación decidió que se debía emprender una revisión global del funcionamiento de la facultad. Esto comprendía seis fases principales:

a) preparación de un documento detallado que describa el funcionamiento de la facultad;

b) entrevistas con el personal respecto a este documento y a otros asuntos relacionados, seguidas por la preparación de un informe;

c) celebración de un seminario, formado por 10-12 personas de dentro y fuera de la universidad, que estudie las consecuencias de la revisión;

d) realización de reuniones en la facultad para identificar y discutir los puntos de intereses y las preocupaciones;

e) identificación por grupos de trabajo de la facultad de los puntos claves, así como su discusión; y

f) preparación de un informe final.

Las principales diferencias entre una revisión típica del CAPRC y la revisión de esta Facultad de Educación son, primero, la duración del proceso, y segundo, la intención del decano de educación de que puedan realizarse algunos cambios sustanciales en el funcionamiento de la facultad durante la revisión antes de esperar a que el proceso de revisión se haya completado. Este enfoque fue aprobado por el CAPRC, quien tendrá en cuenta su utilidad en otras posibles revisiones de otras facultades departamentalizadas, tales como las de artes, medicina y ciencias, una vez que se hayan completado o estén próximas a completarse las revisiones de sus departamentos. La revisión de una facultad debería enfocarse desde estos aspectos: su misión, el papel y la función de la oficina del decano y de las unidades de apoyo (p. ej., audio-visual, investigación y biblioteca), el papel y el trabajo realizado por los comités de la facultad, la diferenciación y la integración de los departamentos, así como de algunas de las materias relacionadas con las revisiones de los departamentos de la facultad.

Resultados

El CAPRC ha identificado en sus informes anuales algunos resultados importantes en relación con las revisiones de programas. Estos y algunos otros resultados conexos los presentamos a continuación bajo los encabezamientos de comités para la revisión de una unidad, informes, problemas temporales, renovación de la unidad y del personal, unidades de apoyo, revisiones y planificación, orden jerárquico de las unidades, carga docente, relación enseñanza-investigación, departamentos profesionales, revisión de programas de postgrado y participación de los estudiantes. En nuestra opinión, estos resultados son importantes en la actualidad para cualquier puesta en marcha de una revisión de programas o para el proyecto de introducir las revisiones en una institución.

Comités para la Revisión de una Unidad

En 1980 se evidenció la preocupación porque el personal de las unidades pudieran nombrar a sus amigos para que formasen parte de los CRUs, de modo que las revisiones no fueran tan críticas como debieran. No obstante, hemos observado la regularidad con la que sensatos y expertos revisores, realizando juicios de toda confianza, han producido informes equilibrados que pueden ayudar a la mejora de las unidades. Existe también preocupación acerca de la disponibilidad de expertos. Sin embargo, aunque hemos visto que las personas más cualificadas están muy solicitadas, es usualmente posible conseguir reunir las en el campus durante tres días. Finalmente, nuestros procedimientos reúnen juntos en un equipo revisores internos y externos, mientras que otras universidades generalmente los sitúan en comités separados. Estábamos algo preocupados con respecto al éxito del funcionamiento de esta estrategia, pero nuestras reservas estaban infundadas porque, según nuestra experiencia, la combinación de revisores internos y externos en un solo comité se ha comprobado que es de gran utilidad.

Informes

A menudo las unidades y los CRUs se preguntan sobre cómo de largos deben ser sus informes. Según el CAPRC, ésta es una magnitud con una gran varianza, aunque el promedio de páginas de un informe

de autoestudio es de unas 90, del de un informe del CRU es de 20 y del de la respuesta de la unidad es de 10. Gran parte de los informes de autoestudio contienen apéndices, que en algunos casos son más extensos que el propio texto. Desde varios puntos de vista el informe de autoestudio es el documento más importante, aun cuando las discusiones formales tiendan a enfocarse más sobre el informe del CRU.

Generalmente, las reacciones de las unidades a los informes del CRU son positivas; sin embargo, en algunas ocasiones las unidades han tenido una reacción negativa, especialmente cuando los comentarios del CRU parecen ser duros o cuando los informes incluyen errores objetivos. No obstante, una vez transcurrido algún tiempo, las unidades suelen darse cuenta de que las críticas expuestas suelen estar justificadas hasta cierto punto y de que garantizan las acciones adecuadas para corregirlas. Las unidades tienen absoluta libertad de expresarse en contra de las declaraciones y recomendaciones del CRU, y, si bien deben ser consideradas cuidadosamente todas las recomendaciones, no tienen que ser necesariamente aceptadas por la universidad o por la unidad. Como los miembros del CRU sólo están en el campus por un corto espacio de tiempo, sólo obtienen una «instantánea» del funcionamiento de la unidad. No obstante, los miembros del CAPRC siempre quedan admirados de la gran cantidad de cosas que los CRUs son capaces de descubrir en tan corto período de tiempo.

Problemas relacionados con el tiempo

La mayor parte de las revisiones transcurren durante una visita de tres días al campus. Normalmente, los CRUs utilizan una cantidad de tiempo semejante en la unidad, fuera de la unidad, y en las discusiones del comité. Para que el tiempo de trabajo de los CRUs sea utilizado de un modo efectivo es necesario un programa cuidadoso, unos informes de autoestudio bien preparados y una presentación eficiente. Con una duración razonable de la visita al campus, por ejemplo tres días, conseguiremos atraer revisores de calidad y minimizar el cansancio.

Hay otro problema en relación con la necesidad de encontrar un intervalo de tiempo mutuamente aceptable antes de la visita del campus del CRU. Suelen transcurrir de tres a cuatro meses entre la presentación de las candidaturas del CRU y la revisión *in situ*. En consecuencia, si

una revisión va a realizarse antes de abril, el informe de autoestudio y la selección del CRU deberán estar terminados antes de año nuevo.

Renovación de la unidad y del personal

Las universidades, al igual que otras instituciones, deben estar en continua renovación si desean conservar su vitalidad. En esta renovación puede ser de ayuda un programa efectivo de revisión de procesos. Incluso en períodos de estabilidad y de moderación financiera cuando la motivación para el cambio puede ser mínima. También hay que considerar que los niveles de empleo del personal por departamentos fueron frecuentemente determinados en años anteriores, cuando las condiciones eran posiblemente diferentes. Las recomendaciones del CRU con respecto a los cambios de los empleos del personal pueden servir de ayuda a la universidad para identificar las acciones adecuadas en aspectos como la sustitución y la recolocación del personal.

Según señala Craven (1980), las mejoras de un programa requieren en algunas ocasiones el haber incrementado la asignación de recursos. Nuestros procedimientos tienen en cuenta esta posibilidad mediante la reserva de fondos que permitan al presidente incrementar los recursos a cualquiera de las unidades revisadas en caso de un empleo adicional de personal o para otros propósitos.

Unidades de apoyo

Nosotros consideramos que las revisiones de unidades de apoyo son necesarias y que es apropiado, efectivo y eficiente tener un comité único que organice tanto las revisiones académicas como las de unidades de apoyo. Nuestro modelo ha funcionado bien en ambos tipos de unidades. No obstante, las unidades de apoyo generalmente son más complejas y diversificadas que las unidades académicas. Además, como las unidades de apoyo, tales como la secretaría, la oficina de servicios del personal, la biblioteca y los servicios de informática afectan a todas o casi a todas las unidades académicas, generalmente se puede apreciar en todo el campus un mayor interés por sus revisiones que por los de la mayoría de los departamentos académicos.

Muchas universidades parecen haber incorporado a su funcionamiento revisiones de las unidades de apoyo. Esto debe ser recomenda-

do. Según nuestros estudios, el mantenimiento de revisiones de unidades de apoyo no sólo nos ha conducido a mejoras, sino también a un mejor entendimiento en todo el campus de los objetivos y del funcionamiento de esas unidades.

Revisiones y planificación

Frecuentemente se menciona el interés de las revisiones de programas para la planificación institucional. Esta postura es especialmente válida cuando las revisiones en todas las unidades del campus —académicas y/o administrativas o de apoyo— son llevadas a cabo en un periodo de tiempo limitado de, por ejemplo, uno o dos años. Resulta más difícil la integración de los resultados de una revisión de programa en un plan global cuando estas revisiones se realizan a lo largo de una década. Sin embargo, esto es factible si se adopta el enfoque de una planificación en continua revisión.

En la Universidad de Alberta reconocemos que nuestro sistema de revisiones de programas es inadecuado por sí sólo para proporcionar información y dirección para una planificación global. Por lo tanto, en 1983 se formaron cuatro grupos de planificación para ocuparse de las siguientes materias: capacidades y aptitudes, informática, estudios para postgraduados y programas nuevos. El plan global aún está siendo terminado, pero ya se ha conseguido una mayor credibilidad, lo que se ha manifestado en un sustancial incremento en la proporción de matrículas de estudiantes de postgrado.

Ordenación jerárquica de las unidades

En algunas ocasiones, se requiere que el CAPRC solicite a los CRUs que ordenen jerárquicamente las unidades que van a ser revisadas, esto es, que indique cómo se ordenarían en relación a unidades similares en Canadá y en otros lugares. El CAPRC se ha resistido a esta petición porque según su punto de vista un procedimiento de este tipo interferiría con el objetivo principal de la mejora del programa. Las bases con las cuales el CRU realizaría una ordenación son además poco claras, y proporcionar una ordenación de este tipo no nos daría una información de la unidad que debemos mejorar, a menos que las bases para la ordenación se hubieran explicitado claramente. Es más, incluso con

los mejores revisores posibles, sería difícil hacer tales juicios para todas las disciplinas, y la inversión de tiempo en deliberaciones sobre las ordenaciones serían improductivas. Además, necesitaría cambiarse el centro de atención de todas las fases de la revisión desde donde se encuentra actualmente; se le daría más énfasis a la ordenación que a la mejora.

Carga docente

Los informes de autoestudio y los informes del CRU hacen referencia frecuentemente a la pesada carga docente de los profesores. La opinión suministrada por el CRU es un suplemento útil para los indicadores numéricos usualmente usados para evaluar la carga docente, los cuales están basados o en el número de estudiantes a los que se enseña o en el número de horas de docencia. Por supuesto, se han de tener en cuenta otros aspectos tales como el nivel de los estudiantes (primeros cursos o no, pregraduados o postgraduados), y los pesos relativos de clases, seminarios, asistencia clínica, prácticas y laboratorios. En la Universidad de Alberta existen claras diferencias en la docencia en los distintos departamentos debidas a factores tales como (a) la naturaleza de la disciplina, (b) la imposibilidad de la universidad de responder rápidamente a los cambios en el tiempo de la carga docente, y (c) el impacto de las financiaciones externas para la investigación y para los estudiantes de postgrado. Ni las horas de clase ni las horas de estudio de los estudiantes solos son por sí solas un indicador digno de confianza. Las horas de clase no permiten hacer variaciones en las matriculaciones en las asignaturas. Por otro lado, las horas de estudio de los estudiantes pueden permitir demasiadas variaciones en las matriculaciones. Se han desarrollado medidas compuestas en la Universidad de Alberta que valoran las horas de clase y las horas de estudio diferenciadamente, así como también tienen en cuenta —pero con una valoración reducida— las actividades programadas aparte de las clases, tales como las horas de laboratorio. Estas medidas son indicadores más dignos de confianza que las medidas simples, pero aún están siendo discutidas en el campus. Debido a los intereses creados, no esperamos un acuerdo total sobre la cuestión de los indicadores de carga docente.

El Departamento de Investigación y Planificación Institucional (DIPI) ayuda al CAPRC en el cálculo de estas medidas compuestas. Además, el vicepresidente académico ha pedido al DIPI y al CAPRC que exami-

nen juntos el funcionamiento de siete departamentos distintos cuyas medidas compuestas de horas de contacto/de estudiantes se encuentran cerca de la media de todos los departamentos. Esta actividad todavía está en fase de organización; los aspectos que serán examinados podrían incluir investigación, publicaciones, participación en comisiones, supervisión de estudiantes graduados y servicios.

La relación enseñanza-investigación

Nuestros informes de autoestudio y los informes del CRU han tendido a enfocar la investigación y la enseñanza como funciones separadas. No obstante, en clases avanzadas de pregrado y en seminarios de postgraduados, las actividades de investigación pueden ser tenidas en cuenta tanto como la enseñanza, de modo que la investigación y la enseñanza empiezan a estar entrelazadas. Aunque una carga docente pesada puede dificultar la productividad investigadora, somos conscientes de que hay profesores con una pesada carga docente y con una productividad alta en investigación, y también hay profesores con una carga docente ligera que son menos productivos en investigación. También deberíamos reconocer que algunos profesores prefieren darle más importancia a sus funciones en servicios que a la investigación. No parece existir ninguna correlación entre productividad investigadora y competencia docente en la enseñanza de pregrado.

Departamentos profesionales

En los departamentos profesionales de las universidades, p. ej., medicina, odontología y derecho, la demanda de servicios modifica las funciones de enseñanza e investigación. En muchas áreas «básicas» tales como las Facultades de Artes y Ciencias, la prestación de servicios se encuentra generalmente subordinada a la enseñanza y a la investigación. La demanda de profesores en facultades profesionales que ofrezcan servicios y que demuestren una competencia actualizada nos lleva a preguntarnos cuestiones tales como ¿qué cantidad de servicios deberían proporcionarse? Las revisiones de programa nos pueden dar información relativa a estas preguntas. La cuestión más importante aquí es que las facultades profesionales, y los profesores en esas facultades, muy a menudo requieren unos tipos diferentes de evaluación de los que tradicionalmente se utilizan.

Revisiones de programas para postgraduados

Nosotros creemos firmemente que las revisiones de programas para postgraduados deberían integrarse en la revisión de cada departamento en particular. En algunas universidades se separan estas revisiones, mientras que en otras se revisan solamente los programas para postgraduados. No obstante, los CRUs deberían prestar atención a aspectos de los programas para postgraduados, tales como el control de los progresos realizados por estos estudiantes y la calidad de la supervisión de su trabajo.

Participación de los estudiantes

Una parte importante de los procesos de revisión necesita de la participación de los estudiantes. En algunos campus, los estudiantes son miembros de los comités. No obstante, nosotros no escogimos este enfoque, sino que en su lugar decidimos pedir que los CRUs obtuvieran la participación de los estudiantes en las revisiones de unidades académicas del modo que los representantes de la junta de estudiantes lo consideren más oportuno en cada unidad en particular.

Reflexiones y recomendaciones

La Universidad de Alberta ha considerado que sus revisiones de programa son de gran utilidad; esta opinión se ha escuchado una y otra vez por parte de decanos y directores. El sistema que nosotros hemos desarrollado sólo ha sido modificado ligeramente desde que fue adoptado en 1980. Otras universidades, por supuesto, utilizarán sistemas que se adapten mejor a sus circunstancias particulares, pero nosotros prevemos que los sistemas regulares de revisión de programas se van a dar cada vez con más frecuencia en las universidades canadienses y creemos que nuestro modelo tiene mucho que ofrecer a aquéllos que están contemplando la posibilidad de iniciar revisiones. Es necesario que las revisiones de programas sean realizadas de un modo cuidadoso y razonable, y los procedimientos se deben establecer cuidadosamente. Si esto se realiza apropiadamente, pueden ayudar a la planificación, y pueden ayudar a las universidades a resolver sus retos financieros. Re-

comendamos que las universidades sigan estos siete procedimientos generales en la revisión de programa:

1. Darle un énfasis especial a la mejora del programa.
2. Utilizar un comité asesor del presidente dirigido por un profesor respetado y capacitado que no sea actualmente administrador.
3. Tener tres informes preparados: un informe de autoestudio, un informe del comité de revisión de la unidad y una respuesta de la unidad.
4. Utilizar los informes de los equipos de revisión con una confidencialidad apropiada, pero asegurándose de que los procedimientos sean abiertos y que se comprendan con claridad.
5. Reservar una cantidad de dinero razonable cada año de modo que algunas unidades puedan conseguir un incremento de fondos como respuesta a una revisión.
6. Nombrar los dos revisores internos y externos para un equipo de revisión.
7. Dar a un gerente la responsabilidad de controlar los progresos hacia la consecución de los logros fijados en las recomendaciones.

Como punto final queremos subrayar la necesidad de veracidad y de buena voluntad, y recomendar que el proceso no se realice apresuradamente. Es importante un comienzo firme en un pequeño número de unidades para que los procesos de revisión continuos y positivos se puedan iniciar y sean aceptados en toda la comunidad del campus.

COMITES DE FINANCIACION E INDICADORES DE RENDIMIENTO EN LA EVALUACION DE LA CALIDAD EN EL REINO UNIDO *

JOHN SIZER

Universidad de Tecnología de Loughborough, Reino Unido

Introducción

En el presente artículo se argumentará que la revisión por colegas no debe ser contrarrestada con los indicadores de rendimiento y que a su vez, en el Reino Unido los indicadores de rendimiento deben y, de hecho lo hacen, aportar información a la revisión por colegas, ya sea en los organismos de financiación como a nivel institucional. En segundo lugar, se intentará demostrar que los enfoques británico y holandés de la evaluación de la calidad en educación superior se asemejan más de lo que suele pensarse. Una de las metas de los *Universities Funding Councils* (UFC) es el animar a las universidades a ejercitar su autonomía al máximo, lo cual consiste en una completa disponibilidad del empleo de los fondos derivados del Comité, tal y como queda reflejado en el Acta de Reforma Educativa (circular 19/89). Otra meta es el minimizar la participación burocrática en los asuntos de las instituciones. No se desea jugar un papel dominante, más bien todo lo contrario. A pesar de que el *University Grants Committee* (UGC) ejercía un papel más dirigista durante la década de los ochenta, nunca desempeñó un rol dominante como el del Ministerio de Educación y Ciencia holandés.

Es cierto que el Reino Unido ha dirigido en Europa el desarrollo y empleo de los indicadores de rendimiento, pero *no* como una alternativa a la revisión por colegas. Tanto los antecedentes como el desarrollo y

* Traducción íntegra del capítulo 11: *Funding Councils and Performance Indicators in Quality Assessment in the United Kingdom* perteneciente a: L.C.J. GOEDEGEBUURE; P.A.M. MAASSEN & D.F. WESTERHEIJDEN (Eds.) *Peer Review and Performance Indicators*. Utrecht: Lemma. pp. 155-181, 1990.

uso de los indicadores de rendimiento y la situación actual son descritos, antes de pasar a considerar sus usos y relaciones con la revisión por colegas.

La esencia del argumento presentado es que los indicadores de rendimiento no sólo informan a la revisión por colegas sino que existen otros muchos motivos por los que los indicadores de rendimiento deben ser, y de hecho lo están siendo, desarrollados y empleados.

1. Contexto nacional del desarrollo y uso de los indicadores de rendimiento

Dos sucesos levantaron el interés por los indicadores de rendimiento, remarcando a la vez su importancia. En primer lugar, la elección de un gobierno conservador en mayo de 1979 y, en segundo lugar, el enfoque altamente selectivo adoptado por el UGC, al aplicar reducciones de fondos en julio de 1981 (Sizer, 1989). Aunque el UGC nunca publicó sus criterios, era del dominio público que se emplearon medidas (*input*) de la calidad de la enseñanza, p, ej., calificaciones a nivel de sobresaliente. Sin embargo, la influencia más significativa fue la llegada de la Sra. Thatcher a Downing Street.

El gobierno Conservador se ha inclinado no sólo por reducir el gasto público y por la economía de mercado —lo que ha llevado a las instituciones a buscar fondos adicionales de origen no gubernamental— sino también a que aquellas justifiquen las actividades y rindan cuentas del empleo de sus recursos y de su rendimiento en términos de las tres *Es*:

- *economía en la adquisición y empleo de los recursos,*
- *eficiencia en el empleo de los recursos y*
- *eficacia en el resultado de los objetivos institucionales, departamentales e individuales, a través de una puesta en práctica de estrategias y planes de acción.*

La relación entre estos evasivos conceptos en la educación superior queda ilustrada en la figura 1. Ello requiere, como se refleja en el informe del *Steering Committee for Efficiency Studies in Universities* (CVCP, 1985), el desarrollo de planes estratégicos que sustenten las decisiones

y estructuras académicas, lo que incorpora la planificación, la distribución de recursos y la rendición de cuentas en un proceso conjunto que une los aspectos académicos, económicos y físicos. Son aspectos importantes de la *eficacia* la sensibilización ante las necesidades de una sociedad post-industrial, avanzada en cuanto a la información y a la tecnología, incluyendo ampliar el acceso a la educación superior dado que disminuye la población de jóvenes de 18 años de edad, contribuyendo a mejorar la economía, compitiendo en la formación de científicos, ingenieros, tecnólogos y técnicos más cualificados y desarrollando la capacidad de adaptación a los cambios (DES, 1987). Además, las universidades deben preparar graduados que sean capaces de superar las sorpresas de un mundo incierto, complejo, volátil y cambiante (Horton, 1990). Otras importantes dimensiones de la eficacia son la selectividad a la hora de distribuir los recursos disponibles para la investigación, una racionalización y evaluación crítica de los departamentos pequeños y una continua búsqueda encaminada a la mejora de la calidad de la enseñanza. Así pues, el gobierno espera que las instituciones rindan cuentas —o sea que se justifiquen económicamente en términos de las tres Es— respecto a la UFC y la PCFC, y, a través de los cuerpos

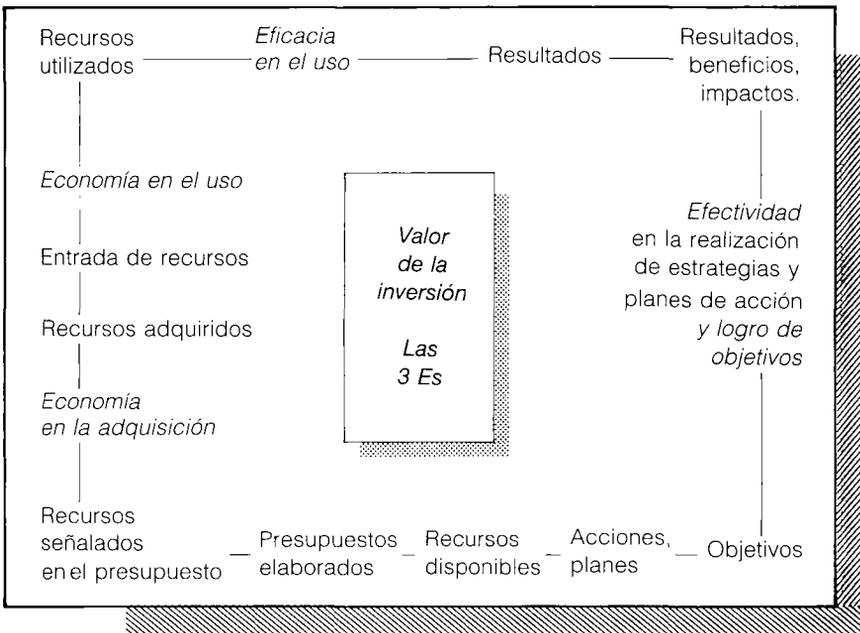


Figura 1. Las tres Es

financieros, al DES, al Tesoro y al Parlamento. Es precisamente en este contexto en el que los indicadores de rendimiento pueden ser considerados sin posibilidad de error como una alternativa a la revisión por colegas más que como un complemento.

3. Enfoque empresarial, economía de mercado, autonomía institucional, metodología de la financiación y rendición de cuentas

El gobierno desea reducir su responsabilidad en la financiación de la educación superior y espera que las instituciones se financien con fondos alternativos, a través de un creciente enfoque empresarial y de economía de mercado. Se argumenta que, reduciendo la dependencia de la financiación del gobierno e incrementando la diversidad de fuentes de financiación, se restablecerá la autonomía institucional que implica la necesidad de fortalecer los mecanismos de rendición de cuentas. Se considera que un excesivo énfasis en la financiación del gobierno y una excesiva dirección por parte de los organismos de financiación, pueden producir una inhibición de la responsabilidad, creatividad, innovación y diversidad de provisión. Por consiguiente, existen presiones sobre las universidades para que cambien su *cultura* y pasen de un modelo de libre oligarquía y consenso, apoyado por los estilos administrativos de la gestión (Kogan y Boys, 1982), a un modelo de corte empresarial y de economía de mercado, con una gestión de estilo ejecutivo (Lee y Piper, 1988; Sizer, 1988). Por consiguiente, el Informe Jarrat (CVPV, 1985: 36) recomienda la elección de los jefes de departamento al Consejo Universitario, bajo la recomendación del Vice-Canciller tras una apropiada consulta; con una clara pormenorización de tareas y responsabilidades respecto a sus departamentos y respecto al empleo de los recursos. Recae precisamente en los directores de departamento y no en los departamentos en su conjunto, la responsabilidad del empleo de los recursos y de una adecuada realización de estrategias y planes de acción para el control de la calidad y el mantenimiento de los estándares.

Una importante y nueva dimensión del proceso de rendición de cuentas es el requisito de los nuevos consejos de financiación, el UFC y el PCFC, de no pensar en las instituciones financieras en términos de paquetes de ayudas financieras y planificación del número de alumnos,

sino en términos de contratos por proporcionar aporten ciertos servicios de docencia e investigación. Los consejos de financiación se mueven hacia la realización de contratos con una combinación de financiación de los estamentos institucionales y planes para ofrecer enseñanza e investigación a todos los alumnos financiados, así como una oferta competitiva para gran parte de la financiación; alta en el caso del UFC y marginal en el caso de PCFC. Los organismos gubernamentales y de gestión serán los responsables de la puesta en práctica de los planes institucionales y la provisión de la oferta educativa. La competencia por la captación de alumnos mantenida entre las instituciones también se ha visto estimulada por el sustancial incremento de las tasas de matrícula para los primeros ciclos, habiéndose producido una considerable reducción en los contratos de financiación de las instituciones.

Es interesante observar que tanto el PCFC como el UFC financiarán contratos educativos (*full time equivalent*, FTE), en lugar de fijar su atención en determinados resultados —graduados— de un nivel de calidad específico. En su documento de consulta *Funding Choices*, el PCFC se refiere a los contratos como «acuerdos escritos entre el Consejo y cada una de las instituciones, que especifican con claridad la provisión educativa que la institución ha aceptado ofertar a cambio de los fondos del PCFC» (1989:26). El UFC ha invitado a las instituciones a presentar sus ofertas: «para ofrecer en el período 1994-95 y para cada especialidad académica, números concretos de plazas de primer y segundo ciclo (enseñanza e investigación)» (Circular 39/89). Probablemente, este énfasis puesto en el proceso de financiación se deba a que el Secretario de Estado, en sus circulares orientativas a los Consejos, define la *rendición de cuentas* en términos de oferta de provisión más que en términos de oferta de resultados. Ello debe influir en los puntos de vista de los Consejos de Financiación acerca del proceso de rendición de cuentas y del papel de los indicadores de rendimiento y mantenimiento de la calidad como parte de dicho proceso.

Además, la filosofía que subyace no es simplemente la de emplear el dominante poder de adquisición para conseguir docencia e investigación al menor coste, sino la de asegurar la rentabilidad en términos de las *Es*, incluyendo responsabilidad, relevancia, innovación, creatividad, selectividad, elecciones difíciles, calidad y excelencia. El documento del PCFC *Funding Choices* ponía a consideración la validez de recompensar la calidad, permitiendo tasas diferenciales o redistribuciones diferenciales de alumnos. Una cuestión clave que se plantea es la siguiente:

¿Debe el PCFC pagar precios más altos por una calidad más elevada del proceso, o bien, bajo un modelo de economía de mercado esta elección debe ser realizada por los consumidores? ¿Deben decidir los consumidores si pagar un elevado precio, debido un alto valor añadido, para completar un curso satisfactoriamente, p. ej., pagar un precio más elevado por la calidad de los resultados o logros más que por la calidad del proceso?

El UFC ha invitado a las instituciones a presentar sus ofertas frente a un precio orientativo por FTE para cada uno de los grupos de disciplinas académicas, pero, en igualdad de circunstancias, (bonita e imprecisa frase) favoreciendo a las universidades que presenten tasas más bajas. El Consejo es responsable de asegurar el mayor incremento posible del número de alumnos de primer ciclo, que sea coherente con el mantenimiento de los niveles académicos. Así pues, la circular (39/89) invitando a las ofertas sugiere:

«En algunas (si no en todas) las materias, las universidades pueden desear tomar ventaja de la puja a la baja, mejorando así las oportunidades de que sus ofertas se acepten al completo, al estipular unas tasas inferiores a las orientativas.»

Estas tasas reducidas pueden ser aplicadas a una parte o a la totalidad de los alumnos. El Consejo está dispuesto a considerar ofertas de tasas por encima del precio orientativo, pero sólo en circunstancias especiales. Sin embargo, la puja de tasas por sí sola no puede ser el único determinante de las decisiones de financiación. El Consejo deberá tener en cuenta la medida en la que los planes futuros sean viables a nivel de financiación y a nivel de satisfacer los objetivos del propio Consejo. La financiación de la investigación continuará siendo altamente competitiva en base a la calidad, pero tenderá también a asegurar adecuadas oportunidades de investigación para los jóvenes académicos que todavía no hayan consolidado una cualificada reputación. Se espera de las universidades que adapten sus planes, cuando sea necesario, con el fin de contribuir a los objetivos y actividades del UFC (como queda reflejado en la circular 19/89), lo cual será tenido en cuenta en la toma de decisiones de financiación para el período 1991-95.

Cualquiera que sea la metodología, está claro que un factor determinante de la eficacia institucional es la *calidad*, tanto en lo que respecta a la docencia como a la investigación. Todo ello implica que la evalua-

ción de la conducta tanto individual como departamental e institucional, no sólo se haya convertido en una importante dimensión de la gestión departamental e institucional, sino que alcanzará aún mayor importancia con el paso del tiempo. Desarrollos muy similares se están produciendo en Australia bajo el gobierno Laborista (Dawkins, 1988). Numerosos países europeos están disminuyendo los controles centralizados sobre los sistemas de educación superior al tiempo que desarrollan estrategias de auto-regulación (Van Vught, 1988), basándose en la premisa de que el incremento de la autonomía institucional aumentará a su vez la eficacia de la instituciones en cuanto a la respuesta ante las cambiantes necesidades de la sociedad en general y de la economía en particular. Esta tendencia hacia una creciente autonomía y diversidad institucionales, como ocurre en el Reino Unido y Australia, también ha generado demandas de una mayor rendición de cuentas, especialmente en lo que respecta al control de la calidad y a la eficacia institucional, por medio de estrategias y planes de acción con resultados positivos. Es importante reconocer que los puntos de partida para incrementar la autonomía institucional fueron básicamente distintos en los Países Bajos y en el Reino Unido, dado el consolidado papel amortiguador del UGC.

4. Indicadores de rendimiento públicamente disponibles

Los indicadores de rendimiento públicamente disponibles, como p. ej., declaraciones, generalmente traducidas a cifras, de los recursos empleados y de los logros de objetivo concretos en áreas relevantes de la empresa (CVCP, 1987: 3), son datos fundamentales en el proceso de la rendición de cuentas, tanto a nivel interno dentro de la institución como externo para los organismos de financiación, los gobiernos y la sociedad. En un modelo de economía de mercado debieran constituir una importante fuente de conocimiento de mercado para los consumidores. El sector universitario ha tomado la delantera en el desarrollo de una serie de indicadores de rendimiento y estadísticas de gestión, amplios y disponibles al público. El PCFC ha establecido un equipo de trabajo para desarrollar una serie de indicadores de rendimiento para su sector, pero se convertirá en un lento proceso dada la carencia de una amplia base de datos.

El trabajo conjunto del Comité Rector UFC (previamente UGC/CVCP), en el desarrollo de los indicadores de rendimiento ha sido descrito ampliamente en numerosas ocasiones (Sizer, 1990). Sin embargo,

es importante recalcar que en el Reino Unido los avances en el desarrollo y la utilización de los indicadores de rendimiento surgen de las conclusiones aportadas por el Informe Jarrat (CVCP, 1985: 22):

«Existe una reconocida necesidad de indicadores de rendimiento fiables y consistentes. Estos deben desarrollarse urgentemente para las universidades en su conjunto y para cada una de ellas en particular, como parte integral de su proceso de planificación y distribución de recursos.»

Esta conclusión originó la recomendación conjunta de la UGC y del CVCP (CVCP, 1985: 36):

«Deben ser desarrollados una serie de indicadores de rendimiento que cubran entradas y salidas (*Input/output*) y que estén diseñados tanto para ser usados en las universidades a nivel individual como para establecer comparaciones entre las instituciones...»

y a la recomendación de que todas las universidades desarrollen *indicadores de rendimiento fiables y consistentes*. Aunque se ha producido una mezcla de estimulación y de presión por parte del gobierno y del UGC/UFC, ha existido desde el comienzo una *capitalización* del trabajo del Comité Rector por parte de los Vice-Cancilleres, y una extensiva asesoría de las universidades en todas las etapas del proceso. Los Vice-Cancilleres han dirigido a los comités, sin que haya existido ningún mandato exterior por parte de organizaciones coordinadoras centralizadas. Esta situación presenta un duro contraste con el fracaso del Ministerio de Educación y Ciencia holandés de asegurar o bien la aceptación institucional de una serie de indicadores de rendimiento propuestos por él mismo o, bien de promover un interés institucional para la elaboración de un enfoque conjunto (Mertens y Bormans, 1990; Vroeijerstijn y Acherman, 1990).

La primera serie de *Estadísticas de gestión e indicadores de rendimiento universitarios* se publicó en 1987, apareciendo después anualmente otras series con indicadores adicionales (CVCP/UGC, 1987, 1988, 1989). La edición de 1989 puede conseguirse en disco. En la lista inicial de 39 indicadores de rendimiento, extendida a 54 en 1988 (véase el cuadro 1) el Comité Rector conjunto CVCP/UFC se ha concentrado en la entrada, en el proceso y en sólo dos series de mediciones de la producción/resultados (destinos de los graduados después de seis meses y tasas de egresados) ya que todavía existe una considerable y árdua tarea en lo que se refiere a la medición de la eficacia, especial-

mente en el área de la investigación, de la calidad docente y de los impactos longitudinales de los resultados de la enseñanza. Así, pues, los indicadores del rendimiento son más útiles en la evaluación de la eficiencia que de la eficacia, pero, por supuesto, una institución sólo puede ser eficiente si es eficaz.

CUADRO 1.

Información de la gestión e indicadores de rendimiento

(C: coste del centro, I: Institución, S: grupo de disciplinas académicas)

1. Gasto por estudiante FTE	C
2. Gasto por personal académico FTE	C
3. Gasto por personal auxiliar del personal académico FTE	C
4. Gasto de equipamiento para el personal académico FTE	C
5. Salarios por investigación del personal académico FTE	C
6. Posgraduados dedicados a la investigación como % de los estudiantes FTE	C
7. Posgraduados a los que se imparte docencia como % de los estudiantes FTE	C
8. Total de posgraduados como % del total de estudiantes FTE	C
9a. Ratio profesor/alumnos (docencia)	C
9b. Indicador 9a: período 1984-85 a 1986-87	C
10. Gasto de la administración central como % del gasto total	I
11. Pagas de la administración central como % del gasto de la administración central	I
12. Gasto de la administración central por estudiante FTE	I
13. Gasto de la administración central por personal académico FTE	I
14. Gastos por biblioteca como % del gasto general	I
15. Gastos en publicaciones como % del gasto en biblioteca	I
16. Gastos por pagas de biblioteca como % del gasto por biblioteca	I
17. Gastos de biblioteca por estudiante FTE	I
18. Gastos de biblioteca por personal académico FTE	I
19. Gastos en libros por estudiantes FTE	I
20. Gastos en publicaciones periódicas por estudiante FTE	I
21. Gastos por servicios informatizados como % del gasto total	I
22. Gastos en pagas por servicios informatizados como % del gasto total en informática	I
23. Gastos en informática por estudiante FTE	I
24. Gastos en informática por personal académico FTE	I
25. Gasto total por instalaciones como % del gasto general	I
26. Gasto por pagas de instalaciones como % del gasto por instalaciones	I
27. Calefacción, agua, electricidad como % del gasto general	I
28. Servicios de limpieza y mantenimiento como % del gasto general	I
29. Reparaciones y mantenimiento como % del gasto total	I
30. Gasto por teléfono como % del gasto total	I
31. Gasto total por instalaciones por estudiante FTE	I
32. Gasto por pagas de instalaciones por estudiante FTE	I
33. Gastos de calefacción, agua, electricidad por estudiante FTE	I
34. Servicio de limpieza y mantenimiento por estudiante	I
35. Gastos de reparaciones y mantenimiento por estudiantes	I

CUADRO 1.

Información de la gestión e indicadores de rendimiento (Continuación)

(C: coste del centro, I: Institución, S: grupo de disciplinas académicas)

36. Gastos de teléfono por estudiante FTE	I
37. Gastos en servicios profesionales por estudiante FTE	I
38. Gastos en sindicatos y sociedades estudiantiles por estudiante FTE	I
39. Destino de los graduados de 1.º ciclo a los seis meses	S
40. Número total de graduados con destino conocido	I,S
41. Desempleo o empleos temporales	I,S
42. Cifras estimadas del indicador 41	I,S
43. Diferencia entre el indicador 41 y 42	I,S
44. Diferencias en %	I,S
45a. Destino de los graduados de 1.º ciclo - desempleados o con empleo temporal	S
45b. Número de egresados	I,S
46. % de los que finalizan sus estudios	I,S
47. Duración del curso	I,S
48. Trimestres de asistencia por éxito?	I,S
49. Años como % de la duración del curso	I,S
50. Calificaciones de ingreso: nivel sobresaliente no's	I,S
51. Calificaciones de ingreso: puntuaciones de sobresaliente	I,S
52. Calificaciones de ingreso: Scottish Highers no's	I,S
53. Calificaciones de ingreso: puntuaciones Scottish Highers	I,S
54. Calificaciones de ingreso: otras calificaciones	I,S

El sub-comité conjunto CVCP/UFC para los indicadores de investigación se ha centrado en métodos bibliométricos de evaluación a nivel de coste del centro o de la materia dentro de las universidades. Inicialmente, un estudio piloto deseaba desarrollar la categorización de publicaciones más adecuada para cuatro disciplinas: física, química, historia y economía; planificando comprobar posteriormente la metodología empleada con otras cuatro disciplinas. Una metodología similar esta siendo comprobada en Australia. El objetivo final es el establecimiento de una flexible base de datos de las publicaciones para cada materia y en cada universidad. El sub-comité ha reconocido que la cuantía de las publicaciones aportan relativamente escasa información acerca de la *calidad* de las publicaciones y que el análisis de la citación puede ser realizado en base a tres contextos distintos:

- revisando la provisión de disciplinas individuales, como en el caso de parte de la reciente revisión de las Ciencias de la Tierra por parte del UGC;

- realizando evaluaciones periódicas de la actuación departamental, bien externamente por parte del UFC y los Consejos de Investigación, o bien internamente por parte de los Comités de Recursos y Planificación institucionales o el Comité de Investigación; y
- como parte de series regulares (anuales) de estadísticas de gestión e indicadores de rendimiento.

El sub-comité publicó un documento consultivo *Aspectos de la evaluación cuantitativa de la investigación departamental*, en el que se plantea la medida en que puede utilizarse el análisis de citas como sustituto de la calidad, y concluye diciendo que queda abierto a la discusión el tema de si las percepciones generadas por el análisis de citas valen la pena el esfuerzo y el coste de recoger y analizar los datos. Además, se debe ser precavido al distinguir entre *calidad, importancia e impacto*, al plantear el análisis de citas. El sub-comité no está convencido de que los costes ocasionados por la publicación anual de series de datos concernientes a cada materia, en todas y cada una de las instituciones, queden compensados por los beneficios derivados de la disponibilidad de dichas series. A pesar de que los resultados de la consulta todavía no se hayan hecho públicos, puede afirmarse que en las universidades británicas existe una falta de confianza en el empleo de los análisis de citas como sustitutos de la calidad. Se está preparando un informe sobre los resultados de la consulta para las universidades, para el CVCP y el UFC. Ello no significa que el análisis de citas no tenga su papel de información en la formulación de la política científica, dentro del DES, el Consejo Asesor para la Ciencia y la Tecnología (ACOST), la Junta Asesora de los Consejos de Investigación (ABRC) y los Consejos de Investigación.

Así pues, el trabajo realizado hasta la fecha por el Comité Rector conjunto CVCP/UFC demuestra que, mientras que la publicación de los indicadores de rendimiento aportan una contribución significativa en el desarrollo del proceso externo de rendición de cuentas por lo que respecta a la eficiencia, así como en la gestión interna de las instituciones, resultan más limitada en lo que respecta al proceso externo de rendición de cuentas, para una gestión eficaz de los fondos de investigación y para el control de la calidad docente. Sin embargo, tal y como han reconocido el PCFC y el UFC, los indicadores de rendimiento serán de utilidad para los organismos de financiación en lo que respecta a la supervisión y realización de los contratos de financiación y planes insti-

tucionales, especialmente teniendo en cuenta que cada universidad ha sido requerida para establecer sus propios objetivos en términos cuantitativos, siempre que sea posible (Circular del UFC 20/89). Pueden desarrollarse medidas progresivas del rendimiento para constatar los resultados con estos objetivos.

5. Utilización institucional de los indicadores de rendimiento

Dentro de las instituciones, la planificación y la distribución de los recursos incluye la elección entre alternativas competitivas y mutuamente excluyentes, cada una de las cuales con su propia combinación de *inputs*, *outputs*, resultados, impactos y beneficios. Los indicadores de rendimiento y las estadísticas de gestión deben informar las decisiones; proporcionan un punto de partida desde el que deben emitirse los juicios académicos y gestores, pero, cuanto menos numerosos sean los indicadores mayor será el espacio que dichos indicadores deben cubrir. Sin embargo, para gran parte de las decisiones políticas y estratégicas, el vacío que debe cubrirse seguirá siendo considerable.

No existe ninguna evidencia para sugerir que los juicios se basen primordialmente en los indicadores de rendimiento. Una vez tomadas las decisiones, una comparación entre los indicadores de rendimiento esperados y los reales debe formar parte de la evaluación del éxito de anteriores decisiones. Las gestiones ejecutivas, experimentadas y eficientes no se basarían en una utilización simplista de los indicadores de rendimiento mientras que las gestiones débiles e ineficaces no podrían hacerlo.

Las decisiones para las que se requiere la información aportada por los indicadores de rendimiento y estadísticas de gestión pueden ser de naturaleza estratégica o a largo plazo, o bien de naturaleza táctica o a corto plazo. Pueden ser a nivel institucional, de facultad o de departamento. En su petición de planes académicos de 1986, el UGC (Circular 12/86), afirmaba esperar que los planes académicos fueran dinámicos en lugar de estáticos y que identificaran los cambios de los planes de docencia e investigación de los departamentos (u otras unidades organizativas) así como las necesarias acciones para llevarlos a cabo. Se aportó una lista de factores a tener en cuenta, incluyendo la propia evaluación de la institución de las posibilidades docentes e investigadoras de los distintos departamentos u otras unidades organizativas (Cir-

cular 12/86). Se esperaba que las instituciones emplearan indicadores de rendimiento —y así lo hicieron—. Algunas de ellas emplearon los modelos *Policy Directional Matrix* al llevar a cabo una evaluación comparativa de los departamentos. (West, 1986: CUA/APA, 1990). Estos modelos se centran en las fuerzas relativas de cada área disciplinar en comparación con otras instituciones y en el atractivo futuro de dicha área para la misión y estrategia de la institución (Sizer, 1982).

Futuros análisis, estrategias institucionales y planes operativos deben conducir a los objetivos y planes de acción de los departamentos. Estas estrategias y planes departamentales a largo plazo deben ser trasladados a planes de acción y presupuestos anuales. A la *rendición de cuentas* le concierne si el responsable del centro, ya sea éste un departamento académico o de servicios, ha llevado a cabo satisfactoriamente el plan al que iban destinados los recursos, p. ej., con lo que se ha conseguido y no con lo que se ha gastado. Los indicadores de rendimiento pueden emplearse tanto en la definición de lo que se planifica alcanzar como en la medición de los logros ¹.

Se ha observado en todas partes que la gestión de algunas instituciones parece confundir la distribución de recursos con la delegación del presupuesto, p. ej., vivir dentro de las propias posibilidades (UGC, 1987). El CVCP también parece tener esta estrecha visión de la rendición de cuentas en sus orientaciones a las universidades, referentes a la estructura y la práctica de la gestión de financiación universitaria (CVCP, 1988). Hace distinción entre dos procesos complementarios de *planificación y distribución de los recursos y rendición de cuentas y control*, y recomienda que éste último debe recaer bajo la responsabilidad del Delegado de Financiación y del Comité de Financiación. La rendición de cuentas y el control financiero constituyen solamente un aspecto de la rendición de cuentas sobre el rendimiento.

En la historia de la evaluación del rendimiento, numerosas instituciones han empleado indicadores de rendimiento cualitativos y cuantitativos como parte de una planificación académica a largo plazo, pero dichos indicadores todavía necesitan convertirse en un sistema de me-

1. Para una mayor información acerca de la experiencia británica en el uso de indicadores de rendimiento para la formulación de estrategias y planes en la educación superior, así como en la posterior evaluación del progreso de su implantación, véase el reciente informe del Grupo de Estudio CUA/APA (CUA/APA, 1990).

didias progresivas del rendimiento, que formen parte de un sistema delegado de administración del presupuesto y rendición de cuentas a nivel de responsabilidad de centro. Se ha constatado una mayor influencia de las medidas de ejecución frente a las de progreso en las decisiones acerca de los futuros presupuestos, p. ej., cuantificación económica de los planes de acción que hacían referencia a las estrategias necesarias para consecución de los objetivos. Las universidades son lentas en reconocer que a la rendición de cuentas le compete lo que se ha conseguido y no lo que se ha gastado, y que los indicadores de rendimiento cualitativos y cuantitativos pueden emplearse tanto para definir lo que se planifica alcanzar como en la medición del rendimiento. Como refleja el cuadro 2, numerosas universidades aún tienen que cumplimentar las recomendaciones del Informe Jarratt, particularmente en lo que respecta al desarrollo y utilización de los indicadores de rendimiento. La rendición de cuentas también implica incentivos, tanto a nivel institucional como departamental, así como recompensas por la buena actuación y sanciones cuando ésta es insatisfactoria. Se ha observado anteriormente que el UFC intenta tener en cuenta el rendimiento en sus decisiones de financiación. Todo ello lleva a plantearse la cuestión de si se ha considerado suficientemente el cambiante rol de los directores de departamento y a sus incentivos, recompensas y desarrollo profesional, así como a los programas de desarrollo de la gestión (Sizer, 1988).

Deben ser reconsiderados con más detalle las responsabilidades de los directores de departamento concernientes a los indicadores de rendimiento. Al igual que ocurrió en el caso de los departamentos clínicos y sus respuestas a los intentos de introducir la gestión de su presupuesto en el Servicio Nacional de la Salud, existe el gran peligro de que los directores de departamento no *vean* que los indicadores de rendimiento departamentales están dirigidos primordialmente para ayudarles a gestionar sus departamentos, para informar sus decisiones y para evaluar su rendimiento. Puede ocurrir que no perciban la *propiedad* del sistema de indicadores de rendimiento, por lo que pueden llegar a considerarlos más como una traba que como una ayuda. Ello se debe a que se produce un conflicto entre el empleo de los indicadores de rendimiento por parte de la gestión institucional en la evaluación de la rendición de cuentas de los directores de departamento en lo que respecta a la docencia, investigación y servicios, y su empleo por parte de los directores de departamento en la gestión interna de sus actividades docentes y de investigación. En el momento presente el énfasis se sitúa en la *rendición de cuentas* más que en su utilización como *herramienta* por parte del

CUADRO 2.
Realización de las recomendaciones del informe Jarratt

(% de respuestas positivas)

<i>I Reducción del tamaño, número y frecuencia de las reuniones de comité</i>	
Revisión de los comités a partir del Informe Jarrat	80
Revisión anual, o de otro tipo, de los comités	30
<i>II Introducción de la delegación del presupuesto</i>	
Progreso realizado a partir del Informe Jarrat	47
Posterior delegación del desarrollo planificado/progresivo	35
Rendición de cuentas a nivel delegado	34
<i>III Costes añadidos y empleo de indicadores de rendimiento</i>	
Indicadores de rendimiento desarrollados y en uso	14
Indicadores de rendimiento en desarrollo	11
<i>IV Cambios en la responsabilidad ejecutiva</i>	
Jefes de departamento con rol y responsabilidades definidas	42
Jefes de departamento con cargo directivo	27
Jefes de departamento responsables de la actividad	23
Jefes de departamento responsables de la gestión de recursos	23
<i>V Creación de un comité de planificación y de recursos</i>	
Establecimiento de un comité de planificación y recursos (PRC)	64
El PRC coordina la planificación académica y económica	45
El PRC es responsable de la planificación corporativa y estratégica	50
El PRC gestiona la distribución de recursos	26
El PRC emprende revisiones especiales, p. ej., planificación del espacio	11
<i>VI Evaluación y desarrollo del personal</i>	
Introducción del esquema de evaluación del personal	24
Introducción planificada o piloto de la evaluación del personal	72
Mayor énfasis en la formación del personal	39
Designación de nueva formación y desarrollo del personal	16

director de departamento. *Los indicadores de rendimiento son mejor interpretados y utilizados por aquellos directores que pueden influir en sus resultados.* No debe prevalecer su utilización, como parte del proceso de rendición de cuentas, en la consecución de los objetivos para los que se destinó el presupuesto. El propósito fundamental de la planificación presupuestaria, el sistema de control, la información para la gestión y los indicadores de rendimiento es que sirvan de ayuda al responsable del presupuesto en la gestión de su propia actuación. Así pues, éste debe poseer la *propiedad* del sistema.

6. Planificación de la investigación institucional, distribución selectiva de recursos y rendición de cuentas

Una de las razones por las que resulta difícil la elaboración de mediciones significativas, dignas de ser publicadas, de la investigación de una institución, es por que la evaluación de la investigación, al igual que ocurre con la evaluación de la calidad docente, debe iniciarse a nivel de individuos y grupos de un departamento (o unidades organizativas) y desarrollarse a través de la evaluación departamental y de la facultad. Por motivos de planificación interna de la investigación, distribución selectiva de recursos, supervisión y rendición de cuentas, es necesario elaborar perfiles individuales, de grupos de investigación y departamentales, referentes a la naturaleza de la actividad investigadora que se está llevando a cabo en un departamento u otra unidad organizativa, e identificar, con la colaboración de todos los miembros, aquellos indicadores de rendimiento que sean relevantes para la evaluación de su actividad investigadora. Puede orientarse a los departamentos en los posibles criterios de evaluación y relevantes indicadores de rendimiento, en cada una de las etapas de la actividad investigadora, p. ej., entrada, proceso, resultado, logro, beneficio e impacto (Sizer y Frost, 1985). Los ejercicios de selectividad de la investigación y los modelos de financiación del UGC/UFC, juntamente con la presión ejercida en las universidades para ser tan selectivas a nivel interno como lo son a nivel externo los organismos de financiación, han dado lugar al reconocimiento de la necesidad de elaborar perfiles de investigación departamentales e indicadores de rendimiento de la investigación. Estos se han visto reforzados posteriormente por el reciente aviso realizado por el UFC de que será necesaria una futura evaluación de la investigación, probablemente hacia 1993-94.

En 1986, el UGC accedió supervisar los planes de las universidades referentes a la elaboración y puesta en práctica de una distribución selectiva de los recursos destinados a la investigación, así como informar anualmente al ministerio sobre sus actividades. Existe, por supuesto, cierto paralelismo con el sistema de financiación condicional de la investigación llevado a cabo en los Países Bajos (Maassen, 1989). El Segundo Informe Anual (UGC, 1988) indicaba que, a pesar de las restricciones económicas, por término general y a partir de 1985 las universidades habían experimentado un considerable avance en la planificación de la investigación y en el empleo selectivo de recursos para la investigación y que se produjo un considerable ímpetu de desarrollo

dentro del sistema. Sin embargo, en numerosas instituciones la selectividad interna de los departamentos y las áreas disciplinares no parecían estar completamente integradas en los mecanismos que produjeron la planificación, la distribución de recursos y la rendición de cuentas en un proceso corporativo, que uniera los aspectos académicos económicos y físicos con el objetivo de asegurar el empleo económico, eficiente y eficaz de los recursos destinados a la investigación. Pocas instituciones demostraron haber desarrollado mecanismos por medio de los cuales el Comité de Planificación y Recursos, asistido por el Comité de Investigación, estaba en posición de supervisar los escasos recursos destinados a la investigación, así como de designar a los directores de departamento como responsables de la realización de los planes de investigación del departamento que habían acordado con el Comité de Planificación y Recursos. Existía evidencia suficiente para afirmar que los mecanismos de planificación de la investigación estaban menos desarrollados que la práctica de una distribución selectiva de los recursos para la investigación, especialmente dentro de los departamentos.

Como se ha mencionado anteriormente, un punto de partida en el desarrollo de los planes de investigación es la elaboración de *perfiles de investigación departamentales*. Con el fin de desarrollar una visión general del empleo de recursos en la investigación y también en la docencia, será necesario elaborar sistemas de rendición de cuentas orientados a los resultados, similares a los que suelen ser supervisados por el Servicio Nacional de la Salud. Este hecho origina complejos problemas de costes de producto que no pueden resolverse fácilmente. Sin embargo, si esto no se soluciona se hace muy difícil la producción de procesos de docencia e investigación así como de indicadores de resultados. El coste de los indicadores FTE incluidos en el informe *Estadísticas e Indicadores de Gestión Universitaria* del CVCP/UGC está distorsionado porque relacionan al alumno FTE con la investigación, las becas y costes de docencia. Es interesante resaltar que, dado el estado del coste de las universidades, el UFC (Circular 39/89) ha declarado que:

«Las tasas por alumno ofertadas por cada universidad durante el período de planificación representarán lo que cada universidad considera como económico para sí misma y competitivo respecto a las ofertas de las restantes universidades. Así pues, reflejarán la declaración por parte de la universidad de su coste docente.»

A petición del Secretariado de Estado, el UFC separará la financiación de la docencia y la investigación a partir de 1990-91. No deja de

sorprender que el UFC haya informado a las universidades de que la intención de separar la financiación de la docencia y la de la investigación no dictaminará el modo cómo éstas deben emplear el conjunto de la subvención o cómo deben rendir cuentas del mismo. Kenneth Baker (1988) consideró como una importante ventaja el hacer una distinción más clara entre la financiación de la investigación y la de la docencia; siendo el principal objetivo el que el dinero destinado a la investigación se gaste precisamente en investigación, lo cual demuestra dar buenos resultados. Baker reconocía que existe un problema en la distinción realizada por los académicos entre el tiempo destinado a la docencia y el destinado a la investigación. Consideraba que una financiación específica animará a las universidades a considerar más sistemáticamente lo que exigen a cada miembro del personal académico, lo cual no puede ni debe convertirse en un complicado sistema de rendición de cuentas. John MacGregor (1989), actual secretario de Estado, reafirmó este punto de vista al decir al CVCP que:

«Sería deseable tener revisiones anuales de los gastos de investigación, pero el UFC no pretenderá una absoluta meticulosidad en términos de rendición de cuentas, lo que por supuesto nadie cree que fuere posible, pero existen múltiples motivos por los que el UFC debe estar interesado en si su propia evaluación de la ayuda que debe conferirse a la investigación se adecua con el modo en que los recursos se aplican activamente en las instituciones. Personalmente opino que si existe una ausencia de este requisito no puede hablarse de una visión razonable de la rendición de cuentas de los gastos públicos.»

7. Control de la calidad docente, juicios comparativos acerca de la calidad docente y comités para los contratos de financiación

Uno de los peligros de un apoyo selectivo a la investigación, de la potenciación de los mecanismos para la gestión docente y la rendición de cuentas, es que se produce un mayor énfasis en la investigación a expensas de la docencia. Sucesivos ministros de Educación y Ciencia han considerado que las universidades necesitan asignar a la docencia el mismo prestigio que tradicionalmente han asignado a la investigación. El Consejo para la Industria y la Educación Superior ha llegado a la conclusión de que la industria valora la enseñanza excelente por encima de la buena investigación.

El DES no está convencido de que se preste la misma atención a la calidad docente en las universidades que en los politécnicos. Existe una mayor tradición de *la calidad docente* en los politécnicos, en parte como resultado de los procedimientos de acreditación del Consejo Nacional de premios Académicos (*Council for National Academic Awards*, CNAA) y de la labor del Inspectorado de Su majestad (Her Majesty's Inspectorate, HMI) (Perry, 1987), pero también en parte porque la investigación no es una preocupación primordial de la mayoría del personal académico. Una alta calidad de la docencia es fundamental y se considera como capital en cuanto a su misión y papel a desempeñar. Será interesante comprobar si la alta calidad de la docencia se convierte en una de las víctimas del proceso competitivo de presupuestos del PCFC y que algunos consideran como una persecución *suicida* del número de alumnos por parte de algunos politécnicos y directores de facultades —especialmente al extenderse como parece el punto de vista de que el precio y la demanda eran los criterios clave empleados por el PCFC en la distribución de las plazas de alumnos en su primera ronda presupuestaria, y que se concedía una menor prioridad a la calidad—. Todo ello reflejaba la limitada capacidad del CNAA y del DES, y, en el caso del CNAA, de la buena voluntad de asesorar acerca de la calidad comparativa.

La principal contribución del CNAA al control de la calidad ha sido el asegurar que los politécnicos y *colleges* desarrollen mecanismos de supervisión y control de la calidad de los cursos. El CNAA no se dedica a la evaluación comparativa de la calidad sino al control de la calidad. Como reconoce el documento *Elecciones de Financiación* del PCFC, si el destino de los informes de HMI es el informar las evaluaciones sobre calidad docente del PCFC, así como el de influir en las elecciones de los alumnos, no solamente debe dárseles una mayor publicidad y apoyo sino que debe subsanarse la falta de publicaciones. El Inspectorado necesitaría un apoyo bastante considerable (véase Booth y Booth, 1989).

Los ministros han ejercido una presión duradera sobre las universidades para que concedan una mayor atención a la calidad docente. El CVCP (1986) ofreció asesoramiento y orientación a las universidades en lo referente al mantenimiento y supervisión de la calidad y los niveles académicos. Se hizo cargo de la supervisión de la medida en que las universidades seguían su consejo, de la publicación de informes en 1987 y 1988 sobre los cambios acaecidos y de ofrecer un posterior asesoramiento si fuera necesario. También aceptó sistemas más efecti-

vos de formación y evaluación del personal y promover el desarrollo de indicadores de rendimiento sobre la calidad docente. En 1988 estableció un grupo de trabajo bajo la dirección del profesor Stuart Sutherland, con el fin de crear una Unidad de Estándares Académicos, con la finalidad de supervisar la puesta en marcha de procedimientos para el control de la calidad docente por parte de las universidades. Esta acción fue considerada por algunos como un mecanismo de defensa encaminado a desanimar al DES a enviar HMI a las universidades. Existe cierto paralelismo con lo sucedido en los Países Bajos. El CVCP (1989) ha aceptado las recomendaciones del Comité Sutherland y se ha puesto en marcha una Unidad de Auditoría Académica (*Academic Audit Unit, AAU*).

El UGC no encontró un medio de evaluar la calidad docente a nivel nacional. Si el UFC debe tener en cuenta la calidad docente al establecer contratos y al supervisar su desenlace satisfactorio debe también desarrollar una metodología adecuada. ¿Debe apoyarse en el sector universitario, como ocurre en los Países Bajos, para crear sistemas en los que pueda confiar y que merezcan también la confianza del público, sin el refuerzo de una inspección universitaria existente en los mencionados Países Bajos? Se trata de una cuestión clave en lo que respecta a la rendición de cuentas de la calidad docente. El AAU puede supervisar los procedimientos empleados en las universidades en la auto-evaluación y la revisión de programas, pero no puede esperarse que evalúe la calidad comparativa de la docencia en cada una de las áreas disciplinares de todas las universidades. Existe una importante distinción entre un *mínimo control de la calidad* y *juicios comparativos de la calidad*.

Si se producen ofertas excesivas al UFC por parte de las universidades para estudiantes de negocios y gestión, como es natural que ocurra, puede al AAU informar que la calidad docente de la universidad *X* es superior a la de la universidad *Y*? ¿Puede una unidad de este tipo asesorar acerca de si la calidad docente se verá perjudicada si se acepta una puja a la baja de la universidad *A*? ¿Puede asesorar sobre el intercambio entre calidad y precio si la oferta de la universidad *A* ha de compararse con las pujas y al máximo precio de las universidades *X* e *Y*, y la oferta de la universidad *B* de una graduación intensiva de dos años de duración? Ciertamente no, especialmente si se tiene en cuenta que el PCFC reconoce que el HMI no puede aportar una asesoría de este calibre sin disponer de recursos adicionales y que, la mayoría de

los que respondieron al documento *Elecciones de Financiación* estuvieron de acuerdo en que la calidad debía ser un elemento importante a considerar en la financiación pero que mantenían sus dudas acerca de cómo podía llevarse a cabo con credibilidad (por ejemplo, Booths y Booths, 1989). Estas dudas parecen haber sido reforzadas por la experiencia del primer ejercicio de ofertación. También resulta extremadamente dudoso si la unidad sería capaz de confirmar si se ha ofrecido la calidad que se había contratado. Por consiguiente, tal y como se reconocía en el documento consultor del UFC, deberá establecerse un adecuado mecanismo para asesorar sobre estos temas, pero ¿dispondrá el UFC de recursos suficientes para ello? Ya que el sub-comité del UGC fue incapaz de elaborar tales juicios, es muy poco probable puedan lograrlo asesores especializados a nivel individual. Puede que comités *ad hoc*, dirigidos por asesores especializados similares a los utilizados en el ejercicio de investigación, sean capaces de discernir si una universidad está capacitada para cumplir con la provisión que ha ofertado, pero es dudoso que puedan establecer juicios comparativos fiables acerca de la calidad docente. Además, ¿cómo puede asegurarse el UFC de que se ofrece la calidad que ha contratado? Un adecuado mecanismo podría a su vez entorpecer el deseo del UFC de que exista la mínima participación burocrática del Consejo en los asuntos de las instituciones.

Otra interesante cuestión que se plantea es la siguiente: ¿en qué tipo de contratos participan las instituciones cuando aceptan otro tipo de alumnos con tasas reducidas, p. ej., alumnos extranjeros con tasas reducidas o alumnos que solamente pagan la tasa de matrícula? ¿Cuáles son las implicaciones de la oferta de la calidad a los alumnos de los contratos de financiación y, quién asegura a los alumnos y a los consejos de financiación que la calidad no se verá mermada por el excesivo número de alumnos? No es de extrañar que el UFC haya reconocido que los ejecutivos deben aplicarse más detalladamente en la determinación de los procesos y procedimientos que serán incluidos en la puesta en práctica de los principios de financiación adoptados por el Consejo. Se sugirió (Sizer, 1989a) que un camino a seguir podría ser el de adjudicar a las instituciones, especialmente a aquellas que ofrecen pujas inferiores o inusuales, así como aquellas que admiten un considerable número de alumnos con sólo tasas de matrícula o extranjeros, la responsabilidad de demostrar su adecuación a los objetivos de la oferta, p. ej., que la puja por la calidad, y por consiguiente la calidad contratada, se han visto cumplidas. La auto-responsabilidad implica la auto-regulación

(Horton, 1990). De hecho, la invitación de la UFC a aportar ofertas específica que:

«Cuando una universidad está ofreciendo plazas de alumnos por debajo del precio orientativo, la planificación debe contar con la suficiente información para convencer al Consejo de que se están realizando las medidas necesarias en cuanto al asesoramiento, mantenimiento y potenciación de los niveles académicos.»

También se requiere de las universidades que describan su política docente y discente en su Informe de Planificación, incluyendo los niveles académicos y la calidad docente. Todavía queda por resolver el problema de los juicios comparativos sobre la calidad de áreas como Negocios y Gestión.

8. Indicadores de rendimiento y tradición de la calidad docente

Si las instituciones han de ejercitar al máximo su autonomía, la responsabilidad de la calidad docente deberá ser gestionada por los directores de departamento, como gestores responsables, como parte integrante de la formación del personal y de la evaluación y rendición de cuentas. El Vice-Canciller/Director ostenta la importante responsabilidad de establecer en la institución una *tradición de calidad docente* que incluye revisiones periódicas de la estructura y contenido de los cursos, métodos docentes, cantidad y calidad de los recursos docentes, calidad del curso, formación del personal y procedimientos de evaluación, grado de satisfacción del alumno y del empleador, así como acreditación y evaluación externa. El Vice-Canciller/Director debe asegurar que los procedimientos de promoción de las universidades no estén dominados por criterios de investigación (*output*) en detrimento de una docencia de alta calidad. Por consiguiente, debería existir un proceso de *rendición de cuentas* para los alumnos, empleadores, organismos de financiación, departamentos del gobierno y público en general, que, comenzando a nivel de cada profesor llegara hasta el Vice-Canciller/Director a través de los directores de departamento, con el fin de asegurar de que no solamente existe una *tradición de calidad docente* sino que ésta se ha ofrecido realmente. Se exhorta a las universidades para que hagan públicos sus informes AAU, al igual que ocurre con los informes CNAAs y HMI. La cuestión clave dentro de una economía de mercado es si dichos informes llegan de modo efectivo a los futuros clientes (los alumnos).

Pueden desarrollarse varios indicadores de rendimiento, tanto cualitativos como cuantitativos y no necesariamente punitivos, relacionados con cada uno de los componentes de una *cultura de calidad de la enseñanza*, en términos de adecuación de la provisión, calidad de la misma, economía y empleo de lo adquirido, eficacia, etc. En la década de 1970 se identificaron una serie de indicadores de rendimiento parciales, que se centran en los distintos elementos de un rendimiento satisfactorio: disponibilidad, conocimiento, posibilidad de acceso, extensión, adecuación, eficacia, eficiencia, logro/beneficios/impactos y aceptabilidad (Sizer, 1979). Todos ellos se muestran en el cuadro 3. Se propusieron algunos tests de adecuación para ser aplicados a cada uno de los indicadores, los cuales fueron utilizados por el Equipo de Trabajo del CVCP/UGC en su segundo informe (CVCP, 1987): relevancia, verificación, ausencia de sesgo, cuantificabilidad, viabilidad económica, aceptación institucional y aceptación externa ².

La cuestión clave que se plantea es la siguiente: ¿cómo puede asegurarse de que existe un adecuado equilibrio entre los aspectos cualitativos y cuantitativos de la evaluación del rendimiento? La calidad del rendimiento de los profesores en el aula puede ser más importante que las ratios profesor/alumnos y sus componentes, promedio de horas de clase, promedio el tamaño del grupo y promedio de horas de contacto con el el alumno, pero las ratios profesor/alumnos muy elevadas tendrán un impacto inevitable en la calidad de la docencia ³. Estos indicadores de rendimiento *parciales* pueden emplearse para informar los procedimientos de auto-evaluación, la revisión externa por colegas y los procedimientos para asegurar una *cultura de calidad de la enseñanza*. Deben además informar las evaluaciones de la *calidad del entorno de aprendizaje, la calidad de la experiencia de aprendizaje y los resultados del aprendizaje*. El breve listado propuesto por el CVCP incluye cierto número de indicadores de rendimiento, como p. ej., tasas de abandono, que el equipo auditor recibe en una sesión informativa previa a su visita a la universidad, además de otros indicadores cuantitativos y material descriptivo.

2. Han sido elaborados a partir de Cave y *et al.*, 1988.

3. La cuestión del equilibrio fue tratada ampliamente en el artículo de Sizer (1979a).

CUADRO 3.

Propiedades de los indicadores de rendimiento en la educación superior

<i>Enfoque</i>	<i>Contenido conceptual</i>	<i>Informa acerca de</i>	<i>Ejemplos de indicadores de rendimiento</i>
- Disponibilidad	- Cantidad y tipo de cursos, facilidades de investigación o servicios centrales disponibles	- Lo que puede obtenerse	- Lista de servicios disponibles en el <i>Careers Advisory Service</i> ; facilidades y oportunidades de investigación de un departamento; número y capacidad de salas de clase y seminarios
- Conocimiento	- Conocimiento por parte de los usuarios de lo existente, rango y condiciones de admisión o utilización de los cursos, etc.	- Quién tiene conocimientos acerca de lo disponible	- Conocimiento por parte de los futuros alumnos de los cursos ofrecidos por un departamento académico. Conocimiento por parte de los futuros usuarios de los servicios ofrecidos por un centro de informática
- Accesibilidad	- Indica si los servicios pueden ser obtenidos por grupos adecuados	- Facilidad de alcanzar y usar las instalaciones (servicios)	- Facilidades de fotocopias; ubicación de los aparcamientos; tiempo medio para la búsqueda de bibliografía en el servicio de biblioteca
- Alcance	- Compara la cantidad de los servicios ofrecidos con la capacidad disponible o la demanda potencial	- <i>Cuanto pero no Cómo de bien</i>	- Estudiantes matriculados en los cursos comparados con las cuotas de los cursos; clientes del centro médico, porcentaje de alumnos del último curso que utilizan el servicio de orientación profesional

<ul style="list-style-type: none"> - Adecuación 	<ul style="list-style-type: none"> - Corrección en cuanto al tipo y la cantidad de servicios ofrecidos, cursos ofertados o investigación llevada a cabo 	<ul style="list-style-type: none"> - ¿Es la cantidad y/o calidad de los servicios ofrecidos la adecuada? 	<ul style="list-style-type: none"> - Demanda de cursos; discrepancias entre las facilidades informáticas requeridas y las disponibles; comparación del tamaño de los grupos con la capacidad de las salas de clase y los seminarios.
<ul style="list-style-type: none"> - Eficiencia 	<ul style="list-style-type: none"> - Compara la entrada de recursos con el producto de los mismos 	<ul style="list-style-type: none"> - Cuantos recursos se han empleado 	<ul style="list-style-type: none"> - Coste por servicio al cliente en el centro médico; coste por estudiante FTE por curso; coste por búsqueda de bibliografía
<ul style="list-style-type: none"> - Eficacia 	<ul style="list-style-type: none"> - Compara el cumplimiento de los objetivos (o de lo que se pretendía); cualitativo/comparativo 	<ul style="list-style-type: none"> - Características: duración, contenido, efecto, proporciones servidas, varianzas del presupuesto/estándares 	<ul style="list-style-type: none"> - Comparación entre lo planificado y lo real: % de utilización de las salas de clase y seminarios/número de estudiantes graduados/empleados, utilización de ordenadores, coste del servicio central, coste por FTE por curso, contenido del curso
<ul style="list-style-type: none"> - Resultados, beneficios, impactos 	<ul style="list-style-type: none"> - Indica beneficio social o económico 	<ul style="list-style-type: none"> - Efectos económicos; efectos no económicos 	<ul style="list-style-type: none"> - Incremento de ingresos debido a la graduación; beneficios para la sociedad de investigación exitosa; beneficios del programa cultural para la comunidad local; patentes y copyrights registrados
<ul style="list-style-type: none"> - Aceptabilidad 	<ul style="list-style-type: none"> - Evalúa la relación entre los resultados del servicio/curso/investigación con las preferencias de los participantes 	<ul style="list-style-type: none"> - Satisfacción del usuario con los servicios; satisfacción del estudiante con los cursos, satisfacción del cliente con el resultado de la investigación patrocinada 	<ul style="list-style-type: none"> - Demanda de cursos; número de quejas al bibliotecario; evaluación del curso al finalizar el programa de clases; repetición de la patrocinación de la investigación

9. Posibilidad de publicación de los indicadores de rendimiento de la calidad docente institucional

La posibilidad de publicación de los indicadores de rendimiento de la calidad docente institucional es muy escasa (Bourke, 1986). Como han observado Booth y Booth (1989), desde la perspectiva de los politécnicos resulta relativamente fácil el elaborar un listado de medidas y observaciones parciales. La dificultad estriba en trasladarlo a algo verdaderamente útil para una institución o un organismo de financiación. La misma dificultad puede surgir al llevar a cabo la serie de variables cuantitativas y cualitativas propuestas por el Grupo Limburg de Investigación (Dochy, Segers y Wijnen, 1990c). Barnett ha argumentado que existe una barrera entre los indicadores de rendimiento y la calidad docente, de tal modo que la calidad docente resulta tener una entidad un tanto *misteriosa*. Concluye diciendo que los indicadores de rendimiento numéricos pueden aportar alguna información acerca de la *calidad del aprendizaje del alumno*, lo cual a lo sumo puede dar una imagen parcial y posiblemente distorsionada, pero para una imagen fundamentada en la información, debemos aumentar el conocimiento interno de los procesos educativos de nuestras instituciones. Sin embargo, los indicadores de rendimiento parciales sugeridos anteriormente pueden aportar una valiosa visión parcial de *la calidad del ambiente de aprendizaje*, que a su vez puede influir tanto en la *calidad* como en los *resultados de la experiencia de aprendizaje*. Para más información esperamos el informe del Comité de Investigación de la Calidad Docente, establecido por el PCFC bajo la Dirección de la Baronesa Warnock.

El Comité Directivo Conjunto CVCP/UFC consideró los posibles indicadores, pero llegó a la conclusión de que algunos indicadores empleados en las instituciones, tales como menciones de honor, son inadecuados para establecer comparaciones entre las instituciones. Considera que es necesario que las universidades realicen una auto-evaluación y una valoración formal de la docencia como práctica, incluyendo el desarrollo de la evaluación individual del profesor y de los cuestionarios de los alumnos. El Comité retomará el factor de los indicadores de la calidad docente a finales de 1990, no sólo debido al compromiso de promover tales indicadores por parte del CVCP, sino también debido a su importancia como parte del proceso de *rendición de cuentas* dentro de un sistema de economía de mercado y para asegurar que los consumidores están plenamente informados. Con toda seguridad el Comité Con-

junto establecerá una estrecha relación con el Comité Sutherland. Se ha propuesto además que distinga entre:

- indicadores de rendimiento que *informan* las evaluaciones de: (i) la calidad del entorno docente, (ii) la calidad de la experiencia de aprendizaje, y (iii) los resultados de la experiencia de aprendizaje;
- indicadores de rendimiento cualitativos y cuantitativos; y
- aquellos indicadores de rendimiento que podrían ser utilizados solamente a nivel interno y aquellos que podrían ser incluidos en la *Estadísticas de Gestión Universitaria e Indicadores de Rendimiento*. La palabra clave es *informar*.

10. Contribución de la revisión por colegas

Los indicadores de rendimiento son solamente un *input* y ciertamente no el más importante dentro del proceso de valoración de la calidad, control de la calidad, rendición de cuentas y evaluación del rendimiento; la revisión por colegas tiene un importante papel que desempeñar, por no decir fundamental. Tiene lugar en cada uno de los estadios del proyecto de investigación. En lo que respecta a la enseñanza tiene lugar cuando los cursos son acreditados por el CNAAB y los estamentos profesionales, examinados externamente e inspeccionados por el HMI. La selección de la investigación del UFC para el ejercicio de 1989, consistió principalmente en un proceso de *revisión por colegas informada* (UFC, 1989). No se utilizó el análisis de citas y la información contable sobre las publicaciones totales no resultó ser de gran ayuda. El panel de expertos coincidió en señalar que la cantidad no igualaba a la calidad. Los miembros, por lo general, se basaron en la información detallada acerca de las publicaciones, p. ej., dos por cada miembro del personal.

Resulta difícil vislumbrar en el Reino Unido una publicación anual de evaluaciones externas por revisión de colegas y valoraciones de la calidad de los departamentos o los cursos, que pueda informar los mercados estudiantiles. También resulta dudoso si los costes sobrepasarían a los beneficios o si se dispondría del suficiente personal de alta cualificación para llevar a cabo las revisiones. Las valoraciones de las revisiones por colegas del *Suplemento de Educación Superior del Times* eran

metodológicamente defectuosas, especialmente en lo que respecta a la calidad de la enseñanza. Numerosas universidades han introducido revisiones periódicas de los departamentos a nivel interno, de entre tres a cinco años, y la mayoría de ellas, si no todas, emplean evaluadores externos. Durante muchos años han existido presiones para la creación dentro del sector universitario de un cuerpo similar al CNAA, que acreditara la institución y los procesos que dan lugar a las revisiones internas. Este rol lo desempeñará la nueva Unidad de Auditoría Académica del CVCP. Resulta interesante destacar que el DES está llevando a cabo una investigación dentro del CNAA y que existen presiones para crear una pequeña unidad de auditoría académica transitoria, combinando un reducido CNAA con la Unidad de Auditoría Académica del CVCP (*The Times Higher Education Supplement*, 1990).

Aunque resulta difícil identificar la medición de los resultados de la calidad docente, un proceso tal de evaluación de los procedimientos institucionales elevaría el nivel de concienciación de la importancia del control de la calidad docente, a nivel departamental e institucional. Si una evaluación resulta insatisfactoria, surge la siguiente cuestión: ¿qué sanciones serían aconsejables? La publicación no sólo perjudicaría la imagen de una institución en una economía de mercado, también tendría un impacto negativo en el contrato de financiación de la institución con el UFC. Sin embargo, las auditorías del CVCP no cubren el mismo terreno que los informes del HMI o las revisiones por colegas de la calidad de las universidades holandesas, y quedan dudas además de que los procedimientos de acreditación del CNAA informe al mercado acerca de la calidad *comparativa*.

Parece ser que el UFC está desestimando el desarrollar cualquier mecanismo formal para evaluar la calidad y para establecer valoraciones comparativas de la calidad docente: tiene otros muchos asuntos que resolver. ¿Está esperando ver los resultados del trabajo de la AAU? La influencia de la unidad puede dar origen a un empleo más exhaustivo de la revisión externa por colegas y posiblemente, pero no pausiblemente, la *propiedad* del CVCP de un sistema formal externo de revisión por colegas, similar al que opera en los Países Bajos, en el que el UFC desempeñaría el papel del Ministerio de Educación y Ciencia. Debe observarse que la revisión externa por colegas del contenido del curso no aparece en el esbozo del listado. Una de las pruebas de la efectividad de la Unidad en cuanto al sugerimiento de aspectos para el control de la calidad o la supervisión de métodos para obtener el control de

dicha calidad, que se beneficiará de posteriores desarrollos, se efectuará cuando aparezca en el listado de las mejores prácticas la revisión externa por colegas, juntamente con la acreditación profesional.

Si existe una revisión externa y detallada de la calidad de la enseñanza de los departamentos o materias, los indicadores de rendimiento informarán los juicios cualitativos de los colegas; como ilustra el trabajo de Dochy *et al.*, las variables cualitativas y los juicios son más importantes que los indicadores de rendimiento cuantitativos. Las tendencias de los indicadores de rendimiento y las comparaciones inter-institucionales indican dónde difiere el rendimiento, pero no necesariamente por qué difiere o cómo mejorarlo. Sin los indicadores de rendimiento los elementos de juicio pueden ser demasiados amplios; la dependencia en los indicadores de rendimiento, en ausencia de valoraciones de colegas es extremadamente peligrosa. La revisión por colegas y los indicadores de rendimiento deben complementarse mutuamente.

11. Indicadores del estado financiero de la institución

Las declaraciones anuales del estado financiero de una institución son documentos clave que conciernen tanto a la propiedad y responsabilidades del cuerpo de gobierno como a la *salud* financiera de la institución. La importancia de una buena gestión financiera, de las previsiones de financiación y de los indicadores de estado de salud de la economía de la institución, se incrementa cuando la política de la educación superior de un país se orienta hacia una mayor autonomía institucional, hacia un incremento de la competencia y diferenciación entre las distintas instituciones, hacia una reducción de la dependencia de la financiación estatal y hacia una mayor diversidad de fuentes de financiación. Bajo estas condiciones existe la posibilidad de que una institución llegue a estar financiada inadecuadamente y que, por tanto, sea incapaz de mantener la cantidad o la calidad de la docencia y la investigación que se ha comprometido a ofertar. Pueden llegar a declararse insolventes, como han puesto de manifiesto la crisis financiera del *University College* de Cardiff y en menor grado, la de la Universidad de Aberdeen y el recorte financiero de algunas facultades de la Universidad de Londres. De ahí las solicitudes por parte del UGC/UFC de previsiones económicas cada vez más detalladas y las solicitudes paralelas del PCFC. Dada la importancia de una gestión económica saneada, parece oportuno el incluir indicadores de salud económica en las declaraciones eco-

nómicas anuales, para que sean transmitidos a los consejos de financiación y para la publicación de comparaciones inter-institucionales. Un prerrequisito esencial es el establecimiento de un estándar consensuado de rendición de cuentas para la publicación de los informes económicos. El Comité de Estándares de Rendición de Cuentas (*Accounting Standards Committee*) ha enviado a las universidades un modelo para la práctica de la rendición de cuentas: *A Statement of Recommended Accounting Practice: Accounting in UK Universities (CVCP, 1989a)*, y un grupo conjunto del CVCP/UFC está elaborando una serie de indicadores de salud financiera para su inclusión en su documento *University Management Statistics and Performance Indicators*. Se planifica publicar un documento consultivo que contenga una serie de once indicadores de rendimiento recomendados, en mayo o junio de 1990.

12. Conclusiones

Se ha argumentado que los indicadores de rendimiento informan y complementan la revisión por colegas. Los británicos precedieron a los holandeses en el desarrollo y utilización de los indicadores de rendimiento dado que se ha conseguido la *propiedad* del Vice-Canciller, lo que no ocurre en los Países Bajos. En la revisión externa por colegas, las universidades holandesas han precedido a las británicas, pero en menor medida en lo que se refiere a los politécnicos, ya que en ellos la *propiedad* es del sistema. Sin embargo, la creación de la Unidad de Auditoría Académica del CVCP puede conducir hasta el Vice-Canciller y a la *propiedad* de sistema formal y externo de revisión por colegas. De no ser así, el UFC deberá crear su propio sistema. Existen numerosos paralelismos entre el sistema holandés de financiación condicional de la investigación y la financiación selectiva de la investigación del UGC/UFC, basada en evaluaciones por colegas.

La diferencia clave entre ambos sistemas se encuentra en el deseo del gobierno británico de aplicar en la educación superior su filosofía orientada al mercado, acompañada por cambios a estilos ejecutivos de gestión dentro de las instituciones y por una década de perentoria financiación gubernamental, en términos reales. De ahí el movimiento suscitado para recabar ofertas de plazas estudiantiles y para incrementar las tasas del primer ciclo, así como las inquietudes de los consejos de financiación acerca del *control de la calidad y la evaluación de la calidad*. Se ha cuestionado la capacidad de los consejos de financiación,

especialmente de UFC, para realizar *juicios comparativos* de la calidad docente en la firma de contratos y en la supervisión de la calidad contratada. El UFC debe hacer recaer la responsabilidad de demostrar la adecuación de lo ofertado en las instituciones, especialmente en aquellas que ofrecen pujas a la baja, ofertas inusuales o que admiten cantidades enormes de alumnos con tasas sólo de matrícula. Sin embargo, a largo plazo este requisito no resolverá la cuestión de *los juicios acerca de la relativa calidad docente dentro de las disciplinas*. Debe identificar aquellas instituciones cuya calidad real o cuya habilidad para mantener la calidad son sospechosas. Pueden elaborar la base de una evaluación de la calidad docente basada en una *revisión por colegas informada*. Puede que no le resulte posible el desarrollar la capacidad de elaborar juicios abiertos acerca de si la calidad puede darse o se ha dado, sin disponer de recursos adicionales, pero a la vez sin forzar el aparato burocrático del UFC. De ahí lo atractivo de permitirle al CVCP dar el primer paso para comprobar si el CVCP es capaz de hacerse con la *propiedad* de un sistema similar al existente en los Países Bajos.

Al final puede ocurrir que si el gobierno acomete una economía de mercado y empresarial en la educación superior, se apoyará más en las fuerzas de mercado y en las acciones de los clientes individuales para forzar a las instituciones a tomarse en serio la necesidad de evaluar y ofrecer calidad docente, que en los contratos financiados por compradores del FTE de las instituciones sin poder compensatorio como sustitutas de los mercados competitivos. En ese caso debe preguntarse a sí mismo: ¿Deben sustituirse los consejos de financiación por un cuerpo cuya principal responsabilidad sea la de asegurar que el mercado funciona con eficacia, regulando y estimulando la información, incluyendo los indicadores de rendimiento y la revisión por colegas, destinada a los futuros alumnos que votan a través de sus tasas regulares o bonos? ¿Puede plantearse una cuestión similar en los Países Bajos?

Existen numerosos paralelismos entre los desarrollos del Reino Unido y de los Países Bajos, aunque los británicos tienden a ir a la zaga de los holandeses. Ello puede reflejar las dimensiones relativas de ambos sistemas de educación superior o la falta de capacidad de análisis político por parte del DES comparado con el de La Haya. Sin embargo, la diferencia clave radica en la filosofía política. Se reflejan por todo el mundo los desarrollos de las financiaciones contractuales y la rendición de cuentas; el uso de indicadores de rendimiento como parte del proceso de rendición de cuentas; la información de la planificación institucio-

nal, la distribución de recursos, así como la rendición de cuentas y supervisión de la salud económica. El creciente interés por el control y la evaluación de la calidad debe mejorar la eficacia de la docencia y la investigación de ambos países. Queda por ver si los movimientos ocurridos en el Reino Unido hacia una economía de mercado aplicada a la educación superior, acompañada por estilos ejecutivos de gestión mejorarán o perjudicarán la eficacia.

SELECCION DE INDICADORES DE RENDIMIENTO UNA PROPUESTA COMO RESULTADO DE UNA INVESTIGACION *

FILIP J.R.C. DOCHY, MIEN S.R. SEGERS y WYNAND H.F.W. WIJNEN
Universidad de Maastrich, Países Bajos

Introducción

En la mayoría de las universidades, el hecho de utilizar datos durante la elaboración de juicios acerca de la docencia y la investigación ha constituido una práctica habitual desde hace bastante tiempo. En ellas existe una estructura institucional, de sólidos y aceptados principios académicos, en cuyo contexto dichos datos han sido considerados.

Tanto los indicadores de rendimiento como los datos de información para la gestión deben ayudar a las decisiones; ofreciendo un punto de partida sobre el que deben fundamentarse los juicios académicos y directivos. No existen sustitutos para tales juicios pero, cuanto menos abundantes sean los indicadores mayor será el vacío que deba llenarse. Una vez tomadas las decisiones e identificadas las apropiadas medidas de progreso del rendimiento, una comparación entre las medidas esperadas y la presentes —basadas en los indicadores de rendimiento— debe formar parte de la evaluación acerca de si las anteriores decisiones se han llevado a cabo satisfactoriamente (Dochy, Segers y Winjen, 1990 a). Los indicadores de rendimiento, adecuadamente definidos e interpretados, pueden desempeñar un importante papel en la determinación de la política a seguir, la dirección de una institución y la optimización de la calidad educativa.

* Traducción íntegra del capítulo 10: *Selecting Performance Indicators. A proposal as a result of research*, perteneciente a L.C.J. GOEDEGEBUURE; P.A.M. MASSEN y D.F. WESTERHEIJDEN (EDS), *PEER REVIEW AND PERFORMANCE INDICATORS*. Utrecht: Lemma, pp. 135-153, 1990.

El desarrollo de una serie de indicadores de rendimiento para la enseñanza superior holandesa puede describirse como un proceso que comenzó en la década de 1970. El desarrollo social y económico requería innovaciones directivas en la enseñanza superior. Se introdujeron indicadores de rendimiento, en conexión con la «nueva filosofía dominante».

El proyecto de investigación a que nos referimos intenta contribuir a este proceso de desarrollo. Por consiguiente, se analiza el concepto de indicadores de rendimiento y los términos relacionados con el mismo. Pueden identificarse distintos temas centrales. En este sentido, son de capital importancia para el diseño y la ejecución de un sistema de control de calidad.

Aparte de elaborar esta estructura conceptual, el proyecto de investigación pretendía proponer una serie de indicadores de rendimiento en base a distintos criterios. El criterio principal fue la validez empírica de los indicadores potenciales. Los resultados del estudio de validez se compararon con otros criterios. Los hallazgos de estos estudios de investigación se clasificaron de acuerdo con las funciones primordiales de la educación superior, el modelo entrada-proceso-resultado y el criterio de medición. Finalmente, aparecen algunos comentarios acerca de la posibilidad de aplicación de los indicadores de rendimiento.

2. Indicadores de rendimiento y conceptos relacionados

Debido al creciente interés por los indicadores de rendimiento, aparecen en la bibliografía numerosas definiciones y términos relacionados. Todo el mundo está de acuerdo en señalar que los indicadores de rendimiento aportan información acerca de las actividades de una institución (Frankmann y Muffo, 1988; Cave *et al.*, 1988; Dochy *et al.*, 1987, 1990). Sin embargo, se toman distintas posiciones en cuanto a las siguientes cuestiones: ¿Son los indicadores de rendimiento cualitativos y/o cuantitativos? ¿Son los indicadores de rendimiento señales o guías más que medidas absolutas? ¿Son sinónimos de los conceptos «estadísticas de gestión» e «datos de información para la gestión»?

En la investigación sobre indicadores de rendimiento que se describe a continuación, se hace una distinción entre estadísticas de gestión, información para la gestión e indicadores de rendimiento. El cuadro 1

Cuadro 1. Definición de los tres conceptos

- *Estadísticas de gestión*: datos cuantitativos.
- *Información para la gestión* los datos cuantitativos o cualitativos que están relacionados entre sí y estructurados como información de dirección. Indicadores de rendimiento: datos empíricos —cualitativos y cuantitativos— que describen el funcionamiento de una institución y el modo por el cual la institución persigue sus objetivos. Ello implica que están relacionados con el contexto y con el tiempo.
- Los *indicadores* son signos más generales que pueden convertirse en un número de características más específicas, denominadas variables.

ilustra la definición de los tres conceptos. Un ejemplo podría clarificar este esquema. El número total de matriculación de alumnos en las instituciones es un dato cuantitativo y por consiguiente una estadística de gestión. El número total de graduados en relación con el número total de matriculaciones puede describirse como un dato de información para la gestión. Esta ratio es un indicador de rendimiento si su mejora se persigue como una meta institucional. Numerosos datos relativos a los costes son considerados como estadísticas de gestión. También pueden ser empleados en las funciones de gestión y de control pero no pretenden evaluar el rendimiento de una manera global. (Cave *et al.* 1988, 1990).

En la definición de indicadores de rendimiento adaptada por nosotros, se emplea el término «funcionamiento» de una institución en lugar del de «calidad» dado su carácter más neutral y concreto.

Los indicadores del rendimiento son datos empíricos de naturaleza cuantitativa o cualitativa. Estos datos se convierten en indicadores de rendimiento si expresan las metas de los actores observados. Ello significa que poseen una importancia contextual al mismo tiempo que temporal. Pueden ser operativizados por una serie de variables; estas son unas características más específicas.

Una relación entre indicadores del rendimiento y objetivos implica que ciertos datos puedan ser para un actor estadísticas o información de dirección, mientras que para otro actor pueden ser interpretados como indicadores de rendimiento o variables. Por ejemplo, desde el punto de vista del gobierno la ratio hombre-mujer dentro de una población estudiantil es, respecto al objetivo de la emancipación, un indicador del rendimiento de una institución (Ministerio de Educación y Ciencia,

1988). Sin embargo, este objetivo gubernamental no es necesariamente una meta institucional y por consiguiente, la ratio puede ser interpretada como una estadística de la dirección o una información que, en sí misma no es un indicador del rendimiento o del funcionamiento de una institución. Finalmente, el funcionamiento de una institución de enseñanza superior (como cualquier otra institución social) difícilmente puede ser expresado en términos absolutos. Por consiguiente, los indicadores de rendimiento deben ser interpretados como señales o guías. Ello implica que pueden servir de gran ayuda a la hora de emitir un juicio pero nunca pueden ser considerados como sustitutos del mismo.

3. Condiciones primordiales para el control de calidad

Respecto al concepto del programa de control de calidad, es necesario aclarar la diferencia entre «control de calidad» y «medida de la calidad» de la educación superior. Se ha comprobado que, con frecuencia, el control de calidad es un proceso continuo de supervisión. La medida de la calidad es principalmente una evaluación puntual y una parte del proceso del control de calidad. Por lo general, se da preferencia al control de calidad. Jessee (1984) señala que, para que un control de calidad sea efectivo, cualquier sistema debe incluir tres componentes interdependientes: supervisión, medición y mejora. Uno de los principales problemas de la mayoría de los programas de control de calidad es el subdesarrollo de los aspectos de supervisión y mejora. A pesar de los numerosos obstáculos que deben sortearse, parece interesante el indagar las reglas básicas del control de calidad. Pero, aun antes de que se clarifiquen las reglas del juego, es necesario que los diferentes «participantes» (en el caso de la educación superior) se pongan de acuerdo respecto a la meta y al funcionamiento del juego. ¿Lo emplearán para mejorar la calidad de la enseñanza superior o para la redistribución de recursos? ¿Funcionarán los indicadores propuestos como un instrumento para supervisar, evaluar, planificar o estimular el diálogo?

3.1. Principales condiciones para la garantía de la calidad

Tanto al diseñar como al trabajar con un sistema de garantía de la calidad deben tenerse en cuenta diversas condiciones indispensables.

El carácter cíclico del proceso que garantiza la calidad requiere ser puesto de relieve. Debe ser un proceso constante de revisión en el que los distintos temas sean analizados a intervalos regulares, p. ej., cada tres años.

El proceso de garantía de la calidad implica, al menos, tres etapas diferenciadas: actividades de supervisión, una etapa de medición y una etapa de mejora. La falta de desarrollo de cualquiera de los tres estadios reduce drásticamente la efectividad del conjunto.

Dentro de dicho proceso, debe llegarse a un acuerdo acerca de la división de responsabilidades así como de las cualificaciones requeridas. Es esencial el informar de manera continuada al personal académico y a otros protagonistas de la educación superior acerca de los cambios y resultados concretos, con el fin de asegurar su colaboración para la consecución de un efectivo sistema de control y garantía de calidad.

Los mecanismos de supervisión deben servir como un claro sistema de alarma para accionar una amplia evaluación del origen de las deficiencias que se descubran.

La evaluación necesita ser desarrollada en base a unos indicadores y variables previamente establecidos.

Debe realizarse una distinción entre los criterios y los ideales o esquemas en relación a la calidad. Un ideal sólo puede alcanzarse en un cierto nivel. La escasez de recursos, los gastos, la confrontación de opiniones y otros factores obstruyen una consecución total. Sin embargo, dadas estas restricciones, pueden establecerse criterios realistas.

No se debe caer en la tentación de usar gran cantidad de datos evaluativos. Es necesario emplear un número limitado de indicadores con un alto grado de validez para detectar las deficiencias de una calidad satisfactoria en una amplia área de investigación y docencia. Pueden requerirse aspectos más específicos o «variables» para establecer procedimientos evaluadores profundos y clarificadores.

Deben crearse procedimientos para un control de calidad a corto plazo mientras la investigación educativa sigue su proceso. Posiblemente, aparecerán en la superficie fallos referentes a la calidad de la investigación o de la docencia, durante las actividades profesionales del per-

sonal académico. Algunos de ellos, puede y deben ser evitados tan pronto aparezcan. Resulta de gran utilidad el crear procedimientos para descubrir estos fallos a tiempo.

3.2. *Un proceso cíclico para garantizar la calidad*

Tras examinar la bibliografía internacional, se encontró que la elaboración de un proceso cíclico para asegurar el control de calidad podría ser de gran valor. El ciclo del proceso propuesto (figura 1), en realidad un proceso de constante revisión, consiste en cinco fases. Este modelo permite tomar en cuenta las condiciones fundamentales para asegurar la calidad. El ciclo posee un carácter continuo y tiene también las etapas de supervisión y medición (fase 3) así como una fase de mejora (fases 4 y 5).

4. **Requisitos y criterios esenciales para la selección de los indicadores de rendimiento**

Ante todo, es necesario establecer como punto de partida los requisitos de los indicadores de rendimiento y los criterios esenciales para

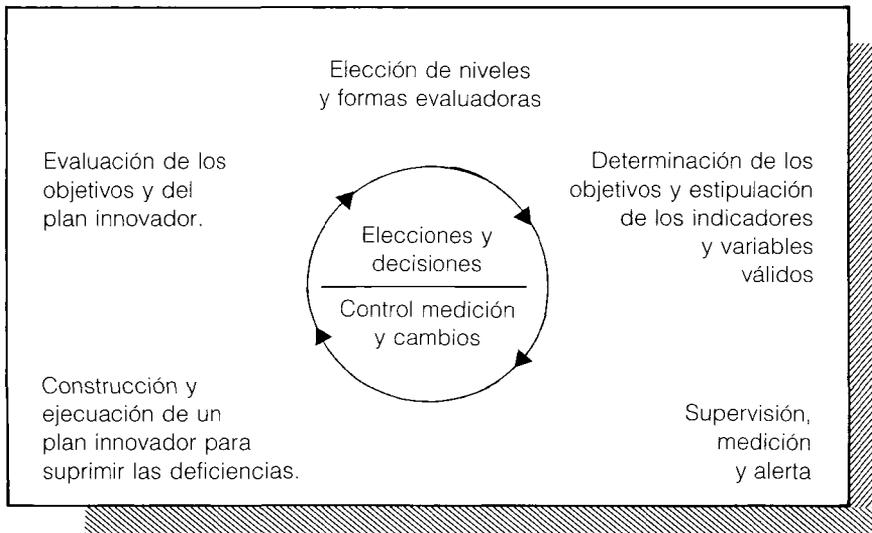


Figura 1: Ciclo para garantizar la calidad.

seleccionarlos. Un primer requisito es que deben estar claramente relacionados con las definidas funciones de la institución. Un segundo requisito es el reconocimiento de que solamente son lo que su nombre señala, indicadores de la medida en que se alcanzan las metas institucionales. Un tercer requisito es que deben ser constataciones válidas de lo que pretenden indicar y que pueden ser medidos e interpretados de un modo correcto y fiable. Finalmente, dado que sus valores son más relativos que absolutos, su incorporación en la maquinaria de la «tecnología directiva» debe ser realizada con extremada precaución.

4.1. *Criterios de selección*

Si uno quiere hacerse una idea del estado de salud de las instituciones de educación superior, parece evidente que lo que se pretende averiguar es por medio de qué aspectos de la educación puede lograrse, lo más completa y correctamente posible. Llegados a este punto, surge el tema de la operativización, o sea, la traducción de «condición (o calidad) de la educación».

Operativizar es trasladar una característica general a fenómenos perceptibles, que, para ese específico y explorado sistema o fenómeno social parecen representar lo mejor posible, en un momento y lugar dados, las características pretendidas. Como un primer estadio de la operativización, la mayoría de constructos pueden dividirse en distintos aspectos o indicadores. El rendimiento de las instituciones de enseñanza superior puede dividirse en tres dimensiones, correspondientes a las funciones de la educación superior: educación (docencia), investigación y servicios públicos. Cada uno de estos indicadores se operativizará en variables. La operativización del «funcionamiento de la educación superior» puede esquematizarse tal y como muestra el cuadro 2.

En los Países Bajos, al igual que en otros países (p. ej., el Reino Unido) han surgido críticas acerca de las listas presentadas de indicadores de rendimiento. Sin embargo, la crítica más importante concierne a la validez de los indicadores de rendimiento. La demanda de una validez de contenido se presenta en este artículo. Segers (1977) afirma que la *validez de contenido* "significa que existe una conexión entre el concepto teórico y la variable empírica que es argumentada por mostrar que el «contenido» del indicador corresponde al «contenido» del concepto teórico. La validez empírica es el juicio sobre el contenido mani-

Cuadro 2. Operativización del «funcionamiento de las instituciones de educación superior»

<i>Concepto teórico: funcionamiento de las instituciones de educación superior</i>		
<i>dimensión 1</i> indicadores de <i>docencia</i> 1,2,3,4,...	<i>dimensión 2</i> indicadores de <i>investigación</i> 1,2,3,4,...	<i>dimensión 3</i> indicadores de <i>servicios</i> 1,2,3,4,...
variables 1.1, 1.2, 1.3,...	variables 1.1, 1.2, 1.3,...	variables 1.1, 1.2, 1.3,...
2.1, 2.2, 2.3,...	2.1, 2.2, 2.3,...	2.1, 2.2, 2.3,...

fiesto de los indicadores como representativos del concepto o dimensión teóricos. Ello significa, según Segers, que se necesita algo más que un juicio a primera vista. El consenso del panel científico acerca de la validez de un indicador hace que la validez empírica, desde un punto de vista metodológico, signifique algo más que una pura subjetividad. El juicio realizado por expertos (colegas de investigación) también modifica la validez del contenido. Sin embargo, la validez del contenido puede alcanzarse solamente cuando el investigador es consciente del hecho de que es imprescindible trabajar con instrumentos sólidamente diseñados.

Segers (1977) distingue cuatro pasos esenciales en el diseño de un instrumento con validez de contenido:

- reflexión teórica y clarificación acerca del significado o contenido del concepto en cuestión;
- distinguir y especificar las dimensiones teóricas conceptuales;
- una bien meditada elección de los indicadores para cada dimensión;
- evaluación del medio en el que los distintos indicadores pueden agregarse para una valoración.

Boesjes-Hommes (1974) coincide con Segers en que la validez del contenido puede lograrse mediante la utilización de jueces. En este sentido, la autora se refiere a tests consistentes que pueden emplearse

para investigar la «ecuación» entre el significado semántico de un concepto y la deducción operativa.

Es posible elegir entre una serie de indicadores de rendimiento en base a distintas consideraciones. El objetivo primordial de la investigación, la validez de los indicadores de rendimiento, implica un primer criterio. Además del principal criterio, los indicadores de rendimiento pueden ser analizados en relación a las siguientes cuestiones;

- ¿Qué indicadores se presentan en la bibliografía nacional e internacional y en los documentos de gestión como datos esenciales?
- ¿Qué indicadores son mensurables?
- ¿Qué indicadores se refieren a las principales metas de la educación superior?

Otras cuestiones requieren una posterior investigación:

- ¿Qué indicadores pueden ser mensurados de un modo fiable?
¿Qué es lo que significa, en términos de energía, la detección y provisión de los datos necesarios para las personas e instituciones implicadas?
- ¿Qué información es accesible en la actualidad y por medio de qué indicadores deben desarrollarse los instrumentos de evaluación?

El objetivo de la siguiente sección es la evaluación de los indicadores y variables en base a tres criterios: validez, experiencia internacional y mensurabilidad. Debemos tomar nota del hecho de que la validez de un indicador es por ahora el único criterio científico. La mensurabilidad es un criterio puramente pragmático, que conduce fácilmente a conclusiones incompletas y erróneamente interpretadas. Finalmente, un criterio como «empleo en un contexto internacional» es todavía más arbitrario. Frecuentemente, en algunos países la elección de ciertos indicadores se basa en criterios pragmáticos (cuyos datos están disponibles), en preferencias personales de uno o varios individuos, etcétera. Rara vez son el resultado de una investigación teórica o empírica.

5. **Un estudio de validez de los indicadores de rendimiento**

5.1. *Objetivo, plan y metodología de la investigación*

El objetivo principal del proyecto de investigación «indicadores de rendimiento en los Países Bajos: un estudio de validez», consiste en examinar la validez de los indicadores de rendimiento (Segers y *et al.*, 1989a, 1989b, 1990). El nivel de agregación es la institución en relación con el gobierno. Los indicadores de rendimiento se interpretan como elementos para el diálogo entre el gobierno y las instituciones. El plan de investigación consta de diversas etapas:

- un estudio de la bibliografía sobre los conceptos de calidad y evaluación de la calidad, así como un análisis internacional de los distintos sistemas de evaluación (Dochy *et al.*, 1987);
- conceptualización de los términos relevantes: indicadores de rendimiento (y términos asociados) y validez;
- elaboración de un inventario de los potenciales indicadores de rendimiento, como resultado de un estudio sobre la bibliografía internacional y sobre las instituciones de educación superior de los Países Bajos;
- validación de los potenciales indicadores de rendimiento
- estudio de la viabilidad.

Al considerar los estudios preliminares sobre la bibliografía, los datos relativos a la calidad de la evaluación se clasificaron en datos generales (indicadores) y datos más específicos (variables). Los indicadores para la docencia, investigación y servicios se muestran en el Cuadro 3, cada uno de estos indicadores se convierte en una serie de variables potenciales. Con el fin de averiguar la validez de los indicadores y las variables, se les pidió a los protagonistas de la educación superior que respondieran un cuestionario. Las preguntas estaban dirigidas a detectar la validez de las variables relacionadas con un indicador, así como la importancia de los indicadores respecto a las tres funciones de la educación superior (docencia, investigación y servicios).

Se agrupó la población de acuerdo con un modelo de constitución múltiple (Connolly, Conlon y Deutsch, 1980) y según el enfoque «partici-

Cuadro 3. Indicadores potenciales para la docencia, investigación y servicios

<i>Educación</i>	<i>Investigación</i>	<i>Servicios</i>
<ul style="list-style-type: none">- alumnos de nuevo ingreso- alumnos en curso- graduados- personal- currículum- ambiente educativo- asuntos financieros- facilidades de infraestructura y de material- cursos ofrecidos- normativa- cooperación entre instituciones	<ul style="list-style-type: none">- publicación, datos aportados- medios materiales para la investigación- personal- organización y puesta en práctica- normativa referente a la investigación- méritos potenciales de los resultados de la investigación- colaboración con otros grupos de investigación	<ul style="list-style-type: none">- servicios internos- servicios externos

pantes» (Freeman, 1984; Stoner y Wankel, 1986; Keuning y Eppink, 1986). El forum lo constituían representantes de puntos de vista tanto internos como externos (Dubin, 1976), elaboración de normativas o influencias del proceso (Pennings y Goodman, 1977) e individuos relacionados con actividades institucionales (docencia, investigación y servicios) (Rawls, 1987; Keeley, 1987). Los participantes en el forum son: alumnos, padres, personal académico y no académico, vice-rectores de las universidades e instituciones HBO, el Ministerio de Educación y Ciencia, empresarios industriales y algunos expertos en el tema de la evaluación de la calidad. La muestra (a nivel de institución) se estratificó en siete áreas disciplinares: ciencias naturales y tecnológicas, ciencias sociales y educación, ciencias médicas, agricultura, ciencias financieras y económicas, derecho y ciencias de la cultura.

Los resultados de la investigación de este estudio de validez (n=359)¹ se clasificaron de acuerdo al grupo de pertenencia del respon-

1. Los participantes en el proyecto de investigación fueron (en paréntesis el número de respondentes): gobierno (20), presidentes del Consejo Ejecutivo (universidad) (7), presidente del Consejo ejecutivo (HBO) (38), profesores titulares (38), personal universitario (35), personal HBO (41), estudiantes universitarios (73), estudiantes de HBO (38), empleadores (21), expertos (29), sector agrícola (14), sector de ciencias sociales y educación (61), sector de la salud (33), sector de ingeniería (55), sector empresarial y de derecho (26), sector de actividades financiero-económicas (30), sector de cultura (8).

diente, a los sectores y subsistemas o disciplinas (sector universitario y HBO).

En la bibliografía no aparecen criterios para realizar la distinción entre los indicadores más o menos válidos. Carmines y Zeller (1970) afirman que, en este contexto: «no existe acuerdo sobre los criterios para determinar el grado en que una medición ha alcanzado la validez de contenido». En la presente investigación, con el fin de operativizar los indicadores se ha empleado el siguiente criterio: si una variable está indicada por más del 60% de los respondientes, se interpreta como una variable muy válida. Una variable elegida por menos del 40% de los respondientes se considera menos válida. Una variable es controvertida cuando es mencionada por entre el 40 y el 60% de los respondientes, para determinar los indicadores que sobresalen de entre los demás se ha empleado el siguiente criterio. Aquellos indicadores que obtienen un valor del 75% o superior por parte de más del 60% de los respondientes toman ventaja sobre los demás (prioridad 1). Aquellos indicadores que obtienen un valor entre el 50 y el 75% por parte de más del 60% de los respondientes son válidos, pero permanecen en segundo lugar (prioridad 2). Los indicadores menos válidos son los que tienen un valor menor del 50% por parte de más del 60% de los respondientes.

5.2. *Una serie de indicadores de rendimiento: resultados de la investigación*

Esta propuesta pretende ser una contribución al desarrollo de una serie de indicadores de rendimiento, como un elemento de diálogo entre el gobierno y las instituciones de educación superior. Está basada en los resultados de los distintos estadios de la investigación. Incluye los siguientes criterios para la elección de los indicadores de rendimiento y las variables:

- son válidos para el gobierno y para el nivel institucional central (Vice-Rectores)
- están relacionados con los principales objetivos de la gestión de la educación superior
- se les reconoce su importancia en la bibliografía y debates internacionales

- los resultados del estudio de validez no muestran ninguna contradicción fundamental con otros niveles (inferiores) de agregación

En primer lugar, los indicadores de rendimiento propuestos —que son los más válidos— se clasifican de acuerdo con las funciones primordiales de la educación superior: docencia, investigación y servicios. En segundo lugar, adoptamos el marco conceptual convencional que considera la educación superior como un *proceso* que transforma las *entradas* —*inputs*— en *productos* —*outputs*— (véase cuadro 4). Surgen por todos los lados los debates sobre qué información usar con el fin de cribar el rendimiento en una institución de educación superior. De hecho, la cuestión olvidada en este tipo de debates es la siguiente: ¿en qué nivel de agregación se usará esa información? Un sistema directivo que se centre en los resultados parece ser que elegirá el operar con indicadores de producto en la relación entre el gobierno y las instituciones.

Ello no significa que sean olvidados tanto los indicadores de entrada como de proceso. Bormans *et al.* (1987) afirmaron que «los comentarios válidos sobre el rendimiento de una institución se basarán por lo general en una combinación de estos tres indicadores». Anteriormente, lo hemos expresado del modo siguiente (Segers y Dochy, 1988): «Las interpretaciones del rendimiento institucional en base a los resultados, implican de forma inequívoca la medición de la entrada y del proceso, en relación con el producto».

Existe consenso respecto a las definiciones de los conceptos. Sin embargo, el paso de las definiciones a indicadores y variables hace aflorar distintos puntos de vista. Al relacionar los indicadores de entrada,

Cuadro 4. El modelo entrada-proceso-producto

<ul style="list-style-type: none">— <i>Entrada</i>: fuentes empleadas (p. ej., facilidades financieras y materiales, estudiantes, personal, ambiente de trabajo) y metas (gubernamentales e institucionales).— <i>Proceso</i>: aspectos organizativos y administrativos (dirección) para las tres tareas primarias de las instituciones de educación superior (p. ej. la organización de un curso, la dirección financiera, etcétera).— <i>Producto</i>: resultados de investigación, docencia y servicios (p. ej., realización profesional, cantidad de publicaciones, cantidad de actividades educativas permanentes para organizaciones externas, etcétera).

proceso y producto, se clarificarán la eficiencia, eficacia y productividad del rendimiento institucional (figura 2).

Finalmente, hemos clasificado los indicadores y variables más válidos de acuerdo con el *criterio de mensurabilidad*. La demanda de mensurabilidad es una condición a la hora de elegir los indicadores del rendimiento, formulados en informes de gestión, especialmente a efectos prácticos. Para hacer posibles las comparaciones inter-institucionales, los datos deben ser comparables. Ello implica que han de ser cuantitativos. Por otra parte, hemos afirmado que los aspectos cuantitativos y cualitativos requieren atención (Dochy, Wijnen y Segers, 1987). La primera cuestión que debemos autoformularnos es: ¿qué significa mensurabilidad?. Son posibles tres interpretaciones:

- una variable puede ser formulada con una cifra: es una variable cuantitativa;
- una variable puede ser formulada con una pregunta que se responda con un sí-no, y por lo tanto puede cuantificarse de ese modo, o bien puede formularse como una pregunta con múltiple elección de respuestas. Ello puede llevarse a cabo a expensas de todo tipo de molestias;
- una variable es una operativización demasiado vaga de un indicador. Una operativización más concreta facilita la cuantificación en una de las formas mencionadas anteriormente.

La formulación en cifras resulta complicada para algunas variables dado que la operativización concreta y el juicio de valor requieren una

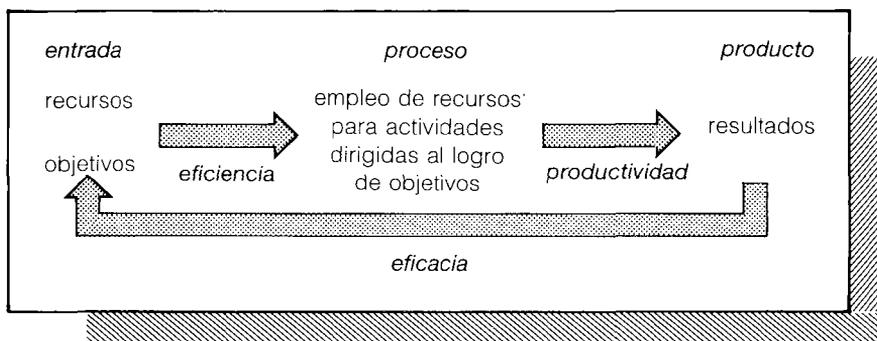


Figura 2: Eficiencia, eficacia y productividad

posterior investigación. ¿Cuáles son los criterios que pueden formularse para, por ejemplo, un efectivo proceso educativo?

La cuestión que se plantea es si se pueden obviar justificadamente aquellos aspectos que no son cuantificables (por el momento), o que lo son de manera muy limitada. Tanto en Estados Unidos como en el Reino Unido, se emplea la revisión por parte de colegas como un indicador para evitar este problema. Sin embargo, la revisión de los colegas es un método, no un indicador del rendimiento. La revisión de colegas puede emplearse si no existe duda acerca de qué indicadores serán el sujeto de dicha revisión.

En resumen, empleamos una categorización basada en cuatro criterios, que pueden sintetizarse del modo siguiente:

- las funciones de la educación superior: docencia, investigación y servicios;
- el tipo de indicador: indicador de entrada, proceso o resultado;
- la mensurabilidad: cualitativa o cuantitativa;
- la validez: indicadores y variables, medidas del contexto y clarificación de los datos.

En la sección siguiente de este capítulo presentaremos la serie de indicadores propuesta para la función docente. Para una visión general más completa de los indicadores válidos de investigación y servicios públicos en la educación superior, véase Dochy *et al.* (1990). En los tres cuadros siguientes (5, 6 y 7) resumimos nuestras propuestas.

Cuadro 5. Indicadores de las variables educativas de entrada

<i>Indicadores</i>	<i>Variables cuantitativas</i>	<i>no cuantitativas</i>
personal	<ul style="list-style-type: none"> – estructura de la edad – % años experiencia laboral – movilidad del personal 	<ul style="list-style-type: none"> – calificaciones educativas – motivación docente
clima educativo	<ul style="list-style-type: none"> – cantidad de tiempo medios y resultados de las innovaciones educativas 	<ul style="list-style-type: none"> – filosofía educativa
infraestructura, medios materiales	<ul style="list-style-type: none"> – cantidad de <i>hardware</i> (ordenadores) – cantidad de <i>software</i> (audio-visual, ordenador) 	<ul style="list-style-type: none"> – calidad del espacio docente

Cuadro 6. Indicadores y variables del proceso educativo

<i>Indicadores</i>	<i>Variables (no cuantitativas)</i>
personal oferta educativa	<ul style="list-style-type: none"> - Política de selección de personal - Posibilidades de libre elección del currículum (<i>asignaturas, módulos</i>) en relación con el volumen del currículum obligatorio - rango de elección de las principales materias - oferta de áreas de estudio: posibilidades regulares, gratuitas y experimentales de formación a tiempo parcial, formación permanente y post-académica.
currículum	<ul style="list-style-type: none"> - eficacia de los procedimientos para el desarrollo de la revisión de los programas existentes o de nuevos cursos - claridad en el contenido y descripción de los objetivos del curso - relevancia de los objetivos del curso en relación a la realidad profesional - acuerdo entre contenido y metas del currículum - realización de las funciones del currículum - corrección, plausibilidad, totalidad y nivel de los objetivos del currículum
normativa	<ul style="list-style-type: none"> - sistema institucional de control de calidad - acuerdo con otras instituciones referente a la distribución y concentración de tareas - realización democrática de la normativa

Cuadro 7. Indicadores y variables del producto educativo

<i>Indicadores</i>	<i>Variables cuantitativas</i>	<i>No cuantitativas</i>
permanencia de estudiantes	<ul style="list-style-type: none"> - % que superan exámenes propedéuticos - obtención del doctorado - tiempo empleado por área de estudio 	
egresión de estudiantes	<ul style="list-style-type: none"> - % de abandonos - media de intervalo entre la graduación y el primer empleo - % de desempleados - % de doctores tras x años (universidades) - media de publicaciones por departamento - % de personal doctorado 	<ul style="list-style-type: none"> - relevancia de la educación en relación con las actividades profesionales x años después de la graduación (HBO) - oferta/demanda de plazas cualificadas, tareas asignadas por el mercado laboral (HBO)

5.3. *Experiencia internacional*

Aparte de las variables más o menos válidas, la investigación reveló algunas variables cuya validez parecía ser controvertida. Una comparación con indicadores y variables propuestos en un contexto directivo de otros países (Reino Unido y Australia) da pie a la cuestión de si esas variables son de alguna utilidad. Especialmente en el Reino Unido se presentaron multitud de advertencias respecto a dichas variables. Describiremos estos datos como medidas del contexto, características explicativas e indicadores esenciales que son difíciles de operativizar.

Medidas del contexto

Algunas variables se mencionan repetidamente en la bibliografía nacional e internacional y en los documentos de gestión, en relación con el desarrollo de un sistema de control de calidad. Sin embargo, dichas variables no constituyen una indicación directa del rendimiento de una institución. Resultan útiles para interpretar los indicadores de rendimiento. Denominamos dichos datos como «medidas del contexto»: aportan la visión del marco de referencia en el que una institución completa de sus tareas. Estos datos no pueden ser manipulados por una institución —o solamente dentro de ciertos límites. Se les denomina con frecuencia como datos de entrada. Una primera categoría de medidas del contexto es cuantitativa y puede describirse como estadísticas de la gestión (véase el Cuadro 8). Además de estas características, existen ciertas medidas cualitativas del contexto que deben ser tenidas en cuenta (Cuadro 9).

Datos clarificadores

Algunas variables están indicadas en la bibliografía nacional e internacional como información de utilidad para obtener una visión en determinados «cuellos de botella». Aportan información valiosa que hace posible la interpretación de los cambios, medidos por los indicadores de rendimiento, de un modo correcto. Quedan reflejados en el cuadro 10.

Cuadro 8. Medidas cuantitativas del contexto

<i>Indicadores</i>	<i>Variables</i>
ingresos de alumnos	<ul style="list-style-type: none"> - número total de ingresos por institución tiempo completo y tiempo parcial - idem por facultad, curso principal - estructura de la población estudiantil por edad - porcentaje de distribución estudiantil por sexo - nota media exámenes finales enseñanza secundaria
personal	<ul style="list-style-type: none"> - ratio profesor-alumno - n.º de horas lectivas por tutor - estadística de días laborables perdidos por baja médica del profesorado
temas financieros	<ul style="list-style-type: none"> - costes de matriculación comparados con otros costes - porcentaje costes para equipamiento - porcentaje coste para biblioteca, % costes para centros audiovisuales - % costes para proyectos innovadores - ratio del presupuesto total en relación con el presupuesto para innovación - % presupuesto para la formación permanente - análisis de costes de las metas educativas, supervisión de presupuestos por tipos de coste - gastos presupuestados frente a gastos reales
infraestructura y medios materiales	<ul style="list-style-type: none"> - espacio docente medio por facultad e institución

Cuadro 9. Medidas cualitativas del contexto

<i>Indicadores</i>	<i>Variables</i>
ingresos de estudiantes	<ul style="list-style-type: none"> - procedimiento de admisión en la educación superior - calidad de los conocimientos pre-universitarios de los estudiantes - origen geográfico de los estudiantes - conocimiento previo de los estudiantes
personal	<ul style="list-style-type: none"> - facilidades de profesores para refuerzo educativo
asuntos financieros	<ul style="list-style-type: none"> - gestión financiera - gasto presupuestado frente a gasto real
infraestructura y medios audiovisuales	<ul style="list-style-type: none"> - centros audio-visuales, medios de comunicación como ayuda a la tutoría

Cuadro 10. Datos clarificadores

<i>Indicadores</i>	<i>VARIABLES</i>
permanencia de estudiantes	<ul style="list-style-type: none">- motivos de abandono- % distribución normativa en contraposición a la duración empírica de los estudios por curso

Indicadores sin una suficiente operativización

Tanto en Estados Unidos como en el Reino Unido y Australia, la «producción» es considerada como uno de los principales indicadores. La operativización de la producción ha dado lugar a gran calidad de debates e investigación. Con frecuencia, las variables son mencionadas como: juicio por parte de los empleadores acerca de los «graduados» —el resultado— en relación al mercado laboral (esto es un método, no un indicador), actividades profesionales después de la graduación, tiempo medio de lapsus entre la graduación y el primer empleo por cada campo de estudio. Algunas cuestiones siempre vuelven a plantearse en este tipo de discusiones. En primer lugar, ¿no es el «valor añadido» un mejor indicador de la calidad de la educación que el porcentaje de graduados o los exámenes aprobados? En segundo lugar, el «valor añadido» no solamente concierne una «ganancia de conocimiento» sino también «una adquisición de habilidades y actitudes». ¿Qué habilidades y actitudes son esenciales para funcionar correctamente en la sociedad o en un empleo? ¿Cuál es la importancia relativa de los distintos aspectos de la personalidad? Además de estas cuestiones sobre la importancia de los datos del producto aparece otro problema: ¿cuán fiable es la medición y cuáles son los costes para desarrollar y operar con dichas medidas?

Además de estos problemas, es importante invertir en una posterior investigación. Partiendo de experiencias internacionales, puede resultar de cierta utilidad el presentar como posible, la siguiente operativización del indicador «producción» (véase Cuadro 11).

La variable «valor añadido» puede ser representada como una especie de contribución al desarrollo personal (Dochy y Bouwens, 1990). Además de este indicador de «producción» otras variables son vagamente operativizadas. Ello significa que no son mensurables de un mo-

Cuadro 11. Indicadores y variables de egresión o salida

<i>Indicadores</i>	<i>Variables</i>
egresión	<ul style="list-style-type: none">- tiempo medio entre la graduación y el primer empleo por área de estudio- relevancia de la educación en relación con actividades profesionales x años tras la graduación (HBO)- % de graduados con tesis presentadas tras x años por facultades (universidades)- adecuación de los graduados respecto al mercado laboral y la sociedad- demanda y oferta de plazas cualificadas y tareas de graduados para el mercado laboral (HBO)- ratio desempleados/graduados por campo de estudio

do concreto. Tales variables son: política de selección de personal, calificaciones educativas, motivación por la docencia, grado de innovación o grado de orientación innovadora y política de selección de los investigadores.

La complejidad de las variables también juega su papel. Ello explica porque además de las variables concretas (estructura de personal según la edad, movilidad del personal, porcentaje de personal con un número dado de años de experiencia laboral fuera de la actual institución), se sugiere en la bibliografía una variable más vaga en cuanto a su formulación (la política de selección).

6. Viabilidad del uso de los indicadores de rendimiento

Evidentemente, los indicadores de rendimiento pueden desempeñar un papel significativo en el diálogo entre el gobierno y las instituciones, como parte de un sistema externo de evaluación de la calidad. Los criterios para la selección de los indicadores de rendimiento están ampliamente aceptados. Sin embargo, no existe consenso en cuanto al peso relativo de los distintos criterios. Nuestro proyecto de investigación tomó como punto de partida la validez empírica de los indicadores de rendimiento. Aquellos indicadores que pudieran aparecer como controvertidos se han comparado con la experiencia internacional. Muchos de estos indicadores pueden definirse como características propias (medidas del contexto) y datos de clarificación. Estos datos debieran formar

parte de un sistema de evaluación de calidad. Sin embargo, quedan por responder algunas cuestiones, teniendo en cuenta la idea de que la serie de indicadores de rendimiento ha de ser fácil de operar, por ejemplo, dentro de unos límites aceptados de costes y energía de los distintos participantes.

Algunas variables resultan todavía demasiado generales debido a su complejidad. Sería deseable mayor volumen de investigación sobre indicadores de rendimiento tales como «valor añadido», aunque no siempre parece prudente trasladar a cifras cada uno de los indicadores. Un enfoque mecanicista de este tipo tiende a perder con frecuencia muchos aspectos de la realidad.

Aunque puede proponerse una serie de indicadores de rendimiento, en base a ciertos criterios, es necesaria una reflexión acerca de sus posibilidades y limitaciones. Es especialmente necesaria la reflexión acerca de estas cuestiones cuando se trata de elaborar un sistema de información para la gestión.

En primer lugar, es deseable la reflexión y la aceptación del rol de los indicadores de rendimiento por parte de los distintos participantes. Solamente entonces puede desarrollarse un sistema fiable de información. Los indicadores del rendimiento pueden interpretarse como señales o guías. Indican los problemas y constituyen, de alguna manera, el punto de partida de la planificación. En segundo lugar, pueden servir de ayuda para emitir juicios. Finalmente, son un elemento esencial en el diálogo entre el gobierno y las instituciones de educación superior.

Los indicadores de rendimiento pueden desempeñar un importante papel solamente si forman parte de una serie coherente de los indicadores de entrada, proceso y producto. Los *inputs*, los indicadores del proceso y los datos del contexto son importantes en la medida que contribuyen a una interpretación válida y fiable. La cuestión respecto a si los indicadores han de ser cuantitativos o cualitativos da pie a respuestas contradictorias. Cave *et al.* (1990) señalan que los indicadores más importantes (p. ej., alumnos egresados) son los más difíciles de operativizar en variables cuantitativas. Aparecen las mismas dificultades en los resultados de la investigación aquí aprotados. Por el momento, resulta todavía imposible de expresar en cifras todos los aspectos que influyen en el funcionamiento de las instituciones de educación superior. Queda abierta la cuestión de si resultará posible alguna vez y, en ese caso, si

sería deseable hacerlo. En cualquier caso se requiere más investigación para operativizar en variables más concretas los indicadores vagamente señalados y para desarrollar instrumentos de medición más fiables. Teniendo en cuenta este problema, resulta peligroso el prestar demasiada atención a aquellos objetivos fáciles de medir. Además de un posible exceso de énfasis en los objetivos mensurables, debe prestarse atención a la explicación de los objetivos por parte del gobierno mediante consultas bilaterales.

Cada institución debe definir sus objetivos en conexión con los establecidos a nivel central y en relación a su propia misión. Tanto la interpretación como el peso relativo de los indicadores de rendimiento deben ser tenidos en cuenta en esta misión. Proponemos tomar los objetivos institucionales como punto de partida para la elaboración de un sistema de control de calidad. Ello implica una reflexión crítica sobre dichos objetivos, tanto desde la perspectiva del gobierno como desde la relevancia social. Además, los indicadores de rendimiento deben iluminar las tendencias o desarrollos y no deben ser solamente flashes aislados. Si se desea una información comparativa, deben tenerse en cuenta los diferentes objetivos de las distintas instituciones; solamente es aceptable una comparación de igual a igual. Finalmente, la puesta en práctica de una serie de indicadores de rendimiento sólo es posible si se limitan a un número reducido.

Indicadores de rendimiento y revisión de colegas

La evaluación se ha desarrollado durante mucho tiempo en base a unos criterios previamente establecidos pero con frecuencia implícitos. La sociedad actual requiere una serie de indicadores explícitos previamente establecidos. Más aún, dado que los indicadores del contexto y los datos de clarificación son fundamentales para la interpretación y dado que no es posible (y quizá no lo será nunca) el expresar todos los indicadores como números, es necesario un acuerdo acerca de los métodos de evaluación. La revisión de colegas puede ser un método aceptable, aunque lento, para los indicadores que no son mensurables y para los indicadores esenciales del contexto.

En nuestra opinión, es evidente que si los métodos institucionales pueden ser operativizados en variables cuantitativas que aporten una información válida acerca de uno o más indicadores del rendimiento, es

aconsejable el utilizarlos. También parece interesante el reunir estos datos de forma sistemática, una año tras otro. De este modo, se hace visible el progreso hacia las metas propuestas. Sin embargo, lo que resulta necesario es completar el cuadro, por medio de la reunión de información de la gestión para interpretar las variables y los indicadores. No resulta inteligente el emplear la revisión de colegas bajo las circunstancias adscritas, ya que se trata de un método demasiado lento y costoso. Pero, debido a que resulta difícil el operativizar algunos objetivos institucionales y algunos indicadores en variables solamente cuantitativas, ciertamente la revisión de colegas puede desempeñar un papel en el sistema de control de calidad. Por supuesto, la revisión de colegas, aunque es un método costoso, también es un medio razonable de obtener información acerca de las variables cualitativas y las medidas del contexto que pueden desempeñar un papel crucial en la interpretación de los datos referentes al rendimiento.

UTILIZACION DE INDICADORES EN LA EVALUACION DE LA DOCENCIA UNIVERSITARIA

MARIO DE MIGUEL DIAZ
Universidad de Oviedo, España

El trabajo que presentamos pretende abordar la evaluación de la unidad de docencia. Es decir, partiendo de un planteamiento de evaluación institucional, *pretendemos clarificar cuáles son los factores o aspectos relacionados con la docencia en el ámbito universitario y determinar los criterios a tener en cuenta para su evaluación*. Tratamos de establecer una visión global sobre todos los procesos implicados en la docencia universitaria y delimitar los principales indicadores que podemos utilizar para conocer su funcionamiento.

El planteamiento institucional que asumimos en este trabajo trata de evitar los reduccionismos con que habitualmente se utiliza el término docencia referido al nivel universitario. En primer lugar queremos establecer una clara distinción entre lo que supone la unidad de docencia universitaria y la actividad docente que realiza un profesor. Los aspectos y procesos que conlleva la función docente en una Universidad en modo alguno pueden ser reducidos a la actividad que desarrolla un profesor en clase. Por consiguiente plantear la evaluación de la docencia universitaria a través de la evaluación del profesor constituye un reduccionismo que está generando bastante confusión. Por ello entendemos que antes de abordar el tema de la evaluación previamente se deben clarificar los componentes de la unidad de docencia y delimitar las competencias específicas de cada órgano que interviene en la misma. La evaluación cobra mayor significado en la medida que cada elemento interviniente es valorado partiendo de su cuota de responsabilidad.

Otro reduccionismo que conviene igualmente evitar cuando se habla de docencia universitaria es limitar esta función a la actividad que se desarrolla en el aula. Del mismo modo que no se puede identificar la

unidad de docencia con la función docente del profesor, tampoco se puede considerar que todas las actividades relacionadas con la docencia pueden quedar restringidas a la clase. El ejercicio de la función docente supone desarrollar diversas tareas y en distintos momentos que deben ser integradas y contempladas como un todo donde no cabe destacar una sobre otra, dada la interrelación entre ellas. Por ello antes de proceder a establecer un proceso de evaluación de la docencia debemos construir el modelo que enmarque todas las dimensiones implicadas en esta actividad. De lo contrario, se ofrecería una visión muy parcial de lo que supone esta actividad en la Universidad.

Finalmente el tercer reduccionismo que intentamos evitar con nuestro planteamiento se refiere a la finalidad de la función evaluadora en la enseñanza superior. Nuestra postura al respecto es que la evaluación de la docencia como parte o unidad de la evaluación institucional es algo más que un proceso de recogida de información orientado a tomar decisiones con fines sumativos, sino que constituye un proceso orientado hacia el cambio y la mejora de los procesos de enseñanza en la institución. Ello significa que, previamente al diseño del modelo de evaluación, debe existir la adecuada información y consenso entre los elementos integrantes sobre lo que se espera de ellos y los medios y ayudas que tienen a su disposición para conseguirlo. Sin la fase de advertencia y monitorización no cabe establecer ningún tipo de evaluación formativa, y sin esta tampoco cabe hablar de cambio institucional.

1. La unidad de docencia. Principales componentes

La evaluación de la institución universitaria, como cualquier otro tipo de organización; puede ser abordada desde diversas perspectivas teóricas y metodológicas. La aproximación más común es estructurar esta evaluación según las actividades de producción propias de esta organización, a saber: docencia, investigación y servicios que ofrece. (Algunos autores consideran que se debe añadir un cuarto factor —la unidad administrativa— dada su implicación directa sobre los anteriores). Aunque toda ordenación entre estas actividades puede ser criticable, la opinión más generalizada entre el profesorado (Belanger, 1990; De Miguel, y Rodríguez, 1990) coincide en destacar a la docencia como la finalidad de la enseñanza superior sobre la que debe recaer mayor peso tanto en el interés como en la actividad. Es lógico, por tanto, que se

aborde prioritariamente la evaluación de la docencia como unidad de producción de la institución universitaria bien desde una perspectiva interna —evaluación de los procesos de enseñanza— como externa, es decir analizando los resultados de dichos procesos en función de sus costos (*accountability*).

Desde la perspectiva interna, antes de establecer cualquier tipo de consideración sobre la evaluación de la docencia en relación con las enseñanzas relativas a un título, programa o curso, debemos proceder a determinar cuáles son los principales aspectos o componentes que intervienen en todo proceso de enseñanza a nivel institucional y cuáles son las unidades u órganos que tienen la responsabilidad de su gestión. Como ya hemos dicho, la dimensión docente en la Universidad no se limita a la enseñanza que realiza un profesor concreto en el aula, sino que implica la *actuación de una serie de órganos e individuos que toman decisiones sobre qué cosas se van enseñar, cómo, cuándo y por quién van a ser impartidas las enseñanzas y con qué objetivos o finalidad*.

Tomando como referencia un modelo de producción, se podría decir de manera simplificada que la Universidad en primer lugar establece una oferta académica —planes de estudio relativos a los títulos y cursos que ofrece— en relación a los cuales obtiene una demanda de alumnado con determinadas características (*input*). En un segundo momento, la institución organiza las enseñanzas y desarrolla una serie de actividades a fin de que la demanda aceptada adquiera los conocimientos y habilidades correspondientes (procesos). Finalmente, fruto de estos procesos y actividades se obtienen unos resultados tanto a nivel de los sujetos como de la propia institución (producto). Todo lo cual se realiza, a su vez, dentro de una organización específica —con una cultura e infraestructura determinada— que envuelve y mediatiza, no sólo la concepción de la docencia sino también el proceso de producción (contexto). Una especificación de estos componentes se refleja en el cuadro adjunto.

A su vez, cada uno de estos componentes de la docencia son administrados dentro de la institución universitaria a través de una serie de unidades básicas de gestión. Estas unidades se definen como los elementos constituyentes más pequeños de las instituciones académicas con responsabilidad propia por lo que respecta a la política académica y presupuestaria (Premfors, 1989; Bercher, 1989). Aunque pueden tomar variedad de formas según el contexto de cada país —departamentos, institutos, servicios, programas, cátedras, etc.— estas unidades básicas

CUADRO N.º 1: Componentes de la unidad de docencia a nivel institucional

CONTEXTO	INPUT	PROCESO	PRODUCTO
Condicionantes de la institución.	Características de las enseñanzas a evaluar.	Variables relativas a los procesos enseñanza-aprendizaje.	Evaluación de los resultados.
Infraestructura: — Recursos — Alumnado — Profesorado	Reconocimiento social de las Enseñanzas. Prestigio del Profesorado. Nivel de entradas de los alumnos.	Organización de la docencia. Elaboración de Programas. Actividad docente en el aula.	Rendimiento de los alumnos. — Durante las enseñanzas. — Al finalizar los estudios. — A largo plazo.
Cultura Organizacional	Adecuación de edificios e instalaciones. Dotaciones y equipamientos necesarios.	Actividades académicas fuera del aula.	Resultados a nivel institucional — Sobre eficacia. — Sobre eficiencia. — Cambio institucional.

de gestión de la docencia se pueden agrupar en tres niveles en función de sus competencias en relación con las cuestiones académicas, a saber: competencias a nivel de individuos (profesorado), a nivel de subunidades (departamentos, centros, institutos,...) y a nivel de la Organización o Institución.

Dado que cada una de estas tres unidades de gestión tienen distinta participación en las cuestiones relativas a la docencia debemos intentar delimitar sus funciones a fin de que cada una de ellas pueda ser evaluada desde su cuota de responsabilidad. A título indicativo —puede variar en cada país y Universidad según sus Estatutos— los ámbitos de competencia de estas tres unidades se pueden establecer en estos términos:

A. Competencias a nivel de la Organización

Los órganos de Gobierno en una Universidad deben asumir el mayor grado de responsabilidad sobre la calidad de la docencia. De ellos depende, entre otras cuestiones, la propuesta y/o aprobación de las titulaciones y planes de estudio, la elaboración de criterios para la organización de la docencia, la política a seguir en materia de selección, formación y perfeccionamiento del profesorado, y el sistema de evaluación y mejora de la institución universitaria. Las decisiones que se tomen en relación con estas cuestiones es fundamental, entre otras razones por su repercusión sobre las otras unidades de gestión.

B. Competencias a nivel de Subunidades

El siguiente órgano de responsabilidad se sitúa a nivel de las subunidades académicas (centros, departamentos, instituciones, etc...). La delimitación de funciones entre ellas se establece normalmente a través de la reglamentación estatutaria de cada Universidad, aunque de forma genérica se podría decir que los Departamentos tienen a su cargo lo relativo a las cuestiones académicas relativas al profesorado —elaboración de programas, distribución de la docencia, coordinación entre el profesorado según niveles y ciclos, criterios de evaluación, etc.—, mientras que los centros se ocupan de estas cuestiones desde el prisma de los alumnos: organización de la actividad docente, horarios, prácticas, tutorías, evaluaciones, etc. No obstante queremos insistir que es imposible establecer una distinción clara al respecto ya que las diferencias entre las competencias de estas subunidades es patente entre los países y Universidades.

C. Competencias a nivel individual

El tercer nivel de competencia o responsabilidad de la docencia universitaria recae sobre el profesorado. Al margen de las decisiones que en cada caso se arbitren en relación con la docencia universitaria no cabe duda que finalmente es el profesor el que asume la tarea de enseñar, y que puede desempeñar dicha tarea con más o menos dedicación e interés, más o menos preparación, y mayor o menor grado de habilidad para conectar con los alumnos. Las diversas características y

factores que intervienen en las funciones —instructor, organizador y orientador— que debe desempeñar, y en las actividades a desarrollar antes de enseñar, durante el acto didáctico, y en la reflexión-evaluación posterior, determinan el éxito o fracaso de su papel y buena parte de la actividad universitaria.

Del análisis anterior se deduce que, desde la perspectiva interna, la función docente en una Universidad se puede estructurar en una serie de aspectos o componentes que intervienen en los procesos de enseñanza y en una serie de órganos e individuos encargados de gestionar estas enseñanzas. Abordar la evaluación desde un planteamiento institucional supone tener en cuenta todos estos aspectos de manera global. Este es el propósito que guía el presente trabajo. Pretendemos ofrecer un proyecto marco de lo que supone la evaluación de la docencia desde una visión comprensiva de todos los factores que intervienen en la misma a nivel institucional, marco del que posteriormente se podrían establecer programas de evaluación concretos elaborados según la diversas unidades básicas de gestión a las que hemos aludido.

2. Evaluación de la docencia: Criterios e indicadores

La evaluación de la docencia universitaria es una tarea compleja dada la peculiar situación que esta actividad tiene en el ámbito universitario. De una parte, es una actividad infravalorada, que la mayor parte de las veces se ejerce por obligación y de manera rutinaria, y sobre la que se tiene la impresión de que no rentabiliza el tiempo que se le dedica, especialmente cuando se compara con los esfuerzos dedicados a la investigación. De otra, las condiciones en las que habitualmente se desarrolla —masificación, métodos de enseñanza tradicionales, escasa renovación de los contenidos, etc.— han dado lugar que el interés por la docencia haya ido progresivamente disminuyendo —tanto en hombre como en mujeres— al igual que en el resto del profesorado de los otros niveles del sistema educativo.

Por otro lado, las referencias actuales sobre evaluación institucional intentan superar esquemas reduccionistas centrados sobre los elementos personales —profesores y alumnos— y tratan de abordar la evaluación de la actividad universitaria desde una perspectiva global. La búsqueda de modelos y estrategias para evaluar la calidad de la enseñanza superior constituye una de las líneas más sólidas dentro de la investigación educa-

tiva actual, que cuenta con múltiples trabajos (Dunkin y Barnes, 1986; Goedegebuure *et al.*, 1990). A título de síntesis se podría decir que la literatura al respecto se agrupa en torno a dos tendencias: *evaluación externa de las instituciones en base a indicadores de rendimiento* —línea potenciada por las autoridades gubernamentales de muchos países— o *partiendo de un sistema de expertos (peer review)*, —sistema que se considera más apropiado en los ámbitos universitarios— y *evaluación interna (autoevaluación o autorregulación)*.

La evaluación en base a indicadores y a un sistema de expertos — aunque conllevan procedimientos de trabajo distintos— no debe entenderse como modelos antagónicos, dado que, en la práctica, son las *normas de referencia sobre la calidad* de la enseñanza las que definitivamente configuran esta evaluación (Cave *et al.*, 1988; Dochy, Segers y Wijnen, 1990).

Esta es la razón por la que, antes de cualquier otro planteamiento sobre la evaluación, debemos determinar los criterios e indicadores que vamos a utilizar para estimar la calidad de un título o programa. En nuestro caso, se trata de buscar una serie de patrones y medidas que permitan una valoración lo más objetiva posible de la actividad docente desarrollada en una Universidad. Este problema, en apariencia simple, conlleva concretar previamente dos aspectos bastante complejos, a saber: *definir cuáles son los criterios de calidad para evaluar la docencia en la enseñanza superior, seleccionar qué tipo de indicadores son apropiados para evaluar sus distintos aspectos o dimensiones, y determinar que tipo de procedimiento consideramos en cada caso apropiado para efectuar la evaluación. En definitiva, establecer los criterios, indicadores y procedimientos a través de los cuales vamos a realizar la evaluación institucional de la docencia.*

Se conoce como *criterio de evaluación* un objetivo de calidad respecto al cual puede ser valorado un determinado fenómeno o proceso. Ello significa que la selección de criterios nos remite al concepto de calidad, en nuestro caso de calidad docente. La definición que se adopte sobre este término no sólo determina los objetivos a lograr sino también la valoración de los resultados y los procesos que se establezcan para alcanzarlos. No es que existan conceptos distintos de calidad, sino que las personas tienen diferentes experiencias y expectativas en torno a ella, que les lleva a plantear formas de apreciación diferentes. Ello determina que el *concepto de calidad de la docencia sea un concepto relativo*

y multidimensional en relación a los objetivos y actores del sistema universitario (Wersterheijden, 1990a). Su análisis debe hacerse en el contexto de los procesos sociales y políticos en el que interaccionan objetivos y actores. Desde un punto de vista práctico la concreción de este concepto queda supeditada a la definición operacional que se adopte a partir del enfoque evaluativo desde el que se aborda la evaluación de la calidad.

Por ello, dada la multiplicidad de elementos y procesos que intervienen en la docencia universitaria, no es fácil determinar criterios claros para evaluar su calidad, dificultad que se acusa en comparación con el proceso de evaluación de la investigación (De Weert, 1990). Además, los condicionantes contextuales en los que se desarrolla la docencia determinan que los criterios seleccionados difícilmente sean aceptados por toda la comunidad universitaria, especialmente cuando no existe normas objetivas que orienten el proceso de valoración.

A título de aproximación algunos autores (Halpern, 1987) se plantean una definición en términos de producción y consideran como «calidad» la mayor diferencia que es posible establecer entre el input y el output con el menor costo. Desde esta perspectiva, calidad docente equivaldría al «provecho educativo» estimado al comparar el rendimiento que presentan los alumnos cuando inician las enseñanzas relativas a un título, programa o curso, respecto al que presentan cuando finalizan estos estudios. Una institución es más eficaz en la medida que a través de su docencia logra que los alumnos obtengan mayores «provechos educativos», es decir, mejores rendimientos. Sin embargo esta definición no resuelve el problema ya que el concepto de rendimiento es a su vez un término relativo, que puede operativarse de formas muy distintas: —a) como conjunto de conocimientos adquiridos, b) como tiempo que invierte el alumno desde que ingresa en la Universidad hasta que obtiene un trabajo estable, c) como rentabilidad económica que obtiene un egresado, d) como satisfacción del cliente, etc.— dando lugar a diversos criterios de *eficacia*.

Además, al igual que todo producto debe ser evaluado en función de sus costos, la actividad docente de una institución debe ser valorada a partir de los medios y recursos con que cuenta. Cabe, pues, establecer criterios sobre la *eficiencia* que demuestra una institución en la racionalización de los procesos y recursos empleados en la gestión de las enseñanzas que imparte. La utilización de este tipo de criterios no sólo posibilita un nuevo concepto de producto —el máximo «valor añadido» para

los mismos «inputs»— sino que permite establecer comparaciones entre instituciones que imparten el mismo tipo de enseñanzas.

Finalmente toda institución educativa, en tanto que organización, necesita revisar y replantear constantemente sus fines, estructuras y normas de funcionamiento, no sólo para obtener mejores resultados, sino también para que sus objetivos se adapten a las necesidades sociales. Es decir, necesita situarse en un *proceso de cambio e innovación* que le permita hacer cada día las cosas mejor. Por ello, al evaluar la docencia a nivel institucional también debemos utilizar criterios relativos a la calidad de los objetivos y procesos establecidos en relación con la mejora y la innovación de las enseñanzas.

Por otro lado (Guba y Lincoln, 1988), cualquier tipo de objetivo de calidad que se establezca como criterio puede ser evaluado desde una perspectiva intrínseca o extrínseca, es decir, según su mérito o valor. Se dice que determinados resultados son valiosos en tanto que se adecuan a una realidad, es decir, responden a las necesidades o a una demanda de un contexto que es quien aporta el criterio de calidad (valoración extrínseca). Por el contrario, otros resultados tienen mérito por sí mismos, independientemente del contexto social (valoración intrínseca), es decir, su valor reside en su calidad. Ambos sistemas de valoración se superponen en todos los procesos de evaluación y deben ser tenidos en cuenta ya que no es lo mismo juzgar un determinado tipo de enseñanzas desde el punto de vista de la calidad que de la utilidad. A partir de los años setenta cada vez se viene cuestionando con mayor fuerza la aplicación del modelo de producción a la enseñanza universitaria —buenos gestores, buenos productores y buenos resultados—, en favor de un modelo centrado sobre el desarrollo del talento —desarrollo frente a producción— cuya preocupación esencial se dirige a lograr una educación para la excelencia (Astin, 1985). Desde esta perspectiva, la evaluación institucional necesariamente tiene que abordar otros desarrollos del alumno —competencias, conductas y actitudes (Hagerty y Stark, 1989)— y de la organización —cultura (Tierney, 1988)—, que normalmente no se tienen en cuenta.

Los criterios, además, deben explicitar las normas que permiten establecer la valoración de los distintos aspectos que se someten a evaluación. Un criterio sin normas no sólo deja el proceso de evaluación en el terreno de lo opinable sino que dificulta el análisis comparativo. Ahora bien, cuando se utilizan normas cabe distinguir entre aquellas que consti-

tuyen un criterio standar, y como tal son aceptadas colectivamente, y aquellos otros criterios con referencia a normas que no son standar, es decir, que necesitan previamente de un acuerdo intersubjetivo. Así, por ejemplo, la valoración de la ratio de egresados matriculados o el número de alumnos por aula pueden considerarse criterios con referencia normas standar cuando se utilizan en relación con la media global, mientras que el tipo de metodología docente o el sistema utilizado en la evaluación de los alumnos constituyen criterios cuyas normas no son standar y por tanto reclaman previamente un consenso intersubjetivo para que cada observador pueda «leer el mismo valor» a partir de los datos que ofrece un instrumento o indicador. De lo contrario, lo que para unas personas son criterios de calidad para otras simplemente son «datos».

De acuerdo con todo lo anterior, en la determinación de criterios en una evaluación deberemos tener en cuenta los siguientes elementos: a) la naturaleza de los factores a evaluar (relativos al contexto, input, proceso y producto), b) el tipo de objetivos de calidad que se persiguen en cada caso (eficacia, eficiencia y mejora), c) el punto de vista o perspectiva que se asume en la evaluación (intrínseca o extrínseca) en relación con el producto y d) el tipo de referencias que se utilizan para establecer la evaluación: sin normas o con normas (estándar o no estándar).

Una vez establecidos los criterios de evaluación procede seleccionar los indicadores a través de los cuales vamos a obtener información pertinente en relación con el fenómeno que constituye objeto de evaluación. Ello supone trasladar las dimensiones generales que configuran el fenómeno que es objeto de estudio —en este caso la docencia universitaria— a un conjunto de variables observables empíricamente. Es decir, operativizar las dimensiones del constructo docencia en términos que

CUADRO N.º 2: Aspectos a tener en cuenta en la elaboración de criterios de evaluación

Factores	Objetivos a evaluar	Perspectiva de valoración	Criterios de Comparación
Contexto Input Proceso Producto	Eficacia Eficiencia Mejora	Mérito/Desarr. Valor/Utilidad	— normas estándar — normas no estándar — sin normas previas

puedan ser observados. Lógicamente la coherencia entre las dimensiones del fenómeno y las variables seleccionadas para su análisis determina la validez de nuestra estimación. Todo este proceso es necesario para que los datos empíricos —cuantitativos o cualitativos— que recojamos a partir de las variables seleccionadas nos informen adecuadamente sobre la situación en la que se encuentra la docencia en una institución. Estos datos empíricos constituyen los indicadores, es decir, unas señales o pistas que permiten estimar dentro de una institución las diferencias entre objetivos previstos y resultados, y comparar distintos niveles de logro entre instituciones que con medios parecidos persiguen metas similares. *De ahí que se entienda por indicador de rendimiento un ítem de información recogido de forma regular que permite averiguar la productividad de un sistema.*

Desde el punto de vista metodológico, la selección de los indicadores a través de los cuales vamos a evaluar un objetivo reclama tener presente diversos aspectos a fin de que los datos que obtengamos puedan utilizarse de forma comparativa. Esencialmente implica los siguientes pasos: a) seleccionar las variables que vamos a observar sobre cada una de las dimensiones o aspectos de la docencia, b) determinar el tipo de media u observación —cuantitativa o cualitativa— que se va a efectuar, c) establecer la modalidad o sistema que se considera más apropiado para efectuar esta observación y recogida de datos (sistema de expertos o indicadores), d) realizar el proceso de valoración de los datos obtenidos en función de las normas explícitas en los criterios o establecer los acuerdos intersubjetivos necesarios para ello, y e) determinar la naturaleza y el valor de las medidas obtenidas a través de la evaluación (datos estadísticos, datos mediatizados por juicios, juicios, opiniones e informes verbales). Un resumen de este proceso puede quedar reflejado en el siguiente esquema (Dochy, Segers y Wijnen, 1990).

Proceso en la definición operacional del concepto docencia

Plano Teórico	Dimensiones	1.	2.	3.
	Indicadores	1,2,3,...	1,2,3,...	1,2,3,...
Plano Empírico	Variables	1.1,1.2,1.3...	1.1,1.2,1.3...	1.1,1.2,1.3...
		2.1,2.2,2.3...	2.1,2.2,2.3...	2.1,2.2,2.3...
		3.1,3.2,3.3...	3.1,3.2,3.3...	3.1,3.2,3.3...

Ultimamente, la búsqueda de indicadores en relación con la docencia ha sido una de las constantes de la investigación institucional. El número de trabajos al respecto ha aumentado considerablemente, especialmente desde que el informe Jarratt (CVCP, 1985) recomendara su utilización (Cave *et al.*, 1988; Kurz *et al.*, 1989; Franke-Wikberg, 1990). Sin embargo, el modelo tiene una serie de problemas implícitos relativos a su definición validez, fiabilidad e intersubjetividad para las distintas audiencias que no son fáciles de resolver (Weert, 1990). Quizás el más importante es que estos indicadores aportan «medidas» que deben evaluarse a partir de unos standars que debe decidir la propia institución. Aunque teóricamente se pueden establecer las dimensiones o factores que condicionan la calidad de la docencia, desde el punto de vista práctico los objetivos a lograr en cada uno de estos factores constituye una decisión a determinar por cada institución. Cuando tales objetivos no están establecidos difícilmente se puede efectuar evaluación institucional partiendo de indicadores de rendimiento.

3. Principales indicadores sobre docencia universitaria

Una vez que hemos establecido el marco que determina los componentes o factores que intervienen en la docencia y los requisitos que conlleva la determinación de criterios e indicadores para evaluar estos factores, vamos a presentar algunos de los indicadores más utilizados en la evaluación de la enseñanza superior. Para ello hemos procedido de la siguiente manera:

- En primer lugar hemos especificado los objetivos fundamentales de la docencia siguiendo el modelo de producción, es decir, tomando como referencia los factores del contexto, input, proceso y producto, relacionados con las enseñanzas correspondientes a un título o curso.
- A su vez, para cada uno de estos objetivos a evaluar, hemos seleccionado una serie de dimensiones o aspectos observables a través de los cuales estimamos que es posible acercarnos a su evaluación.
- Finalmente, una vez establecidas las principales dimensiones de cada objetivo, hemos señalado algunas de las variables que se pueden utilizar como indicadores de calidad, tomando como re-

ferencia trabajos realizados sobre este tema que se relacionan en la bibliografía.

Siguiendo este procedimiento, a continuación presentamos una justificación de los principales objetivos y dimensiones que consideramos necesarios para evaluar la docencia y una relación de algunos de los indicadores que se pueden utilizar para llevar a cabo esta evaluación.

3.1. *Evaluación de los factores contextuales*

El contexto de las instituciones universitarias es muy distinto entre unas y otras. Una evaluación que no tenga en cuenta las diferencias entre las Universidades anula buena parte de la evaluación y falsea todo proceso comparativo. De ahí que necesariamente toda evaluación deba contextualizarse. Ello significa que tanto en el diseño, como en el proceso e interpretación de resultados, la evaluación debe tener en cuenta los factores que condicionan las enseñanzas en cada institución concreta.

El primer factor condicionante de toda actividad universitaria tiene un carácter político. En la medida que depende de un poder central no sólo se compromete la definición de la propia Universidad como organismo autónomo, sino que también se comparte la responsabilidad sobre los resultados. Es lógico, por tanto, que se intente evaluar la calidad de la docencia desde la cuota de autonomía que una Universidad tiene. Por ello un complemento inevitablemente unido a la evaluación es la rendición de cuentas (Vandament, 1989; Thomas, 1990), es decir, la estimación de la eficiencia que demuestran las Universidades en la administración del dinero público. En la medida que los recursos de la Universidad dependen totalmente de un órgano central no se puede hablar de autonomía, ni tampoco se puede sustraer a una evaluación externa.

Existe un segundo factor o condicionante contextual de la docencia universitaria de carácter social que tiene que ver con un problema clave de la propia institución universitaria en tanto que fluctúa entre ofrecer una enseñanza de calidad —restringida a unos pocos alumnos— o extender la universidad al mayor número de sujetos. El dilema entre calidad/selección y extensión/masificación es un problema no resuelto en las Universidades públicas que origina agravios comparativos evidentes. Mientras que a unos centros se les permite establecer numerus

clausus otros se ven obligados a admitir cuantos alumnos desean matricularse. Las diferencias se acusan mucho más en la medida que se tiene en cuenta la disponibilidad de equipamientos y medios. La relación entre los recursos y el número de alumnos constituye un condicionante fundamental de la docencia universitaria del que no se puede prescindir a la hora de efectuar la evaluación.

El tercer factor contextual que igualmente condiciona la docencia es el tipo de cultura sobre la enseñanza que funciona dentro de la organización de un centro concreto. La peculiaridad de las organizaciones educativas —dado el carácter dual de organizaciones estructuradas partiendo de modelos racionales/jerárquicos pero que a la vez funcionan internamente de manera anárquica— tal como sostienen las teorías recientes (de Miguel, 1989), determina que el desarrollo de la actividad docente esté mediatizado por la cultura organizacional que tiene el propio del centro o institución sobre la enseñanzas que imparte (Tierney, 1988). El sentido y utilidad que se otorga a la docencia en las Ingenierías es muy distinto al que se emplea entre las Humanidades. No se pueden evaluar con los mismos parámetros centros o enseñanzas con culturas organizacionales claramente distintas. Ello significa que toda evaluación de la docencia a nivel institucional debe estar referida a un tipo de enseñanza o, si se prefiere, a un título. De esta forma es posible tener en cuenta las «culturas» propias de determinados sectores académicos y a la vez poder establecer comparaciones entre diversos centros que imparten el mismo tipo de enseñanzas.

Desde el punto de vista práctico, estos factores determinan que las Universidades se diferencian unas de otras no sólo en su «infraestructura» sino también por su «cultura». Las diferencias son tan patentes que necesariamente deben utilizarse indicadores que permiten establecer comparaciones entre aquellas instituciones que realmente son comparables en su infraestructura para evaluar enseñanzas —títulos— que son igualmente equiparables por su «cultura». Algunos de estos indicadores son los siguientes:

CUADRO N.º 3: Objetivos e indicadores relativos al contexto de una institución

OBJETIVOS A EVALUAR	ASPECTOS Y DIMENSIONES	PRINCIPALES INDICADORES
Infraestructura/ Recursos	Económicos Instalaciones Equipamiento	<ul style="list-style-type: none">– Presupuestos asignados a la institución– Media de costos por alumno– Porcentaje de gastos dedicados a docencia– Metros cuadrados por alumno en el campus– Tasa de equipamientos especiales (ej. ordenadores) por alumno– Dotación de laboratorios, medios audiovisuales, Ayudas a la docencia, Biblioteca, etc.– Dotación para proyectos de innovación
Características del Profesorado	Número Dedicación	<ul style="list-style-type: none">– Ratio de alumnos por profesor– Estructura académica del profesorado– Número de horas dedicadas a docencia por profesor– Días lectivos durante el curso escolar
Características del Alumnado	Efectivos Admisión	<ul style="list-style-type: none">– Número total de alumnos en la institución– Estructura social de la población estudiantil– Procedencia geográfica de los estudiantes– Procedimientos de admisión en la institución

3.2. *Evaluación de los planes de estudio*

Los títulos y cursos que se imparten en una Universidad concreta constituyen la oferta docente que dicha Universidad ofrece. Dicha oferta se hace además con una finalidad específica y de una manera concreta que se plasma en los planes de estudio. En la medida que las Universidades tienen autonomía pueden elaborar los planes relativos a títulos nacionales —siguiendo las directrices oficiales— e igualmente diseñar títulos y cursos propios (postgrado, master, especialización, etc.) de acuerdo con normas establecidas al efecto.

Las decisiones que se tomen en relación con estos planes son vitales para toda la actividad docente universitaria. No es lo mismo optar

por un modelo centrado sobre la calidad —ofertar solamente títulos y cursos de alto nivel científico, restringido a pocos alumnos— que utilizar criterios sociales para extender la enseñanza superior al mayor número de personas. Igualmente también condiciona la elaboración de los planes las decisiones que se establezcan sobre si las enseñanzas deben organizarse en función del desarrollo del conocimiento científico o en función de las necesidades del entorno social donde se halla enclavada la Universidad. En cualquiera de los casos, los planes pueden ser evaluados en función de los recursos —profesorado, instalaciones, equipos, etc.— de que se dispone para llevarlos a cabo. Antes de implantar unas enseñanzas se debería estimar si la institución está acreditada para ello. Esto significa que, independientemente del modelo que se adopte, aunque toda oferta académica puede ser evaluada por su interés científico y social, desde el punto de vista práctico, la evaluación viene determinada por el reconocimiento social que tiene el título que ofrece, el prestigio de su profesorado y el nivel del equipamiento e instalaciones relacionados con las enseñanzas concretas, características que podemos observar a través de la demanda del alumnado. La demanda constituye el indicador más representativo para evaluar «a priori» la calidad de las enseñanzas relativas a un título o curso. Es cierto que las restricciones impuestas legalmente a la movilidad de los alumnos en muchos países y otras dificultades de orden económico, no permiten a los alumnos seguir las enseñanzas donde quieren, sino donde les dejan, o de donde pueden; pero aún contando con estas limitaciones podemos utilizar criterios relativos a la matrícula y a los procesos de admisión a efectos de establecer comparaciones entre centros que imparten el mismo tipo de enseñanzas. Ambos factores ofrecen indicadores que deben ser tenidos en cuenta para evaluar las características diferenciales de la población que accede a unos estudios concretos.

CUADRO N.º 4: Objetivos e indicadores relativos al input en la evaluación de la docencia

OBJETIVOS	ASPECTOS O DIMENSIONES	INDICADORES
Oferta Académica	Reconocimiento social del Título Prestigio del Profesorado Buenas instalaciones	<ul style="list-style-type: none"> - Número de programas y materias acreditadas por expertos y colegas de otras Universidades - Tasa de empleo entre titulados - Opiniones de antiguos alumnos, profesionales y empresarios sobre las enseñanzas impartidas - Cualificación del profesorado (premios y publicaciones) - Tiempo dedicado a la docencia en relación a investigación - Calidad del espacio dedicado a la docencia - Dotación bibliográfica en relación con las enseñanzas
Demanda del alumnado	Matrícula Criterios de Selección	<ul style="list-style-type: none"> - Porcentaje de alumnos admitidos en relación con las solicitudes - Tendencia de las matrículas durante los últimos diez años - Nivel de las calificaciones exigidas para acceso a los estudios - Adecuación de los conocimientos previos del alumnado al nivel de las enseñanzas

3.3. Evaluación de los procesos

Una vez que se han decidido las enseñanzas a impartir, las Universidades establecen una serie de procesos y actividades a fin de que éstas puedan llevarse a efecto. De manera sintética se pueden resumir las tareas a realizar durante esta fase en las siguientes: organización de la docencia, elaboración de los programas, realización de la actividad docente en las clases y otras actividades académicas fuera del aula.

3.3.1. La Organización de la docencia

Una de las primeras decisiones que debe asumir la Universidad una vez que ha implantado unas enseñanzas es determinar quién va a im-

partir una materia concreta, cuándo y dónde la va a impartir —horario y aula— y cómo va a desarrollar dicha actividad (clases teóricas, prácticas, seminarios, etc.). Asimismo, deberá decidir las normas a través de las cuales distribuye y coordina los alumnos por grupos y asigna la carga docente a los profesores respectivos. Como es lógico, las normas establecidas quedan reflejadas en los planes docentes elaborados por las subunidades correspondientes. La preocupación por la calidad de la docencia se manifiesta, igualmente, en todas las decisiones que en cada Universidad se toman en relación a la formación y selección del personal, el perfeccionamiento del profesorado y el funcionamiento de los órganos de gestión en los centros y departamentos, etc... dada la inmediata repercusión que tienen estos factores sobre la enseñanza.

3.3.2. *La Elaboración de programas*

Todas las enseñanzas relativas a un título o curso se concretan en una serie de programas de estudio en los que se detallan los objetivos y contenidos de una materia o disciplina, así como las estrategias metodológicas y criterios de evaluación que se estima apropiado utilizar. La elaboración de los programas no es una tarea fácil ya que necesita conjugar una selección de objetivos y contenidos coherentes con la filosofía del título o curso, una delimitación y secuenciación científica de los currícula, y una organización didáctica de los procesos de enseñanza-aprendizaje. La elaboración de los programas es algo más que la simple enumeración de temas. No sólo requiere concretar qué se va a enseñar —contenidos— y porqué —objetivos—, sino también los recursos y medios que se van a utilizar en dichas enseñanzas y cómo se va a proceder a su evaluación. La elaboración del programa y materiales de apoyo constituye una de las actividades que mayor repercusión tienen sobre la docencia por lo que debe ser objeto prioritario de evaluación.

3.3.3. *La Enseñanza en el aula*

El núcleo fundamental de toda la acción docente recae sobre la actividad que el profesor ejerce en el aula. Esta es la razón por la que frecuentemente se identifica evaluación de la docencia con evaluación del profesor. Sin embargo, como ya hemos dicho, ni la evaluación de la docencia es igual a evaluación del profesor, ni tampoco ésta se puede

limitar a su actividad en la clase. El profesor realiza otras muchas actividades antes de la clase y después, que son tan importantes como las que efectúa al enseñar e igualmente además de enseñar (instructor) tiene que asumir otros roles relativos a toma de decisiones organizativas (gestor) y en relación con la ayuda y tutela de los alumnos (orientador). Desde el punto de vista de la enseñanza que desarrolla en el aula, los aspectos más destacados de su función que deben ser objeto de evaluación son los siguientes: el nivel en el desarrollo de las clases, el tipo de métodos de enseñanza que utiliza, el sistema de evaluación que emplea y el clima que logra establecer a través de sus relaciones con los alumnos.

3.3.4. *Las Actividades académicas fuera del aula*

La actividad docente fuera del aula no sólo es posible sino que puede ser tan enriquecedora como la enseñanza en el aula. El hecho de que normalmente los centros no se preocupen por estimular estas actividades académicas (conferencias, jornadas, seminarios, profesores invitados, etc...) no significa que tanto éstas como otras que se pueden realizar fuera del centro (estancias, intercambios, prácticas de campo y en empresas...) no pueden considerarse como acciones docentes importantes para la formación del alumno. Igualmente, desde la óptica de los alumnos, todo tipo de acción orientada hacia la ayuda individualizada en la enseñanza —dirección de trabajos, clases prácticas de ejercicios, etc.— como hacia la orientación personal —tutorías, información académica y profesional, servicio de ayuda y consultas, etc.— pueden entenderse como actividades docentes y como tal deben ser evaluadas. La no inclusión de estos aspectos académicos refleja una concepción muy sesgada de los procesos docentes en la enseñanza superior.

CUADRO N.º 5: Objetivos e indicadores en relación a los procesos de la Docencia

OBJETIVOS	DIMENSIONES	INDICADORES
Organización de la docencia	Distribución Asignación del Profesorado Dedicación Gestión	<ul style="list-style-type: none"> - Grupos con número reducido de alumnos - Distribución de la carga lectiva según modalidades de enseñanza (clases teóricas, prácticas, laboratorios, etc.). - Sistema establecido para asignar la docencia según materias, ciclos y especialidades - Sistema establecido para efectuar la sustitución y suplencias de los profesores - Número total de horas impartidas y/o ausencias por profesor - Tiempo dedicado a la coordinación y revisión de las enseñanzas - Criterios en la selección del personal - Funcionamiento de las comisiones de docencia - Reglamentos y normas de funcionamiento de los centros
Elaboración del Programa	Objetivos Contenidos Recursos	<ul style="list-style-type: none"> - Definición de objetivos ajustados a título o curso - Ajuste de objetivos a posibilidades de los alumnos - Relaciones entre objetivos de docencia e investigación - Contenidos actualizados y adaptados a necesidades - Interdisciplinariedad de la enseñanza - Grado de opcionalidad en la elección del currículum - Tiempo dedicado por el profesor a preparar las clases - Elaboración de materiales para la docencia - Procedimientos para revisar programas y nuevas enseñanzas
Evaluación de la docencia en las aulas	Desarrollo de las clases Métodos de Enseñanza Sistemas de evaluación	<ul style="list-style-type: none"> - Sistemas de control de la actividad del profesor en el aula - Valoración de la competencia del profesor por los alumnos - Métodos de enseñanza utilizados en la clase - Usos de medios técnicos en el desarrollo de las clases - Criterios de evaluación establecidos - Análisis de las pruebas utilizadas en los exámenes

OBJETIVOS	DIMENSIONES	INDICADORES
	Clima del aula	<ul style="list-style-type: none"> - Comunicación y relaciones con los alumnos - Actividades organizadas por los alumnos del aula
Actividades académicas fuera del aula	Actividades docentes en el centro	<ul style="list-style-type: none"> - Actividades académicas paralelas (conferencias, seminarios, profesores invitados, etc.) - Sistemas alternativos de Enseñanza (estudiantes, trabajadores a tiempo parcial, etc.)
	Actividades fuera del centro	<ul style="list-style-type: none"> - Relaciones con el mundo profesional (prácticas en Empresas) - Programas para estudiantes fuera del campus (estancias en extranjero, intercambios, becas)
	Tutoría de Alumnos	<ul style="list-style-type: none"> - Tiempo dedicado a dirección de trabajos, consultas, repasos, etc. - Atención individualizada a los alumnos
	Servicios de Información y Ayuda	<ul style="list-style-type: none"> - Servicios y Programas de Información y Ayuda al estudiante - Actividades de acogida y orientación con los alumnos de nuevo ingreso.

3.4. *Evaluación de resultados*

Finalmente la docencia universitaria debe ser evaluada a partir de sus resultados. Son los productos los que en definitiva dictaminan la política de docencia que lleva a cabo una Universidad. La dificultad radica en determinar cuáles son estos productos dada la diversidad de objetivos y filosofías educativas que subyacen tras los planes de estudio. En una primera aproximación se puede establecer una clasificación entre resultados observados a través de los alumnos y resultados a nivel institucional.

La evaluación de resultados en función de los alumnos habitualmente toma como criterio el rendimiento académico. La mayoría de los estudios al respecto utilizan las calificaciones como el indicador explícito del progreso o provecho académico que el alumnado adquiere a través del proceso de docencia tanto a lo largo de los estudios como al final de los mismos. Otros trabajos utilizan un criterio de rendimiento a largo

plazo con el fin de estimar la influencia que los resultados en determinadas enseñanzas tienen sobre la vida laboral y social del individuo.

Este tipo de evaluación centrada sobre rendimiento académico se justifica desde una concepción de la enseñanza basada en un modelo de producción, pero lógicamente ofrece un planteamiento restrictivo de los efectos de la docencia cuando ésta se analiza desde los modelos orientados hacia el desarrollo. Desde esta perspectiva y la evaluación debe orientarse a estimar las competencias y actitudes que el sujeto adquiere a través de un proceso de enseñanza-aprendizaje orientado a desarrollar al máximo sus propias posibilidades antes que centrarse sobre medidas de conocimientos nocionales y puntuales. Trabajos recientes (Astin, 1985; Hagerty y Stark, 1989) ponen énfasis en que es necesario considerar otras dimensiones del desarrollo del alumno —habilidades, conductas, maduración— así como la adquisición de competencias y actitudes específicas en relación con el mundo profesional.

Por otro lado la evaluación de la docencia también debe efectuarse a nivel institucional, ya que no sólo deben operarse los cambios en los individuos sino también en las organizaciones. La Universidad debe someter a control su propia organización con el fin de observar la eficiencia de sus procedimientos y la eficacia en los resultados. La evaluación del rendimiento en estos dos aspectos —objetivos/resultados y resultados/medios según se establezca la comparación en función de los inputs— constituye el punto de partida para detectar los aspectos institucionales que deben ser mejorados, para decidir las innovaciones que es necesario introducir a efectos de potenciar un funcionamiento más adecuado de la organización, y para estimar las medidas de política educativa y profesorado aconsejables dentro de una «cultura de cambio». Si la Universidad no evalúa su propia organización da más posibilidades a que las fuerzas sociales y políticas ejerzan ese control.

4. Propuesta de un modelo

Una vez que hemos efectuado un análisis sobre los componentes de la unidad de docencia y enumerado los principales indicadores que se pueden utilizar para evaluar los distintos aspectos o factores implicados, vamos a presentar una propuesta concreta. Tratamos de ofrecer un mo-

CUADRO N.º 6: Objetivos e indicadores sobre resultados de la docencia institucional

OBJETIVOS	DIMENSIONES	INDICADORES
Rendimiento de los alumnos	Progreso durante las enseñanzas	<ul style="list-style-type: none"> - Resultados de las Evaluaciones realizadas a lo largo de las enseñanzas - Grado de satisfacción del alumno en las enseñanzas que recibe - Relación entre la carga docente elegida y el progreso académico
	Resultados finales	<ul style="list-style-type: none"> - Rendimiento de los alumnos en prueba final de carrera - Media de tiempo entre graduación y primer trabajo - Porcentaje de alumnos que continúan estudios (tercer ciclo)
	Resultados a largo plazo	<ul style="list-style-type: none"> - Tasa a admisión en pruebas de carácter profesionales - Categoría profesional de las actividades desempeñadas por titulados. - Nivel económico (salario) dentro de las actividades elegidas
Resultados a nivel Institucional	Eficacia Objetivos/ Resultados	<ul style="list-style-type: none"> - Promedio de cursos empleados en relación con los establecidos en los planes de estudio - Tasa de egresados y/o abandonos en relación con la matrícula - Relación entre objetivos previstos y alcanzados
	Eficiencia Resultados/ Medios	<ul style="list-style-type: none"> - Número de cursos y/o programas realizados - Nivel de utilización de los medios destinados a docencia - Relación entre costos y resultados. Tasa por alumno
	Cambio Institucional	<ul style="list-style-type: none"> - Innovaciones introducidas en relación con la docencia - Sistemas establecidos para controlar la calidad de las enseñanzas y su adecuación a las necesidades sociales - Programas presentados al desarrollo profesional de los docentes (año sabático, licencias, etc.)

delo que permita al mismo tiempo estructurar la evaluación de la docencia universitaria de forma global y a su vez posibilitar planteamientos de evaluación parcial en función de las distintas unidades de análisis. Nuestra propuesta se limita a efectuar:

a) Una delimitación de los factores implicados en la docencia según las competencias de las distintas unidades de gestión.

b) Una especificación de algunos indicadores de calidad que se pueden utilizar como base para evaluar la docencia en la Enseñanza Superior.

c) Unas sugerencias sobre el método o procedimiento apropiado para iniciar y llevar a cabo dicha evaluación.

4.1. *Competencias según las unidades de gestión*

Toda evaluación institucional de la docencia debe partir de una delimitación de los factores implicados en las enseñanzas relativa a un título o curso. La actividad académica relacionada con una enseñanza específica constituye la unidad de análisis de todo programa de evaluación, programa que, a su vez, se puede estructurar en función de diversas unidades de gestión (Universidad, centros, departamentos, profesorado). El marco de la evaluación institucional viene determinado por el cruce entre los factores intervinientes en los procesos de docencia y las competencias que en relación con los mismos asumen las distintas unidades de gestión. Por ello constituye una tarea prioritaria a todo programa de evaluación determinar qué unidad tiene la responsabilidad directa sobre cada uno de los componentes de la docencia a efectos de delimitar en cada caso el contenido a evaluar. En la propuesta que efectuamos en el cuadro adjunto señalamos con una cruz (X) la unidad de gestión que —según nuestro criterio— debe asumir prioritariamente la responsabilidad de la gestión y evaluación de cada dimensión, y con dos cruces (XX) cuando su intervención es considerada de forma secundaria. Partiendo de este esquema se pueden realizar programas concretos de evaluación para cada una de estas unidades u órganos de gestión.

CUADRO N.º 7: Competencias sobre la Docencia según las unidades de Gestión

Factores	Dimensiones	Universidad	Centros	Departamento	Profesor
Infraestructura de los Centros	Recursos	X	XX		
	Profesorado	X	XX		
	Alumnos	X	XX		
Input: Planes de Estudios	Oferta	X	XX		
	Demanda	XX	X		
Desarrollo de los Procesos Docentes de Enseñanza/ Aprendizaje	Organización de la Docencia		XX	X	
	Elaboración de Programas			X	XX
	Actividad en el Aula			XX	X
	Actividades fuera del aula		X	XX	
Resultados de la Enseñanza	A nivel de Alumno			XX	X
	A nivel institucional	X	XX		

X. Unidad responsable en primer grado.

XX. Unidad responsable de segundo grado.

4.2. *Propuesta de indicadores de calidad*

Una vez establecido el marco que permite una aproximación global o parcial a la evaluación de la docencia, vamos a efectuar una propuesta sobre los indicadores de calidad que pueden utilizar en el proceso de Evaluación. Partimos de que cada institución debe estimar de forma regular la calidad de todos los aspectos implicados en su organización y funcionamiento a través de un sistema de evaluación interna y externa.

Consecuentemente nuestra propuesta tiene un carácter institucional presentando indicadores relativos a todos los factores y dimensiones que intervienen en los procesos de docencia, distinguiendo entre aquellos que pueden ser estimados de forma objetiva —cuantitativos— y aquellos otros de naturaleza cualitativa que normalmente requieren la intervención de expertos. Tanto unos como otros pueden ser valorado

en función de normas estándar y/o intersubjetiva —indicadores de rendimiento— o simplemente limitarse a constatar una información —datos/juicios/opiniones— que difícilmente puede ser utilizada a efectos de comparación.

La selección que hemos efectuado al respecto en modo alguno pretende ser algo definitivo, constituye un esquema abierto a la discusión que puede ser útil para iniciar programas de evaluación con carácter institucional. Nuestra propuesta se limita a sugerir aquellos que nos parecen más importantes desde una perspectiva general, pero lógicamente deberán ser matizados y desarrollados a medida que se vayan realizando evaluaciones concretas a partir de las diversas unidades de análisis.

CUADRO N.º 8: Indicadores de calidad según objetivos a evaluar

Factores relat.	Cuantitativos	Cualitativos
Contexto/ Infraestructura	<ul style="list-style-type: none"> - Tasa de espacio útil por alumno - Tasa de equipamiento especializado (ej. ordenadores, libros, etc.) por alumno - Ratio de alumnos por profesor - Presupuesto por alumno 	<ul style="list-style-type: none"> - Buen mantenimiento de edificios e instalaciones (campus agradable) - Equipamiento avanzado en laboratorios, bibliotecas y recursos - Programa de actividades formativas para el alumnado - Profesorado que dedica más tiempo a la docencia que a la investigación
Input: Planes de Estudio	<ul style="list-style-type: none"> - Porcentaje de estudiantes admitidos entre los que solicitan en relación con otras Universidades - Porcentaje de alumnos procedentes de distintas regiones del país - Nivel de Rendimiento previo (selectividad) de los sujetos que acceden a unas enseñanzas específicas - Porcentaje de doctores que imparten enseñanza en 1.º y/ o 2.º ciclo 	<ul style="list-style-type: none"> - Enseñanzas acreditadas según expertos y colegas de otras Universidades - Opiniones sobre calidad de las enseñanzas de antiguos alum., profesión. y empresarios - Prestigio del profesorado (premios y reconocimientos externos) - Vinculación de las enseñanzas con la sociedad y con la empresa
	<ul style="list-style-type: none"> - Ratio entre las horas de docencia dedicadas a clases teóricas y prácticas - N.º de horas dedicadas a la docencia en el curso en relación con las previstas 	<ul style="list-style-type: none"> - Estrategias de revisión de los planes docentes de los Centros y Departamentos - Organización y Funcionamiento de las Comisiones de docencia

Factores relat.	Cuantitativos	Cualitativos
Procesos de Enseñanza	<ul style="list-style-type: none"> - Tiempo dedicado por el profesor a la docencia (actividades en el aula y fuera del aula) - Tiempo medio dedicado por el alumno al estudio en el centro y fuera del centro - Ratio de asistencia a las clases - Consulta de fondos bibliográficos y utilización de equipos - Tiempo dedicado a la coordinación y revisión de Enseñanza - Número de actividades paralelas a la docencia organizadas en el centro - Número de alumnos que utilizan los Servicios de apoyo a la docencia 	<ul style="list-style-type: none"> - Revisión de los programas y materiales elaborados por el profesor - Métodos de Enseñanza utilizados por el profesor en clase - Uso generalizado de medios en la docencia laborat., ordenad., medios audiov. etc.) - Revisión de los criterios y modelos de exámenes - Actividades organizadas por los alumnos dentro y fuera del campus - Programas y Servicios de Información y Ayuda al estudiante - Sistemas establecidos para efectuar el seguimiento de la docencia y la revisión de las enseñanzas
Producto/ Resultados	<ul style="list-style-type: none"> - Calificaciones obtenidas por el alumnado a lo largo de los cursos - Número de cursos empleados en relación con los establecidos en los planes de estudio - Tasa de egresados en relación con los matriculados - Rendimiento en pruebas de grado y exámenes profesionales (ej. MIR) - Tiempo entre que finalizan estudios y desempeño de actividad profesional - Tasa de costo por graduado 	<ul style="list-style-type: none"> - Relación entre objetivos previstos y alcanzados. - Prestigio de los titulados en la Institución - Programas de innovación en la docencia - Política en materia de perfeccionamiento del personal y el desarrollo de los recursos humanos - Mejoras institucionales en función de los rendimientos

4.3. *La Autoevaluación como procedimiento*

Una vez que hemos determinado el marco de la evaluación institucional y hemos señalado alguno de los indicadores que se pueden utilizar con este fin, nos resta efectuar algunas indicaciones sobre los procedimientos. A este respecto nuestra postura es clara: consideramos que —independientemente de los procesos accountability (evaluación externa)— la finalidad de toda evaluación es la mejora y el cambio institucional. Ahora bien esta mejora no será posible mientras las unidades de gestión a las que hemos aludido no asuman su cuota de responsabilidad y comienzan su propia autoevaluación. Por ello cualquier sistema

que se utilice —indicadores o expertos— o es asumido por la unidad correspondiente o constiye simplemente una ficción de la evaluación. De ahí que hoy todas las recomendaciones sobre evaluación institucional incidan sobre este mismo hecho (Acherman, 1990; Kells, 1990; Cheps, 1990).

De acuerdo con ello, el modelo que proponemos se puede implantar siguiendo un proceso en tres fases:

A. Fase de Advertencia

Se considera una fase previa a la implantación de un proceso de evaluación institucional durante la cual lo importante es crear un clima favorable hacia la misma, sensibilizar a cada una de las unidades de gestión de sus competencias y responsabilidades, y consensuar la finalidad y los procedimientos de los programas que se establezcan. Se trata de que todos los miembros y órganos de la comunidad universitaria comprendan la necesidad de someter su actividad a control, conozcan lo que se espera de ellos y, puedan regularizar su actividad en función de estas expectativas. Entendemos que previamente a iniciar un proceso de evaluación institucional las unidades correspondientes deben conocer las competencias respecto a las cuales va a ser evaluada su actividad y deben establecer los objetivos —standars— que, en cada caso, van a ser utilizados como criterios de evaluación. De no efectuarse esta fase, el programa de evaluación difícilmente puede tener éxito.

B. Fase de Autoevaluación

Una vez decidido el tipo de programa a realizar, la unidad de gestión correspondiente comienza a evaluar el ejercicio de sus competencias en base a un sistema de indicadores de calidad. Inicialmente se puede partir de aquellos que tienen carácter cuantitativo y que pueden ser estimados y analizados directamente por la unidad correspondiente dejando para un segundo momento los de tipo cualitativo que reclaman la participación de expertos de otras Universidades. En este último caso, es oportuno que la unidad correspondiente efectúe también su autoevaluación siguiendo una guía que previamente establecen los expertos. Posteriormente los datos obtenidos en los análisis efectuados por la unidad correspondiente se contrastan con las observaciones que reali-

zan los expertos sobre el terreno. De esta forma los miembros de la organización o unidad —profesores, alumnos, administradores, empleados— pueden analizar e interpretar la información recogida y llegar a «consensuar» un conjunto de observaciones realistas sobre los factores implicados en la docencia.

C. Propuesta de Innovación y Seguimiento

Finalmente a partir de los datos recogidos a través de la fase anterior se debe efectuar —con la ayuda de expertos— unas propuestas relativas a los cambios e innovaciones necesarias en relación con la docencia, así como la delimitación de la unidad u órgano a quien compete su gestión. La experiencia nos dice que no sólo basta señalar los fallos sino que también es necesario estimular y facilitar modelos para introducir las mejoras. Algunos autores entienden que la causa esencial que inhibe todo proceso de innovación es la falta de una adecuada monitorización que estimule, oriente y tutele los procesos de cambio, para lo cual es necesario crear Servicios de Apoyo y Seguimiento de la Docencia dentro de la propia Universidad que se ocupen de estas tareas. Constituye ésta una recomendación que cada día se considera más necesaria.

A título de resumen quisiera resaltar que a lo largo de este trabajo hemos sostenido que la evaluación de la docencia en la Universidad debe enfocarse a nivel institucional tomando como referencia las enseñanzas relativas a un título o curso. Igualmente este proceso de evaluación debe partir de las diversas unidades que intervienen en la gestión de la docencia a efectos de que cada una sea evaluada desde su cuota de participación en el proceso institucional. Sólo si tomamos como referencia en nuestro análisis un título y una unidad de gestión cabe la posibilidad de establecer evaluaciones pertinentes que a la vez permitan la comparación entre distintas Universidades que imparten el mismo tipo de enseñanzas.

En relación con los procedimientos la alternativa actual es inequívoca: o se plantean los procesos desde la autoevaluación o realmente condenamos al fracaso cuantas estrategias se elaboren al respecto. Todas las discusiones en relación con los instrumentos y medios a utilizar en la evaluación tienden hacia las mismas conclusiones: aunque los sistemas de evaluación basado en indicadores objetivos pueden mejorar el diálogo entre el gobierno y las universidades, y el sistema de

expertos se considera apropiado para efectuar una valoración contextualizada de los centros, ambos procedimientos resultan inoperantes para transformar la organización institucional. Cada vez es más necesaria la autoevaluación como estrategia de cambio. Parece que no hay elección, o asumimos la autoevaluación institucional o los políticos efectuarán los cambios por nosotros.

UN EXAMEN DE LAS POLITICAS EUROPEAS EN MATERIA DE EVALUACION DE LA ENSEÑANZA SUPERIOR *

SERGE CUENIN

Centre Nationale de la Recherche Scientifique, Francia

Introducción

En Europa, los centros de enseñanza superior son financiados principalmente, a veces exclusivamente, con fondos públicos. El papel de los Estados —centrales y locales en los países con estructura federal— es, pues, muy importante.

Desde los años cincuenta, la enseñanza superior europea ha experimentado un gran desarrollo en materia de efectivos escolarizados. Este fenómeno ha venido acompañado de una intervención creciente de los Estados para adaptarse a esta nueva situación, creando nuevos centros con fines más tecnológicos (Politécnicos en el Reino Unido e Institutos Universitarios de Tecnología en Francia, por ejemplo) o incluso adaptando la legislación que rige el sistema universitario.

El intervencionismo del Estado se acentuó cuando los recursos disponibles para la enseñanza superior se vieron reducidos por dificultades económicas (segunda crisis del petróleo), la concurrencia de necesidades en otras funciones públicas (salud, por ejemplo) y del deseo de establecer controles obligatorios. Se trataba de no gastar más, sino mejor.

En el espíritu de los gobiernos, la enseñanza superior debía contribuir a la mejora de la situación, en primer lugar en el plano económico, mediante la investigación y la formación de una mano de obra cualifica-

* Este artículo es traducción de la ponencia presentada por el autor en la *Reunión del Proyecto Columbus*, Valparaíso, Chile, agosto 1990.

da en especialidades con futuro. El Estado, garante de los dineros públicos, debe asegurar que éstos son utilizados de acuerdo con sus objetivos.

Al comienzo de los 80, como consecuencia de las restricciones presupuestarias que tienen lugar en numerosos países, el Estado tiene que decidir dónde deben efectuarse los recortes presupuestarios, y otras veces dejar esta tarea en manos de centros u organismos reguladores encargados de repartir los fondos.

Más recientemente, en algunos países, esta estrategia ha dado lugar progresivamente a una actitud más flexible en la que el Estado fija las grandes orientaciones y deja más autonomía a los centros para administrar sus recursos adaptándose así a la evolución de la sociedad, en particular, a sus componentes económicos y tecnológicos.

Por supuesto, los Estados se reservan la posibilidad de evaluar los resultados alcanzados por los centros, los indicadores de rendimiento son un instrumento privilegiado para esta evaluación.

Por otro lado, los propios centros utilizan, cada vez más, instrumentos que les permitan efectuar una autoevaluación con el fin de emplear mejor los recursos disponibles y elaborar planes de desarrollo.

La mayor parte de los datos citados en las dos primeras partes de este informe están sacados de una encuesta realizada en 1985-86 en el marco del Programa de Gestión de Centros de Enseñanza Superior (Programa IMHE) de la OCDE en unas 70 universidades pertenecientes a 15 países, de los cuales trece eran europeos. Este informe recoge, además, ciertas consideraciones sobre los indicadores de rendimiento que nosotros habíamos publicado en la Revista del Programa IMHE de la OCDE (Cuenin, 1987).

Consideraremos tres puntos:

- ¿Qué indicadores son utilizados por las universidades?
- ¿Con qué fines son utilizados?
- Un análisis de algunas experiencias europeas en materia de evaluación, a partir de indicadores.

I. Los indicadores

Un indicador de rendimiento puede ser definido como un dato numérico que permite evaluar el rendimiento cualitativo y cuantitativo de un sistema.

El término «rendimiento» es con frecuencia interpretado por extensión: muchos indicadores de rendimiento no miden, en realidad, rendimiento. Por esta razón, preferimos emplear sistemáticamente el término «indicador» más que indicador de rendimiento, reservando esta calificación a ciertos indicadores bien definidos.

Aunque los indicadores se pueden referir a las diversas funciones de la universidad, sólo la enseñanza y la investigación han sido tomadas en consideración en la encuesta citada anteriormente. Los servicios públicos, los servicios a los estudiantes, etc., han sido omitidos.

Después de establecer una clasificación de los indicadores mencionados por las universidades examinaremos los más utilizados en relación con la enseñanza.

A. Clasificación de los indicadores

A fin de caracterizar los indicadores, los hemos agrupado según cinco criterios que, al combinarse, permiten identificarlos mejor. Los cinco criterios de agrupación son los siguientes:

1. La *naturaleza* de los indicadores

Desde este punto de vista se pueden distinguir tres tipos de indicadores:

a) Indicadores *simples* que se presentan frecuentemente de forma absoluta y que tienen por finalidad descubrir de una manera bastante neutra un hecho, un proceso, etc.

b) Indicadores de *rendimiento* que, respecto a los anteriores, implican una referencia: norma, objetivo, juicio, centro comparable, etc. Se trata, pues, de razonar de una forma relativa y no absoluta. Por ejemplo,

el número de diplomas será un indicador simple, pero la relación número de egresados/número de alumnos que comenzaron la carrera es un indicador de rendimiento.

El indicador simple es más neutro. Se podría decir que puede llegar a ser un indicador de rendimiento después de un juicio de valor. En ocasiones hay, pues, ambigüedad. Por ej., la relación número de profesores titulares/ no titulares puede ser un indicador simple relativo a los enseñantes, pero puede ser también un indicador de rendimiento si se introduce un juicio sobre la conveniencia de tener muchos profesores titulares o, por el contrario, tener pocos.

En general, a fin de evitar al máximo la ambigüedad, se establece como norma que: si cuando la situación es mejor el indicador varía en un sentido y cuando es peor varía en el sentido contrario, entonces el indicador es un indicador de rendimiento. Esta apreciación deberá poder hacerse de forma evidente.

c) A los indicadores precedentes se podrían añadir indicadores *muy generales*, casi siempre obtenidos fuera de los centros y que no son indicadores en un sentido estricto: son frecuentemente opiniones, resultados de encuestas o estadísticas generales. Aunque no responden realmente a la definición de indicador dada antes, son, sin embargo, tenidos en cuenta en la toma de decisiones.

Muy esquemáticamente, se puede decir que estos tres tipos de indicadores se repartían en nuestro estudio del siguiente modo:

- Indicadores simples 23 % de las citas.
- Indicadores de rendimiento 67 % de las citas.
- Indicadores generales 10 % de las citas.

2. El ámbito de aplicación de los indicadores

En este trabajo sólo hemos utilizado indicadores relativos a la enseñanza, la investigación, y sin distinción. Esta limitación es consecuencia de la opción efectuada al diseñar la encuesta.

3. El *lugar de obtención* de los indicadores

Los datos se obtienen en «el interior» o en el «exterior» de los centros.

El término «exterior» debe interpretarse, en nuestra opinión, de manera amplia: se trata de indicadores obtenidos en base a informaciones, o por la intervención de personas u organismos exteriores al centro (ej.: estadísticas del mercado de trabajo, número de empresarios haciendo visitas de contratación en los centros, etc.). Esta concepción es, por tanto, más amplia que la utilizada en el cuestionario destinado a los centros. Hemos querido, introduciendo esta distinción, ser más selectivos en la clasificación de los indicadores citados. En efecto, las subvenciones, las invitaciones para dar conferencias, los desplazamientos, el número de profesores elegidos como expertos, los créditos recibidos del exterior, etc., se dan en función del prestigio del personal del centro.

4. El *momento* del proceso productivo a que se refieren

- Cuando el indicador mide los gastos relativos al personal y a los medios necesarios para la realización de los objetivos, se les conoce como indicadores *de entrada*. Entre éstos también se pueden incluir algunos indicadores relativos a los estudiantes.
- Si permiten calificar las condiciones en las que se desarrolla la actividad de los Centros se les denomina indicadores de *actividad o de proceso*.
- Por último, los indicadores de *resultado* se utilizan para evaluar el producto alcanzado a través de una o varias actividades.

5. Los *componentes* del centro

Con frecuencia, a estos indicadores se les relacionó con los *agentes* que intervienen en los centros —estudiantes y profesores— y con los *medios* empleados.

Estos criterios pueden ser interrelacionados a fin de poner de manifiesto las frecuencias significativas de los datos aportados por la en-

cuesta. De la combinación de tres criterios puede obtenerse mucha información tal como se desprende del Anexo I.

A la vista de este Anexo I se pueden hacer las observaciones siguientes:

1. Los indicadores citados se relacionan mucho más con la enseñanza que con la investigación: 72 % frente a 24 %, el 4 % restante pueden concernir a la enseñanza o a la investigación.

2. Los indicadores de rendimiento en sentido estricto, tal como han sido definidos anteriormente, suponen el 91 % de las cuestiones relativas a investigación, el 62 % de las de enseñanza y el 20 % de las indeterminadas. Parece, pues, que éstos se han utilizado más para evaluar la investigación que la enseñanza.

3. En cuanto al criterio «interior/exterior», se constatan dos cosas: en principio, 7 de cada 10 indicadores mencionados hacen referencia a elementos que conciernen solamente a la Universidad. Además, es en los indicadores simples en los que la preponderancia del criterio «interior» es más manifiesta (98 %). Por el contrario, los indicadores generales están más volcados hacia el exterior (94 %) y los indicadores de rendimiento ocupan una posición intermedia (70 %).

4. Analizando detalladamente el reparto según el criterio «interior/exterior», destacamos tres puntos: en primer lugar, los indicadores relativos a investigación se refieren mayoritariamente a elementos que conciernen al exterior de la Universidad (66 %), lo que confirma que en materia de investigación somos más exigentes en cuanto a los indicadores a utilizar. En segundo lugar, en la enseñanza se observa una situación inversa: 17 por ciento solamente de los indicadores citados se refieren al exterior. Por último, los indicadores «sin distinción» entre enseñanza e investigación se reparten por igual entre el «interior» y el «exterior».

B. Breve análisis de los indicadores más utilizados en materia de enseñanza

1. Las tasas de matriculación

Este indicador se expresa en forma de una ratio: número de alumnos/profesor (raramente se incluye otro personal). Los alumnos suelen contabilizarse ponderadamente según su nivel de estudio. A veces, estas ponderaciones se establecen a nivel nacional.

Las tasas de matriculación suelen utilizarse para la asignación de personal y en algunos centros se tienen también en cuenta para la planificación, e incluso para fijar un número clausus. Hay que subrayar que, en muchas universidades, este indicador sólo se utiliza como un indicador general. De ahí que su importancia sea menor. Sólo lo citan dos universidades de cada diez. Concretamente, cuando se compara la situación de unidades administrativas —normalmente los departamentos— con normas nacionales o incluso con otras unidades administrativas de la misma universidad.

Las limitaciones del mismo se sitúan en varios niveles:

- A nivel de la *elaboración* de este indicador. Es difícil contabilizar a los estudiantes: pueden existir grandes diferencias entre el número de estudiantes matriculados y los efectivos reales. Además, este indicador no incluye el trabajo real de todos los profesores, ya que algunos no siempre son tenidos en cuenta: monitores, auxiliares, docentes a tiempo parcial.
- A nivel de su *interpretación*, que es difícil cuando no existe normalización a nivel nacional. Incluso es ambiguo: puede sugerir que faltan enseñantes cuando hay demasiados estudiantes. Por último, es frecuentemente poco preciso y ganaría al ser construido a partir de categorías más finas.
- Su efecto también es considerado reductor por muchas razones:
 - Enfatiza la idea de que las necesidades de personal pueden ser medidas con ayuda de un ratio.
 - Disminuye la flexibilidad en la asignación de los recursos.

- Resalta la cantidad más que la calidad.
- Ignora los esfuerzos de investigación de los departamentos.

Por último, su *utilidad* es limitada cuando los departamentos tienen gran libertad para emplear sus fondos (personal, otros gastos corrientes y de pequeño equipamiento): se acuerdan recursos para contratación de personal y se emplean para otros fines.

2. *Los costes por unidad*

Se han enumerado muchos tipos de costos en los centros (una media de dos por respuesta). Se les puede agrupar así:

- *Costo por estudiante.* Este concepto se cita una vez cada dos. Se calcula por departamento o por disciplina. Existen numerosas expresiones de este costo: costo corriente, costo por equipamiento.
- *Costo relativo a los profesores:* equipamiento, libros, costo total o incluso costo relacionado con el *salario*; material fungible, costo de personal no docente.
- Se han utilizado también *otros tipos de costos:* coste del m², de la hora de enseñanza, de la hora de tutoría, etc.

Este tipo de indicadores se usan sobre todo para la asignación de recursos. Generalmente son criterios a tener en cuenta por los comités encargados de distribuir los recursos entre las unidades administrativas. Se pueden utilizar para reorientar fondos (supresión de cursos demasiado costosos), identificar los recursos de los departamentos a fin de repartir más equitativamente los créditos, o para responder a las peticiones de ayudas suplementarias. El aspecto comparativo se subraya con demasiada frecuencia: el costo unitario no parece útil más que cuando se compara con lo que sucede en situaciones similares. En resumen, los costos unitarios se conciben más como una base de referencia que influye en la toma de decisiones, sin que esta influencia sea definitiva.

En las Universidades del Reino Unido ha sido, sobre todo, donde se encontraron más dificultades en el uso de este tipo de indicadores. De aquí que en los comentarios que siguen exista un marcado carácter

nacional. En este país, las comparaciones de costo/unidad se hacen a nivel nacional, a partir de datos recogidos por disciplinas. La diversidad de observaciones recogidas al respecto son, sin embargo, suficientemente generales que pueden ser aplicadas a otros contextos nacionales:

- En cuanto a la construcción de este indicador, la clasificación de los gastos varía según los centros y es difícil la separación entre conceptos de enseñanza e investigación.
- Los plazos de publicación de los análisis comparativos son demasiado largos y son, además, poco operativos por estar insuficientemente detallados. Es el caso del concepto «costo de programa».
- A nivel de interpretación de estos indicadores, frecuentemente no se tienen en cuenta factores exteriores al establecimiento: edad y grado de los enseñantes, por ejemplo.
- Por último, un centro no británico ha subrayado que este indicador excluía la noción de calidad, lo que podía hacer que su interpretación se deslizara hacia la mediocridad.

3. *La incorporación al mundo del trabajo*

Este es el indicador de rendimiento más citado cuando se hace referencia a los resultados de las actividades de los centros a nivel de estudiantes. Se obtiene, generalmente, por encuesta de los centros a los antiguos alumnos y, en ocasiones, analizando encuestas nacionales para observar tendencias en la evolución de la demanda de ciertas carreras.

Este indicador se utiliza en numerosos casos, entre los que cabe citar los cuatro siguientes:

- Cuando se trata de modificar los sistemas de enseñanza o incluso los programas de estudio: adaptar, reconducir, desarrollar, crear.
- En la distribución de los recursos financieros.

- Para facilitar informaciones al exterior del centro, por ejemplo, publicidad a favor de la universidad.
- Para determinar las cuotas de admisión.

Algunas de las dificultades en relación con la construcción y utilización de este indicador son las siguientes:

- A nivel de la encuesta, es difícil contactar con los antiguos alumnos; generalmente responden pocos y hay un importante riesgo de sesgo, por lo que resulta lenta y costosa para la universidad.
- La mayor parte de las informaciones se refieren al primer empleo. Es insuficiente para juzgar una enseñanza (los titulados suelen desempeñar frecuentemente empleos temporales en tanto encuentran uno más acorde con su formación).
- El indicador es demasiado reductor, puesto que se concreta en un solo aspecto: el paro.
- Hay una importante limitación en adecuación temporal de este indicador, ya que el mercado de trabajo de los diplomados de enseñanza superior es fluctuante.
- Es difícil traducir las demandas de empleo (cuando se les puede conocer!) en programas.
- La información nacional no se difunde con suficiente rapidez; la información universitaria es demasiado fragmentaria.
- Este indicador interesa mucho a las personas encargadas de cursos profesionales. La sensibilización de otros enseñantes es más difícil.

II. **Ambitos de aplicación**

Los indicadores tienen muchas aplicaciones, pero no procede detallarlas todas aquí. Vamos a destacar solamente las principales agrupando los indicadores en dos grandes campos de aplicación: los utilizados para la toma de decisiones dentro del Centro y los utilizados en la comunicación con organismos exteriores.

A. *Utilización de los indicadores en el interior de los centros*

Los indicadores han sido utilizados para la toma de ciertas decisiones; pero no resulta fácil, porque si bien un indicador puede utilizarse en muchas tomas de decisión, también una decisión puede requerir varios tipos de indicadores. Por otra parte, ciertos indicadores no se utilizan jamás para tomar una decisión. Examinaremos, a continuación, tres grandes tipos de decisiones: la asignación de recursos, la planificación y la gestión pedagógica.

1. *La asignación de recursos*

En las instituciones estudiadas se encontraron dos tipos de asignación: los presupuestos asignados a las facultades y/o departamentos y los destinados al personal, incluyendo el docente y el no docente. La tercera parte de las cuestiones se refiere a estos dos tipos de asignación, lo que demuestra la importancia de los indicadores en este ámbito.

a) *Subvenciones destinadas a las facultades y/o departamentos*

En lo relativo a la enseñanza, lo primero a tener en cuenta es la *carga de trabajo*. Puede ser medida simplemente teniendo en cuenta el número de horas impartidas, o de una manera más precisa ponderando los cursos según diversos métodos. Lo segundo a tener en cuenta son los *efectivos escolares*, también con posibilidad de introducir ponderaciones. Por último, *los resultados de los exámenes, las tasas de matrícula, los costos unitarios, la naturaleza de las enseñanzas* (disciplinas más o menos costosas) se han citado también muchas veces. Pocos, entre estos indicadores, son realmente indicadores de rendimiento.

La asignación de recursos para la investigación, por el contrario, hace más referencia a los indicadores de rendimiento. En primer lugar, *las subvenciones externas*, indicador muy general que se interpreta de muchas maneras, como veremos más adelante. *Las publicaciones de los docentes* también se citan con una contabilización más o menos selectiva. Además, se incluyen *la plantilla de profesores, el número de estudiantes aprobados, las tesis* (dirigidas o leídas), *los proyectos presentados*, etc. Sin embargo, es a propósito de las «subvenciones» donde se han planteado más cuestiones.

b) *La asignación de recursos en personal*

Para la asignación de docentes se repiten mucho dos indicadores: la *tasa de matrícula de alumnos* calculada a partir de numerosas variables y la *carga de docencia*, medida también de muchas formas, sea a nivel de facultad o a nivel de departamentos. Las asignaciones de recursos para personal no docente requiere sobre todo *ratios entre el personal* (docente/no docente) a partir de respetar normas del centro o nacionales.

2. *La planificación*

La mayor parte de los indicadores citados en este campo hacen referencia a la planificación relativa a los estudiantes. En primer lugar, se toma en cuenta la admisión en el centro, por lo que se citan indicadores como *el nivel de acceso, demandas de admisión o evolución de las tendencias del mercado*. Por otro lado, juega un papel importante el progreso de los alumnos en el centro utilizando indicadores como: *tasa de éxito y de abandonos*. Incluso el número actual de estudiantes y su distribución en la universidad son objeto de seguimiento. Por último, el *número de egresados* y su *origen geográfico o étnico* son también indicadores citados muchas veces. Toda planificación se traduce en una previsión de efectivos para los próximos años. Esta previsión puede ser estricta —en los países donde las universidades pueden fijar números clausus— o aproximativa, en aquéllos donde la admisión es libre.

En lo referente a la planificación de la financiación se han contabilizado pocas citas: entre los indicadores que se utilizan para hacer previsiones financieras de futuro, con frecuencia se cita el número de *estudiantes extranjeros* que se van a poder admitir, dado que reporta fondos suplementarios en algunos países.

Por último, la planificación en cuanto a personal sólo se cita con una frecuencia menor del 10 %. Casi siempre para prever la sustitución de las jubilaciones.

3. *La gestión pedagógica*

De ella depende, principalmente, la creación, modificación y supresión de cursos. También incluye, en parte, la planificación, porque si

bien las modificaciones se hacen lentamente en función de la evolución de tendencias, las decisiones, muchas veces, se toman a corto y medio plazo. Son varios los indicadores citados a este respecto y casi todos ellos con semejante frecuencia. Destacaremos la *inserción de los licenciados* en el mercado de trabajo, medida de muchas formas; el *prestigio de las materias impartidas*, estimando mediante las peticiones dirigidas al centro; la *evaluación de los profesores* realizada por los estudiantes; la *carga horaria de trabajo de los estudiantes*; la *duración media de los estudios* o incluso *la evaluación de los egresados*.

B. *La comunicación con los organismos exteriores*

Es este campo donde es importante el papel de los indicadores. Aunque la encuesta no se ha centrado en esta cuestión, algunos centros han facilitado informaciones interesantes sobre determinados indicadores que pueden ser muy útiles al respecto.

1) *Indicadores y organismos de financiación*

En la mayor parte de los países estudiados se encontraron cinco tipos de estadísticas solicitadas por los servicios gubernamentales, los órganos de financiación u organismos diversos. Estas conciernen a:

a) *Los estudiantes*, clasificados por disciplinas y por nivel de estudios. Ciertos países han mencionado una clasificación por nacionalidad.

b) El *personal*, las demandas más precisas conciernen a los docentes. A veces el centro debe facilitar la relación personal docente/no docente.

c) Los *aspectos pedagógicos*: número de diplomas expedidos (el acento se suele poner sobre los doctorados), número de alumnos por profesor.

d) Los *locales*: superficie de los locales, superficies no edificadas.

e) Los *aspectos financieros*: fondos recibidos a través de contratos (financiación no recurrente), principales gastos del centro en relación con algunos aspectos tales como el costo de la energía.

Los organismos exteriores no demandan forzosamente indicadores. Parece que, en ciertos casos, ellos mismos construyen sus propios indicadores a partir de los informes recogidos. Aunque cada país tenga un sistema particular, los aspectos señalados reflejan las preocupaciones de las autoridades cuando piden estos informes.

2) *Utilidad de estos indicadores*

Las opiniones de los centros en este punto no coinciden. Sin embargo, hay una tendencia que domina: el escepticismo. No obstante, es posible clasificar estas opiniones en una escala según su utilidad:

a) Son útiles:

La utilidad de estos indicadores se pone de manifiesto a través de opiniones como ésta:

- Un buen criterio: ¡sin ellas no hay financiación!
- Sin comentarios.
- Subrayando la posibilidad de hacer comparaciones entre establecimientos: costo de financiación, recursos, costos por departamento. Se puede incluso proceder a análisis de regresión.
- Una institución ha señalado que los indicadores eran útiles para ella.

b) Quizá sean útiles para el gobierno.

Se sabe lo que el gobierno quiere hacer, pero la Universidad no está en condiciones de dar una opinión pertinente sobre la influencia de la información por ella facilitada.

c) Son poco útiles.

Se ha podido constatar, a veces, que el gobierno establece globalmente la variación de los créditos acordados para las universidades. En otros casos se observa que las subvenciones son reconsideradas partiendo de los presupuestos anteriores con ajustes basados en las variaciones del número de estudiantes.

d) Son insuficientes.

Algunos centros señalan las dificultades de interpretación de los indicadores: no son suficientemente adecuados para los fines perseguidos o incluso es necesario completarlos mediante discusiones, visitas y otras informaciones.

e) No hay diálogo entre la universidad y los organismos externos.

Esta constatación se ha hecho en el Reino Unido y, en particular, en Australia. En el primer país, la provisión de indicadores genera malestar por el tipo de medida o porque no se informa sobre cómo han de utilizarse. Hay una clara demanda en favor de un diálogo universidad/JGC a fin de ponerse de acuerdo sobre la interpretación de los datos brutos. Se han indicado algunas acciones en este sentido a fin de introducir más elementos cualitativos en las decisiones.

En resumen, no hay rechazo a la utilización de indicadores en el diálogo con los organismos exteriores, pero los centros temen, sobre todo, que los indicadores se utilicen mecánicamente.

3) *Indicadores más importantes*

Cuando se les invitó a citar los indicadores que ellos consideraban más importantes, 7 de cada 10 centros han dado —como media— dos respuestas por centro. A partir de estas respuestas se pueden destacar:

En primer lugar, existe un gran consenso en cuanto a los indicadores considerados más importantes: casi el 60 % de los centros han citado cuatro indicadores (ver anexo 2).

En segundo lugar, se pone de manifiesto la existencia de criterios semejantes en ciertos países. En efecto:

- En Australia, las cuatro universidades han citado el número de estudiantes, en dos de ellas utilizando valores ponderados.
- Sólo en el Reino Unido se ha citado el indicador sobre el nivel de acceso de los estudiantes.

- En Francia, la mayor parte de las Universidades han señalado la importancia de las superficies del campus como indicador.
- Los costos unitarios sólo han sido citados por los países anglosajones: Reino Unido e Irlanda.

En resumen, un rápido recuento demuestra que los indicadores citados se refieren sobre todo a *la docencia* (dos de cada tres, si se considera que los costos unitarios se relacionan con esta actividad). La investigación no recoge más que el 15 % de las citas.

III. Dos estudios de casos universitarios

1) *El caso de la Universidad de Strathclyde (Reino Unido)*

Se trata de una universidad tecnológica creada en 1964, compuesta de cuatro facultades y que ha experimentado desde principio de los años 1980 un fuerte aumento de los efectivos, que pasaron de 1.000 a 7.000 en cinco años, mientras que sus recursos de origen público disminuían un 16 % en términos reales. Ante esta situación, la Universidad que quería impartir carreras nuevas no sólo ha tenido que buscar financiación complementaria, sino también elaborar un plan estratégico para distribuir los recursos de una forma selectiva en favor de las especialidades con más porvenir. Esto se hizo en 1985.

Con el fin de efectuar esta reorientación de los recursos en las mejores condiciones, la Comisión encargada de esta tarea ha utilizado una «matriz de orientación de la acción de la Universidad» adoptada de la industria a la Universidad por J. Sizer. Esta matriz reagrupa los puntos fuertes y débiles de la universidad en un determinado campo.

Cada miembro de esta comisión de planificación, formada por miembros de las cuatro facultades y de la dirección de la Universidad, disponía de 11 indicadores de rendimiento (8 internos y 3 externos) para completar su juicio sobre cada departamento a fin de situarlo en una de las nueve casillas de la matriz. Las respuestas se comunicaron al secretariado de la Comisión en sobre anónimo.

En 1985 los indicadores utilizados no estaban ponderados y no hacían referencia más que a un año. Posteriormente fueron calculados sobre

Interés de la disciplina concreta

		BAJO	MEDIO	ALTO
Puntos fuertes de la disciplina	BAJO			
	MEDIO ¹			
	ALTO			

¹ Prestigio de la Universidad en esa disciplina.

tres años introduciendo una ponderación que primaba el último. Además, fueron considerados dos indicadores externos suplementarios relativos a las opiniones de organismos exteriores sobre ciertos aspectos de la actividad de la Universidad.

En el plano técnico, uno de los problemas encontrados fue poder utilizar independientemente los indicadores relativos a los dos ejes (interno y externo). La evaluación según este método ha señalado dos cosas:

- En primer lugar, la existencia de un consenso real en cuanto a medidas a tomar: las elecciones hechas por los miembros de la comisión eran bastante parecidas.
- En segundo lugar, el hecho de pronunciarse en el anonimato y con ayuda de indicadores de rendimiento ha permitido evaluar más objetivamente los departamentos.

El plan propuesto por la Comisión fue adoptado por la Universidad sin grandes modificaciones.

2) *El caso de la Universidad libre de Amsterdam (Países Bajos)*

Se trata en este caso de un ejemplo de evaluación interna. Elaborado en 1985, este modelo supuso una combinación de dos métodos (psicológico y administrativo) para mejorar la calidad de la enseñanza.

El modelo incorpora cierto número de elementos que permiten evaluar la enseñanza: el progreso de los estudiantes por año y disciplina, condiciones de la evaluación de los conocimientos, y modalidades de enseñanza, principalmente. Se parte de un supuesto teórico sobre progreso de los estudiantes en los cursos universitarios según el cual se estima que el 75 % aprobarán su primer examen, si es necesario repetirán una vez, y el 70 % obtendrán el diploma final al término del cuarto nivel de estudios, repitiendo como máximo dos veces.

Si el progreso real de los estudiantes no está de acuerdo con este esquema de referencia, los profesores deberán explicar por qué y también demostrar que la situación es imputable a los estudiantes más que a ellos. En cada facultad se ha encargado una comisión de analizar los resultados del modelo con toda libertad. Esta comisión puede incluso demandar de los estudiantes información mediante cuestionarios normalizados relativos a la enseñanza que han recibido. Estas encuestas se vienen efectuando desde hace varios años en ciertas materias. Las respuestas permiten clasificar a cada profesor en relación con sus colegas.

Según sus autores, este modelo combina un método en principio psicológico, pues los resultados son comunicados a los interesados — profesores en primer lugar— a fin de que ellos puedan interpretarlos y proceder a una autoevaluación que les lleve a mejorar su enseñanza y a discutir todo ello con las instancias universitarias correspondientes.

También es un método administrativo, puesto que es la administración la que define reglamentariamente las condiciones de promoción o sanción, y porque todas estas medidas constituyen un medio externo de estimular al profesor a mejorar su enseñanza.

Si bien el primer método se ha venido considerando preferible, la observación de la situación en la Universidad ha conducido a utilizar el segundo como complementario.

En resumen, nos parece que este tipo de evaluación no puede aplicarse siempre, pues con frecuencia los profesores no aceptarán tener que

justificarse frente a un modelo mecánico, por una parte, y a las evaluaciones hechas por los estudiantes, por otra. El sistema parece funcionar en esta universidad porque los profesores aceptan los resultados del análisis «no sólo porque los datos son objetivos y pueden ser interpretados inequívocamente, sino porque son todos sometidos, sin excepción, a la misma evaluación».

Este tipo de evaluación constituye una excepción en Europa.

IV. **Dos experiencias europeas**

Nos proponemos examinar aquí dos contextos diferentes en los que se intentan utilizar indicadores de rendimiento para financiar las universidades. Se trata en primer lugar del Reino Unido y, en segundo lugar, Francia.

A. *El caso británico*

El examen de la situación en el Reino Unido es muy interesante porque en el curso de los años 80 se han producido cambios importantes en la enseñanza superior.

Hasta 1979, el Reino Unido se caracterizaba por un sistema de enseñanza superior bastante liberal: las universidades recibían una subvención global del *University Grants Committee* (UGC), organismo en el que las Universidades eran mayoritarias. Los créditos eran concedidos por cinco años, lo que permitía a las instituciones llevar una política de medio plazo. El costo por estudiante era un concepto muy utilizado en las negociaciones entre el UGC y el gobierno para fijar la asignación global destinada a todas las Universidades y para la distribución de esta asignación, distribución que hacía la UGC.

Desde 1979 se ha empezado a seguir una política de austeridad en los centros de enseñanza superior —supresión de ciertas subvenciones, disminución del 17 % de las subvenciones públicas entre 1981 y 1984—; además, otras medidas de carácter no económico (creación de un organismo nacional encargado de dar su opinión en materia de planificación del sector no universitario de la enseñanza superior, publicación de direc-

trices precisas en cuanto a gestión de universidades, etc.). El gobierno esperaba así intervenir mucho más fuertemente que antes a nivel de la enseñanza superior para mejorar la relación coste/eficacia con el fin de reducir la financiación pública y garantizar el nivel y la adaptación al mercado de trabajo de las enseñanzas impartidas.

Unido a este intervencionismo creciente, se ha abordado desde la administración central de la educación una reflexión sobre los indicadores de rendimiento. En efecto, hasta esa fecha, el interés y los esfuerzos de la administración se centraban en los indicadores «de economía» que comparaban los recursos reales con los deseados. Progresivamente, unido a la definición de una política de enseñanza superior para los años 90 (libro verde), se han asignado objetivos más precisos a los centros de enseñanza superior y se han propuesto indicadores en dos direcciones:

a) En materia de resultados

Se le asignaron los objetivos principales a la enseñanza superior: 1) la formación de una mano de obra muy cualificada, 2) la investigación, y 3) la aportación de otras ventajas en el plano social. En realidad, las reflexiones no se han dirigido más que al primer objetivo respecto al cual se han seleccionado tres indicadores:

- *La experiencia profesional de los nuevos diplomados* (medida por las variaciones a corto plazo de la tasa de empleo de estos nuevos diplomados).
- *La proporción de diplomados en la población activa* (medida con ayuda de los censos).
- *La tasa de rentabilidad de los diplomas.*

b) Indicadores de economía

Permiten medir si los egresados se producen a un costo razonable. Se han hecho comparaciones entre países con *costos por estudiante* semejante, pero las situaciones son muy diversas y se pueden sacar pocas conclusiones de tales comparaciones. El otro indicador tenido en

cuenta es el *coste por egresado*, pero también en este caso deben ser recogidos muchísimos informes para poder avanzar conclusiones pertinentes.

En realidad, en el Reino Unido, como en otros muchos países, la utilización de indicadores de rendimiento está limitada por la falta de datos: una estructura de indicadores de rendimiento reclama bancos de datos considerables y la cuestión del costo de estos bancos se plantea siempre en relación a las mejoras esperadas.

Paralelamente, el gobierno ha encargado a una Comisión (Comisión Jarrat) analizar la gestión de las universidades y hacer propuestas para mejorarlas. Esta Comisión, compuesta principalmente por universitarios, ha mostrado las insuficiencias de esta gestión principalmente en período de restricciones económicas y, entre otras cosas, ha recomendado elaborar indicadores que puedan ser utilizados a la vez dentro de las universidades y para establecer comparaciones entre ellas.

Después del informe de esta Comisión, por una parte, se han iniciado trabajos a partir de 1985 para conocer el estado de la cuestión sobre la utilización de indicadores de rendimiento en las Universidades y, por otra, para elaborar y seleccionar un número de ellos que pudieran aplicarse a todas las Universidades del Reino Unido. Estos trabajos se refieren a las principales actividades de las Universidades —docencia e investigación— y los indicadores tenidos en cuenta (alrededor de 25) han sido clasificados en función de las posibilidades de utilización (inmediata, próxima o futura), así como según los niveles en los que serían utilizados (departamentos, universidad). Los indicadores seleccionados figuran en el anexo 3. Estos trabajos han dado la ocasión para profundizar sobre los indicadores de rendimiento que fueron examinados desde muchos ángulos: características, campos de aplicación, unidad, peligros, decisiones a tomar.

Por último, las recientes reformas emprendidas en materia de financiación por las Universidades van a poner en evidencia el papel de los indicadores. En efecto, la *Universities Funding Council* (UFC), que ha reemplazado a la UGC en abril de 1989, está encargada de firmar los contratos con las universidades: las subvenciones van a ser acordadas según modalidades más controladas que antes, tomando los indicadores una gran importancia. Es aún demasiado pronto para hacerse una opinión sobre este nuevo procedimiento, pero será muy interesante evaluarlo,

puesto que van a ser utilizados los indicadores que hasta ahora han sido objeto de numerosos debates.

B. El caso francés

Centralizado desde hace mucho tiempo, el sistema universitario francés está evolucionando, lentamente, hacia una mayor autonomía de las Universidades.

Hasta 1970, la centralización era muy importante y el Ministerio de Educación tenía mucho poder tanto en el plano pedagógico (los programas eran nacionales y la pedagogía cuasi-uniforme) como en el plano financiero: los medios financieros y de personal eran concedidos en función de normas variables según las disciplinas. Desde 1970, la autonomía pedagógica se ha desarrollado más rápidamente que la autonomía sobre los medios financieros y humanos, ya que éstos son todavía asignados por el Ministerio en función de ratios que él mismo define. Además, la gestión de estos medios escapa en gran parte a las Universidades, como indica el cuadro que sigue.

Estructura de la financiación de la enseñanza en las universidades francesas (1986)

Origen de la financiación	Forma de financiación	Principales conceptos	%	Forma de gestión
ESTADO	Con presupuestos del Estado	Personal ⁽¹⁾ 17 83	80 2	Con presupuestos del Estado
	Subvenciones a las Universidades		7	Con presupuestos de las Universidades
Recursos propios de las Universidades		Derechos universitarios	2	
		Otras fuentes ⁽²⁾	9	
TOTAL			100	

(1) La integridad de las remuneraciones del personal se ha visto afectada por la docencia. En toda lógica sería necesario reducir el costo de la docencia y asignar más fondos a investigación.

(2) Una parte de estos recursos puede ser afectada por la investigación.

Así, en 1986, las universidades administraban la quinta parte del gasto total de enseñanza.

La situación experimenta, por tanto, una inflexión notable desde hace algunos años con la aparición de los indicadores de rendimiento a dos niveles: a) a nivel ministerial en materia de concesión de medios financieros y b) en un plano diferente con la creación, en 1985, del Comité Nacional de Evaluación.

1) *A nivel ministerial: una política contractual*

La política ministerial basada en contratos es considerada como el elemento fundamental para lograr una mayor autonomía de los centros de enseñanza superior. Se la puede resumir muy esquemáticamente diciendo que pretende ser *global* (tiene en cuenta todas las actividades del centro y todos sus componentes), *progresiva*—los primeros contratos cuatrianuales fueron firmados con muchas universidades a principios de 1990, a los que siguieron otros— y, por último, *transparente*, puesto que las reglas de atribución de los medios son conocidas y tienen un cierto carácter objetivo.

El Ministerio se asigna un papel de impulso y de coherencia. La negociación de los contratos entre el Ministerio y los centros supone la existencia de una «fotografía» de los mismos a partir de las características generales del centro y de *indicadores*, de una parte, sobre la utilización de medios, y, de otra, sobre los resultados obtenidos en los cuatro aspectos concretos establecidos oficialmente:

- 1) los primeros ciclos,
- 2) la inserción profesional de los estudiantes,
- 3) la formación de docentes de 1.º y 2.º grado.
- 4) la investigación.

Se trata, pues, de una nueva postura: cada centro podrá, en adelante, negociar particularmente sus medios, pero dentro de un marco definido por el Estado: «los objetivos deben ser establecidos tanto en materia de formación como de gestión a plano nacional, y servirán de enmarque a

los proyectos presentados posteriormente por los centros» (discurso del ministro de Educación 29/9/1988).

2) *El Comité Nacional de Evaluación*

Este organismo se creó en mayo de 1985. Se compone actualmente de diecisiete personalidades, de las cuales once son representantes de la comunidad científica y universitaria. Este comité está ubicado en el Ministerio de Educación, pero dirige sus informes directamente al presidente de la República; elige libremente los centros a evaluar (no sólo Universidades, sino también cualquier establecimiento público de carácter científico o cultural). No tiene capacidad legal, pero emite juicios de valor sobre la forma en que las Universidades cumplen sus misiones. El Ministerio de Educación Nacional puede apoyarse en este juicio para tomar decisiones.

Hasta la fecha, los centros evaluados han solicitado voluntariamente participar en esta evaluación, que ha sido efectuada por expertos nombrados por el Comité. Desde su creación el Comité ha evaluado unos treinta centros.

A medida que se realizaban estas evaluaciones ha surgido progresivamente entre los miembros del Comité la necesidad de disponer criterios e indicadores. Aunque hay que subrayar que estos indicadores no se puede utilizar en la evaluación como si fueran un «modelo» mecánico, sin embargo apoyan el juicio que facilitan los expertos acerca de la estrategia puesta en práctica en cada centro evaluado.

La renovación del Comité, en 1989, fue la ocasión de «levantar acta» sin concesiones sobre las Universidades francesas y de hacer recomendaciones a los poderes públicos. Los antiguos responsables aprovecharon la ocasión para demandar con fuerza una verdadera autonomía para las Universidades y el establecimiento de nuevas relaciones entre el Estado, las regiones y las Universidades. Especialmente, se sugiere un nuevo reparto de competencias entre el Estado y las regiones y, sobre todo, la potenciación del poder de los presidentes en las Universidades. Estas recomendaciones van en la línea de incrementar el uso de los indicadores en la gestión de las universidades, incluso cuando la óptica del Comité no sea exactamente la misma que la del Ministerio de Educación.

Conclusión

Cada vez se habla más de evaluación en los sistemas educativos y, especialmente, en la enseñanza superior. La escasez de recursos para financiar la formación superior de un número creciente de jóvenes y adultos hace que los poderes públicos, como principales financiadores, se interesen cada vez más por la forma en que son empleados los créditos y en función de qué objetivos. Al no poder englobar la actividad de los centros en un único modelo de optimización, el recurso de los indicadores aparece como un medio de mejorar la eficacia y la eficiencia del dinero que se entrega a los centros de enseñanza superior.

Por su parte, los centros deben adaptarse a este contexto más difícil y, también, desarrollar la utilización de indicadores para mejorar el juicio de los responsables en todos los niveles de la jerarquía administrativa.

Hace unos veinte años, muchos de los órganos decisorios esperaban que los modelos de distribución de recursos les fueran de una gran utilidad. Han debido decepcionarse porque la actividad universitaria es demasiado compleja como para prestarse fácilmente a tentativas de optimización. Los indicadores, en particular los relativos al rendimiento, constituyen una nueva tentativa más pragmática para ayudar a la toma de decisiones.

ANEXO 1
Frecuencia de las citas según tres criterios

Naturaleza del indicador	Relativo a: (en %)	Lugar de obtención		Global (en %)
		Interior	Exterior	
Simple (23 %)	Docencia: 91	98	2	Interior: 98 Exterior: 2
	Investigación: 1	NS	NS	
	Sin distinción: 8	100	0	
Rendimiento (67 %)	Docencia: 67	88	12	Interior: 70 Exterior: 30
	Investigación: 32	34	66	
	Sin distinción: 1	NS	NS	
General (10 %)	Docencia: 63	3	97	Interior: 6 Exterior: 94
	Investigación: 20	17	83	
	Sin distinción: 17	0	100	
Total de Items	Docencia: 72	83	17	Interior: 70 Exterior: 30
	Investigación: 24	34	66	
	Sin distinción: 4	NS	NS	

NS: No Significativo.

ANEXO 2
Indicadores más importantes

Indicadores	Número de citas	% del total
N.º de estudiantes	13	24
Tasas de matrícula	7	13
Costos unitarios	6	11
Superficies (edificadas o no)	6	11
Actividad de los laboratorios	4	7
Contratos de investigación	3	5
Nivel en el acceso de los estudiantes	3	5
Otros indicadores	13	24
TOTAL	55	100

ANEXO 3
Indicadores de rendimiento seleccionados por los universitarios
ingleses (1986)

INDICADORES	Fecha de utilización		
	1986/87	Futuro próximo	Futuro
<i>Docencia e investigación</i>			
- Costo por estudiante a tiempo completo (TP)	X		
- Beneficios obtenidos por la investigación	X		
- Contribución a la formación del postgraduado y profesional	X		
- Porcentaje de estudiantes postgraduados preparando la tesis	X		
- N.º de estudiantes investigando y subvencionados	X		
- Ocupación de los egresados después de 12 meses			
- Tasa de abandono de estudiantes de 1.º ciclo		X	
- Ocupación de los egresados después de 5 años			X
- Análisis de las publicaciones, contratos, patentes, derechos de autor			X
- Índice de citas			X
- Evaluación por colegas			X
- Edición de revistas, pertenencia a sociedades científicas			X
- Pertenencia a consejos de investigación			X
- Costo por diploma			X
- N.º de estudiantes TP por docente TP	X		
- Costo de equipamiento por enseñante TP	X		
<i>Otros</i>			
- Costo administrativo por estudiante TP	X		
- Costo de locales por estudiante TP	X		
- Costo de biblioteca por estudiante TP	X		
- Costo de servicios por estudiante TP	X		
- Costo de las instalaciones deportivas por estudiante TP	X		
- Otros costes centrales por estudiante TP	X		
- Relación entre personal no docente y docente	X		

LECCIONES QUE DEBEN APRENDERSE DE UNA DÉCADA DE DISCUSIONES SOBRE INDICADORES DE RENDIMIENTO *

EDGAR FRACKMANN
Hochschul-Information System GMBH, Alemania

Introducción: cuestiones cambiantes

Vamos a volver la vista atrás sobre una década de discusiones y proyectos relacionados con los indicadores de rendimiento. Estos parecen ser una obsesión para la comunidad universitaria, o al menos un problema sin resolver, que se manifiesta de vez en cuando, probablemente en situaciones o períodos en los que es necesario tomar decisiones generales y hacer cambios importantes ¹. Los problemas surgen si tenemos que empezar desde cero cada dos, tres o cuatro años para determinar cuáles son los pros y los contras de los indicadores de rendimiento en los sistemas de educación superior, volviendo a dejar sin solución los problemas que no se habían podido resolver anteriormente y no poniendo tampoco en práctica los resultados de los proyectos que se habían iniciado. Las personas relacionadas con estos proyectos pueden abandonar frustradas y transmitir el mismo sentimiento de frustra-

* Este artículo apareció originalmente en *International Journal of Institutional Management in Higher Education*, 11, 2. Julio 1987.

1. Comparar:

- La quinta reunión de trabajo especial sobre indicadores de rendimiento para instituciones de educación superior IMHE (1978).
- La cuarta conferencia general IMHE (1978).
- La quinta conferencia general IMHE (1980).
- El proyecto CERI: «Kennzahlenvergleich and deutschen Vororthochschulen».
- El proyecto en Bavaria «Wirtschaftlichkeitsreport», Alemania.
- Los proyectos de Costes y Contabilidad en las universidades de Braunschweig y Tübingen, Alemania.
- El proyecto reciente sobre indicadores de rendimiento CERI/WRK en Alemania.

ción a sus sucesores en los futuros proyectos sobre indicadores de rendimiento.

¿Realmente hemos aprendido algo gracias a las experiencias y fracasos obtenidos con los proyectos de indicadores de rendimiento y hemos profundizado en la visión que tenemos sobre su uso y abuso en la educación superior?

Según mi opinión, uno de los resultados más relevantes del proyecto alemán CERI no fue la gran diferencia o varianza de los valores para los mismos indicadores de rendimiento que eran utilizados en los departamentos y programas de las seis universidades alemanas, que participaron (Mertens, 1978). Lo que sí fue relevante fue la respuesta a una pregunta que se realizó a los representantes de las universidades que participaban. Se preguntó si el hecho de conocer los valores de los indicadores de rendimiento para sus universidades y compararlos con los valores estimados para las otras universidades participantes influiría significativamente en las decisiones de la universidad en ese año o en el año siguiente a la investigación. Las seis universidades, sin ninguna excepción, respondieron negativamente (*Westdeutsche Rektorenkonferenz*, 1980, p. 159).

Las personas interesadas en los proyectos de indicadores de rendimiento no solamente tienen que enfrentarse a que se ignoren los resultados de los proyectos, si no que también, como se pudo observar en el caso alemán de 1980, a la desgana y a las estrategias ideadas para obstaculizarlos. La primera pregunta que pueden hacerse los investigadores es la de cómo han de definirse y estructurarse los indicadores para que indiquen aquello que se supone que han de indicar. Este problema ha de enfocarse técnicamente considerando la relación entre medios y objetivos. Después de esto si los investigadores han tenido que enfrentarse a la ignorancia e incluso a la hostilidad han de volver a preguntarse cuales son los requisitos previos para la utilización de sus indicadores de rendimiento. Es una situación desesperante la de intentar ajustar el entorno a los objetivos del proyecto. Pero tras la experiencia del fracaso de un par de proyectos, al menos en lo que a sus objetivos se refiere, creo que deberíamos darle más importancia a una serie de cuestiones diferentes: ¿Bajo qué circunstancias son útiles o necesarios los indicadores de rendimiento? ¿Cuáles son las condiciones que obligan a usar los indicadores de rendimiento? ¿Por qué son demandados los indicadores de rendimiento? ¿Quiénes promueven el uso de los

indicadores de rendimiento? ¿Quién o qué debe ser el objetivo de los indicadores de rendimiento?

La primera lección que deberíamos haber aprendido tras una década de discusiones sobre los indicadores de rendimiento debería ser el haber llegado a la conclusión de que hay que reanudar la discusión de un modo más fundamental e intentar investigar el origen de nuestro debate y de los principales problemas por resolver.

La controversia subyacente

El escepticismo y las críticas a las que han tenido que enfrentarse los indicadores de rendimiento entre la comunidad universitaria pueden resumirse ahora, tras diez años de discusión, del siguiente modo:

- Los indicadores de rendimiento intentan ser medidas cuantitativas, aunque el resultado del proceso de la educación superior es cualitativo.
- Los indicadores de rendimiento intentan comparar algo que no es comparable: los procesos de producción en la educación superior no aspiran a producir masivamente productos idénticos.
- Los indicadores de rendimiento están simplificando la complejidad y diversidad de la educación superior y de la ciencia.
- Los indicadores de rendimiento son variables aproximativas (*proxies*), es decir, sustitutivas de una realidad compleja, pero no representan los hechos reales en sí mismos. Existe el peligro de que esas variables se conviertan en un objetivo por sí mismas (problema de reificación).

Podemos afirmar que estos argumentos que hemos descrito para desacreditar a los indicadores de rendimiento caracterizan perfectamente el propósito y el funcionamiento de estos indicadores.

Dado que existe una realidad compleja, y que existe el deseo o la necesidad de dar forma, influenciar, organizar y gestionar esta realidad, se deberá reunir y clasificar la información más relevante, estructurándola, desde la mera clasificación hasta la cuantificación, de modo que

pueda ser considerada como indicativo de la compleja realidad subyacente. Sobre las bases de estas «abstracciones» o incluso «alienaciones» de la realidad compleja se hacen evaluaciones, comparaciones y se toman decisiones sobre alternativas diferentes. De este modo, la realidad compleja se hace «manejable». Cuanto más compleja es la realidad, más prevalece la necesidad de «abstracción», siempre que se considere necesaria la manejabilidad. La historia económica de nuestras sociedades modernas, con el cambio del trueque de mercancías por el uso del dinero, proporciona un ejemplo de lo dicho.

La actitud crítica hacia los indicadores de rendimiento en la comunidad universitaria, así como la discusión sobre ellos que surge repetidamente, revela la existencia de una controversia subyacente. Por un lado la idea predominante es la de que la educación superior, las artes y la ciencia no pueden tener en absoluto un tratamiento gerencial y no podrían ser gestionadas por otros que no fueran los propios científicos. Estos individuos se encuentran intrínsecamente muy motivados y se esfuerzan por conseguir la mayor calidad sin que exista para ello ningún incentivo externo. Estos individuos tienen la responsabilidad y la capacidad de dirigir sus propias investigaciones y sus actividades docentes en el contexto de la ciencia y la sociedad, y dan la debida importancia a las actividades adecuadas. Los científicos como individuos piden libertad para la investigación científica y para la docencia (*Wissenschaftsfreiheit, Lehrfreiheit*) al mismo tiempo que piden recursos sociales para llevar a cabo sus investigaciones científicas y sus actividades docentes.

La otra parte afirma que en el sistema de educación superior las decisiones deben tomarse en función del flujo de recursos sociales al sistema, y las decisiones prioritarias deben estar basadas en algún criterio que reduzca la complejidad. Se duda también acerca de si la libertad de investigación y de docencia del científico deben mantenerse como un valor dominante, dado que los individuos están utilizando recursos de la sociedad. Se está convencido de que la sociedad y la ciencia se han hecho tan complejas que los individuos no pueden tomar decisiones globales desde la posición de sus propias actividades de investigación y de docencia (Luhmann, 1968). Parece ser necesario algún tipo de organización, al menos en lo que se refiere al funcionamiento de los mecanismos de autorregulación de la comunidad científica, o quizá algún incentivo externo a la propia comunidad. La libertad de investigación y de docencia se debería redefinir como la libertad de la comunidad científica, cuya organización es necesaria y esperada por la socie-

dad. Para esta organización son necesarias abstracciones, comparaciones, indicadores, en resumen, información simplificada que facilite la toma de decisiones.

Ahora podemos resumir la segunda lección de la discusión: los indicadores de rendimiento se establecen para resolver el problema de la información en general. Desde fuera de la comunidad científica la sociedad espera información sobre lo que la comunidad científica está haciendo con los recursos sociales que recibe, o sobre si están funcionando los mecanismos de autorregulación y de autoorganización (*accountability information*). Desde dentro de la comunidad científica hace falta disponer de aquella información que es necesaria para poder organizar esa comunidad (*management information*).

La segunda lección nos lleva a la conclusión de que sólo hay tres hipótesis alternativas para los futuros proyectos de indicadores de rendimiento:

- La investigación y la docencia no pueden ser objeto de gestión empresarial y nunca deberían serlo. Por tanto, los indicadores de rendimiento se consideran perjudiciales.
- La investigación y la docencia son objeto de gestión empresarial. En este caso los indicadores de rendimiento son útiles y necesarios.
- Un grupo de indicadores de rendimiento especiales puede ser considerado inadecuado para un conjunto determinado de objetivos de gestión, en cuyo caso aquéllos que los critican han de proporcionar otro sistema funcionalmente equivalente de «información simplificada» para así evitar los casos de utilización inadecuados de los indicadores de rendimiento.

Estabilidad, inestabilidad y contraposición de poderes

Una sociedad (representada por el estado) puede sentirse lo bastante rica como para emplear suficiente dinero en la comunidad científica sin que ésta tenga que rendir cuentas. La sociedad puede confiar en que todas las personas de esa comunidad poseen una motivación intrínseca y unos amplios conocimientos que hacen totalmente innecesario el

pedirles cuentas. Para la asignación de fondos a la comunidad científica se puede utilizar una variedad de juicios y de criterios basados en mejor o peor información. En tanto que nadie plantee quejas ni problemas de información para la toma de decisiones, se puede dar una situación de confianza y de estabilidad entre los poderes que suministran los fondos a la comunidad científica y aquéllos que realizan las investigaciones y las actividades docentes.

La situación de estabilidad se pierde cuando surge o se revitaliza la discusión sobre indicadores de rendimiento, cuando las relaciones entre aquéllos que dan y reparten el dinero y aquéllos que reciben los recursos parecen necesitar el cambio, o parece necesaria alguna modificación general en relación con las asignaciones en todo o en parte del sistema de educación superior.

A menudo es la sociedad, a nivel estatal, quien inicia la discusión sobre los indicadores de rendimiento, y con frecuencia es planteada por aquéllos que dan el dinero, más que por los que lo reciben. La comunidad universitaria se puede sentir más o menos afectada por esta discusión y más o menos vulnerable en la situación de inestabilidad. Son posibles tres estrategias distintas por parte de esta comunidad:

1. *A la defensiva*: la unidad bajo revisión intenta ocultar información y evita la revisión o la comunicación externa. Se consigue mejor este objetivo si no se genera ninguna información con respecto a la unidad. Esta estrategia tendría éxito mientras que la unidad no sea castigada por retener información o no esté penalizada (p. ej., económicamente), según las reglas en vigor.

2. *De colaboración parcial*: la unidad objeto de consideración participa voluntariamente en el trabajo de reunir información relacionada con la unidad. Este trabajo se organiza bajo la idea de «prueba», «experimento» y «algo sin consecuencias», cualesquiera que sean los resultados de la información reunida. Esto conduce a que resulte difícil el reunir información, y a la realización de una labor poco útil, sin ningún resultado. Esta estrategia tendrá éxito mientras que la presión por el cambio no se sienta profundamente por ambas partes y este cambio no esté relacionado con ventajas.

3. *De darle forma al cambio*: la unidad bajo consideración entrega información, ya que lo considera ventajoso para su propia super-

vivencia. Hace sugerencias sobre los indicadores por los que desea ser juzgada. Sugiere los canales de información y su utilización, persiguiendo más el ideal de autonomía y de autorregulación que un modelo de gobierno centralizado. Esta es la estrategia necesaria cuando los cambios no puedan ser evitados.

La tercera lección que debemos aprender de la historia de los indicadores de rendimiento no ayuda a prevenir la frustración de aquellas personas dedicadas a los proyectos de indicadores, pero sí ayuda a explicar los fracasos en los intentos. Tales intentos (investigaciones, pruebas, proyectos piloto) pueden ser considerados como fracasos si los resultados no cambian nada. Estos fracasos muestran que no se siente realmente la necesidad de un cambio de la situación actual por parte de los dos lados de los poderes contrapuestos.

¿Qué se valora y por quién?

Los indicadores de rendimiento son discutidos dentro del contexto de la evaluación del rendimiento de las instituciones de educación superior y de sus subunidades. Diez años de discusión debían haber conducido a una visión mucho más diferenciada sobre qué es aquello que valoran los indicadores de rendimiento y sobre quiénes están haciendo esta valoración.

El rendimiento de la educación superior debe ser dividida en sus elementos ²:

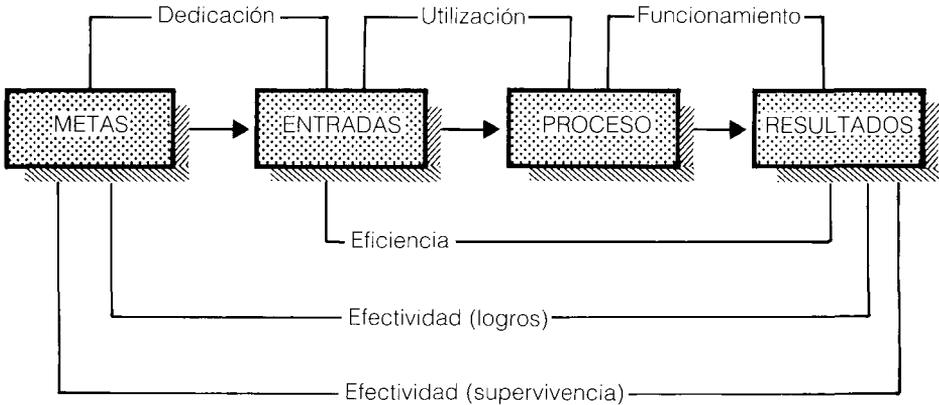
El cuadro 1 nos muestra diez elementos interrelacionados que componen los procesos de la educación superior. Vamos a ver con más profundidad estos elementos basándonos en ideas más bien triviales:

Los indicadores de rendimiento son información sobre el funcionamiento de la unidad sometida a consideración, usada desde el exterior para evaluar el rendimiento y para facilitar la toma de decisiones que afectan a la unidad. El punto de vista externo sobre la evaluación puede ser utilizado por los clientes (aportando recursos, aprobando los resultados del rendimiento), por los que hacen donaciones, por los que toman

2. Basado en Romney *et al.* (1979) y John Calvert en un artículo presentado en el 8 Forum de la *European-AIR*, Loughborough, England (1986).

Cuadro 1

RENDIMIENTO DE LAS INSTITUCIONES DE EDUCACION SUPERIOR Y DE SUS SUB-UNIDADES



las decisiones, o por los que gobiernan en la unidad, y finalmente por la propia unidad (autoevaluación).

Por supuesto que los indicadores de rendimiento no son más que una estructura, un marco para la información. Si la información no es completa, al menos en cuanto al contenido, un cierto número de indicadores dependen de ese marco. Por ejemplo, el indicador de rendimiento «estimación del abandono en los primeros dos años de estudios» no significa nada y no es perjudicial ni provechoso a menos que muestre un número absoluto o un porcentaje (Frackmann, 1983). Esta afirmación puede ser importante, considerando la sensibilidad y la reticencia a la que aún se enfrentan los intentos de construir e investigar en los indicadores de rendimiento.

En la actualidad, el uso de indicadores de rendimiento está de un modo u otro conectado con el flujo de recursos hacia la ciencia y hacia el sistema de educación superior, así como con la utilización de estos recursos. Sobre estas bases, se pueden identificar fácilmente las unidades teóricamente interesadas en el uso de los indicadores de rendimiento, tanto para la evaluación como para la toma de decisiones:

- El estado, en el caso de un sistema de educación superior financiado a este nivel.

- Los organismos «amortiguadores» dentro de la comunidad científica, que tienen la misión de asignar los fondos públicos dentro del sistema de educación superior.
- Los clientes, puesto que deciden sobre el valor de aquéllo que la universidad les brinda.
- Los responsables de la universidad considerada como una entidad organizada.
- Los responsables de los departamentos o de los de otras subunidades.
- Los profesores individualmente como ayudas para la autoevaluación.

Por otra parte, las unidades evaluadas son: el conjunto del sistema de educación superior como un todo, las instituciones de educación superior, el departamento u otras subunidades de la institución, los miembros de la facultad como individuos, y los estudiantes, en la medida en que están participando en el rendimiento de la universidad³.

En el caso de que las decisiones se tomen individualmente, no trascendiendo el marco del individuo (p. ej., la toma de decisión de un cliente o de un profesor), este necesitará obviamente que se reduzca la complejidad y que se simplifique la información sobre la unidad que está bajo revisión, pero las decisiones que se tomen finalmente serán más cercanas a los juicios personales que a criterios bien estructurados y aplicados sistemáticamente. En este caso los juicios no necesitan necesariamente ser transformados en números, en valores de indicadores de rendimiento.

Tan pronto como las decisiones se tomen para fines de las organizaciones y no para propósitos individuales, esto es, cuando los que toman las decisiones no están representando sus propios intereses, sino los intereses de la organización, será necesaria una información bien estructurada por medio de los indicadores de rendimiento. Esta afirmación, por supuesto, es válida sólo allí donde haya un consenso en el

3. La evaluación de los resultados docentes debe ser una evaluación del rendimiento de los estudiantes, pero puede ser una evaluación del rendimiento de la institución o de los profesores.

sistema subyacente de valores, de modo que los criterios bajo los que se toman las decisiones puedan ser conocidos públicamente, especialmente por aquéllos que se ven afectados por las decisiones. Los indicadores de rendimiento deben alcanzar el rango de «procedimientos operativos estandarizados» (Danzinger, 1978), que faciliten el proceso de toma de decisiones sin que sea necesario un continuo comenzar «desde cero». Las decisiones basadas en juicios más o menos fundamentados se deben reducir a «elecciones de importancia marginal» (Wildavsky, 1974).

Debemos distinguir entre tres valores diferentes de los indicadores de rendimiento, que son usados de modo diferente en los procesos de toma de decisiones (Frackmann, 1983):

- *El valor real* («información sobre las cosas tal como son»). Un indicador de rendimiento que no está ligado a ningún valor real no es útil para tomar decisiones. Pero eso mismo también es cierto para un «valor real» siempre que no sea comparado con algún tipo de estimación estándar, con la experiencia pasada o con el nivel al que se aspira de acuerdo con los objetivos y ambiciones que previamente se hayan establecido. El valor real de un indicador de rendimiento puede ser útil para la revisión de viejos valores estándar o para la generación de otros nuevos.
- *El valor estándar* («información sobre las cosas tal y como están especificadas»). Las estimaciones estándar de los indicadores de rendimiento se usan generalmente cuando los responsables de la toma de decisiones sólo pretenden dirigir el funcionamiento de la unidad. Solamente la asignación de fondos es calculada sobre estas bases de valores estándar. No se pretende un control muy detallado de las desviaciones que se pueden producir de los valores estándar, sino solamente de un grado aproximado y general. Que la unidad que recibe los fondos se desvíe en sus resultados de los valores estándar queda como una cuestión de su responsabilidad.
- *El valor deseado* («información de las cosas tal y como deberían ser»). La utilización de estos valores corresponde a las intenciones de la dirección. Se establece un valor estándar, que se espera alcanzar, de modo que en caso de que se produzca alguna desviación de él, la unidad sea penalizada.

Con estas ideas bastante superficiales en mente podemos ahora observar más de cerca los diez elementos del proceso de funcionamiento de la educación superior, y que, pudiendo ser evaluados, constituirán los indicadores de rendimiento. Por conveniencia para posteriores consideraciones vamos a hacer una diferenciación entre el punto de vista interno y el externo. El lado interno es el correspondiente a la institución de educación superior, mientras que el externo es el que corresponde a los organismos externos a las instituciones, como el estado, los clientes, los que hacen donaciones y los organismos «amortiguadores» que asignan fondos. La cuestión que se planteara es la de decidir que es lo que debería ser valorado desde dentro y que debería ser valorado desde fuera de las instituciones.

La institución de educación superior puede ser considerada como una entidad organizada que se esfuerza por sobrevivir en su medio ambiente. Ella atrae recursos hacia su medio y proporciona como resultado algunos productos a su entorno. El medio ambiente de la institución es a primera vista la sociedad en su conjunto, representada por el estado. La sociedad se esfuerza también por sobrevivir y necesita para su supervivencia de lo que produce la institución de educación superior. Mientras que la supervivencia puede ser el objetivo general, tanto para la sociedad como para la institución de educación superior, cada una desde su propia perspectiva, se pueden definir algunos objetivos más concretos que contribuyan a esta supervivencia, cada uno también desde su propia perspectiva. A nivel social pueden ser objetivos tales como el bienestar, el conseguir ciudadanos cultos o el mejorar la competitividad internacional de las industrias nacionales. Para la institución puede ser objetivo de la educación el producir los ingenieros necesarios para una determinada región o la atracción de los estudiantes y de los futuros empleadores mediante una combinación nueva de especialidades en un currículum específico.

Resulta esencial una cierta correlación entre ambos grupos de objetivos —los objetivos de la sociedad con los de la institución— para la supervivencia de ambas. Esta correlación se puede evaluar basándose en alguna información simplificada sobre los objetivos. Pero los objetivos por sí mismos no aportan suficiente información para la toma de decisiones. Lo que es más esencial es el grado en el que cada objetivo es alcanzado por los resultados de la institución de educación superior, lo que se denomina efectividad, con sus diferentes niveles e interpretaciones: el logro de los objetivos de la institución, la supervivencia de la

institución, el logro de los objetivos de la sociedad, la supervivencia de la sociedad. Se puede aplicar lo mismo a las otras subunidades de la sociedad involucradas en la educación superior y en la ciencia: los clientes, los donantes de la institución y los organismos «amortiguadores».

La conclusión que podemos sacar de estas consideraciones es que el principal objetivo de la evaluación de la educación superior y de la ciencia, mediante el uso de los indicadores de rendimiento, deben ser los procesos de funcionamiento (cuadro 1), los objetivos, los resultados y la efectividad, es decir, el logro de los objetivos a través de los resultados de la institución. Existen distintos grados de efectividad que los indicadores de rendimiento tendrán que reflejar, cuando se construyen y usan en los diferentes niveles de observación:

- El estado no puede observar con mucho detenimiento la efectividad de cada institución individual de educación superior, p. ej., los programas, los departamentos, los miembros de la facultad, el logro de los objetivos educativos o de investigación. El foco de la evaluación a nivel estatal está en la efectividad global del sistema de educación superior, de acuerdo con las aspiraciones y objetivos de la sociedad. Los indicadores de rendimiento de uso estatal no deben reflejar necesariamente el rendimiento de una institución individual, sino el rendimiento del sistema global de educación superior.
- Las instituciones «amortiguadoras» tendrán que evaluar el rendimiento institucional con más detalle, basándose en los programas, en las especialidades de estudios y de investigación, en los individuos involucrados en actividades de investigación y en los procesos docentes. La sociedad espera que las instituciones «amortiguadoras» se dediquen a la asignación de los fondos públicos, juzguen y decidan prioridades y alternativas basándose en su nivel de conocimiento y en su visión profunda del desarrollo científico.

Pero el punto de vista de las instituciones «amortiguadoras» sigue siendo sólo una opinión parcial del rendimiento global de la institución, ya que deben permitir que sea la institución la que decida sobre los parámetros que afectan a la supervivencia institucional.

- La evaluación del rendimiento institucional por parte de los clientes y de los donantes está orientada todavía de un modo más parcial: centra su atención sobre la correlación de los objetivos individuales con el logro de los objetivos de la institución.
- Continúa siendo el privilegio de la institución el uso de información simplificada y de los indicadores de rendimiento para juzgar y decidir sobre la efectividad global de la institución, así como sobre cada pormenor de la relación objetivos/resultados dentro de la institución.

Hasta aquí sólo hemos tenido en cuenta la información sobre los procesos de funcionamiento institucional, p. ej., información sobre los resultados, los objetivos y la efectividad, y además hemos descubierto también que hay distintos niveles de información que se usan y que son necesarios para una información más exhaustiva y detallada. La diferencia más importante que hemos descubierto es la que existe entre la institución, con sus necesidades de información sobre su propio rendimiento, y el entorno, con sus necesidades de información sobre el rendimiento de la institución. Es una responsabilidad y un privilegio propio de la institución el evaluar su propia efectividad global, es decir, asegurar su supervivencia institucional, mientras que el enfoque sobre la efectividad institucional por parte del entorno (clientes, donantes, instituciones «amortiguadoras») debe estar orientado de un modo más parcial, o debe considerar la efectividad de la institución sólo como la de una parte de la efectividad social. Es muy importante la distinción entre una valoración «interna» y «externa» de la efectividad, porque esto nos indica el grado de autonomía que la institución disfruta.

Si se realiza esta distinción con el uso de información y de indicadores de rendimiento, la institución conserva la autonomía para decidir cómo compaginar los objetivos de la sociedad, de los clientes y de la ciencia con los objetivos de la institución para garantizar la propia supervivencia de la institución.

Si es válida esta distinción entre indicadores internos y externos de efectividad, es incluso más importante para las otras seis categorías restantes del cuadro 1, lo que nos puede llevar a generar más información simplificada y a la construcción de indicadores de rendimiento. Dado que las demás categorías se refieren a cuestiones de información que se centran en los procesos de funcionamiento «interno» de la edu-

cación superior, la génesis y el uso de información se debería dejar a la discreción de la propia institución. Los agentes o agencias externos no deberían interesarse sobre el modo en que estos factores contribuyen globalmente a la supervivencia institucional:

- La *información sobre entradas* caracteriza qué es lo que la institución ha hecho para transformar los recursos financieros en personal y cómo ha usado los factores trabajo en su proceso de funcionamiento.
- La *información sobre la dedicación* sirve para especificar la información sobre las entradas, mostrando hasta qué punto los diversos factores que entran se dedicaron al logro de los objetivos.
- La *información sobre el proceso* sirve para caracterizar la función de producción escogida para obtener los resultados.
- La *información sobre la utilización* especifica la información de proceso mostrando cómo se usan los factores de entrada en el proceso.
- La *información sobre el funcionamiento* también especifica la información del proceso poniendo de manifiesto aquellas características más importantes por su influencia sobre los resultados.
- La *información sobre la eficiencia* valora la relación entre entradas y salidas de los procesos institucionales.

Es evidente que podría haber excepcionalmente algunas intromisiones «externas» o existir interés desde fuera por la información sobre entradas, sobre el proceso o sobre la eficacia. Por ejemplo, los clientes, los que dan fondos y las instituciones «amortiguadoras» pueden estar interesadas en la capacidad de actuación de la institución, p. ej., la calidad y la reputación de los investigadores, la biblioteca, los servicios de informática y otros servicios de aprendizaje y de investigación, las características de la organización sanitaria de la institución, etc. Pero en este caso, todo esto puede ser considerado como el resultado de actividades institucionales previas, y, por tanto, como un producto de la institución. El estado y las instituciones «amortiguadoras» pueden estar interesados en algunas características del proceso, pueden estar interesadas en que funcione la autogestión institucional en el sentido de autorregulación, o en que se recoja y se use información sobre la efica-

cia y la efectividad de las subunidades institucionales para establecer prioridades entre diferentes decisiones. Esta demanda de información sobre los «procesos» no perjudica realmente a la autonomía institucional.

Pero esta distinción entre «interno» y «externo», generada para simplificar la información y para usar los indicadores de rendimiento que garantice la autonomía institucional dentro del sistema de educación superior y de la sociedad, no debe ocultar el hecho de que la autonomía de la institución de educación superior puede resultar perjudicada por el conocimiento y el uso de información desde el exterior mediante los indicadores de rendimiento. Para prevenir el uso de información que suponga una intromisión en la autonomía institucional no son suficientes los términos «prueba», «ensayo primario», «estudio piloto», o «sin consecuencias» utilizados para describir y para motivar los proyectos de indicadores de rendimiento. Solamente nos conducirá a la obtención de información sobre rendimiento institucional, que no sea perjudicial a la autonomía institucional, los principios y los mecanismos organizativos, o también las tradiciones. Estos principios se tratarán en la próxima sección de este artículo.

La cuarta lección que hemos aprendido tras una década de discusiones sobre indicadores de rendimiento nos revela las razones para la reticencia en la aceptación de los indicadores de rendimiento desde el punto de vista de las instituciones de educación superior. Si se considera que la autonomía institucional es un valor que debe mantenerse, ciertamente surgirán problemas al generar y usar los indicadores de rendimiento, debido a que la información sobre el rendimiento institucional intensifica en general la capacidad directiva de las agencias exteriores (el estado) contra la institución.

Si, por ejemplo, los indicadores de rendimiento se usan (a nivel estatal) como «valores deseados» para recompensar o castigar a una institución en términos de recursos financieros, es decir, indicadores de rendimiento que fijan la relación entre los factores no monetarios de entrada y los resultados de todos los procesos importantes del funcionamiento de la educación superior (información de ingreso y de eficiencia), entonces no es adecuada para la autonomía institucional esta utilización de información simplificada. Si se usara la relación estudiante/profesor o el coste unitario por estudiante como un valor estándar sólo para decidir el importe de las asignaciones estatales de acuerdo a las

cifras de matriculación, esto permitiría a la institución que configurara sus propios procesos docentes.

Limitaciones con las que hay que convivir

Estas últimas observaciones sobre los peligros de los indicadores de rendimiento no nos deben llevar a la conclusión de que los indicadores de rendimiento no deben utilizarse nunca. Por el contrario, si se comparte la convicción de que existe una necesidad de organización dentro de la comunidad universitaria, y entre la educación superior y el entorno social, entonces es necesaria la distribución y el uso de información simplificada.

Pero debemos ser conscientes de que existen muchas limitaciones con respecto a los indicadores de rendimiento que no podemos resolver y con las que tenemos que convivir. Estas limitaciones se refieren obviamente a las críticas mencionadas al principio de este artículo:

- En las instituciones de educación superior la docencia y la investigación se caracterizan por su diversidad. Nadie desea que las instituciones «produzcan» «productos de masas» similares o idénticos. Pero si hay que tomar decisiones, son necesarias las comparaciones, y cada decisión es el resultado de un juicio transformado en números, incluso hecho inconscientemente. Los indicadores de rendimiento comparan la diversidad incomparable y cuantifican resultados cualitativos.
- Una información simplificada significa unos pocos parámetros y una información agregada. Pero cualquier resultado de los procesos de la educación superior será mejor descrita por un buen grupo de indicadores desagregados. Pero cuantos más indicadores son tenidos en cuenta simultáneamente, menos se satisface la intención de simplificar la información.
- La información, incluso si se genera para uso «interno» y para la gestión, una vez utilizada dentro del «público» institucional, no se puede evitar que se alcance publicidad más allá de los límites de la institución.

Estas observaciones de nuevo deben entenderse como una advertencia para tener cuidado con los indicadores de rendimiento, y no co-

mo una crítica hacia su uso. Ello debe dar paso para que hagamos algunas reflexiones sobre qué principios y mecanismos organizativos, o tradiciones, pueden garantizar la autonomía institucional, es decir, cómo usar información simplificada para tomar decisiones en la educación superior sin perjudicar la autonomía institucional. Nos podemos referir a estos mecanismos organizativos, especialmente a los dedicados a la relación entre las decisiones de financiación estatal y el rendimiento institucional, como los mecanismos de «amortiguación» y de «desacomplamiento»:

- Dar las subvenciones a las instituciones para más de un año, es decir, dar a las instituciones suficiente tiempo para que reaccionen frente a los puntos débiles, una vez que éstos han sido descubiertos por la información suministrada por los indicadores de rendimiento. Ellas tendrán que aumentar su rendimiento (efectividad), probar que la información proporcionada por los indicadores es errónea, o tendrán que aceptar el «castigo».
- Dejar a los «iguales» decidir sobre la asignación (instituciones «amortiguadoras» con la función de hacer asignaciones de fondos). Esto garantizará que las decisiones estén más cerca de opiniones informadas que del uso de abstracciones.
- Usar solamente valores estándar en lugar de valores deseados. Esto proporcionará a la institución la necesaria flexibilidad para permitirse desviaciones en alguno de los procesos de su funcionamiento.
- Dar preferiblemente más oportunidades de decisión a los organismos «amortiguadores». Esto concederá más legitimidad de la mayor parte de los indicadores por los que se juzga el rendimiento institucional. Esto permitirá a la institución escoger sus «dependencias» a su discreción.
- No ligar nunca con recompensas económicas o con castigos (siguiendo información simplificada que nos pudiera llevar a esas conclusiones) una conducta, una estructura o una configuración del proceso de funcionamiento de las instituciones de educación superior que sean consecuencia de un mandato obligatorio.

Podemos añadir una quinta lección con la afirmación de que cuantas más instituciones tienen la autonomía concedida, y cuanto más han de

tener cuidado de su propia supervivencia, más deben estar interesadas en la generación voluntaria y en la divulgación de al menos una información simplificada de su actuación (con el objeto de atraer recursos hacia su entorno). Cuanto más instituciones son dependientes directamente de los responsables estatales, más tenderán a ocultar información o a no darla en absoluto.

Algunas observaciones sobre la situación alemana y británica

En 1986 las organizaciones de presidentes de las universidades en Gran Bretaña y en Alemania formularon declaraciones sobre los indicadores de rendimiento (*Westdeutsche Rektorenkonferenz*, 1986; CVCP, 1986). Parece interesante intentar identificar algunas diferencias y similitudes frente a las lecciones que hemos podido aprender de la década de discusiones sobre indicadores de rendimiento.

Vamos a dar una idea sobre la situación en Alemania para explicar la presentación del artículo de los presidentes alemanes sobre los indicadores de rendimiento. En Alemania existe una continua discusión acerca de una mayor competitividad y diferenciación entre las instituciones del sistema de educación superior. Esta discusión pública se inició a nivel de los *Länder*. Una de las ideas fue sin duda la de cambiar la asignación de fondos con una mayor selectividad de acuerdo al rendimiento de las instituciones, las cuales, esto fue lo que se pensó, deberían estar diferenciadas en niveles superiores e inferiores.

El *Wissenschaftsrat*⁴ alemán publicó en 1985 una recomendación sobre la competitividad del sistema alemán de educación superior (*Wissenschaftsrat*, 1985). En esta recomendación se declaró que una precondición para la competitividad debía ser la «transparencia». En otras palabras (que no son usadas muy claramente en este caso), para la asignación de fondos estatales de un modo más selectivo se requiere una información (simplificada) sobre el rendimiento diferenciado de las instituciones o sus subunidades. El *Wissenschaftsrat* recomendó que la Conferencia de Rectores Alemanes (*Weestdeutsche Rektorenkonferenz*: WRK) debía reunir, jerarquizar y publicar esta información sobre el ren-

4. Un comité conjunto del estado y de la educación superior con base nacional. con la tarea de hacer recomendaciones sobre todas las cuestiones relevantes para el sistema de educación superior.

dimiento. Esto, por supuesto, no fue una idea muy buena, dado que la función de la WRK es la de representar los intereses comunes o conjuntos de las universidades, una tarea que tiende más a ocultar que a revelar la diferenciación y la diversidad. Pero el WRK sin embargo sintió la necesidad de expresar algunas opiniones sobre la valoración del rendimiento de las instituciones de educación superior y sobre el uso de los indicadores de rendimiento.

El artículo del WRK alemán es una manifestación muy cuidadosa sobre los indicadores de rendimiento. En efecto, no existe ninguna recomendación entre sus afirmaciones para el uso de ciertos indicadores de rendimiento. Hay una lista ordenada de indicadores en un apéndice bajo el título de «indicadores frecuentes» y no bajo el título de «recomendados». El artículo contiene más bien una advertencia que una descripción de la utilidad de los indicadores de rendimiento. Hace alguna mención de la discusión que tuvo lugar en Alemania sobre la asignación de fondos públicos en el sistema de educación superior según fuera el rendimiento, pero el artículo por sí mismo no se declara sobre si se deben utilizar, cómo deben organizarse o quién debe evaluar el rendimiento y asignar los fondos. La afirmación principal del artículo es la de que los indicadores cuantitativos deben usarse sólo en circunstancias muy excepcionales. Deben prevalecer las descripciones y evaluaciones cualitativas. Esta insistencia sobre la descripción cualitativa implica que, si no es aceptable una transformación de un criterio cualitativo en cifras, al menos en alternativas de «sí o no» —o de «más o menos»—, entonces la asignación de fondos no se puede basar en estos criterios de una forma totalmente visible y transparente.

El artículo básicamente continúa manteniendo la reticencia respecto a los indicadores de rendimiento. Esta actitud puede explicarse dentro del contexto del actual mecanismo de asignación de recursos y dentro de las discusiones sobre los cambios estructurales hacia una mayor competitividad y diferenciación. Actualmente, las universidades alemanas están financiadas por los *Länder*. Las decisiones sobre asignaciones para las respectivas universidades se realizan a nivel de cada uno de los once estados. Ni los indicadores de rendimiento ni información o criterios bien estructurados o públicamente visibles se utilizan para la asignación. La asignación no debe considerarse selectiva.

En la presente discusión sobre una mayor y mejor información sobre el rendimiento de las universidades, no se sugiere de un modo bastante

claro cómo ha de usarse esta información, y por quiénes, y en qué contexto organizativo, como para garantizar la autonomía de las universidades. Sin ningún cambio estructural, con mayor y mejor información basada en los indicadores de rendimiento, sólo podríamos aumentar la capacidad del estado para gobernar las universidades, a expensas de la propia flexibilidad de las universidades. De este modo el argumento (por parte de las universidades) a favor de una información más bien cualitativa, en lugar de los indicadores de rendimiento cuantitativos, debe ser interpretado —a la vista de las estructuras de asignación simples e inmodificables— como un intento de mantener tanto tiempo como sea posible esta gran complejidad para evitar la evaluación desde el exterior del rendimiento de la universidad.

Las estructuras de asignación en Gran Bretaña muestran una mayor complejidad en el sentido de la existencia de mecanismos de «amortiguación» y de «desacoplamiento», que se mencionaron anteriormente:

- La decisión final para la asignación de fondos públicos a las universidades es tomada por «iguales» (esto es, el *University Grant Committee*: UGC).
- Se concede el dinero por tres años, para dar a las universidades suficiente tiempo para reaccionar a las recomendaciones ligadas a la concesión.
- Los fondos se dan como una subvención en bloque; esto significa que la responsabilidad por la flexibilidad en las decisiones de cómo debe gastarse el dinero recae sobre la institución (Carta circular de la UGC (22/85)).

Hay otros dos fenómenos a tener en cuenta por la diferente actitud por parte de las universidades británicas hacia los indicadores de rendimiento en comparación con las universidades alemanas. Hay de algún modo unos canales de información más claramente estructurados entre una universidad y el UGC que asigna los fondos. En 1981, cuando el gobierno impuso un recorte del presupuesto general para el sector universitario, la UGC decidió asignar los fondos reducidos con una fuerte selectividad. Después de estas decisiones, «las universidades se quejaron firmemente», de acuerdo con la UGC (Carta circular de la UGC 4/86), «puesto que el proceso es oscuro»: las universidades no obtuvieron suficiente información relativa al criterio con el que se realizaron las

decisiones de asignación. Por esta razón la UGC comenzó «un enfoque más racional y sistemático» (Carta circular de la UGC 4/84) para las asignaciones de 1986.

A las universidades se les pidió —con la carta circular 12/85— que presentaran algún tipo de autoestudio bajo la hipótesis de una disminución media anual en la subvención para cada institución de un 2 % en términos reales.

El autoestudio debe incluir (UGC 12/85):

- Una corta declaración de sus objetivos globales para el período planificado.
- Una declaración sobre la investigación en la universidad, las áreas tratadas y los costes.
- Una previsión del número de estudiantes.
- Las previsiones financieras.

El procedimiento de cambio de asignación de recursos se anunció en la carta circular 22/85 de la UGC. Nos parece importante repetir una vez más el principio de «desacoplamiento» de las actividades de la UGC: «El propósito de los procesos de asignación es el de permitir al comité establecer la subvención que debe darse a cada universidad, no el de determinar cómo la universidad debe gastarla» (UGC 22/85). Los resultados son el «cálculo de un número de cantidades diferenciadas de dinero en relación con las diferentes actividades y producido por una variedad de fórmulas y criterios» (UGC 22/85):

- Distribución de recursos sobre un criterio basado en la enseñanza, esto es, el número de estudiantes y de unidades en los centros de costes.
- Distribución de recursos según un criterio basado en la investigación, cuyas partes se distribuyen según:
 - a) Un criterio relativo a la cantidad de empleados en investigación y al número de estudiantes que están investigando.
 - b) La base de los ingresos de las juntas de investigación y de los organismos benéficos.

- c) La base de ingreso de contratos de investigación.
- d) Juicios de opinión.

Con respecto al último criterio de asignación, quizá el uso más sorprendente de los indicadores sea la transformación de los juicios en números del 1 al 5 para determinar la distribución.

Puede haber un tercer factor, que contribuye a una actitud distinta hacia los indicadores de rendimiento en Gran Bretaña: el estado manifiesta su interés de que tenga lugar una buena gestión dentro del sistema universitario (esto es en el nivel institucional y en el de la UGC) basada en la información: «Una gestión sólida no solamente se basa en una utilización eficiente de los recursos (ingresos), sino también en la efectividad de los resultados alcanzados (salidas). Esto demuestra la necesidad del desarrollo y del uso de medidas de rendimiento» (Libro verde, 1985). El estado siguió la recomendación del denominado Informe Jarrat para la UGC, la CVCP y las universidades: «Debería desarrollarse una clasificación de los indicadores de rendimiento, incluyendo tanto las entradas como los resultados, diseñada para el uso de dentro de las universidades individualmente y para la realización de comparaciones entre instituciones» (Informe Jarrat, 1985). El estado subraya su interés en el desarrollo de indicadores de rendimiento con sugerencias para un primer grupo de indicadores (Libro verde, 1985).

Contra este conjunto de fenómenos estructurales del sistema de universidades británicas, se hizo pública una declaración de un grupo de trabajo común CVCP/UGC sobre los indicadores de rendimiento (CVCP, 1986) que parece que tomó algo más la ofensiva que el artículo alemán. Se acepta la utilidad e incluso la necesidad de los indicadores de rendimiento para la gestión de las universidades, así como para la asignación de recursos a nivel estatal. La única salvedad es la de insistir en que el tipo de información que proporcionan los indicadores de rendimiento no debería dominar los procesos de gestión, pues unos buenos juicios aún debería jugar un papel predominante. Se explica y propone una lista inicial de 24 indicadores, una lista dictada más por disponibilidad de datos o por la existencia del mecanismo para recopilarlos, que por la importancia intrínseca de los indicadores. Se alienta la publicación regular de indicadores y la continuación de los esfuerzos para definirlos.

Conclusiones finales

Los indicadores de rendimiento establecen una información simplificada que es necesaria para la gestión y la organización. Los indicadores de rendimiento deben siempre usarse desde fuera o para la toma de decisiones por los niveles superiores a las unidades sometidas a la revisión.

Los indicadores de rendimiento son por sí mismos peligrosos para las unidades bajo revisión, debido a que tienden a intensificar la capacidad de gobierno de los responsables externos.

Si existe algún sistema de gestión y de organización completamente aceptado por todos para el sistema universitario, entonces es necesaria para la toma de decisiones alguna información simplificada como son los indicadores de rendimiento. Aunque esto está comúnmente aceptado, la cuestión principal no debe ser la definición y el contraste de los indicadores de rendimiento, sino la definición y la construcción de estructuras organizativas, de gestión y de organización que garantice los dos valores esenciales de libertad docente y de la ciencia, y también la autonomía institucional, así como la intensificación del rendimiento y de la calidad de la educación superior y de la ciencia en la sociedad. La definición de indicadores de rendimiento no es más que el toque final a esta consideración.

LA INSUFICIENCIA DE LOS INDICADORES DE RENDIMIENTO PARA LA EDUCACION SUPERIOR: LA NECESIDAD DE UN MODELO MAS GLOBAL Y FORMATIVO *

HERBERT R. KELLS

Universidad de Twente, Países Bajos
Danbury Park Management Centre, Reino Unido

Los ministerios, los organismos encargados de la asignación de fondos, y otras entidades involucradas en el control y desarrollo potencial de la educación superior de varios países acaban de tomar una importante decisión política, o están a punto de tomarla. Esta decisión se refiere a la utilización de un sistema de elementos cuantificables o relaciones numéricas para medir ciertos aspectos de los *inputs* de las instituciones, de su funcionamiento y de los resultados, a menudo para poder publicar estadísticas comparativas y *rankings* de las instituciones. Los representantes gubernamentales de países como Australia (Linke, 1990) y el Reino Unido (Sizer, 1990) y, en menor grado, Dinamarca (von Linstow, 1990) y Finlandia (Niiniluoto, 1990), proponen que los indicadores de este tipo afecten, e incluso que sean el principal factor, de la política de financiación, o de caracterizar el tamaño, la dimensión y los posibles recursos de las universidades y de sus programas de estudios. Dichas tendencias siguen encontrando apoyo entre los que las creen necesarias para el sistema de educación superior, aunque algunos destacados líderes educativos se han opuesto a tales tendencias por considerarlas extremadamente reduccionistas, no siendo representativas de la calidad y del valor intrínseco, ni de los programas de estudios, y ni de las universidades, además de centrarse sólo en lo que es cuantificable. Recientemente los indicadores de rendimiento han sido objeto de un importante estudio, aunque nada halagüeño, por parte de la OCDE (Cuenin, 1987), pero aún no se ha presentado ninguna propuesta que esté de acuerdo con los requisitos de los que exigen la utilización de los

* Artículo aparecido originalmente en *Higher Education Management*, 2, 3, 258-270, 1990.

indicadores de rendimiento, que satisfaga al mismo tiempo las críticas más importantes hechas a este tipo de medidas, y que además pueda ser útil, y no perjudicial, a las instituciones de educación superior. El objetivo de este artículo es, precisamente, proporcionar tal propuesta, y al mismo tiempo ofrecer una visión profunda sobre el contexto más adecuado y sobre el posible funcionamiento de los indicadores de rendimiento que se proponen a la luz de los argumentos que se presentan en su favor.

Los indicadores de rendimiento y su influencia en la vida universitaria

Los sistemas sociales son influidos por muchas fuerzas, externas e internas, que ejercen su influencia durante períodos de tiempo cortos, largos o medios. En el largo transcurrir de la historia, la larga duración (Braudel, 1980), se desarrollan actitudes culturales importantes, ideas críticas y tradiciones que encuentran su expresión más o menos acertada en los pronósticos y estructuras de las organizaciones. En el caso de las universidades algunos ejemplos de tendencias a largo plazo serían la influencia de la iglesia, la noción que la investigación sea una misión primordial, la configuración de la vida universitaria y el movimiento desde una situación de élite hacia otra de mayor acceso.

En el término del medio plazo, los períodos medidos por décadas en vez de por siglos, existen tendencias algo más mecánicas, pero todavía esenciales, que se definen y que ejercen su influencia sobre el sistema, como una respuesta normalmente retardada a los fenómenos más duraderos. Los ejemplos aplicables a la universidad serían la evolución de los ordenadores, el coqueteo generalizado con la industria y la tecnología, e incluso las tentativas de una reindustrialización nacional (véase Scott, 1984). Igualmente se podrían mencionar los cambios en el gobierno de las instituciones como, por ejemplo, la inclusión de la participación estudiantil y el desarrollo del poder del cuerpo docente en algunos países.

A corto plazo son muchas las tendencias, a menudo cuasi-modas, que se pueden observar. La vida de una idea determinada dura normalmente unos cuantos años. Durante las décadas de los sesenta y los setenta se produjeron muchas corrientes de este tipo en la educación superior: se propuso la abolición de la titularidad del profesorado, la

apertura del proceso de admisión y la solución de los problemas urbanos a través de acciones universitarias. Entre los ejemplos más sustanciales y duraderos están los esfuerzos por establecer programas de formación continua del profesorado, por revitalizar la enseñanza universitaria y por crear instituciones «abiertas» y no tradicionales. Entre los movimientos actuales están las tendencias que abogan por la planificación (y quizá gestión) estratégica, por la implantación de un estilo empresarial, por el fortalecimiento de la función ejecutiva de las instituciones en Europa y por la utilización de los indicadores de rendimiento. Las corrientes a corto plazo suelen tener un carácter enérgico y a veces incluso hasta frenético. El modo en que se formulan suele ser precipitado y con un marco conceptual sólo párcialmente desarrollado. Los defensores de los mismos suelen ser muy locuaces y algunos de sus argumentos no tienen en cuenta las características del sistema que los ha de adoptar, parecen haber sido elegidos por su capacidad de coger adeptos sin tener en cuenta las consecuencias para el sistema.

Hay que admitir, por supuesto, que los tres grandes grupos de fenómenos o temas mencionados arriba no son excluyentes entre sí, y que la tendencia natural, aunque obviamente difícil y con muchas complicaciones, es que los temas ascienden y emergen de manera selectiva por la escala que va del corto al largo plazo. En la mayoría de los casos, el fin de cualquier consenso, por no decir de toda actividad, llega demasiado pronto, aunque también pueden observarse recurrencias cíclicas sobre el mismo tema. No obstante, muchos de los fenómenos de corta duración son más útiles que causantes de dispersión, y algo de su esencia suele «pegársele» a las instituciones implicadas, quedando registrado en la memoria institucional para bien o mal. Finalmente, los fenómenos de corta duración suelen adolecer de un fallo terrible. Pueden ser intentos de conceptualizaciones decepcionantemente parciales, intentos de responder de la manera más cómoda a la forma más obvia de una necesidad aceptada por todos. A menudo pierden de vista la naturaleza más importante o fundamental del contexto, del asunto en cuestión y de las instituciones involucradas. Esto parece ser el caso de algunos aspectos de las propuestas o demandas que se hacen para que las instituciones sean descritas y juzgadas de manera básica mediante el uso de los indicadores de rendimiento. Estas propuestas, y la utilización de tales indicadores, se examinarán a continuación y se formulará otra propuesta para incluir los indicadores de rendimiento dentro de un esquema más formativo. De esta manera puede que entren dichos indicadores en la categoría más aceptable y sustancial de los fenómenos a medio plazo.

El auge de los indicadores de rendimiento y lo que prometen

La triste verdad para el ámbito universitario es que hemos abierto las puertas a la utilización indebida y perjudicial de los indicadores de rendimiento con todos sus posibles daños. No hemos creado, individual ni colectivamente, los sistemas multifacéticos que nos permitan informar a los demás de nuestra eficacia y eficiencia, y que, al mismo tiempo, nos ayudarían a mejorar nuestros servicios y programas de estudio. Esto ha sido así tanto en las buenas épocas como durante los momentos de escasez de fondos, lo que ha motivado las quejas reiteradas de los gobernantes de muchos países. En ausencia de un sistema de evaluación realizado por las propias universidades, los ministerios y organismos encargados de la financiación, o bien han tenido que hacer sus propias valoraciones, y/o bien han tenido que recabar la información que necesitan de la forma más simple y rápida para poder responder a ciertas cuestiones, para adjudicar fondos, recortar presupuestos y tener alguna base que justifique sus acciones. Exigen una información que sea rápida y económica de recoger, sin preocuparse de que refleje la envergadura, variabilidad y sutilezas del sistema. Y la utilizan de forma normalmente directa y a veces con crudeza, ya que falta un cuerpo de análisis que esté centrado en las metas del sistema y de sus instituciones individuales, así como en los aspectos centrales de su función: la enseñanza, el aprendizaje y la investigación.

Hay que reconocer sin titubeos que los ministerios y los organismos encargados de la financiación sí que necesitan, al igual que las propias universidades, unos indicadores generales de la productividad (la cantidad de cosas realizadas y los recursos gastados en su realización), de la eficiencia (el coste de cada cosa o conjunto de cosas realizadas o producidas), y de la efectividad, junto con los mecanismos de control de todo ello. Se trata de dimensiones que afectan a todo el sistema, o de aspectos que cualquier persona razonable estaría de acuerdo en aplicar a las universidades en general. Son necesarios para explicar y, de alguna manera, justificar el sistema y la adjudicación general de recursos destinados al mismo; algo que las universidades no han hecho bien hasta el presente. Se presupone, por supuesto, que las medidas describen las dimensiones y los productos básicos de la universidad, pero el problema surge cuando las medidas describen sólo una fracción de los proyectos y resultados de una universidad o de un sistema. El uso de un conjunto parcial de indicadores como base para tomar decisiones críticas en cuanto a la financiación o la prórroga de los progra-

mas y a la publicación de un *ranking* de las instituciones sería, sin duda, un uso indebido de tales descriptores y medidas aproximativas de la calidad. Estos usos inadecuados podrían, por el afán de rapidez o por ignorancia, poner en peligro el tejido mismo de unas instituciones centenarias. Estas instituciones, aunque pequen de falta de autoanálisis y de no presentar estadísticas, merecen un sistema mejor para controlar y utilizar la información referente a su rendimiento. La construcción de un sistema adecuado de evaluación es un reto que ha de ser asumido por las mismas universidades de acuerdo con los organismos gubernamentales que les asignan los fondos.

El uso indebido de los indicadores de rendimiento puede tener consecuencias graves. En Gran Bretaña el Grupo de Trabajo sobre los indicadores de rendimiento del *Committee of Vice Chancellors and Principals* y del *University Grants Council* (CVCP, 1986a) definió los indicadores de rendimiento como «informes normalmente cuantitativos de los recursos empleados y de los logros obtenidos en las áreas relevantes de los objetivos particulares de la empresa», y sugirió que fueran utilizados como señales o guías, y no como valores absolutos. No obstante, a pesar de las constantes advertencias y consejos acerca de su uso cauteloso, propiciados por los propios defensores de los indicadores de rendimiento, aparecieron en la prensa listas de programas e instituciones ordenados por indicadores, justo en el preciso momento en el que se estaban realizando recortes draconianos en el apoyo financiero a la enseñanza superior británica. Sizer (1989a-1990b) alega que la opinión generalizada en la Reino Unido fue que los mayores recortes presupuestarios se basaron en el uso comparativo de una característica de *input* (las notas medias del final de la enseñanza secundaria), más que en la calidad relativa de los distintos programas o en medidas de la eficiencia. Dado que la limitada información que aportan los actuales indicadores de rendimiento ha sido (en ausencia de información exhaustiva de los «logros diarios» de las instituciones británicas) el instrumento central (o al menos ha influido profundamente) de la evaluación por colegas y de otros juicios utilizados para realizar recortes, es muy probable que los recortes que se hicieron fueran inapropiados. El reciente «ejercicio de selectividad» (*The Times Higher Education Supplement*, 1989) empleado en Gran Bretaña, en el que algunos expertos recibieron un listado de las publicaciones y de los fondos externos de cada profesor de la nación, junto con otras informaciones y, en base a esos datos, hicieron una ordenación jerárquica de todos los departamentos para influir en su grado de financiación, parece haber sido un

ejemplo en el que la celeridad fue empleada en un ejercicio público de gestión con la complacencia de los propios miembros de la comunidad académica. La publicación en la prensa de los resultados por áreas e instituciones multiplica cualquier error, dañando la reputación de algunas instituciones y programas en relación con su imagen frente a mecenases, estudiantes potenciales y empleados, creando así una espiral descendente de la cual difícilmente se recobrarán. Si los problemas son sutiles y complejos, igualmente deberán serlo los instrumentos que se emplean para abordarlos. Argumentar que incluso los indicadores de rendimiento inadecuados son mejores que las valoraciones privadas e informales que se realizan entre colegas «viejos compadres», (un sistema igual de rechazable) no es convincente. Ambos pueden ser usados de forma más eficaz.

Así, los indicadores de rendimiento se han defendido en la educación superior durante los últimos diez años porque reducen la complejidad (Frackmann, 1987) y permiten a los que son ajenos a la universidad «medir lo difícil de medir» (Cuenin, 1987) gracias al uso de valores numéricos como indicadores, de tal forma que facilita las comparaciones e implica una evaluación. Examinemos estos asuntos y otras implicaciones críticas de forma directa y detallada.

Contradicciones y fallos de la concepción actual de los indicadores de rendimiento

La compilación más exhaustiva y útil de los argumentos a favor de los indicadores de rendimiento fue hecha recientemente por Sizer (1989a). Los argumentos los proporcionan el sector gubernamental, algunos ejecutivos de las instituciones y algunos miembros del profesorado. Destaca la necesidad legítima de disponer de información, tanto general como específica, respecto a la productividad y costes de las instituciones. Pero tal sistema se queda muy corto al no reflejar la verdadera efectividad de una institución o de sus programas de estudio. Es una manera limitada, aunque relativamente barata y accesible, de medir algunos de los aspectos de una situación compleja, particularmente de aquéllos que pueden tener relativamente poco que ver con la verdadera efectividad: la capacidad para alcanzar un nivel aceptable, o mejor, la capacidad para alcanzar los objetivos expresos de la organización y para funcionar correctamente de forma continuada, a pesar de los desafíos y otras condiciones discordantes.

Cave *et al.* (1988) realizaron una revisión exhaustiva y crítica de las actuales propuestas de indicadores de rendimiento, incluyendo los 39 indicadores presentados por el Grupo de Trabajo del *Committee of Vice Chancellors and Principals* y del *University Grants Council* en 1987 (que incluía 32 medidas referidas a los costes, cuatro referentes al número y proporción entre estudiantes y plantilla, y una dedicada al tipo de empleo conseguido por los estudiantes seis meses después de graduarse) y los que se usan en otros países. La conclusión a la que se llegó fue que hay muchos problemas con el uso «crudo» de los indicadores de rendimiento y que se hace necesaria una evaluación de la calidad, incluyendo posibles visitas de colegas. Los autores presentan una lista tentativa de catorce indicadores y procedimientos, incluyendo «valor añadido» y «revisión por colegas», para las áreas docentes e investigadoras, pero aportan poco para realizar un planteamiento más equilibrado y formativo.

Estas formulaciones, y las descritas por Ball y Halwachi (1987), así como las de un buen número de grupos de trabajo en otros países, le parecen al presente analista que están fallando en diversas cuestiones importantes. Las más destacadas recibirán aquí atención.

La intención de reducir complejidad

Se debe reconocer claramente que para poder estudiar las entidades complejas a menudo se tienen que emplear variables aproximativas. Pero su uso no elimina la responsabilidad de tener muy en cuenta las cuestiones centrales si se quiere tomar una decisión sobre su carácter, funcionamiento o valor. La noción de reducir la complejidad es aceptable si tal reducción no elimina ni reduce nuestra capacidad de juzgar el verdadero valor, o si el foco de la cuantificación se sitúa sobre las materias centrales y las incluye adecuadamente. Las listas de indicadores de rendimiento que se presentan estudio tras estudio hacen poca o ninguna referencia a las intenciones (objetivos) de la organización que va a ser descrita, y virtualmente ninguna referencia a la calidad de los programas en relación con los resultados específicos docentes e investigadores. En su lugar, algunos de los indicadores hacen referencia, sin paralelismos ni explicaciones, al «valor añadido» (un concepto muy exigente y costoso, usado poco o apenas nada, que requiere de medidas detalladas durante todo el aprendizaje cognitivo, de las dimensiones afectivas o de las habilidades psicomotrices de grupos de personas) y

a la «revisión por colegas», otro proceso importante, que, por supuesto, no es un indicador de rendimiento sino la recopilación sistemática de las opiniones de los expertos externos respecto a la calidad de los programas y su funcionamiento. Volveremos sobre esto más tarde. En cualquier caso, la cuestión de la reducción de la complejidad puede fácilmente convertirse en un argumento, desafortunadamente convincente, sobre todo para que los burócratas utilicen medidas superficiales e inapropiadas, o al menos inadecuadas.

La ausencia de definiciones y enfoques adecuados

Mientras la mayoría de los observadores han comentado la necesidad de una definición adecuada de los indicadores de rendimiento, sólo unos pocos hacen una adecuada distinción entre los tres diferentes tipos de medidas. El tipo de medida del rendimiento más reivindicado es aquél que se refiere a los objetivos del gobierno. La mayor demanda de medidas del rendimiento en Inglaterra y Holanda se han dado en relación a cómo responden las instituciones a las expectativas marcadas por el gobierno en cuanto a acceso, eficiencia de costes, tiempo de estudios, producción de graduados y similares. Son estas cifras, a menudo sin relación con metas institucionales específicas, las que se demandan, tabulan y a veces ponderan y publican. Las medidas del rendimiento (niveles de logros) de los objetivos institucionales, idealmente elaborados por las propias instituciones, constituyen la segunda categoría de indicadores. Su contenido es mayor que el de las cifras buscadas por los funcionarios gubernamentales. Se emplean medidas tanto cuantitativas como cualitativas y muchas más cuestiones que las que suelen tener en cuenta los defensores de las medidas de rendimiento, ya que los objetivos de las instituciones y de los programas son variados, y a menudo no pueden someterse a métodos reduccionistas. En tercer lugar, se debe reconocer la existencia de un amplio grupo de medidas o informaciones sobre investigación institucional que son usadas en la toma de decisiones para la gestión interna de las instituciones referentes a un amplio abanico de temas. Interesa destacar que varios observadores, en su afán de apoyar el concepto de medidas de rendimiento relacionados con el gobierno, señalen a las instituciones de EE.UU. como usuarios de tales baremos, pero los ejemplos que presentan son la carga laboral del profesorado, el coste de los programas y las tendencias en la elección de asignaturas, medidas que son usadas por la administración universitaria para gestionar sus departamentos y facultades. La

gran mayoría de esta información permanece en la propia institución y no es transmitida al gobierno, ni es publicada. La cuestión del propósito, y por tanto del tipo de medida de rendimiento, es una cuestión central, y debe enfocarse según el uso que se le quiera dar.

La noción de «ranking»

Tras décadas en las que la evaluación formal de la efectividad de los programas no se consideraba pertinente en algunos países, algunos funcionarios en esos mismos países parecen ahora estar dispuestos a que los programas y las instituciones no sólo sean evaluados, sino que, además, sean ponderadas mediante indicadores de rendimiento dentro del contexto de instituciones similares. Al no tener en cuenta la complejidad de las organizaciones ni la variedad y madurez de sus objetivos, establecer un *ranking* mediante indicadores de rendimiento burdos parece inapropiado o incluso perjudicial. El peligro consiste en que una simple lista de los costes de las instituciones o de los programas, por ejemplo, no representa la calidad de los programas. No nos permite saber si se han alcanzado las metas, ni si la calidad, sea cual sea su definición, ha sido alcanzada. Desafortunadamente tales listas o *rankings* son tomados como representativos de la calidad o del valor de los programas: que un bajo coste, por ejemplo, es bueno. Las contribuciones a los costes (escala de salarios, historial laboral del grupo, efectos del mercado, intentos especiales de crear ciertas prioridades en el programa) no son tomadas en consideración. Con el uso de un indicador, que tan frecuentemente se ha propuesto, la proporción de graduados de cada especialidad que encuentra empleo, se pueden causar obvias injusticias, dado que los efectos del mercado de trabajo nacional o regionales se salen de la influencia de la universidad. Podemos ver cómo la «reducción de la complejidad» puede traspasar un nivel aceptable al enfrentar crudamente a una institución o programa contra otras, un fenómeno en sí mismo complejo que producirá malentendidos y quizá sólo pueda ser usado por los «recortadores» de presupuestos. Finalmente, la publicación de los resultados en forma de *ranking* puede dañar a las instituciones y a los programas de un modo irreparable. Medir y publicar dichos resultados no ayuda necesariamente a mejorar las instituciones, y, sin embargo, puede dañarlas. Cuando se considera además que las medidas posiblemente no sean válidas, entonces el daño es aún más grave.

Es interesante destacar que en los EE.UU., considerando el mercado educativo por excelencia, no se usan los *rankings* basados en ninguna de estas medidas para las decisiones que realmente importan, como la distribución de los presupuestos, la revisión de programas, acreditación y revisiones estatales de la calidad. Las instituciones son juzgadas según hayan alcanzado o no sus objetivos y por un conjunto de estándares generales, pero no por los indicadores de rendimiento. Los estudios sobre *rankings* por campos de conocimiento, tan a menudo citados, aunque sólo ocasionalmente realizados, no son usados por las agencias que asignan los fondos, y tales estudios incluyen sólo algunos de los programas de las 50 universidades más destacadas de entre un sistema con más de 3.300 instituciones. Mientras que en la revisión interna de las universidades a menudo se cita si los colegas piensan que el programa está entre los 5 o 10 más destacados de su área, no se usan *rankings* basados en medidas cuantitativas. Existen revisiones de programas en Norteamérica que proporcionan a cada universidad la posibilidad de evaluar sus propios programas, de forma interna y no pública, mediante estimaciones de la calidad percibida, con el objeto de reasignar sus recursos marginales, pero están muy lejos de ser listas ordenadas, por instituciones o programas, basadas en los índices de gastos u otros indicadores. La carencia del aspecto formativo (que contempla las mejoras) en los indicadores de rendimiento que se han defendido hasta la fecha es quizá su mayor debilidad. El planteamiento de estos indicadores es primordialmente de crítica y no de avance, de modo que la reacción emocional hacia los mismos es el miedo y la rabia.

El equívoco respecto a la necesidad de los sistemas competitivos y la incompreensión sobre otros sistemas

Mientras que parece claro que el movimiento a favor de los indicadores de rendimiento nació de la frustración de los funcionarios ministeriales y de los organismos de asignación de fondos durante las crisis de presupuestos de los setenta y ochenta, la impresión que se percibe de los argumentos presentados durante los últimos diez años en defensa de su uso es que los líderes de la educación superior deberían hacer lo que hace la industria (medir y considerar «lo esencial») y estar preparados para competir en un entorno regido por el mercado. Si el objetivo de usar los indicadores de rendimiento, tal como están formulados, es proporcionar tales condiciones, entonces los que los han propuesto han malinterpretado seriamente el modo en que operan las empresas con fines lucrativos, como también los procesos usados por aquellas institu-

ciones de educación superior que tienen que existir en un mercado altamente competitivo como lo es el de los EE.UU. Las empresas con fines lucrativos usan datos parecidos a los indicadores de rendimiento, pero también usan medidas detalladas del *output* al igual que medidas de eficiencia, *input* y productividad. Miden todas las dimensiones apropiadas y controlan el rendimiento objetivo a objetivo. Pero más importante todavía es que hacen uso de un proceso más amplio, formal o informalmente, que conjunta los datos sobre el rendimiento de otros hechos y opiniones recabadas desde dentro de la organización, utilizando de forma regular las opiniones de los clientes y expertos externos. Utilizan todas estas fuentes para controlar sus sistemas, para servir de base a la planificación y para introducir mejoras. Estos otros aspectos y la intención vital de mejora están ausentes en las propuestas para el uso de los indicadores de rendimiento. Aunque esas actitudes y planteamientos que están ausentes han sido propuestos por sensatos observadores en Gran Bretaña (Lockwood y Davies, 1985), los últimos diez años han visto pocos esfuerzos por llevarlos a la práctica. Son descartados o descritos como «subjetivos» y «costosos» de acuerdo con Cave *et al.* (1988). El uso de opiniones de expertos externos es la base del nuevo sistema evaluativo en Francia (Staropoli, 1987, 1989), pero el comité francés y sus empleados quieren indicadores de rendimiento en vez de usar evaluaciones internas con carácter formativo. En Suecia se está intentando la autoevaluación, pero en su sistema todavía falta la evaluación externa (Bauer, 1989). Puede que pronto se incluya apoyada en una base comparativa internacional (Nilsson, 1989). Las descripciones de la situación en los EE.UU. y en los Países Bajos, que aparecen en la literatura especializada en indicadores de rendimiento (véase Cave *et al.*, 1988), sólo son parciales y a veces desorientan.

Tal como se había indicado previamente, los indicadores de rendimiento similares a lo que se proponen y que están ganando popularidad en la Europa Occidental no son usados ampliamente en los EE.UU. Lo que se usa, o al menos lo que pretenden las agencias involucradas en estos temas, es un conjunto de procedimientos mucho más formativos, agrupados bajo diferentes denominaciones (revisión de programas, autoestudio, acreditación, cambios planificados, planificación basada en el estudio, planificación estratégica). Combinan la recopilación sistemática y el uso de datos respecto a *inputs*, funcionamiento, y cada vez más, respecto a *outputs* de importancia crítica («análisis de resultados» o estudios de cumplimiento de metas) mediante procesos cíclicos y regulares de autoevaluación que incluyen la recopilación de opiniones

y de hechos, con un proceso de análisis externo: expertos que se reúnen cada cinco años más o menos (equipos locales, equipos de revisión de programas, equipos de acreditación). Esto es muy distinto a un sistema que se limita a recopilar y considerar datos numéricos parciales relacionados con sólo algunas de las metas de un sistema y de sus costes, y que publica a continuación los *rankings* resultantes, e incluso es muy distinto de un sistema que tan sólo usa indicadores de rendimiento cuantitativos para informar a los expertos que deben realizar la revisión.

En Holanda algunos agentes del gobierno han mostrado un interés continuado por los indicadores de rendimiento, particularmente por los que contrastan el rendimiento de las metas gubernamentales, y de hecho, tal como sucede en EE.UU. y Canadá, se usan datos de costes y medidas de productividad en determinados procesos de adjudicación de fondos, conjuntamente con información sobre la calidad de los programas y otros datos. Pero se está produciendo un gran empuje para que en la evaluación de los programas (en las universidades) y de las instituciones globalmente (en los politécnicos) se incluyan las autoevaluaciones sistemáticas, cíclicas y formativas que incluyen visitas regulares a los centros por colegas de la misma especialidad, así como por meros ciudadanos (Kells van Vught, 1988). Hasta hace poco esta tendencia ha estado ausente en la literatura sobre los indicadores de rendimiento holandeses. Respecto a las opiniones útiles y recientes véase Bormans *et al.* (1990).

La supuesta necesidad de limitar los juicios de opinión

En su discurso inaugural en un congreso sobre indicadores de rendimiento, Sizer (1989a) presentó los argumentos a favor del uso de los indicadores de rendimiento haciendo énfasis en dos motivaciones. La primera es la necesidad de que las instituciones justifiquen detalladamente al gobierno el uso eficiente y efectivo de los recursos y de que entren en competencia con el resto de las instituciones en cuanto a conseguir estos recursos. El segundo argumento de Sizer fue que los indicadores de rendimiento más detallados y exhaustivos «tienden un puente sobre el vacío», de forma que se podría depender menos de las opiniones en estos asuntos. Ante la complejidad y diversidad de metas, variables de *input*, contexto y proceso, así como la naturaleza de los *outputs* en las organizaciones de educación superior, resulta difícil creer

que la extensión en el uso de valores numéricos y otros indicadores vaya a proporcionar mejores evaluaciones, ya que no se considera la única cosa que se necesita: un mayor uso de juicios por colegas en autoevaluaciones dentro de los propios departamentos y facultades, y una mayor validación externa también realizada por colegas. Los indicadores de rendimiento pueden a veces ayudar a hacer tales valoraciones, pero la comprensión detallada de las metas, los procesos, el contexto y el funcionamiento de cada conjunto es la única forma de favorecer la toma de responsabilidades y de mejorar las universidades. Los indicadores de rendimiento por sí solos no pueden proporcionar el nivel de comprensión que requieren estas materias complejas. La necesidad que parece existir de reducir la complejidad, más allá de todo límite razonable, es un problema constante, y debe ser resistida por aquéllos que se preocupan por las instituciones.

Una formulación más global y formativa

Cada vez está más claro que a la educación superior, al igual que a otros sectores públicos (hospitales, gobiernos locales, etc.), se le exigirá que explique su rendimiento y justifique su parte en el consumo de los fondos públicos disponibles. De forma similar, tal como ya han destacado muchos líderes políticos y observadores, se les exigirá a las universidades una actitud más efectiva, eficiente y de iniciativa propia, al igual que a los politécnicos o los institutos, si es que quieren prosperar en el ambiente de los noventa. Tanto la explicación como el desarrollo creativo y la puesta en práctica de los programas y el uso de los recursos necesitan ser evaluados. La cuestión es qué tipo de planteamiento debería usarse en dichas evaluaciones y qué beneficios deberían producir las mismas. El presente observador no tiene la más mínima duda de que un proceso formativo, ajustado a la naturaleza de la empresa, se impone. Las propuestas respecto a los indicadores de rendimiento tan sólo satisfacen una pequeña parte de dichas necesidades, y podría argumentarse incluso que, ante la ausencia de un mejor planteamiento que está más orientado hacia las mejorías, el uso de indicadores de rendimiento puede infligir serios daños al sistema. El afán de reducir la complejidad hace que se ignoren las intenciones, la riqueza y, de hecho, parte de la esencia misma del esfuerzo organizativo. Mediante los procedimientos de *ranking* los anteriores planteamientos minimalistas invitan a comparaciones de valores y juicios inapropiados sobre la finan-

ciación, que aunque sean convenientes, pueden llegar a ser tremendamente erróneos. En vez de formar parte de un estudio continuado orientado hacia las mejoras, dentro de un esquema de planificación que utilice los datos y opiniones procedentes de ficheros, encuestas, empleados, clientes y expertos, el planteamiento de los indicadores de rendimiento, con razonamientos erróneos, evita los juicios «subjetivos», como si los juicios en sí no formasen parte de la evaluación. Evita buenos métodos de análisis porque se les supone un coste elevado. Este último aspecto es la actitud del «sabio con el penique, tonto con la libra» que a menudo se aplica en la evaluación sistemática y en la planificación. Enormes sumas se gastan en construir y mantener los sistemas de educación superior, pero el compromiso de gastar una pequeña fracción del presupuesto anual para evaluar los programas o servicios proporcionados por el propio sistema, de forma que pudieran mejorarse y rendir cuentas públicamente, es muy escaso.

La propuesta

Si queremos cumplir con las exigencias legítimas de los funcionarios públicos respecto a la información sobre el rendimiento, si queremos ir más allá del presente planteamiento de los indicadores para incluir el núcleo mismo del aprendizaje y de la investigación, y si queremos reconducir el sistema desde una cuantificación reduccionista, con su proteccionismo resultante, hacia un enfoque abierto y formativo, varias medidas deberían ser tomadas:

- Se debe mantener y ampliar lo mejor del esquema de los indicadores de rendimiento, la voluntad de examinar, de cuantificar donde sea apropiado, y de explicar.
- Se deben desarrollar e instaurar estudios y revisiones continuas de las instituciones (autoevaluación interna y visitas externas) que involucren a profesionales en el proceso de mejorar su organización, que generen los datos y opiniones necesarios y también la propiedad psicológica de las recomendaciones resultantes (véase Kells, 1988, para un tratamiento más exhaustivo de tales procesos).
- Se deben enfocar tanto las necesidades del gobierno y sus medidas como la forma, proceso, estado de salud y logros (vía análi-

sis objetivo a objetivo) de las instituciones individuales, así como de sus programas de estudio. Esto significa que tanto los ministerios como los organismos de financiación deben querer, y deben disponer de más oportunidades para examinar las instituciones individualmente de forma que sean más concordes con la naturaleza y el estatus de las organizaciones involucradas.

- Se debe usar tanto el estudio interno como la validación externa (por colegas) en ciclos regulares y continuos.
- Se debe usar un sistema que pueda formar parte de la gestión (animándola y no distrayéndola) de cada institución, así como de la toma de decisiones. Un sistema que incluya a muchos de los mejores profesionales en la evaluación sistemática, el análisis creativo y la toma de decisiones (planificación) será más fuerte, más capaz de explicarse a los demás y de competir bien, desarrollarse y al mismo tiempo mantenerse.

Para ilustrar los beneficios que pueden producirse en los esquemas de control y de gestión si se aplicara tal sistema presentamos la Tabla 1. De forma resumida y sumaria presenta el tipo de información que podría proporcionar a los comités ejecutivos en toda la institución interesadas y, bajo ciertas condiciones, a funcionarios ministeriales y organismos de financiación. Los ejecutivos de la institución, comités de revisión y responsables de la institución recibirían, por supuesto, los informes completos durante las fases de la evaluación interna y externa. Invitamos aquí al lector a que se fije particularmente en el ítem II de la Tabla 1, en el que se desarrollan los pasos específicos para la evaluación que se sugiere. La información enumerada es tan sólo un extracto de materiales seleccionados, pero es una mezcla rica y global de información cuantitativa y cualitativa basada en los datos y opiniones proporcionados por estudiantes, informes oficiales, bases de datos, plantilla, clientes y expertos invitados. En fuerte contraste, un informe construido siguiendo las recomendaciones de los grupos de trabajo recientes sobre indicadores de rendimiento tan sólo porcionará una parte de la información de la sección III de la Tabla 1.

El informe más útil desarrollado en la Tabla 1 no es, sin embargo, el mayor beneficio del sistema. Lo mejor es que la toma de decisiones respecto a programas complejos y a mejoras sistemáticas en la organización y en los cambios planificados son ayudados por estos procesos.

Tabla 1. Ejemplo de un informe sobre el rendimiento de un programa, disponible para uso institucional y externo

Las Condiciones, efectividad y directivas futuras del programa Bio-tecnológico del Politécnico

I. Entorno y Estatus General

El programa tiene 20 años de vigencia, un tamaño estable (10 profesores de tecnología a tiempo completo; 120 estudiantes a tiempo completo en tres sub-especialidades) y un elemento central (uno de cinco programas técnicos) dentro de la respuesta institucional a las necesidades de investigación aplicada y de formación ambiental e industrial en la región. El 80 % de su financiación proviene del *Funding Council*, y la cantidad restante procede de ayudas y contratos externos. Las instalaciones ocupan 5 laboratorios docentes, 6 salas para servicios y equipamiento, y 12 despachos/laboratorios. Los graduados se convierten en técnicos en los laboratorios industriales y gubernamentales a nivel regional y nacional. El plan de estudios, recientemente reformado, cumple con los estándares actuales de su área según los informes de un equipo revisor.

II. Procedimiento de revisión

El programa incluye la evaluación formativa realizada por los estudiantes y el cuerpo docente cada semestre. Cada año los estudiantes evalúan a cada profesor (con respecto al cumplimiento de los objetivos del curso). Los resultados son estudiados por cada profesor y por el colectivo de la facultad durante unas sesiones anuales de revisión del curso. Los profesores son evaluados también en un proceso de cumplimiento con los objetivos anuales alcanzados que incluye una autoevaluación de los propios objetivos, la investigación realizada y los servicios prestados en base a lo que anualmente ha sido negociado, al mismo tiempo que son revisados por un colega o mentor, además de la revisión que hace el director del programa.

La evaluación continua y periódica del programa se realiza de dos formas. Cada cinco años el conjunto de la plantilla realiza una autoevaluación desde la perspectiva de los objetivos de los programas y los estándares del sector. Se recopilan datos, se recaban las opiniones de los estudiantes y de los profesores, se recibe la visita de un consultor de bio-tecnología y se recopilan opiniones de licenciados y empresarios para analizar el conjunto. Cualquier cambio necesario es propuesto y ejecutado o se planifica como parte de una discusión con el equipo de revisión externo enviado por el *Council of Polytechnics*. El informe de autoevaluación y el informe de los visitantes externos son revisados por la plantilla del programa, así como por el Comité de Revisión del Politécnico de forma que pueda formularse el Plan de Desarrollo revisado para el programa. Los resultados también son empleados para elaborar el presupuesto anual y los ciclos de planificación. Este proceso de revisión interna y externa fue realizado hace dos años.

III. *Medidas generales de rendimiento*
(según las requiere el *Funding Council*)

Estudiantes Nuevos: cualificaciones muy competitivas; gran demanda para ser admitido. Uno de cada tres solicitantes es admitido. Demanda estable a lo largo de los cuatro últimos años. Cualificación media de admisión en el cuarto superior de los admitidos en los programas de bio-tecnología del país.

Resultados de los titulados: una media de 18 graduados por año en los últimos 3 años. Todos aquéllos que quieren encuentran empleo en los 3 primeros meses en el campo de la bio-tecnología. Relación estudiantes/profesor: 11,6/1 sobre una base comparable de dedicación plena. Muy próximo a la media nacional de 11,2/1 para estos programas.

Tasa de finalización de los estudios: los estudiantes concentrados en las tres subespecialidades (fermentación, tecnología del medio ambiente y sistemas de salud) completan los estudios en un 88 % en los tres últimos años. Los márgenes son entre el 83 % al 91 % para las tres especialidades.

Valoración del aprendizaje: los graduados del programa alcanzan sistemáticamente puntuaciones sobre la gama del 90 %, o superiores, en el examen nacional de teoría y práctica para bio-tecnólogos, haciendo que el programa se sitúe de forma continua entre los 3 programas más destacados del país.

Medidas de costes: ratio X: 9287 - ratio Z (*): 11.431.
ratio Y: 16,3 - ratio Q: 7,1.

Nota (*): la ratio Z es alta para la media nacional porque seis de los diez profesores forman parte de la plantilla desde que se inició el programa. Dos de ellos se retirarán el año que viene. Es una plantilla fuerte y competente, altamente estimada en toda la región.

Medidas sobre la investigación: estudiantes de investigación publicaron conjuntamente con miembros de la plantilla seis artículos. El profesorado produjo 21 artículos o ponencias para encuentros regionales o nacionales y 3 para congresos internacionales (18 fueron para revistas referenciadas). Se publicaron dos libros, uno de ellos ganó el prestigioso premio *Bio-Tech*. Se concluyeron dos proyectos subvencionados y para otro se prorrogó la subvención por tres años más. Se aseguraron dos nuevos contratos. Los fondos de fuentes externas para investigación crecieron un 11 por ciento este año.

IV. *Calidad Global y Futuras Directrices*

Este programa está altamente reconocido en toda la nación. Los estudiantes desean participar en el programa y los graduados son buscados por las empresas que sistemáticamente elogian a sus empleados en las evaluaciones. Los miembros del equipo de revisión externa de los últimos cinco años examinaron los logros del programa objetivo por objetivo, revisaron los exámenes de

los estudiantes, leyeron los informes y trabajos de los estudiantes y revisaron la acreditación y los trabajos académicos recientes de la plantilla. El informe del equipo fue altamente favorable y el resultado fue una concesión de mayores fondos discrecionales por parte del Comité de Revisión Interna y de la Oficina del Vicescanciller. Se proporcionaron mayores fondos para proveer de equipamiento necesario al laboratorio de sistemas de salud y para becas de investigación de los estudiantes.

El programa de bio-tecnología está formulando una propuesta para la creación de una cuarta subespecialidad en tecnología de la alimentación basada en un estudio de mercado cuya elaboración ha durado un año. Se están negociando los niveles y las asignaciones a las subespecialidades para la sustitución de los dos miembros de la plantilla que se van a retirar.

El resultado son instituciones más fuertes, con una moral más alta y contando con mayor confianza pública. La combinación de información estadística, u otra información numérica, con juicios cualitativos, proporcionan a quien toma las decisiones en el nivel local, así como a nivel de todo el sistema, una información muy útil cuando hay que tomar decisiones difíciles.

Recientemente se han elaborado sistemas con estos elementos formativos y responsables en Holanda (véase Ackerman, 1988a; Kells y van Vught, 1988), en parte del Canadá (véase British Columbia, 1984; Kells, 1988, y Kells y van Vught, 1988a) y en los EE.UU. (véase Young *et al.*, 1983; Kells, 1988). También han sido elaborados parcialmente en otras partes de Europa Occidental. Para una discusión más centrada y específica véanse las reseñas y propuestas hechas recientemente en Kells (1989b). Cada país debe elaborar su propio sistema de estudio, revisión y planificación para servir a su particular abanico de instituciones y necesidades a la luz de su clima histórico y sociopolítico. La Columbia Británica tiene una gama de prácticas y necesidades un tanto distinta a las de los Países Bajos o EE.UU. Ciertamente el gran reto consiste en motivar a los líderes, tanto nacionales como universitarios u otros, a que consideren las necesidades, beneficios, riesgos, costes y planteamientos posibles en su entorno individual. De no hacerse esto, nuestro sistema de educación no estará óptimamente preparado para prosperar en los noventa, y puede que sea más vulnerable a los daños de los bien intencionados, pero extremadamente limitados, esquemas de control que probablemente debilitarán más todavía nuestro ya castigado sistema.

Es importante reconocer que los indicadores de rendimiento ocupan un lugar en tal sistema. Pueden ser usados de varias formas en la autoevaluación de programas e instituciones para señalar la necesidad de vigilar el nivel de rendimiento o el de costes. Ciertamente pueden ser usados por las agencias gubernamentales si son completados por el conjunto de opiniones institucionales y de expertos respecto a los logros de las metas centrales de la universidad. Para la adjudicación de recursos importantes se deberían considerar ambos tipos de información. Para lograr esto los esquemas de control deben adaptarse de forma constructiva.

Agradecimientos

La idea básica para el presente artículo, escrito a finales de 1989, fue sugerida por Lene Büchert, quien, además, proporcionó muchas sugerencias valiosas. El manuscrito fue leído en galeras por John Davies, Maurice Kogan, Peter Maassen, John Sizer, Vernon Trafford y Frans van Vught. Me hicieron muchas sugerencias valiosas, pero cualquier error en los datos o en los juicios es sólo mío.

EVALUACION DE LA INVESTIGACION CIENTIFICA. UN METODO COMBINADO DE ANALISIS DEL PERFIL *

J. P. DE GREVE
Universidad Libre de Bruselas, Bélgica

A. FRIJDAL
Istituto Universitario Europeo, Italia

Introducción

Las universidades de toda Europa hacen frente en estos momentos a problemas presupuestarios. Además, la política científica, los consejos de investigación, las fundaciones y agencias que proporcionan los fondos se enfrentan al problema de seleccionar los proyectos que más o menos garanticen un resultado cuyo valor científico compense el esfuerzo financiero. Por la misma razón las universidades tienen necesidad de disponer de políticas claras e indicadores de prioridad para desarrollar sus programas de investigación a medio y largo plazo. Mantener el alto nivel de calidad de la investigación será decisivo para el papel que jugará en la próxima década la investigación dentro de las universidades.

La diferenciación de los recursos financieros para los programas de investigación es un hecho. Pero la falta de sincronización entre los diferentes mecanismos de selección conduce a una situación en la que los científicos parecen gastar más tiempo en conseguir fondos o en rellenar solicitudes, que en la propia investigación. El hecho de que tanto tiempo sea dedicado a conseguir fondos debería ser un estímulo para garantizar una regla de asignación de dotaciones correctas y sin asperezas.

* Este artículo apareció originalmente en *Higher education Management*, 1, 1, Marzo 1990.

La evaluación de la investigación científica que se realiza en las universidades e institutos de investigación debe perseguir la calidad y no la posibilidad de resultados útiles. Nosotros queremos desarrollar elementos que clarifiquen los procedimientos de evaluación, que son demasiado a menudo considerados como una «caja negra», y que tiendan a desglosar los procesos de toma de decisiones en los factores que juegan un papel decisivo en la estimación final.

Este proceso tiene lugar de un modo continuo en los diferentes niveles de la actividad académica (Gibbons, 1984):

- Admisión de artículos en revistas con evaluadores.
- Invitaciones a reuniones científicas.
- Nombramientos científicos.
- Premios científicos.
- Concesión de fondos para investigación.
- Obtención de grados de doctor.

El proceso de evaluación

Estrictamente hablando el proceso de evaluación es la fase previa a la toma de decisiones. El proceso de evaluación debe objetivizar los elementos de los casos en los que hay que tomar decisiones. El caso clásico es aquel en el que hay que seleccionar entre varias propuestas dentro de unas limitaciones dadas (dinero, personal, equipamiento). Los criterios de calidad son ubicados dentro de un marco referencial dado. En tomas más tradicionales suele disponerse de mayor número de evidencias para una toma de decisiones correcta; en áreas más de moda la calidad no está tan garantizada.

Ex ante y ex post

La valoración de un caso puede ser desglosada en dos partes: la evaluación de las actividades pasadas (*ex post*) y la de las actividades que se proyectan (*ex ante*). Los juicios sobre las actividades pasadas

pueden basarse básicamente en elementos cuantitativos. La valoración de actividades futuras tendrán que confiar fundamentalmente en los juicios de otros colegas. El equilibrio entre los juicios sobre el pasado y sobre el futuro son decisivos para las perspectivas de la comunidad científica: estáticas, dinámicas o catastróficas.

La evaluación *ex ante* valora un proyecto, grupo o persona con independencia de sus méritos futuros, basándose en el pasado. Como tal es a menudo una proyección *ceteris paribus* en el futuro. En el caso de la evaluación de un científico (para su futura carrera profesional) el elemento básico es la actividad científica en el pasado, a menudo (pero no siempre) acompañada de una propuesta de un programa de investigación. Para la asignación de fondos a un proyecto de investigación el énfasis de la evaluación debe ponerse en el valor científico del futuro proyecto, aunque la capacidad (y, por tanto, las actividades pasadas) también juega un papel importante. Debe ser remarcado que la evaluación *ex ante* consume mucho tiempo y muchos esfuerzos humanos. El examen de las propuestas de investigación, de sus motivaciones o del tipo de equipo que se necesita, y el conseguir las recomendaciones de expertos y colegas demandan un importante esfuerzo administrativo y mucho tiempo.

La evaluación *ex post* sólo juzga el pasado. Si se hace a intervalos regulares y de acuerdo con un nuevo método que permita un uso generalizado, la matriz de rendimiento, pensamos que la evaluación *ex ante* puede ser mejorada.

El sistema de la «revisión por colegas», evaluación de las actividades o planes de los científicos por otros científicos, comúnmente por aquéllos que trabajan en el mismo campo, es uno de los métodos más ampliamente utilizado para la evaluación. También es claramente el más valioso, ya que permite poner el énfasis en la calidad de la investigación. De todos modos, el método también muestra algunas debilidades.

Una de ellas es la elección del evaluador, ya que conocimientos, disponibilidad y racionalidad son demandas a menudo contradictorias. Aspectos tales como el idioma o los aspectos locales de la investigación pueden estrechar el posible círculo de expertos, creándose dificultades por falta de independencia o por la inexistencia de expertos.

Cole, Cole y Simon (1981) investigaron el sistema de revisión por colegas en los Estados Unidos y encontraron que la evaluación de las

propuestas de investigación estaba determinada en una mitad por la calidad de la solicitud y en la otra mitad por factores aleatorios tales como «la suerte en el sorteo del evaluador». Gibbons (1984) afirma que el sistema funciona cuando se satisfacen los siguientes criterios:

- a) La evaluación tiene lugar en disciplinas científicas bien establecidas.
- b) El área tiene unos límites claros.
- c) Existe suficiente dinero disponible.

El sistema funciona peor bajo las siguientes condiciones:

- a) La investigación esta orientada hacia necesidades económicas y sociales.
- b) Se trata de investigaciones multidisciplinarias.
- c) Se evalúa el futuro en vez del pasado.

Usualmente el sistema de revisión por colegas no ofrece mucha información sobre los criterios del juicio realizado, aunque esto puede diferir de un país a otro. Precisamente esta información puede ser muy instructiva para el solicitante cuyo trabajo o proyecto es evaluado, así como para los que toman las decisiones, ya que pueden recibir una información complementaria muy útil de la «defensa» que puede hacer el solicitante (le Pair, 1980). Nosotros creemos que la ausencia de información sobre los criterios utilizados en la evaluación es el resultado global de factores subjetivos utilizados (calidad, originalidad, etc.; véase más adelante «Características de la Evaluación»).

La evaluación *ex post* puede ser nutrida mediante cuantificaciones o medidas de la actividad científica. El acento se pone aquí en los datos bibliométricos. Se examinan a la vez las actividades de comunicación (número de publicaciones, peso de las revistas, tipo de comunicación: libro, artículo o contribución a un congreso) y el impacto (citaciones de las publicaciones por otros autores).

Existe una rica y voluminosa literatura sobre el valor y el peligro del análisis de datos bibliométricos. Otros parámetros relacionados con el impacto son el número de invitaciones a reuniones científicas, el número de investigadores invitados en el laboratorio... A nivel gubernamental un

análisis global de los datos bibliométricos de disciplinas específicas puede proporcionar una perspectiva de la estructura y del potencial de partes específicas de la investigación dentro de un país (van Heeringen, Mommers y van Venetië, 1983).

El tempestuoso éxito de índice de citas de Garfield (*Scientific Citation Index*) fue seguido por un gran debate crítico. Actualmente son bien conocidos los puntos débiles del uso del índice de citas gracias a los trabajos de Price, Garfield, Irvine, Martin, Roy, Rustome y otros. Los aspectos negativos del uso de parámetros cuantitativos son resumidos por Martin e Irvine (1984; ver también Narin, 1978):

a) El número de publicaciones:

1. La contribución a la ciencia difiere de unos artículos a otros.
2. Existen diferencias entre disciplinas con respecto a la actitud respecto a las publicaciones (Schiller, 1981; Roy *et al.*, 1983; Frijdal y De Greve, 1986).

b) El análisis de las citas:

1. Límites técnicos (sólo primer autor, nombres idénticos, variaciones en el nombre...).
2. Variación de la tasa de publicaciones a lo largo de la vida.
3. Citaciones críticas.
4. Efecto «halo».
5. Variación entre diferentes disciplinas.
6. Autocitación.

Los nuevos estudios sobre la utilización del índice de citas tienen en cuenta la mayor parte de estos efectos. De acuerdo con Moed *et al.* (1985) los datos bibliométricos proporcionan un mecanismo de control para la política científica y para la gestión de la investigación universitaria. Pueden también ser usados para examinar colaboraciones a gran escala y trabajos de equipo en campos específicos como la física (Beaver, 1986).

Características de la evaluación

Se debe tener en cuenta que la evaluación de la investigación científica no está sólo caracterizada por elementos directamente relacionados con el material que es evaluado, sino con el proceso de evaluación en sí mismo.

El método y el número de parámetros utilizados en el proceso de evaluación dependen del nivel de agregación. Diferentes niveles son:

- Proyecto, programa, política científica global.
- Científico individual, equipo de investigación, departamento, instituto, disciplina.

Algunos ejemplos pueden ilustrar los problemas implicados. La evaluación por parte de colegas en pequeños países o en grandes investigaciones internacionales (CERN, programas oceanográficos) puede tener el problema de encontrar evaluadores independientes. A nivel de departamento es posible introducir parámetros adicionales en la evaluación, tales como científicos invitados, tesis doctorales, cursos de postgrado, etc. Para casos individuales es importante evaluar la influencia personal en el rendimiento del grupo.

Como hemos mencionado anteriormente es interesante para el solicitante conocer las razones por las cuales una propuesta de investigación es rechazada. Si el rechazo está motivado por los puntos débiles del proyecto, el solicitante puede modificarlo y adaptar el proyecto de acuerdo con las líneas que motivaron el rechazo. Sin embargo, frecuentemente el anonimato de los evaluadores es considerado esencial para mantener el valor del sistema de revisiones por colegas, por lo que este tipo de retroalimentación no la permiten los procedimientos usados en algunos países.

Sin embargo, una salida puede encontrarse mediante el uso de un perfil de calidad del proyecto, como resultado de una evaluación realizada por una comisión de colegas. Este perfil refleja la opinión de los expertos de un modo transparente, pero su presentación en forma de grafo asegura que ni el estilo ni las razones específicas pueden ser atribuidas a ninguno de los expertos en particular.

El método mixto

En vista de las desventajas de los algoritmos de evaluación discutidos anteriormente, una solución puede encontrarse en una combinación de técnicas diferentes. En consecuencia, queremos introducir un método mixto basado en:

- La matriz de rendimiento.
- El perfil del proyecto.

A continuación damos una lista de posibles criterios, cuantitativos así como cualitativos. Cada lista está estructurada de un modo diferente.

La matriz de rendimiento

El rendimiento *ex post* puede ser reconstruido dentro de su propio contexto a través de una matriz cuyos parámetros cuantitativos pueden ser:

- La actitud comunicadora: trabajos, artículos, índice de citaciones, informes, conferencias invitadas.
- Científicos invitados en el laboratorio o el instituto.
- Número de tesis doctorales por año.
- Número de tesis doctorales por investigador.
- Número de estudiantes de postgrado.
- Premios nacionales e internacionales.
- Patentes.
- Contactos industriales.
- Relación entre fondos internos y externos.
- Relación entre dinero externo e investigadores ETC (ETC = Equivalente a Tiempo Completo).
- Estructura de edad del grupo de investigación.
- Pertenencia a comités científicos.

- Becas.
- Visitas de media y larga duración a otros laboratorios.

Un número creciente de países y universidades están recogiendo esta clase de información (Los Países Bajos, OCDE, Universidad Libre de Bruselas...). Todos esos parámetros deben ser considerados en períodos de unos tres años, para evitar grandes fluctuaciones. También deben ser normalizados por las actitudes generales en las diferentes disciplinas (ver, por ejemplo, Frijdal y De Greves, 1986, para la actitud comunicadora en Bélgica). Para revisar las cantidades medias nosotros recomendamos la utilización de una matriz de rendimiento P_{ij} (Frijdal, 1985, ver también el Gráfico 1) que represente la actividad científica. Las columnas j son las diferentes disciplinas (o departamentos, etc., dependiendo de la aplicación), las filas i representan las actividades que pueden ser cuantificadas (véase arriba). La construcción de estas

Gráfico 1

MATRIZ INDICADORA DEL RENDIMIENTO DE LA INVESTIGACION CIENTIFICA

P_{ij}	P_{i1}	P_{i2}	P_{i3}	•	P_{in}
P_{1j}	f_{11}	f_{12}	•	•	f_{1n}
P_{2j}	f_{21}	•	•	•	•
P_{3j}	f_{31}	•	•	•	•
•					•
P_{mj}	f_{m1}	•	•	•	f_{mn}

Columnas: disciplinas.

Filas: actividades cuantificadas tales como publicaciones (subdivididas), participación en reuniones, presentaciones, pertenencias a comités,...

f_{ij} : frecuencia media de una actividad específica i para una disciplina j .

matrices puede ser hecha *ad hoc*, basándose en los elementos disponibles. Esto permite clasificar cada caso, dentro de su entorno científico, sobre la base de su rendimiento, evitando el problema de dar pesos a los diferentes parámetros. La calificación de un caso a través de la confrontación con diversos parámetros permite un buen juicio objetivo del rendimiento, incluso entre disciplinas diferentes.

En el artículo mencionado arriba hemos desarrollado datos que tienen en cuenta esta clasificación. El creciente uso de análisis cuantitativos en varios países hace posible la construcción de estas matrices de rendimiento sobre bases *ad hoc*.

El perfil de la propuesta de investigación

El juicio basado en elementos cualitativos depende mucho de la intuición y el conocimiento de un evaluador experimentado. Sin embargo, este proceso de valoración progresa mediante una serie de criterios que pueden ser descompuestos en diferentes aspectos. No siendo necesario que se le dé un peso interactivo a esos criterios. Una sinopsis analítica en forma de grafo de las puntuaciones de los diferentes aspectos relevantes del dossier promueve la objetividad (Gráfico 2). También permite al solicitante tener una perspectiva de las críticas formuladas a su solicitud. Una confrontación entre los perfiles de los distintos expertos también ayuda a eliminar los puntos débiles del sistema que son posibles en comunidades científicas pequeñas (una universidad, un pequeño país). Nosotros recomendamos, por tanto, un juicio gráfico de cada dossier por dos o tres evaluadores antes de la reunión de la comisión de colegas. Un sistema semejante, con dos evaluadores, puede ser fácilmente desarrollado por las revistas científicas.

Proponemos los siguientes aspectos a tener en cuenta en la valoración:

1. Definición de los objetivos, finalidad de la investigación.
2. Relación con «el estado del arte».
3. Motivación (el porqué) de la investigación.
4. Metodología, hipótesis de trabajo y planificación.

Gráfico 2

PERFIL DE CALIDAD DE UN PROYECTO CIENTIFICO POR UN EVALUADOR O POR UNA COMISION DE COLEGAS

Dosier:		Fecha:				
PERFIL						
	A	B	C	D	E	
1. Finalidad	●	●		●	●	
2. Estatus de la investigación	●		●	●	●	
3. Motivación	●	●		●	●	
4. Metodología/Plan de trabajo	●	●		●	●	
5. Infraestructura existente	●	●	●	●		
6. Necesidades	●	●	●	●		
7. Factibilidad	●	●		●		
8. Originalidad		●	●	●	●	
9. Impacto	●		●	●	●	

A = excelente; B = muy bueno; C = bueno; D = suficiente; E = insuficiente.

- 5. Infraestructura existente.
- 6. Presupuesto.
- 7. Factibilidad.
- 8. Originalidad.
- 9. Impacto sobre los conocimientos existentes.

Estos nueve elementos se valorarán en una escala de 1 a 5: sobresaliente, muy bueno, bueno, suficiente e insuficiente.

La evaluación de las propuestas de investigación y de los proyectos futuros dependerá de los criterios subjetivos/cualitativos como la originalidad y la factibilidad, mientras que el análisis cuantitativo servirá como base fundamental. Es necesario un equilibrio entre la información sobre la investigación en el pasado y sobre la que hay en perspectiva. De todos modos, queremos resaltar que en el pasado se ha dado demasiado peso al rendimiento. A menudo esto ha impedido el nacimiento de nuevas investigaciones como resultado, o basadas, en actividades interdisciplinarias.

Sumario y conclusiones

Un método de evaluación mixto ha sido presentado, usando un análisis cuantitativo y cualitativo del dossier de la investigación. El método tiene tres características: transparencia, la posibilidad de retroalimentación y, a través de la matriz de rendimientos, una interpretación mejor y más objetiva del pasado y un enfoque más sencillo en el caso de interdisciplinaria. Puede ser usado para diferentes propósitos, desde la evaluación de proyectos a gran escala en un país, la adquisición de equipos costosos por una universidad, a la nominación de científicos para puestos de importancia académica. El método es especialmente valioso para la evaluación en pequeñas entidades.

Referencias Bibliográficas

- ABBOT, W.F. (1972): «University and departmental determinants of prestige of sociology departments». *American Sociologist*, 7, Nov, 14-15.
- ABBOT, W.F.; 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. (1988): *Quality Assessment by Peer Review*. VSNU, Utrecht.
- ACHERMAN, H.A. (1988a): «Quality Assessment in a European Perspective». Ponencia presentada en el *10th European Forum of the EAIR*, Bergen, Noruega.
- ACHERMAN, H.A. (1990): «Quality assessment by peer review». *Higher Education Management*, 2, pp. 179-192.
- ADLER, M.J. (1986): *Manifiesto Educativo*. Madrid. Narcea.
- ALCHIAN, A. y DEMSETZ, H. (1972): «Production, information costs, and economic organization». *American Economic Review*, 62, pp. 777-795.
- ALLISON, G.T. (1971): *Essence of Decision*. Boston: Litle, Brown and Co.
- ALONSO, J. (1984): «Educación-Empleo: Datos para una adecuación necesaria». *Revista de Educación*, 273, 129-137.
- ALBACH, PH.G. (1986): «Una nación en peligro: el debate sobre la reforma de la educación en los Estados Unidos». *Perspectivas*, XVI (3), 359-371.
- ARCHER, M.S. (1981): «Educational politics: A model for their analysis», en Broadfoot, P., Brock, C. y Tulasiewicz, W. (eds.): *Politics and Educational Change: An International Survey*, Croom Helm, Londres.
- ARNS, R.G., POLAND, W. (1980): «Changing the university through program review». *Journal of Higher Education*, vol. 51 (3), pp. 268-284.
- ASTIN, A.W. (1962): «Productivity of undergraduate institutions». *Science*, 136, pp. 129-135.
- ASTIN, A.W. (1985): *Achieving Educational Excellence*. Jossey-Bass, San Francisco.
- ASTIN, A.W. y HENSON, J.W. (1977): «New measures of college selectivity». *Research in Higher Education*, 6, pp. 1-8.

- ASTIN, A.W.; SOLMON, L.C. (1979): «Measuring academic quality: An interim report». *Change*, 11, 48-51.
- ASTIN, A.W.; SOLMON, L.C. (1981): «Are reputational ratings needed to measure quality?». *Change*, Oct, 14-19.
- ASTIN, A.W., BLADE, J.H., BOWEN, H.R., GAMSON, Z.F., HODGKINSON, H.L., LEE, B. & MORTIME, L.P. (1984): *Involvement in learning: Realizing the potential of American higher education*. Washington, D.C. National Institute of Education.
- ATKINS, P.W. (1984): *The Second Law*. American Scientific Books, Nueva York.
- AXELSON, L.J. (1960): «Graduate schools and the productivity of their graduates». *American Journal of Sociology*, 66, September, pp. 171-175.
- BAKER, Rt. Hon.K. (1988): *Secretary of State's Speech to CVCP*. 28 September.
- BALDRIDGE, J.V. (1971): *Power and Conflict in the University*. Wiley, Nueva York.
- BALL, R.; HALWACHI, J. (1987): «Performance indicators in higher education». *Higher Education*, 16, 393-405.
- BARAK, R.J. (1982): *Program Review in Higher Education: Within and Without*. National Center for Higher Education Management Systems, Boulder (Colorado).
- BARAK, R.J. (1986): «The role of program review in strategic planning». *AIR Professional File*, vol. 26, pp. 1-7, Association for Institutional Research, Tallahassee.
- BARAK, R.J. y BERDALL, R.O. (1978): *State Level Academic Program Review in Higher Education*. Education Commission of the States, Denver (Colorado).
- BARNETT, R.A. (1989a): «Quality control and the development of teaching and learning», en McVicar, M. (ed.): *Performance Indicators and Quality Control in Higher Education*. Portsmouth Polytechnic, Portsmouth.
- BARRY, B. (1976): «Power: An economic analysis», en Barry, B. (ed.): *Power and Political Theory: Some European Perspectives*. Wiley, Londres.
- 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.
- BEAVER, D. de B. (1986): «Collaboration and teamwork in physics». *Czechoslovak Journal of Physics*, B36, 14.
- BECHER, R.A. (1989): «Educación superior: Estructura de los sistemas», en Hussen, T. y Postletwaite, T.: *Enciclopedia Internacional de Educación*, Tomo IV, Vicens-Vives, Barcelona.
- BECHER, T. Y KOGAN, M. (1988): *Reform of British Higher Education*. Education Reform Group, Londres.

- BELANGER, Ch. (1990): «Faculty performance and workload inequity». *Higher Education Management*, vol. 2, 2, pp. 115-122.
- BELANGER, Ch. y TREMBLAY, L. (1982): «A methodological approach to selective cut-backs». *Canadian Journal of Higher Education*, vol. 12, pp. 25-35.
- BENOIST, H. (1986): «Planning and academic program review». *Planning for Higher Education*, vol. 14, pp. 22-25.
- BENTHAM, J. (1982 edn.): *An Introduction to the Principles of Morals and Legislation*. Londres.
- BERLO, D.K. (1962): *The Process of Communication*. Holt, Rinehart, and Winston, Nueva York.
- BIJLEVELD, R.J. (1987): «The two-tier structure in dutch university education: A first evaluation of a comprehensive innovation». *Centrum voor Studies van het Hoger Onderwijs Beleid*, Enschede. Mimeo.
- BIJLEVELD, R.J. y GOEDEGEBUURE, L.C.J. (1989): «The paralysing effects of a balanced power system», en Maassen, P.A.M. y Vught, F.A. Van (eds.): *Dutch Higher Education in Transition*. Lemma, Culemborg.
- BIRCH, D.W., CALVERT, J.R. y SIZER, J. (1977): «A case study of some performance indicators in higher education in the United Kingdom». *International Journal of Institutional Management in Higher Education*, 1.2, pp. 133-142.
- BLACKBURN, R.T.; LINGENFELTER, P.E. (1973): *Assessing Quality in Doctoral Programs: Criteria and Correlates of Excellence*. Center for the Study of Higher Education, ED 078728, Ann Arbor, Michigan.
- BLAU, P.M.; 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.
- BLEUMINK, E. (1990): «Programma erkenning in het licht van voorgaande internationalisering» (Validación de programas a la luz de la internacionalización en marcha), en Heijnen, G.W.H., Joostens, T.H. y Vroeijenstijn, A.I. (eds.): *Kwaliteitszorg, Waarborg voor Kwaliteit in Het Hoger Onderwijs*. COWOG, Groningen.
- BOESJES-HOMMES, R.W. (1974): *De Geldige Operationalisering van Begrippen (The Valid Operationalisation of Concepts)*. Boom, Meppel.
- 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.
- BOOTH, B. y BOOTH, C. (1989): «Planning for quality: Advice respectfully tendered to the polytechnics and colleges funding council». *Higher Education Quarterly*, 43, 278-288.

- BORMANS, M., BROUWER, R., VELD, R.J. In't y MERTENS, F.J. (1987): «The role of performance indicators in improving the dialogue between government and universities». *International Journal of Institutional Management in Higher Education*, 11, 2, pp. 181-194.
- BORMANS, J.G., et al. (1990): Artículo preparado por el grupo de la OECD/IMHE sobre indicadores de rendimiento y contenido en el *OECD/IMHE Compendium of Performance Indicators for Higher Education*, París.
- BOURKE, P. (1986): *Quality Measures in University*, Commonwealth Tertiary Education Commission, Australia.
- BOWKER, A.H. (1965): «Quality and quantity in higher education». *Journal of the American Statistical Association*, 60, 309, pp. 1-15.
- BOYER, E.M., EWELL, P.T., FINNEY, J.E., & MINGLE, J.R. (1987): Assessment and outcomes measurement –a view from the states. *American Association for Higher Education Bulletin*, 39 (7), pp. 8-12.
- BRAUDEL, F. (1980): *On History*, Weindenfeld and Nicolson, London.
- BRITISH COLUMBIA COUNCIL OF COLLEGE AND INSTITUTE PRINCIPALS (BCCCIP) (1984): *Institutional Evaluation Process for British Columbia Colleges and Institutes*. BCCCIP, Vancouver.
- BRONS, R. (1990): «Middelenallocatie en kwaliteit», en Heijnen, G.W.H., Joostens, T.H. y Vroeijenstijn, A.I. (eds.): *Kwaliteitszorg, Waarborg voor Kwaliteit in Het Hoger Onderwijs*. COWOG, Groningen.
- BROWN, D.G. (1967): *The Mobile Professors*. The American Council on Education, Washington, D.C.
- BUCHANAN, J.M. (1973): «The institutional structure of externality». *Public Choice*, XIV, pp. 69-82.
- CALVERT, J.R. (1979): «Relative performance», en Billing, D. (ed.): *Indicators of Performance*. Society for Research into Higher Education, Guildford, UK.
- CARABAÑA, J. (1984): Sistema educativo y mercado de trabajo en el horizonte del año 2000. *Revista de Educación*, 273, pp. 23-48.
- CARMINES, E.G. y ZELLER, R.A. (1979): *Reliability and Validity Assessment*. Sage, Beverly Hills-Londres.
- CARNEGIE COUNCIL ON POLICY STUDIES IN HIGHER EDUCATION (1976): *A Classification of Institutions of Higher Education*. The Carnegie Foundation for the Advancement of Teaching, Berkeley, California.
- CARNEGIE FOUNDATION (1987): *A Classification of Institutions of Higher Education*. 1987 Edition. The Carnegie Foundation for the Advancement of Teaching, Princeton.

- CARPENTER, R.L. y CARPENTER, P.A. (1970): «The doctorate in librarianship and an assessment of graduate library education». *Journal of Education for Librarianship*, 11, pp. 3-45.
- CARTTER, A.M. (1966): *An Assessment of Quality in Graduate Education*. American Council in Education, Washington.
- CARTTER, REPORT (1977): «The Cartter Report on the leading schools of education, law, and business». *Change*, 9, Feb, 44-48.
- CAVE, M., HANNEY, S., KOGAN, M. y TREVETT, G. (1988): *The Use of Performance Indicators in Higher Education*, Jessica Kingsley Publishers, Londres.
- CAVE, M., KOGAN, M. y HANNEY, S. (1990): «The scope and effects of performance measurement in British higher education», En Dochy, F.J.R.C., Segers, M.S.R. y Wijnen, W.H.F.W. (eds.): *Management Information and Performance Indicators in Higher Education: An international Issue*, Van Gorcum, Assen.
- CLARK, B.R. (1983): *The Higher Education System: Organisation in Cross-National Perspective*. University of California Press, Berkeley/Los Angeles.
- CLARK, B.R. y YOUNG, T.I.K. (1976): *Academic Power in the United States*. ERIC/Higher Education Report, American Association for Higher Education, Washington, D.C.
- CLARK, M.J.; HARTNETT, R.T.; BAIRD, L.L. (1976): *Assessing Dimensions of Quality in Doctoral Education: A Technical Report of a National Study in Three Fields*. Educational Testing Service, Princeton.
- CLARKE, A.M. y BIRT, L.M. (1982): «Evaluative reviews in universities: The influence of public policies». *Higher Education*, 11 (1), pp. 1-26.
- CLEMENTE, F.; STURGIS, R.B. (1974): «Quality of department of doctoral training and research productivity». *Sociology of Education*, 47, 287-299.
- COCHRAN, T.R., HENGSTLER, D.D. (1984): «Political processes in an academic audit: Linking evaluation information to programmatic decisions». *Research in Higher Education*, vol. 10, pp. 181-192.
- COLE, J.R.; LIPTON, J.A. (1977): «The reputation of American Medical Schools». *Social Forces*, 55, 3, 662-684.
- COLE, S.; COLE, J.R. y SIMON, G.A. (1981): «Chance and consensus in peer review». *Science*, 214, 881.
- COMITE NATIONAL D'EVALUATION (1987): *Où Va l'Université?*. Rapport du Comité National d'Evaluation. Gallimard, Paris.
- CONNOLLY, T., CONLON, E.J. y DEUTSCH, S.J. (1980): «Organizational effectiveness: A multiple-constituency approach». *Academy of Management Journal*, 5, 211-217.

- CONRAD, C.F.; 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, Agatons Press, Nueva York.
- CONRAD, C.F.; BLACKBURN, R.T. (1985b): «Correlates of departmental quality in regional colleges and universities». *American Educational Research Journal*, 22, 2, 279-295.
- CONRAD, C.F.; BLACKBURN, R.T. (1986): «Current views of departmental quality: An empirical examination». *The Review of Higher Education*, 9, 3, 249-266.
- CONSEJO ASESOR DE LA ENSEÑANZA SUPERIOR (1989): *Europa 1992 en Het Nederlandse Hoger Onderwijs*. ARHO, 's-Gravenhage.
- CONSEJO ASESOR DE POLITICA INVESTIGADORA (RAWB) (1989): *Advies Inzake het Werktuigbouwkundig en Maritiem-Technisch Onderzoek in Nederland*. (Asesoramiento sobre las investigaciones en el campo de la Ingeniería mecánica y marina). RAWD, Den Haag.
- COOK, C.M. (1989): «Reflections on the American Experience», en *Verslag van de Conferentie Kwaliteitsbewaking Hoger Onderwijs* (Actas del congreso sobre control de calidad). Ministerie van Onderwijs en Wetenschappen, Zoetermeer.
- COUNCIL OF GRADUATE SCHOOLS IN THE UNITED STATES (1979): «Assesment of graduate programs». *Communicator*, vol. 12, pp. 1-2.
- COWAN, J. (1985): «Effectiveness and efficiency in higher education». *Higher Education*, 14, 235-239.
- COX, W. M.; 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.
- CRAMER, D.G. y Assoc. (1990): *College Student Development. Theory and Practice for the 1990s*. Alexandria, VA.: American College Personal Association.
- CRANE, D. (1970): «Scientist at major and minor universities: a study of productivity and recognition». *American Sociological Review*, 30, October, pp. 699-714.
- CRAVEN, E. (1980): «Evaluating program performance», en Jedamus, P. y Peterson, M.W. (eds.): *Improving Academic Management*. Jossey-Bass, San Francisco, pp. 432-457.
- CRAWSHAW, B. (1985): «Contract research, the university, and the academic». *Higher Education*, 14, pp. 665-682.
- CUA / APA STUDY GROUP (1990): *Strategic Choice: Corporate Strategies for Change in Higher Education*. CUA.
- CUENIN, S. (1987): «The use of performance indicators in universities: An international survey». *International Journal of Management in Higher Education*, 11, 2, 117-139.

- CUENIN, S. (1990): «Un examen des politiques européennes en matière d'évaluation de l'enseignement supérieur». Ponencia presentada en la *Reunión del Proyecto Columbus*, Valparaíso, Chile.
- CVCP (COMMITTEE OF VICE-CHANCELLORS AND PRINCIPALS) (1985): *Report of the Steering Committee for Efficiency Studies in Universities*, (Jarratt Report), CVCP, Londres.
- CVCP (COMMITTEE OF VICE-CHANCELLORS AND PRINCIPALS) (1986): *Academic Standards in Universities*, (Reynolds Report), CVCP, Londres.
- CVCP (COMMITTEE OF VICE-CHANCELLORS AND PRINCIPALS) (1986a): *Performance Indicators in Universities: A first statement by a joint CVCP/UGC working group*, CVCP, Londres.
- CVCP (COMMITTEE OF VICE-CHANCELLORS AND PRINCIPALS) (1987): *Performance Indicators in Universities: A Second Statement by the Joint CVCP/UGC Working Group*, CVCP, Londres.
- CVCP (COMMITTEE OF VICE-CHANCELLORS AND PRINCIPALS) (1988): *Guidance in Structure and Practice of Financial Management in Universities*, CVCP, Londres.
- CVCP (COMMITTEE OF VICE-CHANCELLORS AND PRINCIPALS) (1989): *The Teaching Function: Quality Assurance*, CVCP, Londres.
- CVCP (COMMITTEE OF VICE-CHANCELLORS AND PRINCIPALS) (1989a): *A Statement of Recommended Accounting Practice: Accounting in UK Universities* CVCP, Londres.
- 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. Jossey-Bass, San Francisco.
- CHEPS (1990): *Quality Assessment Data Package*. Center for Higher Education Policy Studies, Twente.
- CHISWICK, B.R. (1973): «Schooling, screening, and income». En: L.C. Solmon y P.J. Taubman (Eds.): *Does College Matter? Some evidence on the impacts of Higher Education*. Nueva York: Academic Press (pp. 151-158).
- CHURCH, C.H. (1988): «The qualities of validation». *Studies in Higher Education*, 13, (1), pp. 27-44.
- DANZIGER, J.N. (1978): *Making Budgets. Public Resource Allocation*. Beverly Hills/Londres.
- DAWKINS, J.S. (1988): *Higher Education: A Policy Statement*. Australian Government Publishing Service, Canberra.
- DE ALESSI, L. (1983): «Property rights, transaction costs, and X-efficiency: An essay in economic theory». *American Economic Review*, 73, pp. 64-81.

- DE MIGUEL, M. (1989): «Modelos de investigación sobre organizaciones educativas». *Revista de Investigación Educativa*, vol. 7, 13, pp. 21-56.
- DE MIGUEL, M. (1990): «Indicadores de calidad en la docencia universitaria». Ponencia presentada en el *Congreso sobre Calidad de la Educación Universitaria*, Puerto de Santa María, Marzo.
- 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.
- DES (DEPARTMENT OF EDUCATION AND SCIENCE) (1987): *Higher Education: Meeting the Challenge*. HMSO, Londres.
- DE VREE, J.K. (1982): *Foundations of Social and Political Processes*. Prime Press, Bilthoven.
- DE VREE, J.K. (1990): *Order and Disorder in the Human Universe: The Foundations of Behavioral and Social Science*, Vol. 1-3. Prime Press, Bilthoven.
- DOCHY, F.J.R.C. y BOUWENS, M.R.J. (1990): *The Construction of Knowledge State Tests, Knowledge Profiles and the Measurement of the Value Added*. Educational Technology Innovation Centre, University of Heerlen.
- DOCHY, F.J.R.C. y DE VOLDER, M.L. (1987): «The use of indicators in a cyclical process for quality assurance as a part of new steering conception for higher education in the Netherlands», *Onderzoek van Onderwijs*, 34, Rijksuniversiteit Limburg, Maastricht.
- 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.; Westerheijden, D.F. (eds.): *Peer Review and Performance Indicators*. Uitgeverij Lemma, Utrecht, NL.
- DOCHY, F.J.R.C., SEGERS, M.S.R., y WIJNEN, W.H.F.W. (1990a): «Preliminaries to the implementation of a quality assurance system based on management information and performance indicators. Results of a validity study», en Dochy, F.J.R.C., Segers, M.S.R. y Wijnen, W.H.F.W. (eds.): *Management Information and Performance Indicators in Higher Education: An International Issue*. Van Gorcum, Assen.
- DOCHY, F.J.R.C., SEGERS, M.S.R. y WIJNEN, W.H.F.W. (eds.) (1990b): *Management Information and Performance Indicators in Higher Education: An International Issue*. Van Gorcum, Assen.
- DOCHY, F.J.R.C., WIJNEN, W.H.F.W. Y BOUHUIJS, P.A.J. (1987): *Kwaliteitsbewaking in Het Hoger Onderwijs* (Control de calidad en Educación Superior), Staatsuitgeverij, Den Haag.
- DOLAN, W.P. (1976): *The Ranking Game: The Power of the Academic Elite*. The Nebraska Curriculum Development Center, Lincoln, Nebraska.
- DRESSEL, P.L. y CAMMACK, E.F. (1972): «Conducting self-study», en P.L. Dressel and

- Associates (eds.): *Institutional Research in the University*. Jossey-Bass, San Francisco, pp. 231-291.
- DREW, D.E. (1975): *Science Development: An Evaluation Study*. The National Academy of Sciences, Washington, D.C.
- DREW, D.E.; KARP, R. (1981): «Ranking academic departments: Empirical findings and a theoretical perspective». *Research in Higher Education*, 14, 4, pp. 305-320.
- DUBE, W.E. (1974): «Undergraduate origins of U.S. medical students». *Journal of Medical Education*, 49, pp. 1005-1010.
- DUBIN, R. (1976): «Organizational effectiveness: Some dilemmas of perspective». *Organizational and Administrative Sciences*, 7, pp. 7-13.
- DUNER, A. (1977): *Personal and Occupational Choice*. Lisse, Swets, & Zeitlinger.
- DUNKIN, M. y BARNES, J. (1986): «Research on teaching in higher education», en Witrock, M.C. (ed.): *Handbook of Research on Teaching-Third Education*. MacMillan, Nueva York.
- ECHEVARRIA, B.; RODRIGUEZ, S. (1989): *Proyecto de un Centro de Orientación Académica y Profesional de la Universidad de Barcelona*. I Simposio Nacional de Orientación en la Universidad. Barcelona.
- ELTON, Ch.F.; RODGERS, S.A. (1971): «Physics department ratings: Another evaluation». *Science*, 174, November, pp. 565-568.
- ELTON, Ch.F.; ROSE, H.A. (1972): «What are the ratings rating?». *American Psychologist*, 27, pp. 197-201.
- ENNS, F.; HARRIS, W.E.; HOLDAWAY, E.A.; JOPLING, W.H. y KREISEL, H. (1986): «Program review: The experience of a canadian university». *Higher Education Research and Development* (Higher Education Research and Development Society of Australasia), vol. 5, pp. 31-47.
- ESCOTET, M.A. (1984): *Técnicas de Evaluación Institucional*. MEC, Servicio de Publicaciones, Madrid.
- ESCUADERO, T. (1989): «Aproximación pragmática a la evaluación de la Universidad». *Revista de Investigación Educativa*, vol. 7, 13, pp. 93-112.
- FEASLEY, C.E. (1980): *Program Evaluation*. American Association for Higher Education, Washington, D.C.
- FERNANDEZ, M. (1989): *Así enseña nuestra universidad. Hacia la construcción crítica de una didáctica universitaria*. Madrid: Autor.
- FERNANDEZ, M. (1990): «Avance de resultados de la investigación patrocinada por la

- CAICYT. (Estudio sobre la calidad de la enseñanza universitaria)» En: *Primeras Jornadas Nacionales de Didáctica Universitaria*. Madrid: Consejo de Universidades (pp. 15-36).
- FERNANDEZ, J.; LILLO, B. y MARTINEZ, M.R. (1989): «Proyecto de evaluación de la educación superior española», en Varios: *Consideraciones Metodológicas sobre la Evaluación y Mejora de la Docencia Universitaria*. Servei de Formació Permanent, Valencia.
- FINDLAY, P. (1990): «Developments in the performance indicator debate in the United Kingdom», en Goedegebuure, L.C.J.; Maassen, P.A.M.; Westerheijden, D.F. (eds.): *Peer Review and Performance Indicators*. Uitgeverij Lemma, Utrecht, NL.
- FOSTER, F.; RODGERS, J. (198): «Quality of education and student earnings». *Higher Education*, 9, pp. 21-37.
- FRACKMANN, E. (1983): «Die verwendung von kennzahlen im hochschulbereich (unter berücksichtigung US-amerikanischer verhältnisse)». Statistisches Bundesamt, *Fachgespräch zu Fragen der Ermittlung und Aussagekraft von Hochschulstatistischen Kennzahlen*. Wiesbaden.
- 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.; 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 Mayor. ED 293429.
- FRACKMANN, E. y MUFFO, J.A. (1988): «Quality control, hierarchies, and information. The context of rankings in an international perspective». *Proceedings 10th EAIR Conference*, University of Bergen, Norway.
- FRANKE-WIKBERG, S. (1990): «Evaluating education quality on the institutional level». *Higher Education Management*, vol. 2, 3, pp. 271-292.
- FREEMAN, R.E. (1984): *Strategic Management: A Stakeholder Approach*. Pitman, Boston.
- FRIEDMAN, M. (1953): «The methodology of positive economics», en Friedman, M.: *Essays in Positive Economics*. University of Chicago Press, Chicago/London.
- FRIJDAL, A. (1985): «Research management van het wetenschappelijk onderzoek, een behersmodel». Tesis doctoral, Vrije Universiteit Brussel, Brussels.
- FRIJDAL, A. y DE GREVE, J.P. (1986): «Communication activities in scientific disciplines in Belgium». *Scientometrics*, 9, pp. 37-39.
- FURUMARK, A.M. (1981): «Institutional self-evaluation in Sweden». *International Journal of Institutional Management of Higher Education*, 5, 3, 207-216.

- GALLANT, J. y PROTHERO, J.W. (1972): «Weight-watching at the university: The consequences». *Science*, 175, no. 381, pp. 385-386.
- GARFIELD, E. (1979): *En Citation Indexing: Its Theory and Application in Science Technology and the Humanities*. Wiley, Nueva York.
- GEORGE, M.D. (1982): «Assessing Program Quality». En R. Wilson (ed.) *Designing Academic Program Review*. San Francisco: Jossey Bass.
- GIBBONS, M. (1984): «Methods of evaluation of research». Informe en el *Seventeenth IMHE Special Topic Workshop*, París, 3rd-5th December.
- GILL, W.N. (1975): *Rankings of Graduate Engineering Departments*. State University of New York at Buffalo, Buffalo.
- GILMOUR, J.E. (1980): «Using Marketing to Incorporate Environmental Considerations into Program Review Processes». Artículo presentado en el *Annual Forum of Association for Institutional Research*, Atlanta, Abril.
- GLENN, N.D. y VILLIMEZ, W. (1970): «The productivity of sociologists at 45 american universities». *American Sociologist*, 5, August, pp. 244-252.
- GLOWER, D.D. (1980): «A rational method for ranking engineering programs». *Engineering Education*, 70, 788-794.
- GOEDEGEBUURE, L.C.J.; MAASSEN, P.A.M.; WESTERHEIJDEN, D.F. (1990): «Quality assessment in higher education», en Goedegebuure, L.C.J.; Maassen, P.A.M.; Westerheijden, D.F. (eds.): *Peer Review and Performance Indicators*. Uitgeverij Lemma, Utrecht, NL.
- GOEDEGEBUURE, L.C.J.; MAASEN, P.A.M. y WESTERHEIJDEN, D.F. (eds.) (1990a): *Peer Review and Performance Indicators, Quality Assessment in British and Dutch Higher Education*. LEMMA, Utrecht.
- GOURMAN, J. (1967): *The Gourman Report: Ratings of American Colleges*. The Continuing Research Institute, Phoenix, Arizona.
- GOURMAN, J. (1977a): *The Gourman Report: A Rating of Undergraduate Programs in American and International Universities*. The National Education Standards, Los Angeles.
- GOURMAN, J. (1977b): *The Gourman Report: A Rating of Graduate and Professional Programs in American and International Universities*, rev. 2nd ed. The National Education Standards, Los Angeles.
- GOURMAN, J. (1982): *The Gourman Report: A Rating of Undergraduate Programs in American and International Universities*. The National Education Standards, Los Angeles.

- GOURMAN; J. (1983): *The Gourman Report: A Rating of Graduate and Professional Programs in American and International Universities*, rev. 2nd ed. The National Education Standards, Los Angeles.
- GREGG, R. y SIMS, P. (1972): «Quality of faculties and programs of graduate departments of educational administration». *Educational Administration Quarterly*, 8, pp. 67-92.
- GUBA, EG.; CLARK, D.I. (1978): «Levels of R&D productivity in schools of education». *Change*, May, 3-9.
- GUBA, E. y LINCOLN, Y. (1988): *Effective Evaluation*. Jossey-Bass, Londres.
- GUTIERREZ, A. (1984): «Universidad y empleo». *Revista de Educación*, 273, 65-82.
- HAGERTY, B. y STARK, J. (1989): «Comparing educational accreditation standards in selected professional fields». *Journal of Higher Education*, vol. 60, 1, pp. 1-20.
- HAGSTROM, W.O. (1971): «Inputs, outputs, and prestige of university science departments». *Sociology of Education*, 44, 375-397.
- HALPERN, D.F. (ed.) (1987): «Student outcomes assessment: What institutions stand to gain». *New Directions for Higher Education*, 59, Jossey-Bass, San Francisco.
- HARMAN; G.S. y JOHNSON, R.St.C. (1979): «Academic accountability —courses and programmes—, en Sheldrake, P. y Linke, R. (eds.): *Accountability in Higher Education*, George Allen & Unwin, Sydney.
- HARPEL, R.L. (1986): «The anatomy of an academic program review». *AIR Professional File*, vol. 15, pp. 1-6.
- HARRIS, W.E. y HOLDAWAY, E.A. (1983): «Systematic reviews of university programs and units». *Canadian Journal of Higher Education*, vol. 13, pp. 55-76.
- HEERINGEN, A., VAN MOMBERS, C. & VAN VENETIË, R. (1983): «Science and technology indicators 1983, a comparison of the Netherlands with foreign countries, based on quantitative data». *Background Study* n.º 11 for the Research Council of the Netherlands (in dutch).
- HOBBS, T. (1981) (1651): *Leviathan*. Penguin Books, Harmondsworth.
- HOFSTEE, W.K.B. (1987-1988): «Ex post beoordeling van voorwaardelijk gefinancierde programma's: Misbruik van het wetenschappelijk forum». *Universiteit en Hogeschool*, Jaargang 34, nr. 2.
- HOLDAWAY, E.A. (1981): «The review process in universities». *Australian Journal of Education*, vol. 25 (2), pp. 106-120.
- 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*. Lemma, Culemborg, NL.

- HORTON, R. (1990): *The Funding of Expansión*, Speech at Lancaster University, 11 Enero.
- HOUSE, D.R. y YEAGER, J.H., Jr. (1978). «The distribution of publication success within and among top economics departments: A disaggregate view of recent evidence». *Economic Inquiry*, 16, Octubre, pp. 593-598.
- HOYT, D.P. (1980): *Program Evaluation: Is It Worth It?* Ponencia presentada en *University of Illinois at Urbana-Champaign Conference on Planning and Conducting Program Evaluations and Reviews in Higher Education*, St. Petersburg, Florida, January.
- HÜFNER, K. (1985): «Probleme des entwicklung und anwendung von kennzahlen in hochschulbereich». Diskussionspapiere des Forschungsschwerpunktes. *Ökonomische Theorie der Hochschule*, 3.
- HÜFNER, K.; 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*. Miami University Press, Oxford, OH.
- HUGHES, R.M. (1934): «Report of the Committee on Graduate Instruction». *Educational Record*, 15, 192-234.
- HUSEN, T. (1988): *Nuevo análisis de la sociedad del aprendizaje*. Barcelona: Paidós-M.E.C.
- INSPECTORADO DE LA EDUCACION SUPERIOR (1989): *De Inspectie Hoger Onderwijs En Haar Meta-Evaluatie Taak Binnen Het Stelsel Van Kwaliteitszorg In Het Hoger Onderwijs* (El inspectorado de la educación superior y su tarea meta-evaluadora en el sistema de proteger la calidad de la educación superior). IHO, De Meern.
- 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*. Paris, OCDE, Septiembre.
- JACOBSON, R.L. (1985): «Too much information, too few questions for campus reviews?» *Chronicle of Higher Education*, vol. 26, 2 October.
- JANY, G.O. y KELLER, R.J. (1981): «An Evaluation of a State University Systemwide Academic Program Review Proces». Artículo presentado en el *Annual Forum of Association for Institutional Research*, Minneapolis, May.
- JARRAT REPORT (1985): *Report of the Steering Committee for Efficiency Studies in Universities*. U.K. Committee of Vice-Chancellors and Principals (CVCP), Londres, Marzo.
- JESSEE, W.F. (1984): «Quality assurance systems: Why aren't there any?». *Quality Review Bulletin*, pp. 408-411.

- JOHNES, J. y TAYLOR, J. (1990): *Performance Indicators in Higher Education*. Open University Press, Londres.
- JOHNSON, R.R: (1978): «Leadership among american colleges». *Change*, 10, November, pp. 50-51.
- JONES, L.V.; LINDZEY, G.; COGGESHALL, P.E. (eds.) (1982a): *An Assessment of Research-Doctorate Programs in the United States: Humanities*. (5 Vol.: Biological Science, Humanities, Social and Behavioral Science, Mathematical and Physical Science, Engineering). National Academic Press, Washington, D.C.
- JONES, L.V.; LINDZEY, G.; COGGESHALL, P.E. (eds.) (1982b): *An Assessment of Research-Doctorate Programs in the United States: Social and Behavioral Sciences*. The National Academic Press, Washington, D.C.
- JONES, L.V.; LINDZEY, G.; COGGESHALL, P.E. (eds.) (1982c): *An Assessment of Research-Doctorate Programs in the United States: Biological Sciences*. The National Academic Press, Washington, D.C.
- JONES, L.V.; LINDZEY, G.; COGGESHALL, P.E. (eds.) (1982d): *An Assessment of Research-Doctorate Programs in the United States: Mathematical and Physical Sciences*. The National Academic Press, Washington, D.C.
- JONES, L.V.; LINDZEY, G.; COGGESHALL, P.E. (eds.) (1982e): *An Assessment of Research-Doctorate Programs in the United States: Engineering*. The National Academic Press, Washington, D.C.
- JONES, R.W. (1969): «The student faculty ratio in graduate programs of selected departments of sociology». *American Sociologist*, 4, pp. 123-127.
- JORDAN, R.T. (1963): «Library characteristics of colleges ranking high in academic excellence». *College and research Libraries*, 24, pp. 369-376.
- Juris Doctor Magazine* (1976): «The popular vote: Rankings of top schools». Vol. 6, pp. 17-18, 21.
- KALLEN, D. (1987): «Fracaso y malos resultados escolares: los nuevos iletrados». *Perspectivas*, XVII (2), pp. 227-237.
- KAPLOWITZ, S.A. (1978): «Towards a systematic theory of power attribution». *Social Psychology (Quarterly)*, 41, 131-148.
- KASTEN, K.L. (1984): «Tenure and merit pay as reward for research, teaching and service at a research university». *Journal of Higher Education*, 55, pp. 500-514.
- KEELEY, M. (1987): «A social-justice approach to organizational evaluation». *Administrative Science Quarterly*, 21, pp. 272-292.
- KELLS, H.R. (1983): *Self-Study Processes. A Guide for Postsecondary Institutions*. MacMillan, Nueva York.

- KELLS; H.R. (1988): *Self-Study Processes. A Guide for Postsecondary and Similar Service-Oriented Institutions and Programs*. 3.ª edición, American Council of Education, McMillan, Nueva York.
- KELLS, H.R. (1989a): «The future of self-regulation in dutch higher education», documento presentado en la reunión de la *Association for University Governance and Management* (VUB&M), Utrecht.
- KELLS, H.R. (1989b): «University Self-Regulation in Europe: The Need for an Integrated System of programme Review». *European Journal of Education*, 24-4, pp. 299-308.
- KELLS, H.R. (Ed.) (1990a): *The Development of Performance Indicators for Higher Education. A Compendium for Eleven Countries*. PIMHE-OCDE, París.
- KELLS, H.R. (1990b): «The inadequacy of performance indicators for higher education. The need for a more comprehensive and development construct». *Higher Education Management*, 2, 3, pp. 258-270.
- KELLS, H.R.; 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.
- KELLS, H.R.; Van VUGHT, F.A. (EDS) (1988a): *Self-Regulation, Self-Study, and Program Review in Higher Education*. Lemma, Utrecht.
- KELSO, CH.D. (1975): «How does your law school measure up». *Student Lawyer*, 4, pp. 20-24.
- KENISTON; H. (1959): *Graduate Study and Research in the Arts and Sciences at the University of Pennsylvania*. University of Pennsylvania Press, Philadelphia.
- KEUNING, D. y EPPINK, D.J. (1986): *Management En Organisatie: Theorie En Toepassing* (Gestión y organización: teoría y aplicaciones), Stenfert Kroese, Leiden/Antwerpen.
- KNAPP, R.H. y GOODRICH, H.B. (1952): *Origins of American Scientists*. Russell & Russell, Nueva York.
- KNAPP, R.H. y GREENBAUM, J.J. (1953): *The Younger American Scholar: His Collegiate Origins*. The University of Chicago Press, Chicago.
- KNUDSEN, D.D.; VAUGHAN, 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. (1984): «The political view», en Clark, B.R. (ed.): *Perspectives on Higher Education*. University of California Press, Berkeley.
- KOGAN, M. (1986): «The evaluation of higher education. An introductory review». *International Journal of Institutional Management in Higher Education*, 10, 2, 125-139.

- KOGAN, M. (ed.) (1989): *Evaluating Higher Education*. Jessica K. Publishers, Londres.
- KOGAN, M. y BOYS, C. (1982): «The politics of sectoral change», en Wagner, L. (ed.): *Agenda for Institutional Change in Higher Education*. Society for Research in Higher Education, Guildford.
- KRAUSE, E.D. y KRAUSE, L. (1970): «The colleges that produce our best scientists: A study of the academic training grounds of a large group of distinguished american scientists». *Science Education*, 54, pp. 133-140.
- KUH, G.D. (1981): *Indices of Quality in the Undergraduate Experience*. American Association for Higher Education, Washington.
- KUHN, T.S. (1970) (1962): *The Structure of Scientific Revolutions*, 2nd. enl. ed. University of Chicago Press. Chicago.
- KURZ, R. et al. (1989): «Faculty performance». *Journal of Higher Education*, vol. 60, 1, pp. 43-58.
- LAMOURE, J.E. (1984): *Les Instituts Universitaires de Techonologie*. Institut Européen d'Education et de Politique Sociale, Paris.
- LAPORTE, J. (1990): *Perspectivas de la Universitat a Catalunya*. Barcelona: Servei de Publicacions Universtitat de Barcelona.
- LATOUR, B. y WOOLGAR, S. (1979): *Laboratory Life*. Sage, Beverly Hill.
- LAVENDER, A.D.; MATHERS, R.A. y PEASE, J. (1971): «The student-faculty ratio in graduate programs of selected departments of sociology: A supplement to the Janes Report». *American Sociologist*, 6, pp. 29-30.
- LAWRENCE, J.K.; GREEN, K.C. (1980): *A Question of Quality: The Higher Education Rating Game*. American Association for Higher Education, Washington.
- LE PAIR, C. (1980): «Decision making on grant applications in a small country». *Scientia Yugoslavia*, 6, pp. 137-143.
- LEE, R.A.y PIPER, J.A. (1988): «Organizational control, differing perspectives: The management of universities». *Financial Accountability and Management*, 4. N.º 2, pp. 113-128.
- LEVIN, M.E. y LEVIN, J.R. (1991): «A Critical Examination of Academic Retention Programs for At-Risk Minority». *J. of College Student Development*, 32, (4), pp. 323-334.
- LEVIN, H.M y RUMBERGER, R. W. (1989): «Educación, trabajo y empleo en los países desarrollados: Situación y desafíos para el futuro». *Perspectivas*, 19, 2, pp. 221-242.
- LEWIS, L. (1968): «On subjective and objective rankings of sociology departments». *American Sociologist*, 3, May, 129-131.

- LIBRO VERDE (1985): *The Development of Higher Education into the 1990s*. Secretary of State for Education and Science, HMSO, Londres, Cmd 9524.
- LIESHOUT, R.H. (1984): *Without Making Elaborate Calculations for the Future*. Enschede.
- LIESHOUT, R.H. y DE VREE, J.K. (1985): «How organizations decide». *Acta Politica*, XX, pp. 129-155.
- LINDSAY, A.W. (1981): «Assessing institutional performance in higher education: A managerial perspective». *Higher Education*, 10, 687-706.
- LINDSAY, A.W. (1982): «Institutional performance in higher education: The efficiency dimension». *Review of Educational Research*, 52, 2, 175-199.
- LINKE, R. (1990): Artículo preparado para el grupo de la OECD/IMHE sobre indicadores de rendimiento y contenido en el *OECD/IMHE Compendium of Performance Indicators for Higher Education*, París.
- LITTEN, L. y HALL, A. (1989): «In the eyes of our beholders», *Journal of Higher Education*, vol. 60, 3, pp. 302-324.
- LIU, H.C. (1978): «Faculty citation and quality of graduate engineering departments», *Engineering Education*, 68, Abril, pp. 739-741.
- LOCKWOOD, G. and DAVIES, J. (1985): *Universities: The Management Challenge*, SRHE and NFRE-Nelson, Windsor.
- LODER, C. (ed.) (1989): *Quality Assurance and Accountability in Higher Education*. Kogan Page, Londres.
- LUHMANN, N. (1968): «Selbststeuerung der wissenschaft». *Jahrbuch für Sozialwissenschaft*, 19, p. 147.
- MAASSEN, P.A.M. (1989): «Quality assessment in higher education: The dutch experience», en McVicar, M. (ed.): *Performance Indicators and Quality Control in Higher Education*. Portsmouth Polytechnic, Portsmouth.
- MAASSEN, P.A.M.; 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.
- MAASSEN, P.A.M. y VUGHT, F.A. Van (1989): «Is government really stepping back?», en Maassen, P.A.M. y Vught, F.A. Van (eds.): *Dutch Higher Education in Transition*, Lemma, Culemborg.
- MAASSEN, P.A.M. y WEUSTHOF, P.J.M. (1989): «Quality assessment in dutch higher education: A big leap forward or a trojan horse?», en Maassen, P.A.M. y Vught, F.A. Van (eds.): *Dutch Higher Education in Transition*, Lemma, Culemborg.

- MacGREGOR, Rt. Hon. J. (1989): Secretary of State's Speech to CVCP, 27 September.
- MAGOUN, H.W. (1966): «The Cartter Report on quality in graduate education». *Journal of Higher Education*, 37, 481-492.
- MARGULIES, R.Z.; BLAU, P.M. (1973): «America's leading professional schools». *Change*, 5, Nov, 21-27.
- MARSH, H.W. (1984): «Validity of students evaluations of college teaching: dimensionality, reliability, validity, potential biases and utility». *Journal of Educational Psychology*, 76, pp. 707-754.
- MARTIN, B.R. & IRVINE, J. (1984): *Research Policy*, 13, pp. 183, 247, 311. Un conjunto de artículos sobre indicadores aplicados al rendimiento científico de los aceleradores de partículas. Ver también Irvine, J. and Martin, B.R. (1985): *Scientometrics*, 7, p. 281.
- MBA Magazine (1974): «The 15 top-ranked business schools in the United States». December, pp. 21-25.
- MBA Magazine (1975): «The top 15: The MBA survey of graduates business schools». December, pp. 33-35.
- McCORKLE, C.O. y ARCHIBALD, S.O. (1982): *Management and Leadership in Higher Education*. Jossey-Bass, San Francisco.
- McDONALD, R. y ROE, E. (1984): *Reviewing Departments*. Green Guide n. 1 HERDSA, Kensington, Australia.
- McGILL UNIVERSITY (1982): *Principles and Guidelines for Systematic Cyclical Reviews*. Senate Committee on Academic Policy and Planning.
- McVICAR, M. (ed.) (1989): *Performance Indicators and Quality Control in Higher Education*. UK Polytechnic, Portsmouth.
- M.E.C. (1989): *Hacia una Clasificación de las Universidades según Criterios de Calidad*. Colección Forum Universida-Empresa, Madrid.
- MEDINA, R.I. (1985): «Los fines formativos de la universidad en distintos pensadores y sistemas». *Revista Española de Pedagogía*, 169-170, 345-370.
- MERTENS, F. y BORMANS, M. (1990): «Background to the development of a system of performance indicators in the Netherlands», en Dochy, F.J.R.C., Segers, M.S.R. y Wijnen, W.H.F.W. (eds.): *Management Information and Performance Indicators in Higher Education: An Internationale Issue*, Van Gorcum, Assen.
- MERTENS, P. (1978): «Kennzahlenvergleiche deutscher universitäten» *Betriebswirtschaftliche Forschung und Praxis*, 1, pp. 25-34.
- MERTENS, P. (1979): «Comparative indicators for German universities». *International Journal of Institutional Management in Higher Education*, 3, 1, 155-168.

- MEULEN, B. Van Der y WESTERHEIJDEN, D.F. (1990): *Evaluatie-Onderzoek Verkenning-commissies: Tussenrapportage Inventarisatie* (Comites de evaluación de la investigación. Informe intermedio). Universiteit Twente, WMW/CSHOB, Enschede (no publicado).
- MILLER, R.I. (1979): *The Assessment of College Performance*. Jossey-Bass, San Francisco.
- MILLER, R.I. (1986): «Evaluating Institutional Quality: Some Ways and some Problems». *International Journal of Institutional Management in Higher Education*, 10, (3), pp. 241-251.
- MINISTERIO DE EDUCACION Y CIENCIA (1985): *Hoger Onderwijs: Autonomie en Kwaliteit (HOAK)* (Educación Superior, Autonomía y Calidad), SDU, Den Haag.
- MINISTERIO DE EDUCACION Y CIENCIA (1988): *Hoger Onderwijs en Onderzoekspaan (HOOP)*, (Educación Superior y Plan de Investigación), SDU, Den Haag.
- MINISTERIO DE EDUCACION Y CIENCIA (1989): *Ontwerp HOger Onderwijs en Onderzoek Plan 1990* (Borrador del Plan para la Educación Superior y la Investigación de 1990), SDU, Den Haag.
- MINISTERIO DE EDUCACION Y CIENCIA (1989a): *Wet Hoger Onderwijs en Wetenschappelijk Onderzoek, Wetsvoorstel* (Borrador de las cuentas de la Educación Superior y de la Investigación), SDU, 's-Gravenhage.
- MINTZBERG, H. (1983): *Structure in Fives; Designing Effective Organizations*. Prentice Hall, Englewood Cliffs.
- MOED, H.F.; BURGER, W.J.F.; FRANKFORT, J.G. & VAN RAAN, A.F.J. (1985): «On the measurement of research performance». *Research Policy*, 14, p. 131.
- MORA RUIZ, J.G. (1991): *Calidad y Rendimiento en las Instituciones Universitarias*. Consejo de Universidades, Madrid.
- MORA RUIZ, J.G. (1991): «La evaluación de las instituciones universitarias». *Revista de Investigación Educativa n.º 17 (en prensa)*.
- MORANTES, E. (1987): «A statewide comprehensive outcomes assesment program». *Documento presentado en la Fifth Annual Regents Conference of the State University and Community College System of Tennessee*, Nashville, USA.
- MORGAN, D.L., KEARNEY, R.C. y REGENS, J.L. (1976): «Assessing quality among graduate institutions of higher education in the United States». *Social Science Quarterly*, 57, December, pp. 670-679.
- MORGENTHAU, H.J. (1978): *Politics among Nations*, 5th ed., rev. Knopf, Nueva York.
- MUNSON, C.E. y NELSON, P. (1977): «Measuring the quality of professional schools». *UCLA Educator*, 19, pp. 41-52.

- NARIN, F. (1978): «Evaluative bibliometrics: The use of publication and citation analysis in the evaluation of scientific activity». Informe preparado para la *National Science Foundation*, Filadelfia.
- NATIONAL SCIENCE BOARD (1969): *Graduate Education: Parameters for Public Policy*. The National Science Foundation, Washington.
- NEAVE, G. (1987): Éditorial en *European Journal of Education*, 22, pp. (2), 121-123.
- NEAVE, G. (1987): «Judas goat or brave new world: Technological change and the relationship between government and higher education». Ponencia presentada a la *European Association for Institutional Research*, Enschede, Netherlands, August.
- NEAVE, G. (1988a): «On being economical with university autonomy», en Kogan, M. y Tight, M. (eds.): *Autonomy in Higher Education: Proceedings of the Annual Conference of the Society for Research into Higher Education*. Milton Keynes, Open University Press.
- NEAVE, G. (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.
- NEAVE, G. (1988b): «The making of the executive head». *International Journal for Management of Higher Education*.
- NEAVE, G. y RHOADES, G. (1987): «The academic profession: Divided albeit national», en Clark, B.R. (ed.): *The Academic Profession: Disciplinary, National, and Institutional Settings*. University of California Press, Berkeley/Los Angeles.
- NEAVE, G. y VUGHT, F.A. Van (eds.) (1991): *Prometheus Bound*, Pergamon, Londres.
- NICHOLS, J. (1982): *Old Miss through the 1980s: A Commitment to Goals and a Process for Their Accomplishment*. University of Mississippi.
- NIINILUOTO, I. (1990): Artículo preparado para el grupo de la OECD/IMHE sobre indicadores de rendimiento y contenida en el *OECD/IMHE Compendium of Performance Indicators for Higher Education*, París.
- NILSON, K.A. (1989): Presentación en panel en el «Quality Assessment in a European Perspective», 11th Annual Forum of European AIR, Trier, Germany.
- NISKANEN, W.A. (1971). *Bureaucracy and Representative Government*. Aldine-Atherton, Chicago.
- NUFFIC (1990): *Waardering van Kwaliteit: Internationale Erkenning en Waardering van Het Nederlandse Hoger Onderwijs*. Nuffic, 's-Gravenhage.
- OCDE (1986): *Innovation Policy: France*. OECD, París.
- OCDE (1990): *Review of Educational Policy in the Netherlands*. OCDE, París.

- OECD/IMHE (1990): *Compendium of Performance Indicators for Higher Education*. OCDE, París.
- OIJEN, P.M. van (1990): «Demographic Trends, future needs of society and the implications for higher education planning in the Netherlands». Ponencia presentada en el 29.º *Seminario del Centro para la Investigación e Innovación Educativa de la OCDE* París.
- OLSON, M. (1971) (1965): *The Logic of Collective Action*, 2nd ed., MA, Cambridge.
- OOMKES, F.R. (1986): *Communicatieleer*. Boom, Meppel.
- OROMANER, M.J. (1970): «A note on analytical properties and prestige of sociology departments». *American Sociologist*, 5, Agosto, 240-244.
- OS, W.V.; DRENTH, P.J. y BERNEERT, G.F. (1987): «AMOS: An evaluation model for institutions of Higher Education». *European Journal of Education*, 22, (2), pp. 171-181.
- PAARDEKOOPEL, 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., Westerheijden, D.F. et al. (eds.): *Peer Review and Performance Indicators: Quality Assessment in British and Dutch Higher Education*. LEMMA, Utrecht.
- PCFC (POLYTECHNICS AND COLLEGES FUNDING COUNCIL) (1989): *Funding Choices*. PCFC, Londres.
- PEDRO, F. (1988): «Higher education in Spain: Setting the conditions for an evaluative state». *European Journal of Education*, 23, 1/2, 125-139.
- PENNINGS, J.M. y GOODMAN, P.S. (1977): «Towards a workable framework», en Goodman, P.S. y Pennings, J.M. (eds.): *New Perspectives on Organizational Effectiveness*. Jossey-Bass, San Francisco.
- PEREZ, V. (1981): «Universidad y empleo». *Papeles de economía*, 8.
- PERKINS, D. y SNELL, J.L. (1962): *The Education of Historians in the United States*. McGraw-Hill, Nueva York.
- PERRY, P. (1987): «Accountability and inspection in higher education». *Higher Education Quarterly*, 41, pp. 344-353.
- PETERSON, M.W. (1980): «Analyzing alternative approaches to planning», en Jedamus, P. y Peterson, M.W. (eds.): *Improving Academic Management*. Jossey-Bass, San Francisco, pp. 113-163.
- PETROWSKI, R.; BROWN, E.L.; DUFFY, J. (1973): «National Universities' and the ACE ratings». *Journal of Higher Education*, 44, 7, 495-513.
- PFEFFER, J. y SALANCIK, G. (1978): *The External Control of Organization: A Resource Dependence Perspective*. Nueva York: Harper and Row.

- PIRSIG, R.M. (1974): *Zen and The Art of Motorcycle Maintenance*. William Morrow & Co Nueva York.
- POLLITT, Ch. (1987): «The politics of performance assessment: Lessons for higher education?». *Studies in Higher Education*, 12, 1, 87-98.
- POLLITT, Ch. (1990): «Measuring University Performance: Never Mind the quality, Never Mind the Width?». *Higher Education Quarterly*, 44, PP. 60-81.
- POPPER, K.R. (1965) (1963): *Conjectures and Refutations*, 2nd ed., Harper & Row, Nueva York.
- POTTER, D.L. (1986): *The measurement of student achievement and the assurance of quality in Virginia higher education*. Richmond, VA.: Council of Higher Education for Virginia.
- PREMFORS, R. (1983): «Research and policy-making in Swedish higher education», en Husen, T. y Kogan, M. (eds.): *Educational Research and Policy: How Do They Relate?*. Pergamon, Oxford.
- PREMFORS, R. (1986): «Evaluating basic units: Seven fundamental questions». *International Journal of Institutional Management in Higher Education*, 10, 2, 169-174.
- PREMFORS, R. (1989): «Viability of basic units in universities». *Higher Education in Europe*, vol. XIV, 3, pp. 5-22.
- RAWLS, J. (1987): «Justice as reciprocity», en Gorovitz, S. (ed.): *Utilitarianism*. Bobbs-Merill, Indianapolis.
- RCT (Comite de Reconocimiento de la Teología) 1989: *Rapport van de verkenningcommissie grodgeleerheid*, Zoetermeer: Ministerio de Educación y Ciencia.
- RIP, A. (1988): «Contextual transformations in contemporary science», en Jamison, A. (ed.): *Keeping Science Straight: A Critical Assessment of Science and Technology*. University of Göteborg, Göteborg.
- ROBBINS REPORT (1963): *On Higher Education*. HMSO, Londres.
- RODRIGUEZ, S. (1990): «Problemática y tendencias de la Orientación universitaria». En: *La Reforma Educativa: Un reto para la Orientación*. Valencia: A.E.O.E.P., pp. 107-122.
- ROE, E. y McDONALD, R. (1984): *Informed Professional Judgment. A Guide to Evaluation in Post-Secondary Education*. University of Queensland Press, Brisbane.
- ROE, E.; McDONALD, R. y MOSES, I. (1985): *Reviewing Departments and People. The Maintenance of Quality in Higher Education*. University of Queensland Press, Brisbane.
- ROE, E; McDONALD, R. y MOSES, I. (1986): *Reviewing Academic Performance*. University of Queensland Press, Sta. Lucía, Australia.

- ROMNEY, L.C. (1978): *Measures of Institutional Goal Achievement*. Boulder, Colorado: National Center for Higher Education Management System.
- ROMNEY, L.C.; BOGEN, G.; 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.; ANDERSEN, Ch.J. (1970): *A Rating of Graduate Programs*. American Council on Education Washington.
- ROURKE, F.E. (1969): *Bureaucracy, Politics, and Public Policy*. Little, Brown & Co., Boston.
- RULLMANN, P.M.M. (1990): «A Non-University perspective on quality assessment in the Netherlands, en Goedegebuure, L.C.J., Maassen, P.A.M., Westerheijden, D.F. et al. (eds.): *Peer Review and Performance Indicators*. Lemma, Utrecht.
- RUSSO, J.R., BROWN, D.G. y ROTHWEILER, J.G. (1977): «A model for internal program review». *College and University*, vol. 2, pp. 288-298.
- SAEZ-CARRERAS, J. et al. (1987): *La Relación Universidad-Sociedad: El Consejo Social*. Dirección Regional de Universidad e Investigación, Murcia.
- SANTO, M.A. (1990): «Criterios de referencia sobre calidad del proceso de enseñanza-aprendizaje en la Universidad». En: *Primeras Jornadas Nacionales de Didáctica Universitaria*. Consejo de Universidades, Madrid.
- SCHILLER, A.L. (1981): «Communication in the arts and the humanities». *Journal of Research Communication Studies*, 3, pp. 47-63.
- SCOTT, P. (1984): *The Crisis of the University*. Crook Helm, Londres.
- SCULLY, M.G. (1979): «The well-known universities lead in rating of faculties' reputations». *Chronicle of Higher Education*, 19, 17, January 15, 6-7.
- SEGBERS, J.H.G. (1977): *Sociologische Onderzoeksmethoden* (Métodos de investigación en sociología). Van Gorcum, Assen.
- SEGBERS, M.S.R. y DOCHY, F.J.R.C. (1988): «Nederlands hoop in bange dagen: De output? (La esperanza de Holanda en los días oscuros: ¿resultados?)». *Universiteit & Hogeschool*, 34, PP. 120-128.
- SEGBERS, M.S.R., DOCHY, F.J.R.C. y WIJNEN, W.H.F.W. (1989): *Een Set van Prestatie-Indicatoren voor de Bestuurlijke Omgang Tussen Overheid en Instellingen voor Hoger Onderwijs* (Un conjunto de indicadores de rendimiento para las relaciones administrativas entre el gobierno y las instituciones de Educación Superior). Ministerio de Educación y Ciencia. Zoetermeer.
- SEGBERS, M.S.R.; DOCHY, F.J.R.C. y WIJNEN, W.H.F.W. (1989a): *De Bestuurlijke Hante-ring van Prestatie-Indicatoren in Het Hoger Onderwijs* (El uso de los indicadores de

- rendimiento para la Educación Superior). OTIC/Rijksuniversiteit Limburg, Heerlen/Maastricht.
- SEGERS, M.S.R.; DOCHY, F.J.R.C. y WIJNEN, W.H.F.W. (1989b): «Nog een rozenstruik met doornen? Prestatie-indicatoren in het hoger onderwijs» (¿Otro rosal con espinas? Indicadores de rendimiento en la Educación Superior). *Tijdschrift voor Hoger Onderwijs*, 7, pp. 44-52.
- SEGERS, M.S.R.; WIJNEN, W.H.F.W. y DOCHY, F.J.R.C. (1990): «Performance indicators: A new management technology for higher education?. The case of the United Kingdom, the Netherlands and Austria», en Dochy, F.J.R.C., Segers, M.S.R. y Wijnen, W.H.F.W. (eds.): *Management Information and Performance Indicators in Higher Education: An International Issue*, Van Gorcum, Assen.
- SHATTOCK, M.L. (1990): «The evaluation of the universities contribution to society». En: *Report of the IMHE Study Group on Evaluation in Higher Education*, presentado en la 10th General Conference of Members Institutions of Programme on Institutional Management en Higher Education. Paris.
- SHICHOR, D. (1970): «Prestige of sociology departments and the placing of New Ph. D.'s». *American Sociologist*, 5, May, pp. 157-160.
- SIMMS, J.R. (1983): *The Limits of Behavior*. Intersystems, Seaside (CA).
- SIMON, Ch. (1988): «Un ballon d'oxygène aux universités pour leur crédit de fonctionnement». *Le Soir*, 3 July, Bruxelles.
- SIZER, J. (1979a): «Assessing institutional performance: An overview». *International Journal of Institutional Management in Higher Education*, 3, 1, pp. 49-77.
- SIZER, J. (1979b): «Indicators in times of financial stringency, contraction and changing needs», en Billing, D. (ed.): *Indicators of Performance*. Society for Research into Higher Education, Guildford, UK.
- SIZER, J. (1982a): «Assessing institutional performance and progress», en Wagner, L. (ed.): *Agenda for Institutional Change in Higher Education*. Society for Research in Higher Education, Guildford, UK.
- SIZER, J. (1982b): «Institutional performance assessments under conditions of changing needs». *International Journal of Institutional Management in Higher Education*, 6, pp. 17-28.
- SIZER, J. (1988): «The management of institutional adaptation and change under conditions of financial stringency», en Eggins, H. (ed.): *Restructuring Higher Education*. Milton Keynes: SRHE, Open University Press.
- SIZER, J. (1989): «A critical examination of the events leading up to the UGC's grant letters dated 1st July 1981». *Higher Education*, 18, pp. 639-679.

- SIZER, J. (1989a): «Performance indicators, quality control, and the maintenance of standards in higher education», en McVicar, M.: *Performance Indicators and Quality Control in Higher Education*. Portsmouth Polytechnic, Portsmouth.
- SIZER, J. (1990): «Performance indicators and the management of universities in the U.K.: A summary of developments with commentary», en Dochy, F.J.R.C., Segers, M.S.R. y Wijnen, W.H.F.W. (eds.): *Management Information and Performance Indicators in Higher Education: An International Issue*, Van Gorcum, Assen.
- SIZER, J. (1990a): «Funding councils and performance indicators in quality assessment in the United Kingdom», en Goedegebuure, L.C.J.; Maassen, P.A.M.; Westerheijden, D.F. (eds.): *Peer Review and Performance Indicators*. Uitgeverij Lemma, Utrecht, NL.
- SIZER, J. (1990b): Artículo preparado para el grupo de la OECD/IMHE sobre indicadores de rendimiento y contenida en el *OECD/IMHE Compendium of Performance Indicators for Higher Education*, París.
- SIZER, J. y FROST, R. (1985): «Criteria for self evaluation of department research, profiles, responsible, and responsive universities research project», *Working Paper*, Loughborough University of Technology.
- SLOPER, D.W. (1982): «Recent evaluations in Australian higher education: Context and incidence». *Higher Education*, vol, 11, pp. 405-440.
- SMITH, R.; 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): «New findings on the links between college education and work». *Higher Education*, 10, pp. 615-648.
- SOLMON, L.C. (1989): «Economía de la calidad de la educación». En: T. Husen y T.N. Postlethwaite (Eds.) *Enciclopedia Internacional de la Educación*. Barcelona. M.E.C. y Vicens Vives (vol. 2, pp. 688-695).
- SOLMON, L.C.; 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*, Atherton Press, New York.
- SPAAPEN, J.B. et al. (1988): *De Moeizame Relatie Tussen Beleid en Onderzoek: Evaluatie Vijf jaar Voorwaardelijke Financiering*. Ministerio de Educación y Ciencia, Zoetermeer.
- SPAAPEN, J.B. (1989): «External Assessment of Dutch Research Programmes». En: M. Kogan (Ed.) *Evaluating higher education*. Londres: Jessica Kingsley.
- 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. (1987): «The Comité National d'Évaluation: Preliminary results of a French experiment». *European Journal of Education*, 22, 2, 123-131.
- STERN, M.E. (1990): «The new majority: Impact of older students upon the university today». Ponencia presentada en el 29.º *Seminario del Centro para la Investigación e Innovación Educativa de la OCDE*, París.
- STIGLER, G.J. (1961): «The economics of information». *Journal of Political Economy*. LXIX, pp. 213-225.
- STONER, J.A.F. y WANKEL, C. (1986): *Management*. Prentice Hall, New Jersey.
- STUFFLEBEAM, D.L. (1981): *Standards for Evaluations of Educational Programmes, Projects, and Materials*. Joint Committee on Standards for Educational Evaluation. McGraw Hill, Nueva York.
- STUFFLEBEAM, D. y SHINKFIELD, A. (1987): *Evaluación Sistemática*. Paidós-MEC, Madrid.
- 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.
- THE TIMES HIGHER EDUCATION SUPPLEMENT, (1 September 1989), 989: 1-4.
- THOMAS, H. (1990): *Education Costs and Performance: A Cost-Effectiveness Analysis*. Cassell, Londres.
- TIDBALL, M.E. y KRISTIAKOWSKI, V. (1976): «Perspective on academic women and affirmative action». *Educational Record*, 54, pp. 130-135.
- TIERNEY, W. (1988): «Organizational culture in higher education». *Journal of Higher Education*, vol. 59, 1, pp. 2-21.
- TROW, M. (1987): «The National Reports on Higher Education: A Skeptical View». *Educational Policy*, 1-4.
- UFC (UNIVERSITIES FUNDING COUNCIL) (1989): *Report on the 1989 Research Selectivity Exercise*, UFC, Londres.
- UGC (UNIVERSITY GRANTS COMMITTEE) (1987): *Report to Secretary of State on Universities' Responses to the Jarratt Report*. U.G.C., Londres.
- UGC (UNIVERSITY GRANTS COMMITTEE) (1988): *Second Annual Report on Monitoring Research Selectivity*. U.G.C., Londres.
- VANDAMENT, W.E. (1989): *Managing Money in Higher Education*. Jossey-Bass, San Francisco.
- VAZQUEZ, G. (1985): «Masificación y calidad universitaria: La suerte de la universidad

- entre el igualitarismo y el desarrollo de la excelencia». *Revista Española de Pedagogía*, 169-179, pp. 371-389.
- VOLKWEIN, J.E. (1989): «Changes in quality among public universities». *Journal of Higher Education*, vol.60, 2, pp. 136-152.
- VON LINSTONW, H. (1990): Artículo preparado para el grupo de la OECD/IMHE sobre indicadores de rendimiento y contenida en el *OECD/IMHE Compendium of Performance Indicators for Higher Education*, París.
- VROEIJENSTIJN, A.I. (1989): «Autonomy and assurance of quality, two sides of one coin». Ponencia presentada en el *International Conference on Assessing Quality in Higher Education*, Cambridge, July.
- VROEIJENSTIJN, A.I. (1991): «External quality assessment: servant of two masters?» Ponencia presentada en el Congreso sobre Mantenimiento de la Calidad en la Educación Superior, Hong Kong, julio.
- VROEIJENSTIJN, A.I. y ACHERMAN, H. (1990): «Control oriented versus improvement oriented quality assessment», en Goedegebuure, L.C.J., Maassen, P.A.M., Westerheijden, D.F. et. al. (eds.): *Peer Review and Performance Indicators: Quality Assessment in British and Dutch Higher Education*. Lemma, Utrecht.
- VSNU (Asociación para la Cooperación de las Universidades Holandesas) (1988a): *De Evaluatie van het Project Proefvisitaties* (Evaluación del proyecto piloto del comité de visitas). VSNU, Utrecht.
- VSNU (Asociación para la Cooperación de las Universidades Holandesas) (1988b): *Gids voor de Externe Kwakiteitszorg*. (Guía para Evaluaciones Externas de la Calidad, VSNU, Utrecht).
- VSNU (Asociación para la Cooperación de las Universidades Holandesas) (1988c): *Over de Kwaliteit van Het Nederlandse Universitaire Onderwijs*. (Sobre la calidad de la educación universitaria holandesa), VSNU, Utrecht.
- VSNU (Asociación para la Cooperación de las Universidades Holandesas) (1990a): *Kwaliteitszorg Binnen Het Universitaire Onderwijs*. VSNU, Utrecht, 27 Abril.
- VSNU (Asociación para la Cooperación de las Universidades Holandesas) (1990b): *Voorwaardelijke Financiering, Beoordelingsronde 1989*. VSNU, Utrecht, 8 Junio.
- VUGTH, F.A. Van (1988): «A new autonomy in european higher education?. An explanation and analysis of the strategy of self-regulation in higher education». *International Journal of Institutional Management in Higher Education*, 12, pp. 16-26.
- VUGHT, F.A. Van (1988a): «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*. Lemma, Culemborg, NL.

- VUGHT, F.A. Van (1989): «Higher education in the Netherlands», en Maassen, P.A.M. y Vught, F.A. Van (eds.): *Dutch Higher Education in Transition*. Lemma, Culemborg.
- VUGHT, F.A. Van (ed.) (1989a): *Governmental Strategies and Innovation in Higher Education*. Jessica Kingsley Publishers, Londres.
- VUGHT, F.A. Van (1991): «How evaluate quality in higher education?». Ponencia presentada en el Congreso sobre *Calidad de la Educación Universitaria*. Puerto de Santa María, Marzo.
- VUGHT, F.A. Van, y WESTERHEIJDEN, D.F. (1989): «Political behavior in higher education». Ponencia en el *CHEPS Anniversary Conference*, Enschede.
- WEBSTER, D.S. (1981): «Methods of assessing quality». *Change*, 13, Oct, pp. 20-24.
- WEBSTER, D.S. (1983): «America's highest ranked graduate schools, 1925-1982». *Change*, May/Jun, 14-24.
- WEBSTER, D.S. (1984): «Who is Jack Gourman and why is he saying all these things about my college». *Change*, Nov./Dec., pp. 14-19.
- WEERT, E. de (1990): «A macro-analysis of quality assessment in higher education». *Higher Education*, 19, pp. 57-72.
- WEST, P. (1986): «The use of performance indicators in decision making in the University of Strathclyde». Ponencia presentada en el *OECD/CERI IMHE Programme's Special Topic Workshop*, Paris, 8-10 December.
- WESTDEUTSCHE REKTORENKONFERENZ (1980): *Kennzahlenvergleich an Deutschen Vororthochschulen*. Abschlussbericht, Bonn-Bad Godesberg.
- WESTDEUTSCHE REKTORENKONFERENZ (1986): «Zur beurteilung und entwicklung der ansätze sur leistungsberwertung und —messung von hochschulen». *Stellungnahme des 149 Plenums der WRK*, 1 July.
- WESTERHEIJDEN, D.F. (1988): *Schuiven in de Oosterschelde: Besluitvorming Rond de Oosterschelde 1973-1976*. Universiteit Twente, Faculteit Bestuurskunde, Enschede.
- WESTERHEIJDEN, D.F. (1990a): «Peers, performance and power. Quality assessment in the Netherlands», en Goedegebuure, L.C.J.; Maassen, P.A.M.; Westerheijden, D.F. (eds.): *Peer Review and Performance Indicators*. Uitgeverij, Lemma, Utrecht, NL.
- WESTERHEIJDEN, D.F. (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.
- WEUSTHOF, P. (1989): «Zelfevaluatie: Wat, hoe en waartoe». *Tijdschrift voor Hoger Onderwijs*, nr. 7.

- WHITLEY, R.D. (1984): *The Intellectual and Social Organization of the Sciences*. Oxford.
- WILDAVSKY, A. (1974): *The Politics of the Budgetary Process*, 2nd ed., Boston/Toronto.
- WILLIAMS, G. (1986): «The missing bottom line», en Moodie, G.C. (ed.): *Standards and Criteria in Higher Education*. Society for Research into Higher Education, Guilford.
- WINSTON, J.A. (1984): «Notes on planning for data analysis in evaluative studies (Appendix)», en Roe, E. y McDonald, R.: *Informed Professional Judgment. A Guide to Evaluation in Post-Secondary Education*. University of Queensland Press, Brisbane.
- WISPE, L.G. (1969): «The bigger the better: Productivity, size, and turnover in a sample of psychology departments». *American Psychologist*, 24, pp. 662-668.
- WISSENSCHAFTSRAT (1985): *Empfehlungen zum Wettbewerb im Deutschen Hochschulsystem*. Cologne.
- WITTROCK, B. (1985): «An international perspective on the Leverhulme Report on Higher Education». En Blackstone, T. y Williams, G. (eds.): *Higher Education in a Time of Adversity*. Society for Research into Higher Education, Guildford, U.K.
- WOOD, L. y DAVIS, B.G. (1978): «Designing and evaluating higher education curricula». *AAHE-ERIC Higher Education Report*, n. 8, American Association for Higher Education, Washington, D.C.
- YORKE, D.M. (1987): «Indicators of institutional achievement: Some theoretical and empirical considerations». *Higher Education*, 16, pp. 3-20.
- YOUNG, K. et al. (1983): *Understanding Accreditation*, Jossey-Bass Publishers, San Francisco.
- ZIMAN, J.S. (1987): *Science in a 'Steady State': The Research System in Transition*. Science Policy Support Group, Londres.

