

Evaluación de la calidad universitaria: validez de contenido

Salvador Chacón Moscoso, José Antonio Pérez-Gil, Francisco Pablo Holgado Tello y Ángel Lara Ruiz
Universidad de Sevilla

El cambio de perspectiva sobre la educación superior en España ha posibilitado el desarrollo de nuevos marcos jurídicos que dan mayor autonomía a las universidades. Al mismo tiempo los gobiernos, tanto nacional como autonómicos, necesitan recabar información sobre cómo se gestiona dicha autonomía. Esta situación ha impulsado la necesidad de definir el constructo de calidad en la universidad y de generar procedimientos adecuados para su evaluación. El objetivo de este trabajo es por una parte, justificar cómo los sistemas de indicadores suponen una posible solución a esta problemática, y por otra analizar empíricamente la validez de contenido de un conjunto de indicadores seleccionados a partir de una revisión bibliográfica. El proceso seguido en el estudio empírico fue el siguiente: se seleccionaron indicadores que hacen referencia a distintos aspectos relacionados con la educación superior (enseñanza, investigación y gestión); se consultó a un grupo de expertos sobre la adecuación de dichos indicadores y se utilizó el índice de Osterlind (1989) como una forma de operativizar el consenso entre los expertos respecto a la idoneidad de los indicadores seleccionados para evaluar la calidad universitaria. Se presentan los indicadores obtenidos en el análisis como una alternativa posible para el estudio de la calidad universitaria.

Evaluation of quality in higher education: Content validity. The new conception of higher education in Spanish society and in its national, regional and local governments has generated laws that give more independence to universities. Because of this situation, governments need information about how universities use that independence. This context has stressed the need of defining the construct of quality in higher education and the difficulty that implementing an appropriate evaluation procedure entails. This paper has two aims: first, to justify that systems of indicators are a plausible alternative to both problems, and second, to study the content validity of a group of indicators chosen from the literature. We have carried out the following process in the empirical study: we selected some indicators related to higher education (about teaching, research, and services); we consulted experts about the congruence of these indicators, and finally we used the Osterlind's index (1989) to operationalize the consensus between experts about their judgments. We describe the quality indicators obtained in this analysis as a possible alternative to study quality in higher education.

Desde los años ochenta se han producido importantes cambios en el ámbito universitario, donde tanto los responsables políticos como los agentes sociales empiezan a percibir a la universidad de forma más crítica, planteando la necesidad de justificar sus resultados. En este contexto, se han desarrollado nuevos marcos jurídicos que han dado mayor autonomía a las universidades como estrategia para aumentar la calidad en las mismas. Como contrapartida a este aumento de autonomía surge la necesidad de recabar datos sobre lo que de hecho se lleva a cabo en el seno de las universidades (Consejo de Universidades, 1996, 1998, 1999). Esta situación ha dado lugar a que dicha autonomía haya de estar ligada a la implantación de mecanismos de evaluación institucional que reporten información sobre cómo se están utilizando los recursos aportados (Osoro, 1995; Volkwein et al, 1995).

Esta situación plantea una doble problemática: la dificultad de definir el constructo de calidad universitaria (Oakes, 1986; Mateo y Fernández, 1993; Segers y Dochy, 1996; Schargel, 1997), y la necesidad de generar mecanismos de evaluación válidos que satisfagan las necesidades de información mencionadas (Luxán, 1998; Ginés, 1999; Oliver, Rosel y Jara, 2000).

La creación de sistemas de indicadores de calidad válidos y fiables con capacidad de generalización es una alternativa que puede responder a la necesidad de estructurar el concepto de calidad universitaria desde distintas dimensiones conceptuales (metodología docente, resultados de la enseñanza, recursos, servicios, gestión, resultados de investigación...). Además, estos sistemas de indicadores pueden suponer un elemento relevante para diseñar, junto con otros aspectos, un buen sistema de evaluación de calidad universitaria (Johnes y Taylor, 1990; Van Os, 1990; Blank, 1993; Nuttall, 1995; Osoro, 1995; Segers y Dochy, 1996).

En este marco y desde un punto de vista metodológico, para la elaboración de dichos sistemas de indicadores, el análisis de la validez de contenido se torna un elemento de gran relevancia para garantizar que el conjunto de indicadores seleccionados sea una muestra relevante y representativa del constructo calidad universitaria.

El objetivo general de este trabajo es describir la validación de contenido de un sistema de indicadores de calidad universitaria desarrollado en la Universidad de Sevilla (acción específica del Plan Nacional de Evaluación de la Calidad de la Universidad de Sevilla -B.O.E. 101; 28/4/98-). Para ello, en primer lugar comentaremos brevemente el estado actual en que se encuentra la evaluación de la calidad universitaria. En segundo lugar, nos aproximaremos a la evaluación institucional y cómo se concreta en el uso de sistemas de indicadores en la educación superior. A continuación, nos ocuparemos del concepto de indicador, así como de los requisitos para su uso científico. Y por último, describiremos los resultados empíricos obtenidos en el procedimiento de validación de contenido del sistema de indicadores obtenidos en una revisión bibliográfica.

Evaluación de la calidad en educación superior

En la última década, la evaluación está cobrando un énfasis progresivo extendiéndose por áreas de aplicación que eran impensables hace poco. Entre tales campos de aplicación destacan la evaluación medioambiental, la evaluación en materia de política de inmigración, la evaluación del coste-efectividad del gasto en defensa, o, la evaluación de los programas de ayuda a los países afectados por catástrofes. Siguiendo esta tendencia, junto con el creciente interés de los ciudadanos por mejorar sus condiciones de vida y de los gobiernos por justificar sus resultados, la evaluación de las instituciones ha experimentado un importante desarrollo en los últimos años (Chelmsky, 1997).

De acuerdo con este creciente interés en conocer qué se hace en el seno de las instituciones públicas, encontramos que entre las esferas sociales y políticas se sostiene que la información relativa al funcionamiento de las universidades es relativamente escasa, y esto ocurre en una sociedad donde los distintos agentes sociales demandan información fiable sobre las instituciones que les afectan. Esta circunstancia dificulta tanto la toma de decisiones de las propias instituciones, como las decisiones tomadas en sus relaciones con los agentes estatales (Ginés, 1999).

Normalmente se aduce que la evaluación se torna difícil en el ámbito universitario por sus características peculiares (dificultades en cuanto a su definición, colectivo crítico, difícil control, genera expectativas especiales,...). En definitiva, se argumentan razones poco concretas y en las que subyace la difusa idea de seguir protegiendo la solidez de tales instituciones. Esta actitud ha provocado que se haya avanzado poco en el desarrollo de sistemas de información sobre las universidades, no encontrándose una metodología evaluativa ni objetiva, ni sistemática (Escudero, 1989).

De manera complementaria al creciente interés por conocer cómo funcionan las instituciones públicas en España, desde los años ochenta hemos asistido a importantes cambios en el ámbito universitario (Luxán, 1998). Se han desarrollado nuevos marcos jurídicos que han dado una mayor autonomía a las instituciones universitarias como estrategia para aumentar los niveles de calidad. Pero como contrapartida ha surgido la necesidad de recabar datos sobre lo que de hecho se desarrolla (Consejo de Universidades, 1996, 1998, 1999).

Si bien ni es fácil evaluar la calidad, ni hay un método ideal para ello, sin embargo es necesario evaluarla y medirla, así como desarrollar un sistema válido y fiable, aprender a utilizarlo de manera correcta y constructiva e interpretar sus resultados y medidas con el objetivo de orientarse hacia la mejora continua (Van Os, 1990; Medina, 1996).

La creación y desarrollo de sistemas de indicadores válidos y fiables con capacidad de generalización supone una alternativa que puede responder a la necesidad de estructurar y medir el concepto de calidad. Además, estos sistemas de indicadores pueden suponer un elemento relevante, que junto con otros aspectos, permitan diseñar un sistema de evaluación para dar cuenta a la sociedad del rendimiento académico y científico; proporcionar elementos de juicio para la adopción de decisiones en política universitaria; responder, desde la propia autonomía universitaria, a las exigencias internas de mejora de su calidad; y optimizar los recursos públicos (Johnes y Taylor, 1990; Van Os, 1990; Blank, 1993; Nuttall, 1995; Osoro, 1995; Segers y Dochy, 1996).

Evaluación institucional e indicadores de calidad La evaluación institucional

En España a los puntos anteriores se le intenta dar respuesta desde el desarrollo del Plan Nacional de Evaluación de la calidad de las universidades, lo que ha provocado la creación de Agencias Autonómicas de Evaluación de la Calidad (como por ejemplo, la Agencia para la Calidad del Sistema Universitario de Cataluña o la Unidad para la Calidad de las Universidades Andaluzas -UCUA-), así como la consiguiente creación de unidades de evaluación de calidad en la mayor parte de las universidades españolas. Esta situación es un signo inequívoco de que se ha asumido la mejora de la calidad de la educación superior a través de los programas de evaluación de la calidad (Valera y López, 2000).

Como comentábamos con anterioridad, esta preocupación entronca de manera directa con los nuevos desarrollos en evaluación centrados en las instituciones. Y en este sentido, si la evaluación la podemos entender como el proceso sistemático y científico por el que se llegan a emitir juicios de valor sobre el objeto evaluado (Stufflebeam, 1987; Alvira, 1991; Shadish, Cook y Leviton, 1991; Fernández-Ballesteros, 1995; Anguera y Chacón, en preparación), la evaluación institucional la podríamos delimitar igualmente y de manera muy genérica, como el proceso sistemático y científico basado en una metodología específica, por el que lleguemos a emitir juicios de valor basados en una información válida y fiable sobre las distintas dimensiones institucionales que han sido objeto de la evaluación.

Por otro lado, la evaluación institucional universitaria aun cobra más fuerza debido al cambio de perspectiva que ha sufrido. De ser concebida como un proceso burocrático, ahora es percibida de una manera más crítica como un proceso autónomo y cuya actividad se desenvuelve en una sociedad orientada a métodos de mercado (Ginés, 1991, 1999; Rodríguez, 1998). Por todo ello se hace necesario la implantación de mecanismos de evaluación institucional que garanticen que los recursos aportados se utilizan pertinentemente en función de las demandas y necesidades del estado y de la sociedad (Osoro, 1995).

La situación descrita ha propiciado que en la mayoría de los países desarrollados, tales como Austria, Grecia, Gran Bretaña, Estados Unidos, Finlandia, Francia o Suecia, se hayan puesto en marcha procesos de evaluación institucional en los que la construcción de sistemas de indicadores juega un importante papel (Osoro, 1995). De esta forma, el desarrollo de indicadores válidos se ha convertido en un tema de gran interés, de ello da cuenta el gran volumen de reuniones científicas que se han organizado en los últimos años y que de manera directa o indirecta tratan el concepto de indicador y su relación con la educación superior (por

ejemplo, Proyecto de la OCDE sobre indicadores, 1987, 1989, 1991; Symposium on Developing State Indicator System: From Theory to Practice. San Francisco, 1989; Indicadores en la universidad: Información y decisiones. León, 1999; Jornadas sobre calidad Universitaria: Indicadores de calidad para la universidad: su definición y uso. Universidad de Cantabria, 2000).

A continuación, realizamos una breve sinopsis sobre qué entendemos por indicador en general, y qué aspectos metodológicos pueden ser relevantes en la construcción de un sistema de indicadores. Posteriormente ilustraremos el proceso que hemos seguido mediante la acción específica mencionada anteriormente «Modelos de evaluación e indicadores de calidad» adscrita al Plan Nacional de Evaluación de la Calidad de la Universidad de Sevilla (B.O.E. 101; 28/4/98), para el análisis de la validez de contenido de una selección exploratoria de indicadores.

Indicadores

Desde un ámbito metodológico, cabe destacar el carácter mediador que tienen los indicadores entre lo teórico y lo empírico, entre los «constructos teóricos» y el «mundo externo», es decir, mientras que la definición conceptual establece el significado en términos abstractos, los indicadores se centran sobre los aspectos observables y empíricamente detectables (Anguera, 1989). En este sentido se entiende que los indicadores son un instrumento al servicio de la ciencia y de la conceptualización, posibilitando investigar los conceptos empíricamente y estableciendo cauces de operativización. Desde esta perspectiva, se podría decir que lo indicado por el indicador es el concepto, en tanto que se pretende la operacionalización de la estructura de un concepto (Casas, 1989).

Por otro lado, es indudable que los indicadores intentan medir sus conceptos de referencia lo más significativamente posible, y una forma de incrementar esta significatividad es a través de la combinación de indicadores entre sí de forma que surjan redes o sistemas a través de los cuales delimitar empíricamente lo indicado. De acuerdo con esto, entendemos que la relación establecida entre el indicador y lo indicado es de carácter probabilístico, por tanto cada constructo teórico ha de estar representado por múltiples indicadores, de forma que tal constructo quedará entonces delimitado por lo que de hecho tienen los indicadores en común (McCallum, 1995). En este sentido podemos considerar el constructo como un función descriptora o de clasificación, según sea el caso, obtenida a partir de las relaciones existentes entre los indicadores estudiados (Pérez-Gil y Moreno, 1991).

En trabajos clásicos, como los de Lazarsfeld (1973) o Thorndike y Hagen (1989), se plantean las secuencias generales que permiten la transformación de conceptos en indicadores (Anguera, Pérez-Gil y Chacón, en preparación). En general constan de las siguientes fases:

1. Definición del constructo. Identificación y definición del constructo que se pretende medir.
2. Especificación. Determinación del conjunto de operaciones por medio de las cuales se puede manifestar y percibir el constructo.
3. Construcción de indicadores. Establecimiento de un conjunto de procedimientos o definiciones para convertir las observaciones en enunciados cuantitativos de grado y cantidad.

El objetivo final es establecer juicios razonables sobre el grado en el que las evidencias aportadas por el sistema de indicadores

permiten establecer acciones basadas en el modelo de medida desarrollado (Messick, 1989). En este sentido, no cabe plantearse la cuestión de si el sistema de indicadores es válido o no en sí mismo, sino si es válido para el propósito que pretendemos.

Requisitos para su uso científico

El uso de indicadores ha de reunir una serie de requisitos (Casas, 1989). En este sentido, los indicadores pueden adoptar los roles de descriptor de la realidad, de instrumento guía para la acción, o de instrumento que hipotetiza relaciones dentro de un proceso de investigación.

Los roles, y en consecuencia las funciones que los indicadores desempeñan, han de hacerlo con validez. Podemos considerar la validez como un criterio de calidad relacionado con la adecuación de los indicadores al objetivo que suscitó su aplicación, es decir, que el instrumento que conforma los indicadores tenga utilidad para medir el concepto que pretende medir.

De acuerdo con ello, es necesario dotar a los indicadores de evidencias razonables de validez. En un marco general de estudio de la validez hay tres procedimientos habituales para obtener evidencias de validez, denominados de contenido, de constructo y predictiva (Muñiz, 1998).

La validez de contenido trata de garantizar que los indicadores utilizados constituyan una muestra relevante y representativa del conjunto de todos los posibles indicadores del concepto que pretenden medir. De manera convencional la validez de contenido es definida en términos de representatividad de la muestra de indicadores, de esta forma, se entiende la validez de contenido como el grado en el que los indicadores seleccionados representan de forma adecuada nuestro constructo de interés. En términos estadísticos un sistema de indicadores podría ser considerado como representativo si la muestra es conformada aleatoriamente. Sin embargo, uno de los principales problemas está en delimitar cuál o cuáles son esos constructos referentes poblacionales; de ahí que una aproximación más pragmática es la que entiende por representatividad aquel conjunto de indicadores que «re-presentan» o «re-producen» las características esenciales del constructo de nuestro interés (Suen, 1990).

La adecuación de los indicadores a los criterios de validez de contenido implica por un lado, la construcción de los mismos con base al marco teórico en el que se apoya el concepto a medir, esto es, ha de definirse el concepto de un modo claro y preciso. Y por otro lado, un procedimiento que permita, de una manera operativa, sintetizar las valoraciones que realice un grupo de expertos (investigadores, académicos y/o profesionales) sobre la adecuación de los indicadores seleccionados para medir una determinada dimensión de calidad (Crocker y Algina, 1986; Osterlind, 1989). En este sentido, la información resultante de la valoración de los diferentes expertos puede ser operativizada utilizando el índice de congruencia entre cada indicador y el dominio que mide (Rovineili y Hambleton, 1977). Su expresión formal es la que se muestra en la ecuación 1:

$$I_{ik} = \frac{(N-1) \sum_{j=1}^n X_{ijk} + N \sum_{j=1}^n X_{ijk} - \sum_{j=1}^n X_{ijk}}{2(N-1)n}$$

Ecuación 1. Índice de congruencia

Donde;

X_{ijk} = Valoración del indicador i en el dominio k por el juez j .

N = Número de dominios que se contemplan en el instrumento.

n = Número de jueces que valoran el indicador.

Es interesante señalar que mediante este tipo de índices de congruencia se aporta una operativización del grado de consenso que puedan manifestar los distintos implicados (en este caso en el sistema universitario). Sin la existencia de un consenso mínimo acerca del tipo de información a recoger y la utilidad que esta información pueda tener, cualquier sistema de indicadores tendrá pocas posibilidades de «subsistir».

Por otra parte, es necesario considerar que la validez de contenido, de constructo, y de criterio, son el resultado de distintas aproximaciones al constructo de validez (Suen, 1990). Estas conceptualizaciones sobre la validez aportan las evidencias para aproximarnos al objetivo perseguido, es decir que el sistema de indicadores propuesto sea útil para nuestros propósitos.

Indicadores sobre calidad universitaria: validez de contenido

A continuación describimos el estudio empírico realizado para analizar la validez de contenido de una muestra de indicadores de calidad universitaria.

Metodología

Procedimiento general empleado

El procedimiento empleado en este estudio preliminar puede esquematizarse a través de los siguientes puntos:

En primer lugar, realizamos una revisión bibliográfica sobre los documentos de mayor relevancia propuestos por el Consejo de Universidades en el Plan Nacional de Evaluación de la Calidad de las Universidades (1996, 1998, 1999) para la evaluación de las diferentes titulaciones; también fueron considerados otros documentos de interés. En concreto, los documentos que hemos manejado se describen en el apartado sobre la muestra.

En segundo lugar, a partir de las directrices generales propuestas por el Plan Nacional de Evaluación de la Calidad de las Universidades (op. cit.) adoptamos como referentes de evaluación las áreas de enseñanza, investigación y gestión. Estas áreas son las que propone el Consejo de Universidades para la evaluación de las titulaciones, y además se encuentran de manera mayoritaria en los diferentes documentos manejados.

En tercer lugar, procedimos a sistematizar cada una de dichas áreas en función de una serie de dimensiones. Dichas dimensiones fueron extraídas a partir de los aspectos temáticos contemplados en cada una de las áreas propuestas.

En cuarto lugar, procedimos a realizar una primera selección exploratoria de indicadores en función de su relevancia en cuanto al contenido y de su viabilidad de aplicación. Dicho criterio lo operativizamos al considerar de manera preferente los documentos existentes en torno al Plan Nacional de Evaluación de la Calidad de las Universidades. Posteriormente, los indicadores en cuestión fueron asignados a cada una de las dimensiones propuestas para cada área.

En quinto lugar, realizamos un muestreo de un grupo de expertos para analizar la validez de contenido de los indicadores seleccionados. La muestra estuvo compuesta por expertos tanto en materia de evaluación y medida, como en temas relacionados con la

calidad en la educación superior. La muestra fue intencional y realizada a nivel nacional. Estuvo compuesta por un total de 11 jueces en el área de enseñanza; y 6 jueces en el área de investigación y gestión.

En sexto lugar, remitimos un cuestionario al grupo de expertos seleccionados. En el cuestionario se solicitaba la colaboración de los expertos para que valorasen en una escala de 3 puntos el grado de «idoneidad-congruencia» en la asignación de los distintos indicadores a las distintas dimensiones de calidad planteadas (1 -bajo grado de idoneidad-; 2 -grado medio de idoneidad-; y 3 -alto grado de idoneidad-).

En séptimo y último lugar, procedimos a la selección de los indicadores con mayor puntuación ($I_{ik} > 0,5$) en el índice de Osterlind (1989) para formar parte del sistema exploratorio de indicadores propuestos.

Muestra

Para la estructuración del modelo de trabajo, así como para la selección preliminar de indicadores tomamos como referencia el conjunto de publicaciones que el Consejo de Universidades (1996, 1998, 1999) propone como relevantes a la hora de abordar la evaluación de las titulaciones, así como algunas otras publicaciones de interés. En concreto, fueron las siguientes: Alonso, Echeverría y Martín (1999); Báez, Guarro, Miró, Peñate, Pérez y San Luis (1994); Botril y Borden (1994); Casanueva, Periañez y Rufino (1998); De Miguel, (1999); Escudero (1999); Escudero, Garcés, Palacián, Blanco y Bueno (1996); Guerra, Rueda, Vázquez y Lecue (1999); Johnes y Taylor (1990); Osoro (1995); Roca y Gil (1998); Tricio, Bol, Vilorio, Quindós y Serna (1999). La relación de indicadores que conformaron la población inicial a partir de la cual se realizó la selección de la muestra de indicadores analizados en este trabajo se encuentra en Holgado (2000).

Resultados

Como resultado de la revisión bibliográfica, encontramos que en la mayor parte de los documentos revisados la evaluación de la calidad se realiza desde las tres áreas ya mencionadas de enseñanza, investigación y gestión. A su vez, estas tres áreas se descomponen en general en determinadas dimensiones más específicas. Concretamente el área de enseñanza se descompone en las siguientes ocho dimensiones: a) datos descriptivos generales de la universidad; b) datos descriptivos generales de las titulaciones; c) infraestructura y recursos de las titulaciones; d) aspectos sobre las asignaturas de las titulaciones; e) metodología docente; f) resultados de la enseñanza; g) servicios; y h) otros aspectos relacionados con la enseñanza. El área de investigación se descompone en tres dimensiones, éstas son: a) recursos en general; b) fuentes externas; y c) producción. Por último, el área de gestión está estructurada en las dimensiones de: a) admisión; b) gestión en general, en la titulación; c) recursos humanos, organización y control; y d) servicios de información y documentación.

Para cada dimensión se propone la recogida de determinados datos que se toman como indicadores significativos. Por ejemplo, dentro del área de enseñanza encontramos la dimensión «resultados de la enseñanza», donde a su vez, entre otros aspectos, se realiza el registro de indicadores como la «tasa de rendimiento».

Los resultados que hemos obtenido los presentamos a continuación a través de un bloque de tablas para cada área de evalua-

ción (enseñanza, investigación y gestión). En cada tabla se presentan los indicadores más representativos de las dimensiones particulares de cada una de dichas áreas. Cada tabla contiene una primera columna con la denominación del indicador analizado. Una segunda columna donde aparece el valor del índice de Osterlind en el indicador en cuestión. Y una tercera columna donde aparece el número de expertos que valoraron el indicador. En concreto, presentamos los indicadores que han obtenido una mejor puntuación en cuanto a su grado de idoneidad ($I_{ik} > 0,5$) y que por tanto han sido seleccionados entre los indicadores propuestos.

En la tabla 1, respecto a la dimensión A (datos descriptivos generales de la universidad), se aprecia que aquellos indicadores que hacen referencia a los recursos humanos (alumnos, profesores y personal de administración y servicios) con los que cuenta la universidad, parecen ser los de mayor consenso en cuanto a su representatividad con la dimensión propuesta. Por otro lado, el número de titulaciones de ciclo largo es el indicador que mayor valor obtiene. Esta circunstancia quizás pueda ser debida a la asociación existente entre las universidades de mayor tradición y el número de titulaciones de ciclo largo que ofertan.

<i>Tabla 1</i> Dimensiones sobre calidad en la enseñanza universitaria		
INDICADOR	I_{ik}	N
<i>Dimensión A: Datos Descriptivos Generales de la Universidad</i>		
Nº de estudiantes de 1 ^{er} ciclo	0,545	11
Nº de estudiantes de 2 ^o ciclo	0,545	11
Nº de titulaciones de ciclo corto	0,545	11
Nº de titulaciones de ciclo largo	0,636	11
Nº de personal docente e investigador (P.D.I)	0,545	11
Nº de personal P.A.S.	0,545	11
Nº de estudiantes de 3 ^{er} ciclo	0,545	11
<i>Dimensión B: Datos Descriptivos de cada Titulación</i>		
Nº de estudiantes de 1 ^{er} ciclo	0,545	11
Nº de estudiantes de 2 ^o ciclo	0,545	11
Nota mínima y media de acceso	0,727	11
Nº de P.D.I.	0,636	11
<i>Dimensión C: Infraestructura y Recursos de la Titulación</i>		
Dotación de la biblioteca	0,909	11
Dotación de medios audiovisuales e informáticos	0,727	11
Dotación de los laboratorios	0,818	11
Recursos financieros por partidas en material docente, bibliográfico, otros	0,545	11
<i>Dimensión D: Sobre las Asignaturas de la Titulación</i>		
Estructura del plan de formación	0,909	11
Nº de alumnos por asignatura	0,545	11
Criterios de asignación de profesores a asignaturas por parte de los departamentos	0,636	11
Carga docente del profesorado	0,636	11
Ratio asignaturas/profesor	0,636	11
<i>Dimensión E: Metodología Docente</i>		
Tipo y contenido de exámenes	0,6	10
Tiempo dedicado a la dirección de trabajos, consulta, repasos, otros	0,818	11
<i>Dimensión F: Resultados de la Enseñanza de la Titulación</i>		
Tasa de éxito (Graduación)	0,727	11
Tasa de retraso (Graduación)	0,545	11
Tasa de abandono en el primer año	0,545	11
Tasa de éxito (Rendimiento)	0,727	11
Tasa de rendimiento	0,545	11
<i>Dimensión G: Servicios de la Titulación</i>		
Existe algún mecanismo de información a los alumnos de nuevo ingreso. Cuál	0,8	10
<i>Dimensión H: Otros Aspectos Relacionados con la Enseñanza</i>		
Existen cursos para los alumnos en técnicas de estudio. Cuáles	0,6	10

Respecto a la dimensión B (datos descriptivos de cada titulación), encontramos cierta coherencia con la anterior, es decir, los recursos humanos con los que cuenta la titulación parece que es un indicador bastante representativo de la dimensión propuesta. Sin embargo, observamos que el nivel académico de los estudiantes de nuevo ingreso, medido por la nota de acceso, es el indicador de mayor consenso.

Respecto a la dimensión C (infraestructura y recursos de la titulación), la dotación de la biblioteca aparece como el indicador más representativo para medir la infraestructura y recursos de la titulación. A su vez, la dotación de laboratorios y de medios audiovisuales también muestran su gran relevancia para medir la calidad de la infraestructura y los recursos.

En cuanto a la dimensión D (sobre las asignaturas de la titulación), la estructura del plan de estudios se muestra como el indicador más representativo. Por otro lado, determinados aspectos relacionados con el profesorado, también son aspectos a tener muy en cuenta.

En relación con la dimensión E (metodología docente), el indicador más representativo es el que viene dado por el tiempo medio dedicado a la tutorización de actividades paralelas a la docencia.

Con respecto a la dimensión F (resultados de la enseñanza de la titulación), las tasas de éxito tanto en graduación, como en rendimiento, aparecen como los indicadores más representativos.

En relación con la dimensión G (servicios de la titulación), sólo aparece un indicador como representativo, en concreto el que hace referencia a la existencia de mecanismos de información. A pesar de ello y considerando que la relación entre los constructos y los indicadores ha de tratarse desde una perspectiva probabilística, sería recomendable que esta dimensión fuera medida por más de un indicador, pues los servicios que presta la titulación no se tienen por qué circunscribir de manera única a los mecanismos de información a los alumnos de nuevo ingreso. A pesar de ello, ninguno de los expertos consultados propuso ningún otro indicador alternativo para incluir en esta dimensión.

En último término en relación con la dimensión H (otros aspectos relacionados con la enseñanza), la existencia de cursos en técnicas de estudio es el único indicador que aparece como rele-

vante. De igual forma, en este caso habría que arbitrar procedimientos para aumentar el número de indicadores propuesto.

En la tabla 2, con respecto a la dimensión A (recursos en general), el número de becas y el número de proyectos de investigación solicitados y concedidos son los indicadores que destacan. Con respecto a la dimensión B (fuentes externas de la titulación), el indicador más representativo es el número de contratos de investigación y cantidades contratadas por año. En cuanto a la dimensión C (producción en la titulación), es donde por primera vez encontramos un indicador con un consenso absoluto. En concreto es el que hace referencia al número de publicaciones diferentes que produce el conjunto de la titulación. A su vez, observando los indicadores representativos de esta última dimensión encontramos por un lado, aquellos que hacen referencia a la producción estrictamente científica de la titulación en particular, y por otro lado, otros indicadores tales como el peso de la universidad, que tratan de ponderar la importancia de cada universidad con respecto al resto en cuanto a dicha producción científica.

En la tabla 3, con respecto a la dimensión A (admisión en la titulación), lo que parece determinar la calidad es la existencia de información activa y personalizada durante el proceso de admisión, así como que dicha admisión se rija por un sistema informatizado. En cuanto a la dimensión B (gestión general en la titulación), la existencia de una base de datos informatizada para la gestión de expedientes es clave. Además, la dotación que se dedica a los procesos de gestión y el tiempo medio de espera para los distintos procedimientos, son elementos de vital importancia para medir la calidad. Con respecto a la dimensión C (recursos humanos, organización y control), encontramos que la existencia de directrices claras y explícitas son aspectos muy a tener en cuenta. En último término, en cuanto a la dimensión D (servicios de información y documentación), algunos aspectos relacionados con la gestión de las bibliotecas son los más representativos de la calidad de dichos recursos. Así por ejemplo, el presupuesto invertido por alumno y año en la biblioteca, el grado de informatización de los fondos o la existencia de acceso directo a los mismos, son indicadores que obtienen un consenso absoluto.

Tabla 2
Dimensiones sobre calidad en la investigación universitaria

INDICADOR	I _{ik}	n
<i>Dimensión A: Recursos en General</i>		
Nº de becas de investigación solicitadas y concedidas	0,833	6
Nº de proyectos de investigación solicitados y concedidos	0,833	6
Tasa de éxito becas de investigación	0,666	6
Tasa de éxito proyectos de investigación	0,666	6
<i>Dimensión B: Fuentes Externas de la Titulación</i>		
Nº de contratos de investigación y cantidades contratadas por año	0,8	5
Otras ayudas a la investigación por tipo de entidad subvencionadora	0,666	6
<i>Dimensión C: Producción en la Titulación</i>		
Nº de tesis doctorales	0,666	6
Nº de ponencias presentadas a congresos internacionales	0,666	6
Nº de artículos en revistas científicas	0,833	6
Nº de publicaciones según su tipo	1	5
Peso de la Universidad	0,666	6
Índice de actividad universitaria	0,666	6

Tabla 3
Dimensiones sobre calidad en la gestión universitaria

INDICADOR	I _{ik}	n
<i>Dimensión A: Admisión en la Titulación</i>		
Existencia de un proceso de admisión informatizado	0,833	6
Existencia de información activa y personalización del proceso de admisión	1	6
<i>Dimensión B: Gestión general en la Titulación</i>		
Existencia de un proceso de matriculación informatizado. Fases del proceso	0,8	5
Dotación de recursos materiales para la gestión (ordenadores, limpieza, otros)	0,833	6
Tiempo medio de espera para consultas y plazo de expedición de certificaciones.	0,833	6
Existencia de base de datos de alumnos informatizada para la gestión de expedientes	1	6
<i>Dimensión C: Recursos Humanos, Organización y Control</i>		
Existencia de un órgano o unidad de control interno. Cuál	0,8	5
Existencia de criterios establecidos para la contratación o renovación profesorado	1	5
Existencia de manual de funciones de puestos para el P.A.S.	0,666	6
Existencia de manuales de procedimiento de las acciones administrativas	1	6
Existencia de guías para el usuario	1	6
<i>Dimensión D: Servicios de Información y Documentación</i>		
Existencia de unidades de información administrativa. Nº de ventanillas abiertas	1	6
Existencia de unidades específicas de información administrativa I+D. Cuáles	0,8	5
Existencia de centros de documentación. Cuáles	0,8	5
Existencia de biblioteca central	0,833	6
Presupuesto invertido por alumno y año en la biblioteca	1	5
Existencia de acceso directo a fondos bibliográficos	1	6
Existencia de sistemas de control de préstamos y depósitos. Cuál	0,6	5
Grado de informatización de los fondos bibliográficos	1	6

A modo de reflexión final

Plantear como objetivo la evaluación y mejora de la calidad universitaria supone un problema difícil de abordar, ya no sólo por la multiplicidad de perspectivas desde las que conceptualizar y delimitar la problemática, sino que además nos estamos refiriendo a un ámbito de intervención inestable, donde la toma de decisiones está influida por intereses, a veces, contrapuestos.

En este contexto de trabajo disponer de información de calidad sobre distintas dimensiones de interés en el ámbito universitario, se convierte en un recurso imprescindible para plantear un acercamiento a esa meta de difícil concreción como es el aumento de la calidad universitaria. Plantear sistemas de indicadores, junto con otros procedimientos de registro y evaluación, puede suponer una posible respuesta a esa necesidad de disponer de una información relevante para la evaluación y mejora de la calidad universitaria.

Consideramos que una de las principales virtudes del planteamiento defendido en este trabajo es que aúna de una forma operativa, sintética y directa la problemática de coordinar un marco teórico de interés, el consenso de distintos implicados y la utilidad de la información a registrar mediante los indicadores seleccionados.

En último término, enfatizar que hemos descrito el primer paso a considerar antes de abordar la validez de constructo y de criterio, para seguidamente estudiar las posibilidades de generalización. En este sentido, animamos a cualquier lector implicado en la temática a que utilice, en la medida de sus posibilidades, los indicadores propuestos. Sería de esta forma como podremos avanzar en la validación y generalización de un «buen» sistema de indicadores de calidad universitaria que pueda ser de utilidad para los distintos implicados en el sistema universitario.

Referencias

- Alonso, J., Echeverría, M^ªJ. y Martín, S. (1999). La gestión de las bibliotecas universitarias: Indicadores para su evaluación. En Consejo de Universidades (comp.), *Indicadores en la universidad: Información y decisiones* (pp. 479-491). Madrid: MEC.
- Alvira, F. (1991). *Metodología de la evaluación de programas*. Madrid: CIS.
- Anguera, M.T. (1989). Innovaciones en la metodología de la evaluación de programas. *Anales de Psicología*, 5, 3-42.
- Anguera, M.T. y Chacón, S. (en preparación). Aproximación conceptual. En M.T. Anguera (coor.), *Evaluación de programas sociales y sanitarios: Un abordaje metodológico*. Madrid: Síntesis.
- Anguera, M.T., Pérez-Gil, J.A. y Chacón, S. (en preparación). Instrumentos de evaluación. En M.T. Anguera (coor.), *Evaluación de programas sociales y sanitarios: Un abordaje metodológico*. Madrid: Síntesis.
- Báez, B., Guarro, A., Miró, M., Peñate, W., Pérez, F. y San Luis, C. (1994). *Evaluación y mejora de la calidad docente en la Universi-*

- dad de La Laguna: Una propuesta de actuación. Universidad de La Laguna.
- Blank, R.K. (1993). Developing a system of educational indicators: Selecting, implementing, and reporting indicators. *Educational Evaluation and Policy Analysis*, 15 (1), 65-80.
- Botril, K.V. y Borden, V.M.H. (1994). Using performance indicators to guide strategic decision making. Appendix: Examples from de literature. *New Directions for Institutional Research*, 82, 107-119.
- Casanueva, C., Periañez, R. y Rufino, J.I. (1998). *Calidad percibida por el alumno en el servicio docente universitario: Desarrollo de una escala de medición*. Documento interno de trabajo. Universidad de Sevilla.
- Casas, F. (1989). *Técnicas de investigación social: Los indicadores sociales y psicosociales. Teoría y práctica*. Barcelona: PPU.
- Chelmsky, E. (1997). The coming transformations in evaluation. En E. Chelmsky y W.R. Shadish (comp.), *Evaluation for the 21st century: A handbook* (pp. 1-26). Londres: Sage Publications.
- Consejo de Universidades (1996, 1998, 1999). *Plan nacional de evaluación de la calidad de las universidades, guía de evaluación*. Madrid: MEC.
- Crocker, L. y Algina, J. (1986). *Introduction to classical and modern test theory*. Florida: Holt Reinhardt and Winston.
- De Miguel, M. (1999). La evaluación de la enseñanza. Propuesta de indicadores para las titulaciones. En Consejo de Universidades (comp.), *Indicadores en la universidad: Información y decisiones* (pp. 413-430). Madrid: MEC.
- Escudero, T., Garcés, R., Palacián, E., Blanco, F. y Bueno, C. (1996). *Bases para un plan estratégico de la Universidad de Zaragoza: Encuesta sobre acciones estratégicas*. Zaragoza: ICE.
- Escudero, T. (1989). Aproximación pragmática a la evaluación de la universidad. *Revista de Investigación Educativa*, 7 (13), 93-112.
- Escudero, T. (1999). Indicadores de rendimiento académico: Una experiencia en la Universidad de Zaragoza. En Consejo de Universidades (comp.), *Indicadores en la universidad: Información y decisiones* (pp. 251-262). Madrid: MEC.
- Fernández-Ballesteros, R. (1995). *Evaluación de programas: Una guía práctica en ámbitos sociales, educativos y de salud*. Madrid: Síntesis.
- Ginés, J. (1991). La evaluación de las instituciones universitarias. *Revista de Investigación Educativa*, 9 (17), 27-48.
- Ginés, J. (1999). Indicadores y decisiones en las universidades. En Consejo de Universidades (comp.), *Indicadores en la universidad: Información y decisiones* (pp. 19-30). Madrid: MEC.
- Guerra, C., Rueda, E., Vázquez, P. y Lecue, M. (1999). Un sistema de indicadores para analizar el perfil de titulaciones y departamentos en el ámbito de la universidad generalista. En Consejo de Universidades (comp.), *Indicadores en la universidad: Información y decisiones* (pp. 157-185). Madrid: MEC.
- Holgado Tello, F.P. (2000). *Evaluación e indicadores de calidad universitaria. Validez de contenido*. Trabajo de investigación no publicado. Departamento de Psicología Experimental. Universidad de Sevilla.
- Johnes, J. y Taylor, J. (1990). *Performance indicators in higher education*. Londres: Open University Press.
- Lazarsfeld, P. (1973). De los conceptos a los índices empíricos. En R. Boudon y P. Lazarsfeld (eds.), *Metodología de las ciencias sociales*. (pp. 35-46). Barcelona: Laia.
- Luxán, J.M. (1998). La evaluación de la universidad en España. *Revista de Educación*, 315, 11-28.
- Mateo, M.A. y Fernández, J. (1993). Dimensiones de la calidad de la enseñanza universitaria. *Psicothema*, 5 (2), 265-275.
- McCallum, R.C. (1995). Model specification, procedures, strategies and related issues. En R.H. Hoyle (comp.), *Structural equation modeling: Concepts, issues and applications* (pp. 16-37). Londres: Sage Publications.
- Medina, M.E. (1996). Evaluación de la calidad asistencial en Servicios Sociales. *Intervención Psicosocial*, 5 (14), 23-42.
- Messick, S. (1989). Validity. En Linn, R. E. (comp), *Educational measurement* (pp. 13-102). National Council of Measurement in Education. Series on Higher Education Oryx Press.
- Muñiz, J. (1998). La medición de lo Psicológico. *Psicothema*, 10 (1), 1-21.
- Nuttall, D.L. (1995). Choosing indicators. En R. Murphy y P. Broadfoot (eds), *Effective assessment and the improvement of education – A tribute to Desmond Nuttall* (pp. 214-235). Londres: Falmer Press.
- Oakes, J. (1986). *Educational indicators: A guide for policymakers*. Wisconsin: Center for Policy Research in Education.
- Oliver, J.C., Rosel, J. y Jara, P. (2000). Modelos de regresión múltiple: Aplicación en psicología escolar. *Psicothema*, 12 (3), 487-494.
- Osoro, J.M. (1995). *Los indicadores de rendimiento en la evaluación institucional universitaria*. Zaragoza: ICE.
- Osterlind, S.J. (1989). *Constructing tests items*. Boston: Kluwer Academic Publishers.
- Pérez-Gil, J.A. y Moreno, R. (1991). Una validación de la interpretación no causal del análisis factorial. *Currículum, extra 1/2*, 225-229.
- Roca, V. y Gil, M.T. (1998). *Identificación de los atributos de calidad de servicio en las universidades públicas: El caso de la Universidad Jaime I*. Documento interno de trabajo: Universidad Jaime I.
- Rovinelli, R.J. y Hambleton, R.J. (1977). On the use of content specialist in the assessment of criterion-referenced test item validity. *Dutch Journal of Educational Research*, 2, 49-60.
- Rodríguez, S. (1998). El proceso de evaluación institucional. *Revista de Educación*, 315, 45-65.
- Schargel, F.P. (1997). *Cómo transformar la educación a través de la gestión de calidad total: Una guía práctica*. Madrid: Díaz de Santos.
- Shadish, W.R., Cook, T. y Leviton, L. (1991). *Foundations of program evaluation. Theories of practice*. Nueva York: Sage Publications.
- Segers, M. y Dochy, F. (1996). Quality assurance in higher education: Theoretical considerations and empirical evidence. *Studies in Educational Evaluation*, 22, 115-137.
- Stufflebeam, D.L. (1987). *Evaluación sistemática. Guía teórica y práctica*. Madrid: Paidós-MEC.
- Suen, H. (1990). *Principles of tests theories*. New Jersey: Hillsdale.
- Thorndike, R.L. y Hagen, E.P. (1991). *Measurement and evaluation in psychology and education*. (5ª edición). Nueva York: Wiley.
- Tricio, V., Bol, A., Vilorio, R., Quindós, M^a.P. y Serna, A. (1999). Una experiencia departamental en la evaluación de la calidad y adaptación de los indicadores para la investigación, enseñanza y gestión. En Consejo de Universidades (comp.), *Indicadores en la universidad: Información y decisiones* (pp. 231-249). Madrid: MEC.
- Valera, A. y López, J.A. (2000). Metodología de la evaluación de la enseñanza universitaria. *Psicothema*, 12 (suplemento 2), 553-556.
- Van Os, W. (1990). Strategies for quality assessment: The human factor. *Revista Española de Pedagogía*, 186, 281-293.
- Volkwein, J., Farmer, D., Fernández, T., Hernández, D., Lee, M., Nettles, N. y Patton, G. (1995). *Framework for outcomes assessment*. Albany: Commission on Higher Education.

Aceptado el 17 de noviembre de 2000