

# LOS ESTUDIANTES COMO EVALUADORES DE LA DOCENCIA Y DE LOS PROFESORES: NUESTRA EXPERIENCIA

TOMAS ESCUDERO ESCORZA

## RESUMEN

En este trabajo se describen las experiencias de evaluación de la docencia y de los profesores, por los estudiantes de la Universidad de Zaragoza en las últimas décadas. Se analizan, asimismo, las virtudes y limitaciones de este procedimiento de evaluación.

## ABSTRACT

The aim of this work is to explain the experiences of evaluation about teaching and teachers by the students of the University of Zaragoza in the last decades. The advantages and restrictions of this procedure are also analysed.

## PALABRAS CLAVE

Universidad de Zaragoza, Evaluación, Docencia y profesores.

## KEYWORDS

University of Zaragoza, Evaluation, Teaching and teachers.

Me pide la organización del IX Congreso de Formación del Profesorado que presente y analice la experiencia evaluativa de la Universidad de Zaragoza, y más específicamente, la experiencia relacionada con la evaluación de la docencia y de los profesores por parte de los estudiantes.

Como es lógico, en las siguientes páginas intento responder lo mejor posible a la invitación concreta que se me hace, pero planteando mi análisis desde una perspectiva de alcance más amplio que la específica experiencia que ha tenido y tiene lugar en mi universidad, aunque esta sea el elemento central sobre el que se conduce la reflexión.

También quiero señalar que la experiencia evaluativa en la Universidad de Zaragoza va mucho más allá de la evaluación con las encuestas a los alumnos, sobre todo en los últimos años, tras nuestra integración en el Plan Nacional de Evaluación de la calidad de las Universidades y la puesta en marcha de un Programa de Evaluación global de la Universidad.

Pero nadie puede negar la importancia que tiene en todo este proceso la recogida de la opinión valorativa de los alumnos, aunque pensamos que debe ser integrada como un elemento de lo que denominamos evaluación institucional, y no aislarlo como si se tratara de algo al margen, aunque paralelo, del resto de acciones evaluadoras.

## 1. LAS ENCUESTAS A LOS ESTUDIANTES: USO Y SIGNIFICADO

Es cierto que, actualmente, una amplia mayoría de expertos en evaluación institucional defienden la diversificación y complementación de procedimientos y fuentes en la evaluación, algo que está incidiendo de manera decisiva en los modelos de evaluación que emplean las universidades, pero no es menos cierto que las encuestas de valoración de los estudiantes son el procedimiento de uso más frecuente y generalizado y en demasiadas ocasiones el único que se emplea de manera sistemática.

Hace casi una década nos decía Ory (1991) que tres de cada cuatro instituciones de educación superior americanas y la práctica totalidad de las grandes, utilizaban las encuestas a los estudiantes. Hoy en día, el fenómeno está todavía más extendido y no es fácil encontrar una universidad importante en el mundo que, de una u otra forma, no utilice este procedimiento de evaluación de su enseñanza.

Pero esta utilización masiva no ha evitado la controversia alrededor de este procedimiento, que sigue teniendo muchos detractores, bien es cierto que más veces por su utilización descontextualizada o indebida que por sus debilidades o condicionantes técnico-metodológicos.

También hay que señalar que, con sus virtudes y sus defectos, las encuestas son el procedimiento que tiene tras de sí más estudio, más análisis científico y más soporte metodológico y de resultados contrastado. Como nos decían hace tiempo Dowell y Neal (1982), cualquier otro procedimiento tiene tras de sí muchas más preguntas metodológicas por resolver. Por ejemplo la evaluación por pares (colegas) y expertos, también de gran utilización, tienen mucho menos soporte empírico que la fundamentación que las encuestas a los estudiantes que, durante bastantes años, han sido un tema estrella para los investigadores y las revistas especializadas (Marsh y Dunkin, 1992).

Este panorama no implica que estemos ante un procedimiento sin limitaciones técnicas y sin contradicciones, sino que es válido y útil cuando se utiliza concretamente. En la actualidad, el gran reto de los evaluadores institucionales no está tanto en seguir profundizando en el debate sobre bondades y maldades del procedimiento, sino en centrar adecuadamente su utilización y de manera especial en los temas relacionados con el estatus socio-laboral de los profesores.

Guste o no guste, las universidades van a tener que atender a la voz de sus estudiantes, por lo que resulta crítico canalizarla debidamente, evitando fenómenos de evaluación espontánea descontrolada (Escudero, 1989), que son bastante inútiles y, además, irritantes para los afectados. Si la voz de los estudiantes la integramos debidamente en el marco global de la evaluación institucional estamos, sin duda, ante uno de los elementos de apoyo más poderosos (quizás el que más) para la mejora de la enseñanza universitaria. No podemos olvidar la importancia que tiene en todo proceso de evaluación educativa la perspectiva del cliente y tampoco, que, en este caso, son los estudiantes los principales portadores de la misma.

Naturalmente, estas opiniones no tienen la misma validez en todos los ámbitos y para todos los propósitos; es especialmente valiosa dentro de la evaluación diagnóstica, en la detección de puntos fuertes y débiles desde la perspectiva del cliente-usuario, sobre todo en el ámbito de la docencia y, en general, en el de las interacciones profesor-alumno.

### 1.1. Algunos resultados de la investigación

No tiene mucho sentido extendernos en este trabajo en la descripción de los resultados de la investigación sobre las encuestas a los estudiantes y su utilización (Escudero, 1996), pero sí me parece obligado señalar algunos elementos, a mi entender claves, de la misma.

En primer lugar hay que reconocer que el interés por estudiar la visión de los alumnos, que según Good y Mulryan (1990) empieza ya a finales del siglo pasado, ha ido siempre estrechamente ligado al estudio de la eficacia docente del profesorado. Este hecho, sin duda, ha supuesto la identificación bastante generalizada de las encuestas a los alumnos como procedimiento de evaluación del profesorado. Sin embargo, su utilidad y su utilización es mucho más amplia, pues permiten evaluar la docencia en la mayoría de sus facetas y aspectos, entre lo que está, claro está, la actuación del profesorado.

Cook (1989) afirma que la primera escala publicada fue la "Purdue Rating Scale of Instruction", aparecida en 1926, sobre la que pronto aparecieron trabajos de investigación (Remmers, 1934). Posteriormente y hasta nuestros días, muchos centros e instituciones han ido elaborando sus propias escalas o adaptado otras y muchas páginas de literatura especializada se han dedicado a su estudio y en no pocos casos a la revisión y meta-análisis de los resultados (Cohen, 1981; Marsh y Dunkin, 1992).

Muchos de estos trabajos destacan la utilidad y buenas características de las encuestas a los estudiantes para valorar la docencia, sobre todos si se comparan con otro tipo de juicios valorativos. Pero no se ocultan sus limitaciones y se sugiere que se complementen con otras fuentes de valoración.

Aunque no procede reproducir aquí una síntesis amplia de los resultados de las investigaciones (Escudero, 1996), sí que debemos señalar algunas conclusiones importantes como las siguientes:

- a) No existen pruebas de que los estudiantes evalúen básicamente en función de la popularidad, amabilidad o sentido del humor del profesor.
- b) El juicio de los alumnos es estable y consistente.
- c) Los juicios de los exalumnos muestran una gran validez concurrente con los de los alumnos.
- d) Los alumnos tienden a ser algo más indulgentes si conocen el valor sumativo de sus juicios sobre la estabilidad y promoción del profesorado.
- f) Los juicios valorativos correlacionan con las calificaciones, pero no existen evidencias de incidencia en este hecho de la indulgencia del profesor o de los alumnos. Puede ser consecuencia lógica del superior aprendizaje.
- g) No se confirma la creencia de que se valora mejor a los profesores de clases pequeñas que de grandes.
- h) El interés previo al curso suele generar valoraciones más positivas. Este hecho se refleja en los cursos optativos.

- i) No se han detectado interacciones sistemáticas entre los juicios de los estudiantes y diferentes aspectos como el sexo de alumnos y profesores, el horario del curso, el nivel, el área disciplinar, el rango académico del profesor y la personalidad de los estudiantes.
- j) La validez de los juicios mejora cuando se complementa con otros elementos y otros juicios y cuando se toma la clase y no al alumno como unidad de análisis. Este hecho se hace especialmente relevante en procesos evaluativos con propósito sumativo.
- k) Los cuestionarios reflejan bien lo que el estudiante ve y vive en el aula, de forma que sus juicios son más sensibles a los cambios en la docencia que en su propio rendimiento académico.
- l) No parece discutirse la multidimensionalidad del juicio de los estudiantes sobre la docencia, pero sí se ha planteado una viva polémica sobre su utilización. La valoración global (unidimensional) parece tener una mayor utilidad para decisiones sumativas, mientras que la multidimensional (factorial) la tiene en las decisiones formativas.
- m) Las prevenciones del profesorado ante la evaluación de los alumnos se centra fundamentalmente en la publicidad y utilización de los resultados. Cuando se dan garantías técnicas y se utilizan debidamente los resultados, globalmente, el profesorado no se opone a la evaluación de la docencia, ni la utilización para ello de cuestionarios a los estudiantes.

Sobre la validez de las encuestas, básicamente nos encontramos con dos visiones (Abrami y otros, 1990). Una que defiende que los cuestionarios son válidos si reflejan con precisión las opiniones de los alumnos sobre la calidad de la enseñanza, con independencia de que reflejan o no su aprendizaje y una segunda, que defiende su validez si reflejen de manera precisa la eficacia instruccional, concepto que, como es bien sabido, resulta difícil de acotar.

Las críticas a la validez de los cuestionarios parecen inscribirse en la segunda visión presentada (Escudero, 1993), porque sí que se ha evidenciado que los cuestionarios bien contruidos, utilizados y analizados terminan reflejando bastante fielmente la visión de los alumnos.

En el proceso de validación de estos cuestionarios es habitual el análisis dimensional, a través del análisis factorial, evidenciándose no sólo la multidimensionalidad, sino también un alto nivel de coincidencia en el número, en el entorno de seis, y en el tipo de factores entre diversas escalas (Marsh y Hocevar, 1991). Sin embargo, las coincidencias no son tantas cuando se trata de valorar su utilidad y su utilización. La cuestión es si se deben utilizar los juicios sobre cada dimensión de manera aislada (visión multidimensional), o si se debe utilizar el juicio global, el promedio o una composición ponderada de los diversos factores (visión unidimensional). Naturalmente, este debate es especialmente vivo sólo cuando se habla de evaluaciones con orientación sumativa.

También existe controversia con el asunto de la privacidad de los resultados de las encuestas, sobre todo cuando existen implicaciones para individuos concretos, generalmente para los profesores. El extremo de este peligro está en lo que hemos denominado *evaluación*

*espontánea* (Escudero, 1989), promovida directamente por algún grupo de estudiantes, sin control metodológico adecuado y con una utilización que suele ser contraproducente para la mejora docente. Es una evaluación que puede servir para sacar algunos colores, criticar a ciertos profesores y colmar al anhelo de crítica de algunos alumnos. Los resultados se terminan exponiendo, generalmente, en los pasillos o vestíbulos de los centros. Esta evaluación, a pesar de que suele plantearse con buenas intenciones por parte de los alumnos, es negativa en sus efectos y debe evitarse, porque termina dando argumentos a aquellos profesores que se oponen a cualquier evaluación. Esta especie de *contra-evaluación* es conocida en las universidades americanas como *anti-calendario*, pues se utiliza como documento de consulta para evitar algunos cursos y profesores entre los ofrecidos por el calendario oficial (Goldschmid, 1978).

Parece incuestionable que cualquier procedimiento de evaluación debe utilizarse con cautela, sobre todo cuando incide en decisiones fiscalizadoras y promocionales. Conviene triangular y complementar procedimientos y fuentes diversas, porque cualquier procedimiento aislado tiene limitaciones y riesgos.

Finalmente como sugerencias prácticas a la hora de aplicar los cuestionarios para evaluar la docencia, se pueden destacar las siguientes:

- a) La opinión de los usuarios es especialmente valiosa para las decisiones de carácter formativo. Las cuestiones deben permitir el análisis diferenciado de distintas facetas y aspectos de la docencia.
- b) La previa aceptación, tanto de alumnos como de profesores, es condición indispensable para el éxito de este tipo de evaluación. También es fundamental la existencia de voluntad en el profesorado para estudiar las demandas que se le plantean y la de apoyos institucionales para acciones de mejora.
- c) Los análisis comparados demuestran que los esquemas de evaluación se reproducen de un país a otro y que resulta científicamente apropiada la utilización de instrumentos usados por otras instituciones, siempre que se haya seguido un adecuado proceso de contextualización.
- d) Los formatos de los cuestionarios pueden ser variados, todos ellos con ventajas y desventajas. En teoría, la fiabilidad y la precisión aumentan con el número de puntuaciones diferentes en la escala, pero en la práctica, esto no es siempre posible o aconsejable. Habitualmente se utilizan escalas de cinco puntos en preguntas unidireccionales (nada-mucho) y de siete o más en las bidireccionales (a favor-en contra).
- e) Los cuestionarios deben ser claros y concisos, de forma que puedan contestarse sin ayudas o explicaciones complementarias.
- f) Los cuestionarios largos pueden producir tedio y cansancio; se sugiere la utilización de cuestionarios cuya contestación no requiera más allá de 10-15 minutos.
- g) Tampoco debe ser la evaluación demasiado frecuente, ya que crece el peligro de que se convierta en algo rutinario e inútil. Globalmente, la más razonable parece ser la evaluación anual, al final de curso, antes de la evaluación final de los alumnos.

- h) La consistencia de opiniones entre cursos y entre años sucesivos es un buen argumento para confiar en las mismas a la hora de tomar decisiones correctoras.
- i) Los juicios valorativos deben contextualizarse siempre, especialmente cuando se planteen enfoques comparativos. Las comparaciones deben limitarse a cursos y programas que tienen lugar en contextos similares.
- j) No se debe abusar de la voluntad de colaboración de los estudiantes. Si sistemáticamente no se atienden sus demandas y sugerencias de mejora, tarde o temprano dejarán de colaborar con la evaluación institucional y aumentará el riesgo de procesos de evaluación espontánea

## 2. ESTA EVALUACION EN LA UNIVERSIDAD DE ZARAGOZA

La primera expresión del “pedagogismo activo” que tuvo mucha incidencia entre alumnos y profesores en nuestro sistema educativo en la década de los setenta y parte de los ochenta, tuvo lugar en la Universidad de Zaragoza, en la Facultad de Ciencias, a finales del curso 1971-72 (Escudero, 1989). Un conjunto de reivindicaciones didácticas y curriculares del alumnado, entre otras razones, genera un conflicto académico importante, incluso con cierre de la Facultad de varios meses. En este contexto, los estudiantes elaboraron un informe valorativo de la docencia y de los profesores.

Unos años después, durante el curso 1975-76, un buen grupo de profesores, sobre todo no numerario de la citada Facultad, que en no pocos casos habían participado en los movimientos estudiantiles descritos anteriormente, constituyeron *un Seminario de Pedagogía*, con reuniones y debates semanales de varias horas, con lecturas de libros y con elaboración de documentos. Este Seminario lo apoyamos desde el Instituto de Ciencias de la Educación e incluso tuvo una modesta subvención económica.

Partiendo del “Cuestionario de Evaluación” que aplicaba el ICE a los profesores y titulados que existían a todos sus cursos, el Seminario de Pedagogía diseñó, elaboró y aplicó dos cuestionarios de evaluación de la docencia por los alumnos. El primero denominado *Encuesta de Evaluación de la Clase*, para aplicarse al final del primer trimestre del curso, con el objetivo de “mostrar al profesor cómo está desarrollando su labor docente, con vistas a introducir posibles mejoras”. El segundo cuestionario, *Encuesta de Evaluación del Curso* para valorar globalmente la calidad del curso y del profesorado. Ambos cuestionarios tenían enfoque y contenido diferentes. En el segundo se enfatizaba mucho más cierta orientación sumativa y de graduación del profesorado. (Profesores No Numerarios, 1976 y 1984).

Estos cuestionarios no se aplicaron de manera masiva y sistemática, pero se experimentaron en distintas clases, con profesores diversos en nivel administrativo y en imagen pedagógica. Los cuestionarios mostraron buena validez aparente y capacidad para distinguir buenas y malas acciones docentes.

Este Seminario terminó teniendo mucha influencia en la posterior revitalización del movimiento evaluador en la Universidad de Zaragoza, a resultas de la puesta en marcha de la Ley de Reforma Universitaria (1983).

Tras estas experiencias, se produjo un periodo de calma en nuestra Universidad en cuanto a presión evaluadora general, pero se realizaron un buen número de pequeños estudios y proyectos, generalmente promovidos desde determinadas cátedras o departamentos y apoyados y asesorados por el Instituto de Ciencias de la Educación. Casi todos estos trabajos estaban relacionados con reformas curriculares y problemas didácticos debidos a la masificación. Como ejemplos de estos trabajos pueden citarse los de las Cátedras de Derecho Civil, Radiología, Matemáticas en Ingeniería y Física-Veterinaria (Escudero, 1989).

Alguno de estos trabajos tuvo cierto alcance en materia de evaluación curricular, como uno llevado a cabo durante el curso de 1977-78 (Nieto y otros, 1978), en el que se consultó a 870 alumnos sobre los objetivos y técnicas didácticas de la Anatomía.

Paralelamente a este tipo de trabajos, durante estos años, algunos profesores a título personal o en colaboración con compañeros del departamento, aplicaban periódicamente encuestas a sus alumnos, para detectar su nivel de satisfacción con el curso e introducir cambios sugeridos por tales opiniones. No entramos en algunas investigaciones y estudios llevados a cabo en el ICE durante esta época (Escudero, 1989), que introducían valoraciones por parte de estudiantes de la enseñanza o de algunas facetas, pues el enfoque era global y tenía otro tipo de orientación al que se trata en estas reflexiones.

Tras la aplicación de la LRU en 1983 y seguramente con el impulso de sus previsiones en materia de evaluación y el debate estatutario, tuvieron lugar en nuestra Universidad varias experiencias espontáneas de evaluación de la docencia y de los profesores, promovidas y dirigidas por los propios alumnos, aunque con colaboraciones aisladas de algunos profesores. Quizás los tres ejemplos más significativos tuvieron lugar en la Escuela Superior de Ingenieros Industriales, en la Escuela Universitaria del Profesorado de EGB de Zaragoza y en la Sección de Geológicas de la Facultad de Ciencias.

La experiencia más institucional fue la de la Escuela de Ingenieros, siendo coordinada por una comisión mixta de alumnos y profesores, utilizando como instrumento básico un cuestionario elaborado en la Universidad Politécnica de Madrid. La motivación principal de esta experiencia era la resolución de casos problemáticos, fundamentalmente relacionados con los exámenes. Esta experiencia dio paso a la evaluación de la docencia que se implantó después en toda la Universidad.

En el caso de la Escuela de Formación del Profesorado, el trabajo fue promovido, diseñado y controlado por la comisión de alumnos de la Escuela. Se trataba fundamentalmente de evaluar el plan de estudios, la coherencia de su desarrollo y las relaciones de los profesores con los alumnos. En esta experiencia se utilizó la encuesta de evaluación del curso que había elaborado el precitado Seminario de Pedagogía de Profesores No Numerarios. Hubo algunos problemas de participación porque llevaba bastante tiempo la evaluación de todas las asignaturas y, además, buena parte del profesorado mantuvo una postura contraria a la aplicación, aludiendo a problemas técnicos y falta de fiabilidad.

Los participantes en esta experiencia resaltaron que el cuestionario marcaba mucho las diferencias entre buenos y malos profesores.

También tomó como base las encuestas del Seminario de Pedagogía, el Colectivo Contracultura de Geológicas, en la experiencia llevada a cabo durante dos años en dicha Sección de la Facultad de Ciencias. El segundo año se simplificó mucho el cuestionario, pues resultaba muy largo tal como se detectó en la experiencia anterior. Todos los alumnos

de la sección eran invitados a contestar el cuestionario de evaluación de la clase y tuvieron una participación de alrededor del 25% del alumnado el primer año y cerca al 35% el segundo.

Tras el análisis de resultados se elaboraba un informe por asignatura sobre: valoración global, relación profesor-alumno, criterios de evaluación, prácticas de campo y sugerencias de los alumnos. Estos informes se entregaban a los delegados de curso, a los profesores implicados y se hacían públicos en el tablón de la Facultad. También se aplicaron encuestas a los profesores, pero la participación fue muy escasa. Los implicados en esta experiencia resaltaban el interés de los resultados, pero consideraban que era excesivo llevarla a cabo completa todos los años; habría que profundizar en aspectos distintos cada año. El problema más espinoso era el de la publicación de los resultados, sobre todo si se recogían observaciones concretas sobre los profesores, en algún caso *tremendamente descalificatorias* para su preparación o para su labor.

### 3. LA EVALUACION PREVISTA EN LOS ESTATUTOS

Todas estas experiencias espontáneas dejaron de tener lugar cuando se puso en marcha la evaluación de la docencia anual, prevista en los Estatutos de la Universidad (1985), al desarrollar el mandato del Artículo 45.3 de la LRU, para que los Estatutos dispusieran “los procedimientos para la evaluación periódica del rendimiento docente y científico del profesorado”.

El modelo de evaluación estatutario era desarrollado fundamentalmente en un capítulo titulado “La evaluación y control de la docencia”, con nueve artículos, que fue aprobado por el Claustro en una votación bastante accidentada, por el abandono de la sala, en el momento de la votación, de un grupo de profesores contrarios a dicho modelo de evaluación. Como rasgos principales de este modelo hay que destacar que habla de evaluación de la docencia en vez de evaluación del profesorado, que introduce el informe del propio profesor como elemento de la evaluación y que le da carácter público a los informes de evaluación, pero no se establece la obligatoriedad de su exposición pública.

El órgano central de todo el proceso es la Comisión de Evaluación y Control de la Docencia -una general para toda la Universidad y una en cada Centro- compuesto por igual número de alumnos y profesores.

Se establece una evaluación anual, basada como mínimo en:

- a) Los resultados de las encuestas de evaluación del desarrollo del curso académico cumplimentadas por los alumnos.
- b) El informe del propio profesor del curso, que conocerá previamente el resultado de las encuestas.

Los procedimientos de recogida de información deben ser aprobados por la Junta de Gobierno y la Comisión de cada Centro elevará un informe a la de Universidad, en los casos en los que la evaluación resulte negativa. Los informes de las diversas Comisiones se entregarán a los interesados y tendrán carácter público.



Dos evaluaciones consecutivas negativas sobre un profesor, obligan a la Comisión de Evaluación de la Universidad a elevar un informe al Rector para que lo traslade a la Junta de Gobierno para su conocimiento y efecto.

Si el 25% o más de los profesores de un departamento es evaluado negativamente, la Comisión de Evaluación de la Universidad analizará la situación docente del citado departamento y propondrá las soluciones que considere necesarias.

Además de los casos citados, las Comisiones de Evaluación actuarán cuando se presenten ante ellas reclamaciones o denuncias sobre la actuación docente de un profesor, departamento o centro.

Asímismo, los centros elaborarán una memoria anual valorativa de un funcionamiento docente, que remitirán a la Comisión de Evaluación y Control de la Docencia. Por último, los departamentos elaborarán una memoria anual de actividades y proyectos docentes y de investigación, que remitirán a las Comisiones de Evaluación de la Docencia y de Investigación de la Universidad.

#### **4. EL CUESTIONARIO DE LA UNIVERSIDAD DE ZARAGOZA**

Este modelo de evaluación se aplicó por primera vez con carácter experimental el curso 1987-88, tras lo que la Comisión de Evaluación y Control de la Docencia abrió un periodo de enmiendas y sugerencias en todos los centros y, tras su análisis y discusión, presentó a la Junta de Gobierno un nuevo modelo de cuestionario y procedimiento de aplicación. El nuevo cuestionario (recogido en el Anexo), que se sigue utilizando actualmente, aunque la Comisión está estudiando su reforma, fue aprobado por la Junta de Gobierno en Resolución del 14 de abril de 1989.

Como puede apreciarse en el Anexo, este cuestionario consta de veintiséis preguntas comunes para todos los centros de la Universidad y se abre la posibilidad de cuatro preguntas específicas más, que deben ser diseñadas y analizadas por la Comisión del centro correspondiente.

Sin entrar en un análisis exhaustivo y pormenorizado de las características métricas de este cuestionario, podemos recoger algunos datos relevantes relativos al mismo en materia de fiabilidad y validez (Escudero y otros, 1997), llevados a cabo utilizando los correspondientes programas del paquete SPSS de análisis estadístico (Norusis, 1990).

Utilizando el coeficiente  $\infty$  de Cronbach como medida de la consistencia interna, basada en la correlación promedio de las cuestiones dentro de la propia prueba, para distintos centros y grupos de clase de distintos tamaños, se encontró un coeficiente de promedio de valor 0,84. Además, el rango de variabilidad de estos coeficientes entre centros y cursos era bastante estrecho. Se trata, por tanto, de un coeficiente de fiabilidad muy alto y de un cuestionario que no puede criticarse mucho en esta faceta.

En cuanto a la validez del cuestionario, una forma habitual de inspección es el análisis de los factores que realmente mide la prueba y su aportación relativa al conjunto de la medición. Este análisis, llevado a cabo con seis grandes muestras de origen muy variado, condujo a resultados muy parecidos, con muchísimos elementos comunes entre cada

análisis, resultando fácil establecer los factores que mide el cuestionario y su potencial explicativo del conjunto de la valoración. Estos factores son los siguientes (Escudero y otros, 1997):

- a) Destaca sobre todos un *primer factor*, claro y contundente, con un poder explicativo por sí solo de alrededor del 40% de la varianza, que se centra fundamentalmente en lo que podríamos denominar como *actuación didáctica del profesor*.

Una descripción más detallada de este factor, en función de las preguntas del cuestionario que recoge podría ser la siguiente: el alumno destaca *al profesor comunicador eficaz, ameno, estimulante y potenciador del diálogo* (preguntas 17, 18, 19 y 20), *que utiliza medios y recursos didácticos, es claro y ordenado y está disponible para consultas del alumno* (preguntas 14, 16 y 21) y, *además se prepara las clases* (pregunta 15).

Con alguna frecuencia, en este factor se incluye también el hecho de si se han definido con claridad los objetivos y contenidos de la asignatura (pregunta 2), si se coordina bien la teoría y la práctica (pregunta 10) y si se domina la materia (pregunta 13).

- b) Un *segundo factor*, también claro, que añade al anterior alrededor de un 12% de capacidad explicativa de la varianza, está centrado alrededor de *las condiciones materiales y ambientales de impartición de la asignatura* (preguntas 7, 8 y 9). A esto se añade con alguna frecuencia la valoración del material recomendado por el profesor (pregunta 6) y de la coordinación entre teoría y práctica (pregunta 10).
- c) Alrededor de un 7% más de la varianza explicada la aporta un *tercer factor* que identifica *la extensión y la dificultad de la asignatura y de los exámenes* (preguntas 3, 4, 23 y 24).
- d) Un *cuarto factor*, que añade alrededor de 5% de la varianza, resalta el *cumplimiento profesional* (preguntas 11 y 12), al que se suele añadir la valoración del dominio de la materia que se imparte (pregunta 13).
- e) La pregunta 26 del cuestionario, referida al tiempo transcurrido entre los exámenes y la comunicación de los resultados, constituye por sí sola un *quinto factor*, que añade algo menos del 5% de la varianza explicada.

El resto de la varianza por explicar (alrededor del 30%) se identifica por otros factores que, individualmente, añaden poco al modelo establecido por los cinco factores resaltados anteriormente. Por ello son desechados en el análisis factorial.

Conviene recordar también que los factores II, III, IV, V, aislados del resto, explican un porcentaje de la varianza total mayor del que añaden a la explicada ya por el factor I (el de más capacidad explicativa de todos) y los siguientes, a medida que se van introduciendo en el modelo.

Globalmente, el resultado de estos análisis factoriales es muy similar al que se encuentra para cuestionarios de evaluación empleados en otras universidades españolas y extranjeras. Algo lógico, porque el cuestionario empleado en nuestra Universidad es muy

parecido a los empleados en universidades con gran tradición en materia de evaluación de la enseñanza pues, de hecho, se construyó apoyándose en ellos. Los resultados obtenidos en el análisis factorial y en el de fiabilidad, avalan esta similitud y, al mismo tiempo, certifican que el cuestionario empleado en nuestra Universidad ofrece unas características métricas razonables y equiparables a los de uso habitual en las principales universidades. Ello no implica que no sea mejorable y que no tenga defectos y lagunas. Esto tampoco evita, lógicamente, las críticas que puedan hacerse al procedimiento en cuanto a concepción y a su potencialidad dentro del proceso global de evaluación de la enseñanza.

## 5. RESULTADOS DE LA EVALUACION DE LA DOCENCIA

Además del análisis individualizado de los resultados de la evaluación de cada curso, que debe llevarse a cabo por las diferentes Comisiones y por el profesor implicado, contamos con un análisis exhaustivo de los resultados de los cuestionarios, para toda la Universidad, cada Centro y cada Departamento, con expresión por cada pregunta de la media, la desviación típica y los valores mínimo y máximo, para los cursos académicos 1993-94, 1994-95 y 1995-96. (Escudero y otros, 1997).

No tiene mucho sentido extenderse en el análisis de tales resultados, pero sí recoger una reflexión de síntesis, útil para conocer cuál es el diagnóstico global que hacen los alumnos de la Universidad de Zaragoza de la docencia que reciben. En síntesis es la siguiente:

- 1) Se observa una acusada estabilidad general de los resultados entre los años analizados.
- 2) La media de *asistencia a clase* de los alumnos que responden a los cuestionarios es alta, con una dispersión relativamente baja, lo que indica que, básicamente, siempre según la declaración de los propios alumnos, las encuestas las cumplimentan alumnos que asisten de manera continuada a las clases o, al menos, de manera frecuente.
- 3) La *definición de los objetivos y los contenidos* de las asignaturas se valora bastante positivamente en cuanto a precisión como promedio, aunque los datos reflejan la existencia de casos con imprecisiones en este sentido.
- 4) Los *materiales* recomendados por los profesores tienen una valoración promedio muy centrada entre los valores extremos de la escala (muy deficiente y excelente).
- 5) También se sitúa en el centro de la escala la valoración promedio de las *condiciones materiales* en las que se imparten las clases teóricas y las de problemas. Sin embargo, las clases prácticas y de laboratorio, etc., tienen una valoración algo más baja y, al mismo tiempo, con más dispersión entre clases.
- 6) Como promedio, la *coordinación* entre la parte teórica y el resto de actividades docentes se valora como simplemente correcta, aunque con una ligerísima tendencia de mejora en los últimos años.

- 7) La *asistencia* a clase de los profesores es muy bien valorada por el alumnado, con un alto grado de estabilidad en los promedios en los últimos años. La *puntualidad* del profesorado también es bien valorada, aunque algo menos que la asistencia y con algo más de dispersión entre profesores, aunque ésta parece ir disminuyendo algo en los últimos años.
- 8) Es muy positiva la valoración promedio del *dominio de la materia impartida* por parte del profesorado. También se valora positivamente, aunque en menor medida, *la preparación de las clases*. Sin embargo, en ambos aspectos se detecta una dispersión de opiniones considerable entre unos profesores y otros.
- 9) La *utilización de medios y recursos didácticos*, la *claridad y orden en la presentación de la materia* y la *eficacia en la transmisión de conceptos* se valoran algo por encima del valor central de la escala. Un poco por encima de este último punto se sitúa el promedio de *amenidad*, que es el aspecto de los recogidos en el cuestionario con relación a las características del profesorado que tiene una valoración relativa más baja. En la mayoría de estas características didácticas del profesorado, la dispersión de valoraciones entre profesores es muy alta y claramente la más elevada de todos los aspectos tratados en el cuestionario. Son, sin duda, los aspectos donde los alumnos ven más diferencias entre unos y otros profesores.
- 10) El profesorado, como promedio, obtiene una valoración discreta, aunque con tendencia positiva, por *potenciar el diálogo* con los estudiantes y *por crear interés por la asignatura*. En ambos aspectos, se observa bastante dispersión entre las valoraciones otorgadas a unos y otros profesores. Los resultados son sensiblemente más positivos en cuanto a *la disponibilidad del profesorado para consultas y tutorías*.
- 11) Los *criterios de evaluación* se definen con bastante precisión aunque existan ciertas lagunas de imprecisión. Los *exámenes y pruebas* se consideran en promedio bastante difíciles, notándose un ligero aumento en esta tendencia en los últimos años, así como una mayor dispersión de valoraciones entre asignaturas en este sentido. La *corrección* de los exámenes se valora como bastante dura y la *correspondencia* entre el nivel de exigencia de los exámenes y lo tratado en clase se valora entre buena y regular. *El tiempo que transcurre entre la realización de los exámenes y la comunicación de los resultados* se considera algo más amplio que lo normal, aunque existe bastante dispersión entre profesores en esta sentido.

Globalmente, el diagnóstico de los alumnos tiene más luces que sombras, pero dado el alto nivel de variabilidad en muchas facetas, es evidente que existen ámbitos y casos con mucho por mejorar.

De manera generalizada, lo que podríamos denominar *profesionalidad del profesorado* es el bloque en el que las valoraciones son más altas. Ello no es una condición suficiente de la calidad de la docencia, pero sí se trata de una condición necesaria. Sin un cumplimiento profesional adecuado del profesorado parece imposible una enseñanza de calidad. Con ese cumplimiento no se asegura la calidad pero, sin duda, se establecen las bases para caminar hacia ella. El diagnóstico nos dice que estas bases están consolidadas en la Universidad de Zaragoza y que el esfuerzo principal del perfeccionamiento debe ir por otros terrenos.

El profesorado asiste a sus clases con razonable puntualidad, se las prepara, tiene un buen dominio de las materias que imparte y está disponible para tutorías y consultas. Este es el retrato normal, el típico, aunque los datos de dispersión en las valoraciones nos indican que existen casos donde algunas de estas situaciones no son tan positivas. Es en estas últimas en las que el diagnóstico debe seguir incidiendo y, consecuentemente, promover acciones de corrección.

En el *ámbito curricular y de ordenación docente* el diagnóstico también resulta globalmente positivo, aunque menos optimista que en el aspecto anterior. Además, en este terreno sí que se observan diferencias sensibles de imagen sobre su docencia entre los estudiantes de unas y otras titulaciones, departamentos y centros. Así por ejemplo, la sensación promedio de dificultad y extensión en las asignaturas, de dificultad en los exámenes y exigencia en la corrección, relativamente altas, varían bastante de unos centros y departamentos a otros. Algo parecido sucede con la valoración de la importancia de los contenidos para la formación, en general alta, pero también con acusadas diferencias entre unos y otros casos.

Con independencia de los niveles de exigencia y dificultad, algo sobre lo que resulta *arriesgado emitir un juicio valorativo externo, la correspondencia entre lo tratado en clase y lo exigido en los exámenes* parece no ser la correcta, según los alumnos, en no pocos casos. Además, el tiempo transcurrido entre los exámenes y la entrega de los resultados es considerado como promedio superior a lo considerado como normal, con bastantes casos con un tiempo de espera demasiado amplio.

Los materiales recomendados por el profesorado y la coordinación entre la teoría y el resto de actividades de la asignatura superan justamente lo que puede considerarse como aprobado, pero dada la dispersión observada, este dato nos debe hacer reflexionar y pensar que estas facetas deben ser mejoradas en muchos casos.

Las condiciones de impartición de las clases de teoría y problemas también llegan justamente a la mitad de la escala de valoración, pero se quedan por debajo de este nivel en las condiciones de las clases prácticas y de laboratorio. En este último caso, sobre todo, la dispersión entre asignaturas es alta, pero, en general, existen ciertas diferencias, a juicio de los alumnos, en las condiciones de impartición de la docencia entre unos centros y otros. De cualquier manera, se trata de un ámbito en el que la Universidad de Zaragoza debe mejorar, más en unos centros que en otros y sobre todo en lo relativo a prácticas y laboratorios.

El ámbito propiamente didáctico, de actuación del profesorado en este sentido, ofrece un panorama global ligeramente por encima del aprobado, pero con alta variabilidad entre profesores. No parece haber duda de que la Universidad de Zaragoza debe seguir potenciando el perfeccionamiento de las capacidades docentes de buena parte de sus profesores y que en esta tarea la labor departamental parece absolutamente imprescindible.

Finalmente, es importante resaltar que aunque la estabilidad entre años es muy acusada y los cambios en las valoraciones son escasos, cuando se han producido en los últimos años, han sido más frecuentemente hacia mejores valoraciones que en sentido contrario. Este hecho es positivo, pero no esconde la extendida sensación de que los cuestionarios de los alumnos no tienen el impacto esperado en la mejora de la docencia.

## 6. LA EVALUACION DEL DOCTORADO

Desde el curso 1995-96, la Comisión de Doctorado con la colaboración técnica del Rectorado lleva a cabo una evaluación de los cursos, apoyada en un cuestionario muy sencillo que se envía por correo a los participantes en los cursos de doctorado.

Las respuestas se analizan por departamentos, programas, macroáreas y para toda la Universidad, incidiendo no sólo en los resultados estadísticos, sino también en la multitud de observaciones que se recogen y que son de extraordinaria utilidad en este caso.

Este procedimiento de evaluación requiere perfeccionamiento, sobre todo en lo concerniente a la implicación más directa de los departamentos en el análisis de cada curso y de sus posibles mejoras. Hasta ahora el protagonismo lo ha llevado casi exclusivamente la Comisión de Doctorado, incidiendo fundamentalmente en problemas globales.

Aunque la imagen detectada en la evaluación del curso 1996-97 es más optimista que la que se percibe en la observación asistemática, sensible a la queja particular, existen elementos de preocupación que requieren intervención. En algunos cursos no se ofrecen la totalidad de créditos ofertados. El incumplimiento y el trastoque de horarios no es poco. Finalmente, se ha detectado una excesiva concentración de cursos en el último trimestre que dificulta mucho la organización y el funcionamiento del doctorado.

## 7. VALORACION FINAL

No tengo ninguna duda de la importancia que tiene la opinión de los estudiantes como elemento de evaluación de la docencia, ni tampoco del efecto positivo global que ha tenido y tiene su utilización en la Universidad de Zaragoza. Mis dudas se centran en si esa utilización es la debida y suficientemente eficiente.

Yo entiendo que este procedimiento es un procedimiento de diagnóstico y una fuente de sugerencias con orientación formativa de primera magnitud. Probablemente, para estas funciones, se trata del procedimiento individual más poderoso y útil. Sin embargo, fuera de este contexto, para otras tareas, su potencialidad disminuye bastante y aparecen más claras algunas de sus limitaciones.

Las encuestas a los estudiantes no son el mejor procedimiento de control del *cumplimiento del profesorado*, ni tampoco lo son para la toma de decisiones de tipo fiscalizador. Sin duda que pueden aportar información en tal sentido, pero no podemos olvidar que la perspectiva del alumno es fundamentalmente la perspectiva del cliente, que reclama calidad en el servicio y señala lo que le satisface y lo que no le satisface, y no la perspectiva del inspector o la del responsable académico.

Es una obviedad que las encuestas no deben ser el procedimiento para controlar si un profesor asiste o no a sus clases regularmente y si lo hace o no con puntualidad. Pues bien, este tipo de cuestiones son las que más problemas plantean a la hora de aplicar los cuestionarios y las que más argumentos dan a los profesores críticos para atacar su validez. Este control se ejerce mucho mejor de manera más directa, asumiendo las responsabilidades de control y, en su caso, de penalización, quien tiene la obligación de asumirlas. La utilización de las encuestas en esta tarea tiene algo de escapismo para evitar confrontaciones directas.

También surgen problemas cuando se quieren utilizar las encuestas como argumento único para penalizar al profesor, por ejemplo no concediéndolo un tramo de complemento retributivo docente. El profesor afectado termina judicializando el asunto, entrando en un terreno complejo. Las encuestas tienen muchas limitaciones como argumento judicial por las propias condiciones de su utilización. Si se quiere potenciar su capacidad pedagógica, la metodología de aplicación debe ser relajada y algo alejada de controles de tipo judicial. No se discute, por tanto, la verdad del resultado, sencillamente se argumenta que está obtenido con escaso control legal.

Las decisiones fiscalizadoras las deben tomar personas o comisiones autorizadas, utilizando todo tipo de información, cualitativa y cuantitativa, incluidos los resultados de las encuestas, como consecuencia de una valoración cualitativa del problema. Cargar esta responsabilidad a las encuestas es un error.

Otra crítica, sobre todo por parte del alumnado, es que los resultados de las encuestas no traen consecuencias y que se terminan convirtiendo en algo rutinario. Pero este peligro, que puede ser real, se acentúa sobre todo si se ve el proceso como un instrumento de penalización del profesorado, algo que ya hemos señalado como poco apropiado. En esta misma línea podemos interpretar el interés de algunos estudiantes por hacer públicos los malos resultados de algunos profesores, algo que no sirve más que para producir irritación innecesaria.

Cuando sí tienen razón los alumnos es cuando demandan más atención a sus reclamaciones y sugerencias sobre la docencia. El problema es que esto incide fundamentalmente sobre la responsabilidad individual de cada profesor y sobre la de cada departamento. El verdadero reto está en motivar y responsabilizar a profesores y departamentos hacia la reflexión y mejora pedagógica.

Es cierto que se necesitan más apoyos pedagógicos institucionales para los profesores y los departamentos, pero no es menos cierto que es más crítico que los utilicen aquellos que verdaderamente los necesitan, porque no siempre resulta fácil. Con alguna frecuencia, los cursos de perfeccionamiento pedagógico que ofrecen las universidades a sus profesores no son atendidos por los profesores con problemas en el aula. Es más frecuente que profesores bien valorados sean los que en mayor medida utilizan esta oferta de formación.

## ANEXO

### CUESTIONARIO PARA LA EVALUACION DE LA DOCENCIA UNIVERSITARIA

(Atención, la numeración entre paréntesis sólo indica orden, no valoración del concepto)

1. *Ha asistido Vd. a las clases de este profesor de forma:*

continuada (1)      frecuente (2)      ocasional (3)      mínima (4)

#### I. ASIGNATURA

2. *Los objetivos y contenidos de la asignatura han sido definidos:*

con gran      con bastante      con alguna      de manera      no fueron  
precisión (1)      precisión (2)      imprecisión (3)      vaga (4)      definidos (5)

3. *Esta asignatura le parece:*

muy difícil (1)    difícil (2)    adecuada (3)    fácil (4)    muy fácil (5)

4. *La extensión del programa impartido le parece:*

muy amplia (1)    amplia (2)    adecuada (3)    escasa (4)    muy escasa (5)

5. *Los contenidos de la asignatura le parecen para su formación:*

muy importantes (1)    importantes (2)    de importancia media (3)    poco importantes (4)    nada importantes (5)

**Opine puntuando de 1 (muy deficiente) a 5 (excelente) sobre los siguientes aspectos:**

6. *Material recomendado por el profesor.*
7. *Condiciones materiales y ambientales de impartición de la parte teórica de la asignatura.*
8. *Condiciones materiales y ambientales de impartición de las clases de problemas y ejercicios prácticos.*
9. *Condiciones materiales y ambientales de impartición de las clases de laboratorio, prácticas de campo, taller, clínicas...*
10. *Coordinación entre la parte teórica y el resto de las actividades de la asignatura.*

**II. PROFESORADO**

**Opine puntuando de 1 (muy deficiente) a 5 (excelente) sobre la actuación o características del profesor en:**

11. *Asistencia a clase.*
12. *Puntualidad.*
13. *Dominio de la materia que imparte.*
14. *Utilización de medios y recursos.*
15. *Preparación de las clases.*
16. *Claridad y orden en la presentación de la materia.*
17. *Eficacia en transmitir los conceptos.*
18. *Amenidad.*

**III. RELACIONES PROFESOR-ESTUDIANTES**

**Opine puntuando de 1 (muy deficiente) a 5 (excelente) sobre los siguientes aspectos:**

19. *Potenciación del diálogo.*
20. *Actuación del profesor para crear interés por la asignatura.*
21. *Disponibilidad del profesor para consultas, tutorías, sugerencias, etc.*

**IV. EVALUACIONES**22. *Los criterios de Evaluación de la asignatura han sido definidos:*

con gran precisión (1)    con bastante precisión (2)    con alguna imprecisión (3)    de manera vaga (4)    no fueron definidos (5)

(Caso de no haber realizado ningún examen o evaluación anterior a esta fecha **no conteste a las siguientes preguntas de este apartado**).



## 23. Los exámenes y pruebas realizados le parecen:

muy difíciles (1)    difíciles (2)    asequibles (3)    fáciles (4)    muy fáciles (5)

## 24. La corrección de los exámenes la considera:

muy dura (1)    dura (2)    adecuada (3)    blanda (4)    muy blanda (5)

## 25. El nivel que se exige en los exámenes se corresponde con lo tratado en clase de manera:

excelente (1)    buena (2)    regular (3)    deficiente (4)    muy deficiente (5)

## 26. El tiempo transcurrido entre la realización de un examen y la comunicación de los resultados le parece:

excesivo (1)    demasiado  
amplio (2)    normal (3)    breve (4)    muy breve (5)

## V. PREGUNTA 27 A 31

Estas preguntas, de haber sido definidas por la Comisión de Centro, se entregarán en hoja aparte.

## REFERENCIAS BIBLIOGRAFICAS

- ABRAMI, P.C. y otros (1990). "Validity of student ratings of instruction: what we know and what we do not". *Journal of Educational Psychology*, 82, 2, 219-231.
- COHEN, P.A. (1981). "Student ratings of instruction and student achievement: A meta-analysis of multisection validity studies". *Review of Educational Research*, 51, 282-309.
- COMISIÓN DE DOCTORADO (1998). *Evaluación del Doctorado. Curso 1996/97*. Vicerrectorado de Evaluación y Mejora de la Enseñanza: Universidad de Zaragoza.
- COOK, S.S. (1989). "Improving the quality of student ratings of instruction: A look at two strategies". *Research in Higher Education*, 30, 1, 31-45.
- DOWELL, D.A. y NEAL, J.A. (1982). "A selective review of the validity of student ratings of teaching". *Journal of Higher Education*, 53, 1, 51-62.
- ESCUDERO, T. (1989). "Experiencias evaluativas en la Universidad de Zaragoza. En *Informes de Investigación Evaluativa, n.º 1. Consideraciones metodológicas sobre evaluación y mejora de la docencia universitaria*. (pp. 89-132). Universidad de Valencia.
- ESCUDERO, T. (1993). "Enfoques modélicos en la evaluación de la enseñanza universitaria". *III Jornadas Nacionales de Didáctica Universitaria "Evaluación y Desarrollo Curricular"*. Servicio de Publicaciones, Universidad de Las Palmas, 5-59.
- ESCUDERO, T. (1996). *Proyecto docente e investigador*. Universidad de Zaragoza.
- ESCUDERO, T. y otros (1997). *Nuestra docencia según los estatutos*. Vicerrectorado de Evaluación y Mejora de la Enseñanza: Universidad de Zaragoza.
- GOLDSCHMID, M. L. (1978). "The evaluation and improvement of teaching in higher education". *Higher Education*, 7, 221-245.
- GOOD, T. L. y MULRYAN, C. (1990). "Teacher ratings: A call for teacher control and self-evaluation". En J. Millman y L. Darling-Hamond (Eds.), *The new Handbook of Teacher Evaluation* (pp. 191-215). Sage publications, Newbury Park, Col.
- LEY ORGÁNICA 11/1983, de 25 de agosto, de Reforma Universitaria. (BOE de 1 de septiembre de 1983).
- MARSH, H.W. y HOCEVAR, D. (1991). "The multidimensionality of students' evaluation of teaching effectiveness: The generality of factor structures across academic discipline, instructor level, and course level". *Teaching & Teacher Education*, 7, 1, 9-18.
- MARSH, H.W. y DUNKIN, M.J. (1992). "Students' evaluations of university teaching: A multidimensional perspective". En J.C. Smart (Ed.), *Higher Education: Handbook of Theory and Research* (pp. 143-233), Vol. VIII. Agathon Press: New York.
- NIETO, J.L. y otros (1978). "Sondeo de opinión entre los alumnos de la Universidad de Zaragoza, acerca de los objetivos y técnicas didácticas de la Anatomía". *Actas del X Congreso de la Sociedad Anatómica Española*. Zaragoza, 449-458.
- NORUSIS, M. J. (1990). *SPSS Base System User's Guide*. SPSS Inc. Michigan.
- ORY, T. (1991). "Changes in evaluating teaching in higher education". *Theory into Practice*, XXX, 1, 30-36.

- PROFESORES NO NUMERARIOS (1976): *Seminario de Pedagogía en Universidad*. Documento fotocopiado, Universidad de Zaragoza (Editado posteriormente en 1984 con el título *Sobre Pedagogía en la Universidad*. Secretariado de Publicaciones, Universidad de Zaragoza).
- REAL DECRETO 1271/1985, de 29 de mayo, por el que se aprueban los Estatutos de la Universidad de Zaragoza (BOE de 29 de julio de 1985).
- REMMERS, H.H. (1934): "Reliability and halo-effect of high school and college students' judgments of their teachers". *Journal of Applied Psychology*, 18, 619-630.