

La evaluación de la calidad del aprendizaje en la universidad

MARÍA LUISA PÉREZ CABANÍ, MARÍA REYES CARRETERO,
MONTSERRAT PALMA Y ESTER RAFEL

Universidad de Girona



Resumen

El trabajo que se presenta en este artículo tiene como objetivo conocer los criterios y decisiones de los profesores universitarios para evaluar el aprendizaje de los estudiantes, y analizar la incidencia de estos factores en la calidad del aprendizaje. Para ello se ha contado con la colaboración de un grupo de profesores de la Universidad de Girona que nos han proporcionado las demandas de evaluación que formulan a través de los exámenes y las respuestas escritas de los estudiantes a estas demandas. Complementariamente se ha realizado una entrevista a los profesores colaboradores para conocer sus criterios e intenciones al realizar la evaluación. Entre las conclusiones obtenidas de este estudio destacaríamos la relación que se produce entre la tipología de pregunta y la demanda formulada, la diversidad de criterios de evaluación existente y la complejidad que supone evaluar la calidad del aprendizaje.

Palabras clave: Evaluación del aprendizaje, calidad del aprendizaje, estudios universitarios.

Evaluating quality of learning in the university

Abstract

The aims of the study were to understand the criteria and decisions of university teachers in assessing what students have learnt, and to analyse how these factors affect the quality of learning. A group of teachers at the University of Girona provided us with the assessment demands they formulated in examinations and with students' written replies to these demands. The teachers taking part were also interviewed on their criteria and intentions when assessing students. The study's three main conclusions are that there is a relationship between the type of question and the demand formulated, that evaluation criteria differ widely, and that assessment of the quality of learning is complex.

Keywords: Learning assessment, quality of learning, university studies.

Agradecimientos: Este trabajo se ha realizado con una ayuda de investigación concedida por la Universidad de Girona y queremos manifestar nuestro agradecimiento a las personas que han colaborado con nuestro equipo. A. P. Bosch Codolà y M. R. Terradellas Piferrer por su participación en la recogida y análisis de los datos, a S. Parera Izquierdo por su asesoramiento en los aspectos metodológicos y a los profesores de los Departamentos de Psicología y Pedagogía de la UdG por permitirnos el acceso a los materiales de evaluación de sus asignaturas y responder de manera desinteresada a nuestras entrevistas.

Correspondencia con las autoras: Departamento de Psicología. Universidad de Girona. Pl. Sant Domènec, 9. 17071 Girona. E-mail: mllperez@zeus.udg.es

Original recibido: Abril, 1998. *Aceptado:* Septiembre, 1998.

INTRODUCCIÓN

La evaluación en las instituciones universitarias constituye uno de los componentes fundamentales de su funcionamiento, puesto que proporciona información que permite valorar la calidad de los procesos educativos que se llevan a cabo y de los resultados de los mismos. Por este motivo, el concepto de evaluación en el ámbito universitario tiene una consideración amplia, que afecta a los diferentes estamentos y dimensiones que integran esta institución y que puede ser abordada desde diferentes perspectivas. Algunas de las principales dimensiones que son objeto de evaluación, siguiendo el estudio realizado por Hernández Pina (1997) son los objetivos de la institución, el aprendizaje de los estudiantes, el rendimiento del profesorado, los programas académicos, los servicios, la gerencia, la financiación y las relaciones externas.

De acuerdo con esta perspectiva, desde la Administración educativa se han elaborado, en los últimos años, diferentes instrumentos que permiten recoger información sobre la calidad de las instituciones universitarias a partir del análisis y valoración de los diferentes elementos estructurales y funcionales que las componen. Muestra de ello son el Programa Experimental de Evaluación de la Calidad del Sistema Universitario y el Proyecto Piloto Europeo para la evaluación de la Calidad de la Enseñanza Superior realizados durante los años 1993, 1994 y 1995, por el Consejo de Universidades, y que constituyeron la base para el diseño de un modelo de evaluación que se ha materializado en la primera convocatoria del Plan Nacional de Evaluación de la Calidad de las Universidades en 1996.

Entre las diferentes dimensiones consideradas relevantes en este marco institucional de evaluación en la universidad, la evaluación del aprendizaje que realizan los estudiantes nos parece una variable especialmente relevante, por lo que, sin infravalorar la importancia de las demás, hemos centrado en ella nuestra investigación. Los motivos que nos conducen a profundizar en el estudio de esta dimensión son los siguientes:

(1) El primero de ellos se sustenta en que las conclusiones del informe del Plan Nacional de Evaluación de la Calidad de las Universidades (1997) ponen de manifiesto las carencias de los sistemas de evaluación del aprendizaje y la insatisfacción de los alumnos respecto a los mismos. Entre los argumentos que conducen a estas conclusiones destacaríamos la ausencia de criterios objetivos y generales de evaluación, la falta de coherencia entre los criterios de evaluación de diferentes asignaturas de un mismo estudio, la falta de publicidad de dichos criterios, la centralización de los procesos de control sobre los exámenes y el desajuste entre los contenidos impartidos y los niveles exigidos.

(2) En segundo lugar, la existencia de investigaciones que demuestran cómo el sistema de evaluación que utilizan los profesores para valorar el aprendizaje de los alumnos afecta a la calidad de dicho aprendizaje, llegando a considerar que las intervenciones para mejorar los procedimientos de aprendizaje de los alumnos y su nivel de calidad, deberían empezar por cambiar los sistemas de evaluación (Elton y Laurillard, 1979; Marton, Hounsell y Entwistle, 1984; Biggs, 1993, 1996a).

(3) Un tercer motivo se basa en la constatación de la existencia de diferentes modelos de evaluación, centrados casi exclusivamente en los criterios que aplica cada profesor (Mateo, 1993; Gros y Romañá, 1995). Si a este eclecticismo se suma la falta de reflexión y trabajo en común sobre el tema, el resultado es la coexistencia de demasiados “modelos” evaluativos que conducen a la confusión y desorientación de los estudiantes respecto a como afrontar la demanda que se les formula.

Teniendo en cuenta las consideraciones anteriores, hemos centrado el objeto de estudio de nuestra investigación en la evaluación de la calidad del aprendizaje de los estudiantes universitarios, tomando como referencia los trabajos previamente realizados sobre este tema.

La mayoría de éstos muestran un interés creciente en buscar instrumentos que proporcionen una medida objetiva sobre qué aprenden los alumnos en las aulas universitarias y cómo lo aprenden, que permitan analizar los factores que inciden en la calidad del aprendizaje y que ofrezcan criterios de evaluación e indicadores de calidad que hagan posible optimizar el proceso de evaluación del aprendizaje de los estudiantes (Marton, 1988; Chadwick y Rivera, 1991; Entwistle, 1992; Clariana, Monereo y Pérez Cabaní, 1992; Barbier, 1993; Espín y Rodríguez, 1993; Jornet, 1993; Boulton-Lewis, 1994, 1995; Pérez Cabaní y Terradellas Piferrer, 1996a y 1996b).

Pero un paso previo a la búsqueda de instrumentos de medida es concretar a qué nos referimos cuando hablamos de calidad del aprendizaje: ¿es de más calidad el aprendizaje del alumno que recuerda más información? o, ¿el que consigue aquel otro que sabe explicar con sus propias palabras lo que ha aprendido? o, ¿el que realiza un tercer aprendizaje que es capaz de solucionar un problema que se le plantea?

Marco conceptual de referencia

Existen numerosos trabajos que definen la calidad del aprendizaje a partir de la adopción de diferentes criterios, así, si se consideran los resultados del aprendizaje, la calidad se puede definir como la consecución de metas y objetivos, como adecuación a los propósitos; si se atiende más al proceso de enseñanza, la calidad se puede considerar como consistencia, o calidad como disponibilidad de recursos. Más actual es la consideración de calidad en términos de transformación, cambio y mejora —perspectiva a la que nos adherimos—. Esta nueva consideración lleva implícita una visión del profesor como facilitador del aprendizaje y del alumno como participante activo en este proceso. Como pone de relieve Gallagher (1994), durante los últimos años se ha modificado la concepción de la función de los alumnos, pasando de considerarse receptores pasivos de información a aprendices que toman decisiones —*decision makers*— sobre qué y cómo aprenden. Desde esta perspectiva, el interés se centra en que los alumnos participen más en los procesos de enseñanza y aprendizaje, y en valorar si son capaces de pensar y actuar de forma independiente y articulada después de abandonar la universidad (Barnett, 1992; Harvey, Burrows y Green, 1993; De Miguel, 1995; Hernández Pina, 1997).

En un análisis pormenorizado de la participación y actividad de los alumnos en el proceso de aprendizaje algunos autores (Marton, 1988; Entwistle, 1987; Resnick, 1989; Snow y Swanson, 1992) señalan cómo la calidad del aprendizaje está determinada por el enfoque adoptado para aprender. Así, un enfoque superficial conducirá a un aprendizaje basado en la retención de hechos, un enfoque profundo favorecerá la comprensión y estructuración de la información, y un enfoque estratégico tendrá como objetivo la obtención de buenas calificaciones.

Estas características cualitativas del aprendizaje se pueden entender como los dos extremos de un *continuum*, cuyo análisis requeriría centrar la atención, no en cuántos contenidos ha aprendido un estudiante, sino en qué aprende (a relacionar conceptos, a solucionar problemas, a aplicar un conocimiento en una situación concreta, a recordar hechos...) y cómo lo aprende (qué procedimientos de aprendizaje conoce, cómo los utiliza y con qué objetivo en una situación determinada), y valorar en qué medida, las habilidades cognitivas que intervienen en

el proceso de aprendizaje afectan la calidad de los resultados que se obtienen al finalizarlo.

Pero la adopción de un enfoque ante una situación de aprendizaje no depende únicamente del alumno, sino que está condicionada por diferentes variables relacionadas con el aprendiz, con el profesor y con las características de la tarea y de la situación en que éste se realice. En este sentido, la percepción que el aprendiz tenga de lo que se le exigirá en términos de evaluación será decisivo en la manera como enfocará el proceso de aprendizaje (Biggs, 1993).

Así pues, de lo expuesto hasta el momento se desprende la importancia de la demanda de evaluación que formule el profesor en la calidad del aprendizaje resultante, pero no podemos olvidar que esta será diferente según conciba el profesor la enseñanza, el aprendizaje y la evaluación. Biggs (1995, 1996a) diferencia al respecto dos concepciones —a nuestro entender como extremos de polos opuestos—, que determinan prácticas evaluativas diferenciadas: desde la primera de ellas el aprendizaje se valoraría en función de la cantidad y acumulación de contenido, y la enseñanza como transmisión del mismo, lo que conduciría a evaluar en términos binarios la cantidad y fidelidad de la información reproducida. Desde la segunda perspectiva, más próxima a una concepción constructivista, el aprendizaje se concebiría como un proceso activo de construcción de conocimiento en el que la función del profesor sería guiar y orientar este proceso para conseguir un nivel de comprensión más profundo. En consonancia con esta segunda concepción, la evaluación debería contemplar diferentes dimensiones y valorar los cambios cualitativos en relación a qué se aprende y cómo se estructura el conocimiento.

Estas diferentes maneras de concebir la evaluación y de formular las demandas requieren la activación de habilidades cognitivas de diferente nivel de complejidad, por lo que tendrán una incidencia remarcable en el enfoque y la calidad del aprendizaje de los alumnos. En este sentido es especialmente interesante la aportación de Lohman (1993) cuando diferencia entre demandas de evaluación que requieren la solución de problemas de contenido conocido, en las que hay que activar habilidades cristalizadas, que implican un bajo nivel de transferencia, y demandas que requieren un alto nivel de transferencia en las que hay que activar habilidades “fluidas” que generalmente se adquieren a través de la educación y la experiencia.

Así pues, la diversidad de variables que intervienen en el proceso de enseñanza-aprendizaje —objetivos educativos, características del contenido, conocimiento previo de los alumnos, tipos de interacción que se produce...—, así como la originalidad de cada situación educativa susceptible de ser evaluada, ponen de manifiesto cómo buscar instrumentos para evaluar la calidad del aprendizaje es una tarea compleja. Por consiguiente, será necesario disponer de instrumentos elaborados con criterios amplios y suficientemente flexibles, que ofrezcan la posibilidad de evaluar la calidad del aprendizaje teniendo en cuenta las variables diferenciales de cada situación.

En el caso de nuestra investigación nos hemos centrado, inicialmente, en las situaciones de examen. Somos conscientes de que los exámenes constituyen únicamente un elemento de la evaluación sumativa, y de que existen otras —y quizás mejores— opciones para evaluar la calidad del aprendizaje de los estudiantes universitarios —trabajos, ejercicios prácticos, portafolios...—, pero también es cierto que los exámenes aún son el medio más utilizado —y en muchas ocasiones el único— en las aulas universitarias. Este hecho lo ponen de manifiesto Grós y Romañá (1995) en un estudio realizado sobre qué implica ser profesor en la universidad, cuando constatan que siguen siendo los exámenes —sea cual sea la

nomenclatura utilizada— los que determinan la calificación de los estudiantes en la universidad.

No es nuestro objetivo analizar en profundidad las razones de este hecho, pero constatarlo nos ha conducido a tomar la decisión de centrar nuestro trabajo en buscar instrumentos que nos permitan analizar la calidad del aprendizaje que reflejan los estudiantes universitarios en los exámenes. Hemos escogido, entre los instrumentos creados hasta el momento con este propósito, la taxonomía SOLO (*Structure of the Observed Learning Outcome*), propuesta por Biggs y Collis (1982). (Ver Anexo 1). Esta decisión se sustenta en que existen algunos trabajos previos en los que se ha utilizado la taxonomía SOLO en el nivel universitario, en los que se ha podido constatar su pertinencia para discriminar diferentes niveles en la calidad del aprendizaje de diversas materias (Trigwell y Prosser, 1990; Pérez Cabaní, 1993; Boulton-Lewis, 1994, 1995; Monereo y Pérez Cabaní, 1996).

OBJETIVOS DE LA INVESTIGACIÓN

En base a los argumentos anteriores, los objetivos que guían esta investigación son los siguientes:

- 1. *Conocer la incidencia de la tipología de las preguntas y de las demandas que se formulan en los exámenes en la calidad del aprendizaje que manifiestan los estudiantes en las respuestas a estas preguntas.*

Nuestro interés se centra en conocer en que medida la tipología de pregunta y la demanda que se formula posibilitan diferenciar niveles de calidad del aprendizaje en la respuesta de los estudiantes.

- 2. *Valorar la adecuación de la taxonomía SOLO como instrumento de evaluación de la calidad del aprendizaje de los estudiantes universitarios.*

Como hemos expuesto anteriormente, esta taxonomía ha sido utilizada y validada en diferentes áreas de conocimiento mediante diferentes tipos de actividades. Nuestro objetivo en esta investigación es valorar su utilización para evaluar la calidad del aprendizaje de los estudiantes universitarios a través de las respuestas a las preguntas de examen.

- 3. *Observar la posible relación de los criterios de evaluación de la taxonomía SOLO con los criterios de evaluación utilizados por los profesores universitarios.*

La taxonomía SOLO permite establecer diferentes niveles de calidad del aprendizaje en base a unos criterios preestablecidos: la capacidad, es decir, la cantidad de datos que ha memorizado el sujeto en relación a la cuestión que se le plantea y la relevancia de las interrelaciones que establece entre los datos; las operaciones implicadas en la elaboración de la respuesta (inducción, generalización, ordenación, deducción...), que indican la interrelación lógica que se establece entre la pregunta y las operaciones que el alumno realiza para elaborar la respuesta; y la consistencia de los argumentos y conclusiones que se ponen de manifiesto en las respuestas. Nuestra intención es valorar la adecuación de estos criterios para la evaluación de los estudiantes universitarios, el grado de coincidencia con los criterios de los profesores, y la necesidad de incluir otros criterios complementarios.

- 4. *Analizar la incidencia de las decisiones que toma el profesor respecto a la evaluación en la calidad del aprendizaje de los estudiantes universitarios.*

Entendemos la evaluación como una parte integrante del proceso de enseñanza-aprendizaje. Las variables que intervienen en este proceso están íntimamente interrelacionadas, y las modificaciones que se puedan introducir en alguna de ellas incidirán en como se desarrollen las otras. Así, este objetivo se centra en analizar cómo las decisiones que toma el profesor respecto a la elaboración y rea-

lización del examen pueden incidir en las características de la respuesta que elaboren los estudiantes.

MÉTODO

Muestra

La muestra se compone de 53 preguntas de examen y de 147 respuestas a estas preguntas. Para obtenerla se pidió la colaboración a profesores de la Facultad de Ciencias de la Educación de nuestra Universidad. En total colaboraron de manera voluntaria 10 profesores que nos facilitaron 27 exámenes correspondientes a 15 asignaturas de los diferentes estudios de la Facultad —Psicología, Pedagogía, Magisterio y Educación Social—. De cada examen se seleccionaron 2 preguntas (54 preguntas) de manera que quedaran representadas las distintas tipologías que configuran los criterios teóricos seleccionados así como la totalidad de las tipologías de examen que los profesores utilizan. A lo largo del estudio una pregunta fue desestimada por su imprecisión, quedando definitivamente configurada la muestra en 53 preguntas.

Instrumentos

— El instrumento utilizado para evaluar la calidad del aprendizaje, tal como hemos puesto de manifiesto anteriormente, es la taxonomía SOLO, creada por Biggs y Collis con el objetivo de analizar la organización estructural del conocimiento en función del tipo de respuesta que un estudiante proporciona a una cuestión planteada en diferentes situaciones, contenidos, o áreas de conocimiento. La descripción detallada de esta taxonomía se muestra en el anexo 1.

— El instrumento que hemos seleccionado para analizar la incidencia de las decisiones que toman los profesores de la muestra respecto a la evaluación, en la calidad del aprendizaje de sus estudiantes, es la entrevista semiestructurada. Se ha establecido un guión de los temas a tratar de acuerdo con los objetivos de la investigación y se han formulado unos indicadores para cada punto de la entrevista, propuestos y consensuados por los miembros del equipo.

Esta entrevista se ha utilizado para obtener información sobre las siguientes cuestiones:

- Qué valor otorga el profesor al examen dentro del proceso de evaluación.
- Qué espera cada profesor cuando formula una demanda determinada en una pregunta de examen. Por ejemplo: ¿qué espera de la demanda *explica* ? y, ¿qué espera que respondan los alumnos cuando les pide: *indica brevemente y argumenta* ?
- En qué medida la imprecisión de algunas demandas puede comportar dificultades a los estudiantes para ajustarse al enunciado de la pregunta.
- Qué criterios conducen al profesor a utilizar la misma tipología de preguntas y demandas, o en caso de que haya diversificación, cuáles son las razones del cambio y las situaciones en que se producen.

Procedimiento

Las fases del trabajo que se han desarrollado para llevar a cabo esta investigación se concretan de la siguiente manera:

Primera fase

Con el objeto de conocer la incidencia de la tipología de preguntas y demandas en la calidad del aprendizaje de los estudiantes, se procedió a la distribución de las preguntas de la muestra en base a los criterios teóricos seleccionados: su tipología y la demanda que explicitan en el enunciado.

a) En relación a la tipología de la pregunta, existen múltiples y diferentes alternativas para formular preguntas de examen, así como diferentes criterios para categorizarlas. Muestra de ello es el trabajo de López Feal (1986), en el que se expone un estudio minucioso de la formulación de ítems, se analizan las diferentes alternativas existentes, sus ventajas e inconvenientes. Después de analizar las propuestas de diferentes investigadores (Fermin, 1972; Lauforcade, 1976; López Feal, 1986; Chadwick y Rivera, 1991; Espín y Rodríguez, 1993), hemos tomado como referencia la clasificación que presentamos a continuación, pues aunque no están representadas todas las tipologías descritas por los diferentes autores, todas las preguntas de nuestra muestra quedan incluidas en ellas .

1. Preguntas de ensayo amplio:

Incluye aquellas preguntas en las que no se plantea ningún tipo de limitación en la forma de organizar, seleccionar o presentar el contenido de la respuesta. Ante este tipo de pregunta el estudiante tendrá que organizar la respuesta de la manera que considere más conveniente.

Ejemplo:

Comentar un fragmento o una frase.
Elaborar una propuesta para...

2. Preguntas de ensayo restringido:

Incluye aquellas preguntas formuladas de manera que limitan la forma y/o la proyección de la respuesta que debe dar el estudiante. Los límites pueden ser marcados tanto por el contenido (el alumno tiene que ajustarse estrictamente a lo que se le pide), como por la forma (limitaciones de espacio, tiempo...).

Ejemplos:

Señalar diferencias.
Resumir en 15 líneas.
Explicar siguiendo un orden (cronológico, lógico...).
Explicar ventajas e inconvenientes.
Indicar métodos para...
Aplicar principios.
Interpretar material lingüístico, gráfico...

3. Preguntas de respuesta breve:

Incluye aquellas preguntas que se responden a través de una palabra, frase, número... y aquellas en que los estudiantes tienen que completar una proposición.

Ejemplos:

Quién fue...?
Cómo se llama...?
El autor del libro X es...?

4. Preguntas de opción múltiple:

Incluye aquellas preguntas con un enunciado seguido de diversas alternativas de respuesta, de las cuales solamente una es correcta o mejor que las otras. Pueden adoptar diferentes formas.

Ejemplos:

Ofrecer dos opciones y escoger una (verdadero-falso, sí-no...)
Ofrecer más de dos opciones: a) b) c) d)
Asociar parejas: 1-d, 4-b...

b) El segundo criterio tomado en consideración, como hemos expuesto anteriormente, es la demanda que se explicita en el enunciado de la pregunta. Dar respuesta a la demanda formulada en una pregunta de examen requiere contar con las habilidades cognitivas implicadas en su resolución y conocer y utilizar adecuadamente los procedimientos para dar respuesta a la cuestión planteada (Monereo y col., 1994).

Existen, al respecto, algunos trabajos realizados desde el ámbito de la Psicología de la Instrucción que analizan las habilidades cognitivas implicadas en la resolución de una tarea y las clasifican según diferentes criterios —como el nivel de complejidad o el orden en el procesamiento de la información— (Nickerson, Perkins y Smith, 1987; Monereo, 1991). Pero en relación a nuestros objetivos, debemos tener en cuenta, como exponen Gómez, Jorba y Prats (1997), que las habilidades implicadas en la resolución de una tarea no siempre son excluyentes, al contrario, en muchos casos están íntimamente relacionadas e incluso pueden ser movilizadas a diferentes niveles de complejidad. Considerando pues las aportaciones de estos trabajos y teniendo en cuenta las distintas y complejas relaciones que se pueden establecer entre las diferentes demandas formuladas, hemos optado por contemplarlas separadamente, tal como se formulan en las 53 preguntas que conforman la muestra, y analizar la relación de cada una de ellas con los diferentes niveles de calidad del aprendizaje que reflejan las respuestas de las preguntas analizadas. Las demandas formuladas son las siguientes:

Analizar; Argumentar-Justificar; Calcular; Clasificar; Comentar; Comparar; Exponer-Describir; Definir; Diseñar; Ejemplificar; Explicar; Reconocer-Diferenciar; Recordar; Relacionar.

Las 53 preguntas de examen pueden consultarse en el anexo 2.

Segunda fase

a) Para valorar la adecuación de la taxonomía SOLO como instrumento de evaluación de la calidad del aprendizaje de los estudiantes universitarios, procedimos al análisis y distribución de las respuestas de los estudiantes en los diferentes niveles de la taxonomía —desde el nivel preestructural hasta el nivel abstracto—, en función de los criterios establecidos por la misma. Para situar cada respuesta en el nivel correspondiente y garantizar la fiabilidad se realizó un análisis interjueces —tal como proponen Biggs y Collis (1982)—. Para llevar a cabo este análisis se procedió a la discusión y concreción de los criterios de la taxonomía y se crearon tres grupos compuestos por dos jueces cada uno. Los jueces eran miembros del equipo de investigación y profesores que habían utilizado previamente la taxonomía SOLO en sus evaluaciones. Seguidamente se repartieron las preguntas de la muestra al azar entre los grupos quedando así la muestra de respuestas a analizar repartida de manera equiparada entre los tres grupos. Cada juez situó individualmente cada respuesta de las preguntas correspondientes en uno de los niveles establecidos en la taxonomía, pasando a continuación a la contrastación con el otro juez de su grupo y posteriormente a la discusión entre todos los jueces en los casos en los que hubo desacuerdo.

b) Análisis de la relación existente entre la distribución de las respuestas en los diferentes niveles de la taxonomía SOLO, la tipología de la pregunta y la demanda formulada. Para ello se ha procedido a clasificar la información recogida y a relacionar las variables en términos percentuales.

Tercera fase

Para conocer la posible relación de los criterios de evaluación de la taxonomía SOLO con los criterios de evaluación utilizados por los profesores universitarios, así como la incidencia de las decisiones que toma el profesor respecto a la evaluación en la calidad del aprendizaje de los estudiantes, hemos analizado a través de la información que nos aporta la entrevista, los criterios que utilizan los profesores al tomar diferentes opciones para elaborar los exámenes y formular las demandas de las preguntas, los criterios de corrección y el valor de los exámenes en el proceso de evaluación, y la información que han tenido los estudiantes al respecto.

Los diferentes temas tratados hacen referencia a aspectos generales del proceso de evaluación y también a cuestiones concretas relacionadas directamente con los enunciados de las preguntas seleccionadas. Estos aspectos se han introducido en el guión procurando no inducir al entrevistado a dar respuestas sesgadas, por este motivo, la entrevista consta de una primera parte en la que se plantean algunos aspectos preliminares e introductorios, de una segunda parte en la que se abordan cuestiones específicas en relación a las preguntas seleccionadas, y de una tercera parte en la que se plantean aspectos generales respecto a la evaluación. Las entrevistas se han registrado en audiograbaciones que se han transcrito literalmente para su análisis posterior.

RESULTADOS

En relación a la tipología y la demanda de las preguntas

En la tabla I se muestra la distribución resultante de las preguntas en función de los dos criterios seleccionados.

Aunque la información recogida sobre la tipología y la demanda de las preguntas no nos permite establecer conclusiones, aporta algunos datos interesantes, ya que, como se puede observar en la tabla I existe una estrecha relación entre la tipología de la pregunta y la demanda que se propone. Lo más frecuente es que una demanda corresponda a una única tipología o a tipologías similares (ensayo amplio-restringido, respuesta breve-opción múltiple). La tipología de pregunta que utilizan con más frecuencia los profesores de la muestra es la de ensayo restringido y la demanda más frecuente es reconocer-diferenciar.

a) Mostramos en la tabla II el resultado de los análisis interjueces realizados para garantizar la fiabilidad de la distribución de las respuestas de los estudiantes en los diferentes niveles de la taxonomía SOLO.

Como se pone de manifiesto en la tabla II, el grado de acuerdo de los jueces en la clasificación de las respuestas en los diferentes niveles de la taxonomía SOLO sobrepasa el 90% del total del análisis realizado, lo cual garantiza la fiabilidad y permite establecer las relaciones que mostramos en las tablas siguientes.

b) La tabla III muestra la relación que se establece entre el nivel de aprendizaje que se obtiene en la taxonomía SOLO y la tipología de las preguntas formuladas.

c) La tabla IV muestra la relación que se establece entre el nivel de aprendizaje que se obtiene en la taxonomía SOLO y la demanda de las preguntas formuladas.

Entre los datos recogidos en las tablas III y IV destacaríamos especialmente dos cuestiones. La primera, que las preguntas de opción múltiple y de respuesta breve analizadas, salvo una excepción, generan respuestas que sitúan la calidad del aprendizaje de los alumnos únicamente en los tres primeros niveles de la

TABLA I
Distribución de las preguntas en función de la tipología y la demanda

TIPOLOGÍA ↓	Ensayo amplio	Ensayo restringido	Respuesta breve	Opción múltiple	Total y %
DEMANDA					
Analizar	28	30-44-45-49		12 a)	6 (11,32%)
Argumentar Justificar		13-14-17- 26-37-43			6 (11,32%)
Calcular			12c)-47		2 (3,77%)
Clasificar		20			1 (1,88%)
Comentar	16-33-36	50-51			5 (9,43%)
Comparar			25		1 (1,88%)
Exponer/ Describir		38	22-24		3 (5,66%)
Definir			10-21-23		3 (5,66%)
Diseñar	31-34-35-40	15-18-39			7 (13,20%)
Ejemplificar		32			1 (1,88%)
Explicar	19-41-42		46		4 (7,54 %)
Reconocer- Diferenciar			12b)-48	1-2-3-4-5- 6-7-9-11	11 (20,75%)
Recordar			8		1 (1,88%)
Relacionar		29		27	2 (3,77%)
Total y %	11 (20,75%)	19 (35,84%)	12 (22,64%)	11 (20,75%)	53 (100%)

taxonomía SOLO. La segunda, que las preguntas con respuesta de ensayo que se han analizado, independientemente de que sean de tipo restringido o amplio, generan respuestas de todos los niveles de la taxonomía SOLO.

En relación a las decisiones que toman los profesores respecto a la evaluación a través de exámenes

El análisis de las decisiones que toman los profesores en la planificación y elaboración de los exámenes, así como los criterios que implícita o explícitamente las condicionan, las hemos analizado a través de las siguientes categorías:

TABLA II
Resultados del análisis interjueces

	Número de respuestas a analizar	Número de desacuerdos	Número de acuerdos	% acuerdos
Grupo interjueces 1	49	4	45	91,8%
Grupo interjueces 2	50	4	46	90,0%
Grupo interjueces 3	48	5	43	89,5%
TOTAL	147	13	134	media acuerdos 90,4%

TABLA III
Relación entre la tipología de las preguntas y la calidad del aprendizaje valorada a través de la taxonomía SOLO

TIPOLOGÍA ↓	Ensayo amplio n=31	Ensayo restringido n=65	Respuesta breve n=26	Opción múltiple n=25	Total N=147
NIVEL SOLO					
Preestructural n=8		6 (9,23%)	2 (7,69%)		8 (5,44%)
Uniestructural n=55	12 (38,70%)	15 (23,05%)	12 (46,75%)	16 (64%)	55 (37,40%)
Multiestructural n=48	10 (32,25%)	21 (32,30%)	12 (46,15%)	5 (20%)	48 (32,65%)
Relacional n=34	9 (29,03%)	21 (32,30%)		4 (16%)	34 (23,12%)
Abstracto n=2		2 (3,07%)			2 (1,36%)

a) Premisas de las que parten los profesores para preparar el examen de una asignatura

Los profesores entrevistados toman como referencia para preparar los exámenes tanto los objetivos como los contenidos que se han impartido en la asignatura. También se plantean recoger las ideas principales de los temas tratados y valo-

TABLA IV
Relación entre la demanda de las preguntas y la calidad del aprendizaje valorada a través de la taxonomía SOLO

NIVEL SOLO ↓	Preestructural n=8	Uniestructural n=55	Multiestructural n=48	Relacional n=34	Abstracto n=2
DEMANDA					
Analizar n=13		2 (15,38%)	5 (38,46%)	6 (46,15%)	
Argumentar Justificar n=20	3 (15%)	5 (25%)	7 (35%)	4 (20%)	1 (5%)
Calcular n=3		3 (100%)			
Calsificar n=3		1 (33,33%)	2 (66,66%)		
Comentar n=17	1 (5,88%)	5 (29,41%)	3 (17,64%)	8 (47,05%)	
Comparar n=3		1 (33,33%)	2 (66,66%)		
Exponer/ Describir n=12	1 (8,33%)	3 (25%)	7 (58,33%)	1 (8,33%)	
Definir n=7		4 (57,14%)	3 (42,85%)		
Diseñar n=28		7 (25%)	11 (39,28%)	9 (32,14%)	1 (3,57%)
Ejemplificar n=3		2 (66,66%)	1 (33,33%)		
Explicar n=6		2 (33,33%)	3 (50%)	1 (16,66%)	
Reconocer- Diferenciar n=23	1 (4,34%)	18 (78,26%)	4 (17,39%)		
Recordar n=2		2 (100%)			
Relacionar n=7	2 (28,57%)			5 (71,42%)	

rar la consecución de unos “mínimos” de cada tema o cuestión planteada. Otras cuestiones que se tienen en cuenta son: la manera en que se ha impartido la asignatura y el nivel de profundidad con que se han tratado los contenidos; la elaboración de las preguntas de manera que permitan distinguir el nivel de comprensión de los alumnos; el tiempo disponible para responder el examen; la necesidad de discriminar entre los estudiantes que conocen los conceptos básicos de la asignatura y los que no; el número de alumnos y el calendario de evaluación.

b) Valor que otorgan los profesores al examen en el proceso de evaluación

Los profesores entrevistados otorgan al examen un valor máximo en la evaluación. Incluso en algún caso en que se complementa con otros instrumentos, continúa siendo un elemento imprescindible para superar la asignatura. El valor otorgado se justifica con los siguientes argumentos:

— Los exámenes constituyen la única referencia para aprobar o suspender y son el medio que proporciona a los estudiantes la posibilidad de promocionar en los estudios.

— La gran cantidad de alumnos por asignatura y la organización por cuatrimestres dificulta la utilización de otros instrumentos y metodologías.

— El examen, bien preparado, es un instrumento que permite utilizar unos criterios de evaluación objetivos y fiables —existe cierta preocupación ante la posibilidad de ser injustos o no utilizar un mismo criterio para todos los alumnos—.

c) Elaboración de las preguntas

El proceso de elaboración de las preguntas está muy condicionado por la demanda que se quiere formular y por la tipología de pregunta que se selecciona.

Así, cuando el objetivo de los profesores es que las preguntas conduzcan a los alumnos a analizar e interpretar diferentes situaciones; a poner de manifiesto lo que recuerdan y lo que han entendido; a sintetizar las ideas que exponen; a argumentar con criterios claros el porqué de sus decisiones basándose en los contenidos de la asignatura, y a justificar sus propuestas, la tipología escogida es el ensayo (amplio o restringido).

Cuando las preguntas se elaboran con el objetivo de que los alumnos recuerden unos conceptos, que los seleccionen o los discriminen, la tipología de pregunta escogida es la de opción múltiple o respuesta breve. También, en algunos casos, se escogen estas tipologías cuando se pretende que los examinados apliquen una fórmula o un principio a una situación concreta. El objetivo de los profesores es que las preguntas conduzcan a los alumnos a utilizar un vocabulario básico; a reconocer y memorizar unos conceptos y la relación que se establece entre ellos.

d) Criterios de puntuación y corrección de las preguntas

En las preguntas de ensayo se otorga la máxima puntuación cuando las respuestas están argumentadas, cuando se justifican las decisiones tomadas y cuando se ajustan a la demanda de la pregunta. Se tiene en cuenta la relevancia de los contenidos, el nivel de elaboración de los mismos, la adecuación de la terminología utilizada, etc.

En las preguntas de opción múltiple y respuesta breve el criterio es de todo o nada, por lo que no es posible diferenciar matices y niveles de calidad.

e) Tiempo y espacio del que disponen los alumnos para realizar el examen

Ningún profesor plantea que estas variables sean determinantes en el resultado de los exámenes. En el caso de las preguntas de ensayo, no se considera que dejar más tiempo lleve a responder preguntas más elaboradas.

f) Factores que creen los profesores que influyen en las respuestas que dan los estudiantes a los exámenes

Las opiniones de los profesores al respecto son diversas, aunque se pueden diferenciar claramente dos grupos en función de la tipología de pregunta que se utiliza.

Los profesores que formulan preguntas de ensayo consideran factores de influencia la capacidad de adaptarse a la demanda del profesor y los conocimientos previos al respecto; la asistencia a clase y la participación en las actividades que se realizan en el aula; el interés que se haya puesto en estudiar la materia; la capacidad para resolver problemas; y la capacidad de pensar y relacionar.

Los profesores que formulan preguntas de respuesta breve o de opción múltiple consideran que la única incidencia es el conocimiento de los alumnos “o lo saben o no lo saben”.

DISCUSIÓN DE LOS RESULTADOS

El análisis de la evaluación de la calidad del aprendizaje en situación de examen en el ámbito universitario es un proceso complejo, pero los resultados obtenidos en esta investigación nos aportan informaciones de interés en relación a los objetivos planteados y nos ofrecen indicadores para orientar su continuidad.

(1) Respecto a la incidencia de la tipología de las preguntas y de las demandas que se formulan en los exámenes en la calidad del aprendizaje de los estudiantes, nos parece interesante destacar las siguientes cuestiones:

Una tipología de pregunta va asociada generalmente a unas demandas determinadas. Esto es comprensible, ya que difícilmente se puede hacer la demanda de argumentar o de elaborar una propuesta de actuación en una pregunta de respuesta múltiple, así como tampoco tiene demasiado sentido hacer una demanda de reconocer-diferenciar en una pregunta de ensayo amplia. No obstante, este hecho tiene importantes implicaciones respecto a la evaluación.

En primer lugar, no podemos olvidar que lo que queda reflejado en el examen es solamente un aspecto parcial de la estructura de conocimiento de los estudiantes sobre un tema. Cualquiera que sea la tipología y la demanda de la pregunta formulada estará dirigida y enfocada a un aspecto parcial del conocimiento. Cuanto más focalizada y concreta sea la pregunta, mayor será el nivel de restricción que impone al alumno. Pero en cualquier caso, los estudiantes tienen que disponer de un conocimiento y de unas habilidades que les permitan utilizar los procedimientos más adecuados para dar respuesta a una demanda determinada. Como señala Lohman (1993) la solución de problemas complejos en un dominio requiere, además de la comprensión clara de los conceptos y de las controversias

existentes en el mismo, que los estudiantes dispongan de las habilidades necesarias para organizar y conectar de manera flexible la información relevante, lo que implica, a su vez, un elevado nivel de transferencia a otros contextos cuando el problema plantea situaciones alejadas de la propia experiencia. En este sentido, las decisiones del profesor serán especialmente relevantes en relación a la calidad del aprendizaje que se puede manifestar. Estaríamos de acuerdo con Taylor (1994) cuando expone que algunos formatos y demandas no permiten llegar a niveles complejos de respuesta, sino simplemente a una asociación de datos. Por lo tanto, un profesor que solo plantee un tipo de demanda, estará potenciando y evaluando en los alumnos unas habilidades y unos procedimientos determinados, ignorando otros quizás tanto o más necesarios para adquirir conocimiento relevante sobre un tema.

Centrándonos en los datos obtenidos en nuestro estudio, las habilidades y los procedimientos que “se tienen que poner en juego” para dar respuesta a las demandas de las preguntas que hemos analizado son muy diversos, y comportan un proceso de estudio y de actuación muy diferente durante el examen. No requiere el mismo proceso disponer de unos datos para posteriormente en un examen reconocerlos, que disponer de los criterios necesarios para diferenciar dos teorías y argumentarlas. Estas diferencias se ponen también de manifiesto en un estudio realizado por Tang (1991) —referenciado por Biggs (1995)—, cuando muestra que diferentes formatos y demandas de examen requieren diferente conocimiento procedimental. Según este autor, para dar respuesta a las demandas de alto nivel de complejidad cognitiva, los estudiantes necesitan disponer de estrategias de aprendizaje sin las cuales les sería imposible realizar la tarea. Para adquirirlas, Tang propone que se ponga de manifiesto en las actividades de evaluación que lo importante es pensar en profundidad y desde diferentes perspectivas sobre un tema, no obtener un determinado número de ítems correctos.

Por otra parte, los alumnos son conscientes de esta diversidad, y quizás por este motivo ponen tanto interés —hecho que manifiestan los profesores en las entrevistas— en conocer la demanda del profesor y la tipología de examen. Nos parece especialmente interesante la aportación de Biggs (1996a) al respecto cuando indica que el alumno va construyendo la estructura de una materia en función de cómo interpreta la demanda de evaluación. De esta manera, es la evaluación, más que los objetivos o las intenciones del currículum, la que dirige las decisiones del alumno. Desde esta perspectiva, Entwistle y Entwistle (1992) diferencian cinco niveles jerárquicos de comprensión:

(1) Reproducir el contenido de los apuntes y materiales de clase sin poseer una estructura clara.

(2) Reproducir el contenido utilizando la estructura del profesor.

(3) Desarrollar una estructura propia pero solamente para generar respuestas anticipadas a las preguntas del examen.

(4) Ajustar las estructuras de forma estratégica de manera que representen la comprensión personal, pero controlando los requerimientos del examen.

(5) Desarrollar una concepción individual de la disciplina desde una perspectiva amplia y reflexiva.

Aunque estaríamos de acuerdo con los autores en que la evaluación tendría que potenciar y valorar el último de estos niveles de comprensión, la realidad muestra que no siempre es así. En un estudio realizado por Clariana, Monereo y Pérez Cabaní (1992), sobre la influencia de la percepción de los estudiantes respecto a la demanda del profesor en el resultado del examen, se pudo constatar que los alumnos que tienen una idea más precisa de la demanda que formulará el profesor son los que obtienen puntuaciones más elevadas.

(2) En cuanto a la adecuación de la taxonomía SOLO como instrumento de evaluación de la calidad del aprendizaje de los estudiantes universitarios destacáramos las siguientes cuestiones:

La taxonomía SOLO demuestra ser un instrumento efectivo para evaluar con criterios objetivos preguntas con demandas que permitan abordar un tema desde diferentes puntos de vista y poner de manifiesto distintos niveles de comprensión y profundidad en el tratamiento de un tema o cuestión. Respecto a su relación con la tipología y la demanda de las preguntas, es necesario remarcar que no todas las demandas ofrecen la posibilidad de diferenciar niveles de calidad del aprendizaje, algunas —como por ejemplo enumera, calcula...— sólo permiten poner de manifiesto aspectos muy parciales del aprendizaje realizado. Por este motivo, generalmente en preguntas de opción múltiple y respuesta breve resulta muy difícil discriminar los niveles superiores de la taxonomía. De hecho, Biggs y Collis al describir las condiciones de utilización de la taxonomía ponen de manifiesto que las características de la demanda y el formato de las preguntas en sí mismas pueden ser un factor determinante del nivel jerárquico de calidad de la respuesta. Cuestiones como la claridad en la redacción de la pregunta, la familiaridad que tengan los estudiantes con los detalles de la tarea propuesta o su nivel de complejidad pueden facilitar o dificultar que se produzcan diferentes niveles de respuesta.

Por otra parte, la utilización de la taxonomía SOLO comporta algunas limitaciones que sería necesario solventar: cuando en una pregunta se discriminan todos los niveles de la taxonomía, el nivel intermedio —multiestructural— abarca una amplia gama de respuestas, es decir, si bien es relativamente fácil discriminar las respuestas que se sitúan en los niveles superiores e inferiores, el nivel intermedio es poco discriminativo. Los resultados obtenidos por Boulton-Lewis (1994) al respecto, son similares a los nuestros, ya que la mayoría de la población estudiada —alumnos y profesores universitarios que responden a la pregunta “¿qué significa para ti aprender?”— se sitúa en el nivel multiestructural de la taxonomía. Será necesario pues, buscar alternativas para optimizar la utilización de este instrumento en diferentes materias y estudios universitarios, y complementarlo con otros instrumentos que bajo criterios amplios permitan evaluar la calidad del aprendizaje de los estudiantes universitarios. Para conseguir estos propósitos Biggs (1996b) sugiere la utilización de portafolios de manera complementaria a la taxonomía SOLO, para poder valorar desde la comprensión y la aplicación de un contenido hasta la reflexión profunda y personal sobre el cambio conceptual y la toma de decisiones.

(3) En tercer lugar, de la relación entre los criterios de evaluación de la taxonomía SOLO y los que utilizan los profesores que componen nuestra muestra nos gustaría destacar los siguientes aspectos:

La aproximación al análisis comparativo de los criterios de evaluación de la taxonomía SOLO, y los criterios utilizados por profesores de nuestra muestra, ponen de manifiesto que mientras la taxonomía SOLO propone unos criterios de evaluación orientados hacia la valoración de la calidad del aprendizaje que realizan los estudiantes en función de la complejidad de la estructura cognitiva implicada en una demanda determinada, los criterios que guían la toma de decisiones de nuestros profesores para elaborar el examen responden a premisas muy diversas que ponen de manifiesto, tal como apuntábamos en la introducción de este artículo, la ausencia de un modelo de evaluación consensuado. Las divergencias se producen, básicamente, en relación a tres variables:

(a)Cuál es el objetivo de la evaluación y más concretamente del examen. A este respecto los criterios abarcan desde considerar que el examen sirve para dis-

criminar o medir los conceptos básicos que conocen los alumnos, hasta valorar su utilidad para conocer la comprensión que tiene un alumno de una materia.

(b) Qué aspectos o contenidos se tienen que valorar o a cuáles dar prioridad —nivel de elaboración, adecuación de la terminología—.

(c) Cuál ha de ser el rol del profesor, el del alumno y el del contenido en todo este proceso —valor que se concede a la docencia que se ha impartido, factores que pueden influenciar en la respuesta del alumno, características de la materia que implican un tipo de respuesta...—.

Estas variables coinciden en gran parte con el análisis realizado por Biggs (1995) quien considera que los criterios que utiliza el profesor en la elaboración de un examen responden, al menos, a tres dimensiones: en primer lugar, a la función que se otorga al examen dentro del proceso de evaluación, en segundo lugar, a la consideración del profesor respecto a la naturaleza de aquello que se tiene que evaluar, y en tercer lugar al contexto en que se sitúa el examen. En base a estas dimensiones se pueden diferenciar seis opciones para diseñar un examen, que abarcan desde realizar una medición de manera cuantitativa y descontextualizada, hasta tomar una norma de referencia cualitativa y contextualizada —lugar en el que sería óptimo utilizar la taxonomía SOLO como instrumento para diferenciar niveles de calidad—. Biggs también plantea que existe una desproporción, una tendencia hacia el primero de estos extremos, y que hace falta un cambio. Pero conseguirlo es difícil, no porque los profesores no vean la necesidad de evolucionar hacia modelos más cualitativos, sino porque falta bagaje para interpretar los resultados, y la fiabilidad y validez de las nuevas técnicas no es evidente. Esta última cuestión también es consistente con los resultados de nuestro estudio, ya que se pone de manifiesto, en algunas entrevistas, cierta inseguridad de los profesores ante el proceso de evaluación en general, respecto a los criterios “adecuados” a tener en cuenta y, más concretamente, a tener que justificar ante los alumnos la calificación otorgada. De todo ello se desprende que aún queda mucho trabajo por realizar para resolver algunos de los problemas que plantea la evaluación del aprendizaje de los estudiantes en el ámbito universitario.

(4) Finalmente, de la incidencia de las decisiones que toma el profesor respecto a la evaluación en la calidad del aprendizaje de los estudiantes, nos gustaría destacar las siguientes consideraciones:

Por lo que se desprende de las entrevistas realizadas a los profesores, los factores que determinan las decisiones respecto a la evaluación serían, de manera sucinta, la concepción de qué significa enseñar y aprender —“...yo lo que pido a los alumnos es que no memoricen, sino que sepan explicar...”—, las características de la materia —tipo de contenido que los alumnos deben tener claro—, y algunas condiciones del contexto —número de alumnos y falta de tiempo para corregir—.

Estos factores son de diferente índole y es difícil analizarlos conjuntamente, además, no siempre quedan reflejados en los exámenes analizados, lo que pone de manifiesto que lo que dice el profesor que hace no siempre es coincidente con la demanda que llega a los estudiantes. En este sentido, nos parece interesante la distinción que recoge Biggs (1996b) de un estudio realizado por Argyris (1976) en el que se expone que los profesores generalmente toman sus decisiones en línea a su concepción —explícita o implícita— del aprendizaje y la enseñanza, y distingue entre teorías que guían la práctica profesional —teorías de referencia— y teorías que guían la práctica en una situación concreta —teorías en uso—, las cuales no siempre son coincidentes, lo que explicaría el desajuste que se produce en algunos casos entre las decisiones del profesor y su actuación y,

consecuentemente, la desorientación del alumno respecto a la demanda que se le formula.

Así pues, los datos obtenidos ponen de manifiesto que la evaluación de la calidad del aprendizaje es un proceso complejo, y aunque conceptualmente existe un relativo consenso sobre *qué significa*, se puede constatar que en el momento de realizar una actividad de evaluación es difícil partir de un criterio común y globalizador. Cuando los profesores evalúan la calidad del aprendizaje, a menudo se centran en aspectos singulares que pueden dificultar la percepción de los alumnos, no únicamente de la demanda, sino también de las intenciones del profesor.

CONCLUSIONES

Dedicaremos este último apartado a comentar brevemente algunas reflexiones sobre la evaluación de la calidad del aprendizaje, que hemos ido elaborando y perfilando a lo largo del proceso de la investigación realizada.

En primer lugar, querríamos explicitar nuestro convencimiento de que aún quedan muchos interrogantes por resolver respecto a la evaluación de la calidad del aprendizaje de los estudiantes universitarios. La evaluación tiene una función reguladora del aprendizaje del alumno, ya que muchas de las decisiones sobre qué y cómo aprender vendrán condicionadas por el tipo de evaluación al que los estudiantes tengan que enfrentarse. Cuestiones como qué representaciones sobre la evaluación conducen a un profesor a tomar unas decisiones y a actuar de una manera determinada, qué instrumentos de evaluación pueden ser más adecuados y en qué momentos, o de qué manera inciden los criterios y decisiones del profesor en las decisiones que toma el alumno en su proceso de aprendizaje, son temas sobre los que habrá que continuar investigando.

Por otra parte, también nos gustaría poner de manifiesto que la evaluación del aprendizaje en la Universidad tiene implicaciones sociales y éticas que pesan sobre el profesorado y que llevan a una preocupación constante por la objetividad. Éste es quizás el principal motivo de que los exámenes continúen siendo un instrumento imprescindible y decisivo en la calificación del resultado del aprendizaje, aunque no constituyan por sí mismos una garantía en tal sentido, puesto que no existen unos criterios de calificación ni un modelo de evaluación consensuado.

Por último, no podemos olvidar que el estudio de la evaluación como parte fundamental del proceso educativo y como factor determinante de la calidad del aprendizaje, es un elemento decisivo para afrontar los nuevos retos y cambios educativos que la sociedad de la información y la revolución tecnológica plantean. Contar con una educación de calidad será fundamental para afrontar los retos de la sociedad del futuro, y conseguirla implica no sólo partir de criterios objetivos de calificación sino que la gestión del proceso formativo, a todos los niveles, como ponen de manifiesto Le Boterf *et al.* (1993), adquiera una relevancia que desde el modelo tradicional de transmisión de conocimientos es difícil imaginar.

En este sentido, nuestro análisis de la incidencia de la evaluación en la calidad del aprendizaje de los estudiantes universitarios es una aportación más al estudio de la calidad de los procesos educativos y formativos, y, en definitiva, a lo que consideramos que es —y será cada día más— una preocupación y una necesidad de readaptación de nuestro sistema educativo superior a los nuevos tiempos. Estas cuestiones son las que guiarán la continuidad de nuestro trabajo.

Referencias

- ARGYRIS, C. (1976). Theories of action that inhibit individual learning. *American Psychologist*, 31, 638-654.
- BARRIER, J. M. (1993). *La evaluación en los procesos de formación*. Barcelona: Paidós/MEC.
- BARNETT, R. (1992). *Improving Higher Education: total quality care*. Buckingham: Society for Research into Higher Education and Open University Press.
- BIGGS, J. B. (1988). Approaches to learning and to essay writing. En R. R. Schmeck (Ed.), *Learning Strategies and Learning Styles* (pp. 185-228). New York: Plenum Press.
- BIGGS, J. B. (1993). What do inventories of students' learning processes really measure? A theoretical review and clarification. *British Journal of Educational Psychology*, 63, 3-19.
- BIGGS, J. B. (1995). Assessing for learning: Some dimensions underlying new approaches to educational assessment. *The Alberta Journal of Educational Research*, XLI, 1, 1-17.
- BIGGS, J. B. (1996a). Assessing learning quality: reconciling institutional, staff and educational demands. *Assessment and Evaluation in Higher Education*, 21, 1, 5-15.
- BIGGS, J. B. (1996b). Enhancing teaching through constructive alignment. *Higher Education*, 32, 347-364.
- BIGGS, J. B. & COLLIS, K. F. (1982). *Evaluating de quality of learning. The SOLO taxonomy* (Structure of the Observed Learning Outcome). New York: Academic Press.
- BOULTON-LEWIS, G. (1994). Tertiary students' knowledge of their own learning and SOLO Taxonomy. *Higher Education*, 28, 387-402.
- BOULTON-LEWIS, G. (1995). The SOLO Taxonomy as a Means of Shaping and Assessing Learning in Higher Education. *Higher Education Research and Development*, 14, 2, 143-154.
- CHADWICK, C. B. y RIVERA, N. I. (1991). *Evaluación formativa para el docente*. Barcelona: Paidós.
- CLARIANA, M., MONEREO, C. y PÉREZ CABANÍ, M. L. (1992). Percepción del modelo instruccional del profesor y rendimiento académico. En C. Marcelo y P. Mingorance (Eds.), *Pensamiento de Profesores y Desarrollo Profesional. Formación inicial y permanente* (pp. 453-462). Sevilla: Universidad de Sevilla.
- CONSEJO DE UNIVERSIDADES (1994). *Programa experimental de evaluación de la calidad del sistema universitario: Informe final*. Madrid: MEC.
- CONSEJO DE UNIVERSIDADES (1995). *Programa de Evaluación de la Calidad de las Universidades*. Madrid: MEC.
- CONSEJO DE UNIVERSIDADES (1996). *Plan Nacional de Evaluación de la Calidad de las Universidades. Guía de evaluación*. Madrid: MEC.
- CONSEJO DE UNIVERSIDADES (1997). *Informe sobre los resultados de la primera convocatoria del Plan Nacional de Evaluación de la Calidad de las Universidades*. Madrid: Secretaría General del Consejo de Universidades.
- DE MIGUEL, M. (1995). La calidad de la educación y las variables de proceso y de producto. *Educación*, 8, 29-51.
- ENTWISTLE, N. (1987). *La comprensión del aprendizaje en el aula*. Barcelona: Paidós/MEC, 1988.
- ENTWISTLE, N. (1992). *The impact of teaching on learning outcomes in higher education*. Edimburgh: USDU/Commissioned by the Employment Department Training, Enterprise and Education Directorate.
- ENTWISTLE, A. & ENTWISTLE, N. (1992). Experiences of understanding in revising for degree examinations. *Learning and Instruction*, 2, 1-22.
- ELTON, L. & LAURILLARD, D. (1979). Trends in student learning. *Studies in Higher Education*, 4, 87-102.
- ESPÍN, J. y RODRÍGUEZ, M. (1993). *L'avaluació dels aprenentatges a la universitat*. Barcelona: Universitat de Barcelona.
- FERMÍN, M. (1972). *La evaluación, los exámenes y las calificaciones*. Buenos Aires: Kapelusz.
- GALLAGHER, J. J. (1994). Teaching and Learning: New models. *Annual Review of Psychology*, 45, 171-195.
- GÓMEZ, L., JORBA, J. y PRATS, A. (1997). L'ús de la llengua en situació d'aprenentatge des de les àrees curriculars. *Articles de Didàctica de la Llengua i la Literatura*, 13, 79-90.
- GROS, B. y ROMANA, T. (1995). *Ser profesor. Palabras sobre la docencia universitaria*. Barcelona: Publicacions de la Universitat de Barcelona.
- HARVEY, L., BURROWS, A. & GREEN, D. (1993). *Criteria of Quality: summary papers*. Birmingham: Quality in Higher Education Project, University of Central England in Birmingham.
- HERNÁNDEZ PINA, F. (1997). El aprendizaje de los alumnos en el marco del Plan Nacional de Innovación de las instituciones universitarias. En H. Salmerón Pérez (Ed.), *Evaluación Educativa. 7ª Jornadas sobre la LOGSE*. Granada: Grupo Editorial Universitario.
- JORNET, J. M. (1993). Enfoques de la evaluación universitaria. En III Jornadas Nacionales de Didáctica Universitaria "Evaluación y desarrollo profesional" (pp. 63-86). Las Palmas de Gran Canaria: Universitat de Las Palmas de Gran Canaria.
- LAUFORCADE, P. D. (1976). *Evaluación de los aprendizajes*. Buenos Aires: Kapelusz.
- LE BOTERF, G., BARZUCCHETTI, S. y VINCENT, F. (1993). *Cómo gestionar la calidad de la formación*. Barcelona: AEDIPE.
- LOHMAN, D. F. (1993). Teaching and testing to develop fluid abilities. *Educational Researcher*, 22, 7, 1-23.
- LÓPEZ FEAL, R. (1986). Construcción de instrumentos de medida en ciencias conductuales y sociales. Barcelona: Alamex.
- MATEO, J. (1993). Presentation: En J. Espín y M. Rodríguez, *L'avaluació dels aprenentatges a la universitat* (pp. 11-13). Barcelona: Universitat de Barcelona.
- MARTON, F. (1988). Describing and improving learning. En R. R. Schmeck (Ed.), *Learning Strategies and Learning Styles* (pp. 53-82). New York: Plenum Press.
- MARTON, F., HOUNSELL, D. & ENTWISTLE, N. (Eds.) (1984). *The experience of learning*. Edinburgh: Scottish Academic Press.
- MONEREO, C. (1991). Procesa: un proyecto curricular basado en estrategias de aprendizaje. En C. Monereo (Comp.), *Enseñar a pensar a través del currículum escolar* (pp. 63-96). Barcelona: Casals.
- MONEREO, C. (Coord.), CASTELLÓ, M., CLARIANA, M., PALMA, M. y PÉREZ, M. L. (1994). *Estrategias de enseñanza y aprendizaje. Formación del profesorado y aplicación en la escuela*. Barcelona: Graó.

- MONEREO, C. y PÉREZ CABANÍ, M. L. (1996). La incidencia de la toma de apuntes sobre el aprendizaje significativo. Un estudio en enseñanza superior. *Infancia y Aprendizaje*, 73, 65-86.
- NICKERSON, R. S., PERKINS, D. N. y SMITH, E. (1987). *Enseñar a pensar. Aspectos de la aptitud intelectual*. Barcelona: Paidós/MEC.
- PÉREZ CABANÍ, M. L. (1993). *Ensenyament de l'ús diferencial d'estratègies d'aprenentatge*. Barcelona: Universitat Autònoma de Barcelona. Tesis doctoral publicada en microfichas.
- PÉREZ CABANÍ, M. L. y TERRADELLAS PIFERRER, M. R. (1996a). L'évaluation de la qualité des apprentissages des étudiants universitaires. Une expérience a l'université de Girona (Espagne). Comunicación presentada en: Xè. *Colloque ADMEE Europe. "Dix années de travaux en évaluation. Quelles lignes de Force ? Quelles perspectives ?"*. Université Pierre Mendès. Grenoble (França).
- PÉREZ CABANÍ, M. L. y TERRADELLAS PIFERRER, M. R. (1996b). Enseñar, aprender y evaluar lo aprendido en la Universidad. Comunicación presentada en *Las Jornadas de Formación del Profesorado Universitario y Calidad de la Enseñanza*. Leioa, Universidad del País Vasco, del 21 al 23 de noviembre.
- RESNICK, L. B. (Ed.) (1989) *Knowing Learning, and Instruction. Essays in honor of Robert Glaser*. New Jersey: Lawrence Erlbaum Associates.
- SNOW, R. E. & SWANSON, J. (1992). Instrucciona Psychology: Aptitude, adaptation, and assesment. *Annual Review of Psychology*, 43, 583-626.
- TANG, K. C. C. (1991). *Effects of different assessment methods on tertiary students' approaches to studying*. Unpublished doctoral dissertation, University of Hong Kog.
- TAYLOR, C. (1994). Assessment for measurement or standards: the peril and promise of large scale assessment reform. *American Educational Research Journal*, 31, 231-262.
- TRIGWELL, K. & PROSSER, M. (1990). Using student learning outcome measures in the evaluation of teaching. *Research and Development in Higher Education*, 13, 390-397.
- VAN ROSSUM, E. J. & SCHENK, S. M. (1984). The relationship between learning conception, study strategy and learning outcome. *British Journal of Educational Psychology*, 54, 73-83.

Anexo 1

Breve introducción a la taxonomía SOLO

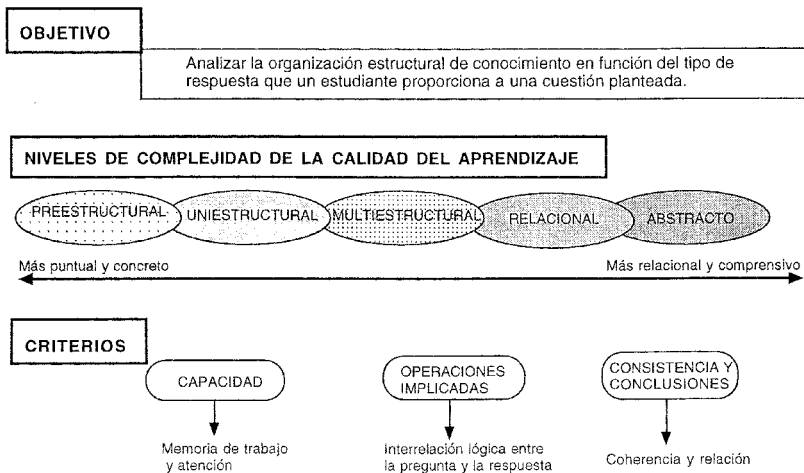
Los estudios que han conducido al desarrollo de la taxonomía SOLO se han realizado con el objetivo de analizar la organización estructural del conocimiento en función del tipo de respuesta que un estudiante proporciona a una cuestión planteada en diferentes situaciones, contenidos, o áreas de conocimiento. En esta organización estructural del conocimiento, los autores distinguen diferentes niveles de complejidad -preestructural, uniestructural, multiestructural, relacional y abstracto-, que permiten analizar la calidad del aprendizaje desde los niveles más concretos hasta los más abstractos y complejos. Los niveles más elevados de la taxonomía corresponden a un aprendizaje más profundo, a una interpretación personal del contenido, que relaciona la tarea con situaciones alejadas del contexto inmediato, que establece relaciones con otros conocimientos relevantes y con materiales procedentes de diferentes fuentes de información. Contrariamente, los niveles inferiores de la taxonomía SOLO corresponden al tratamiento de la información de manera aislada y reproductiva (Van Rossum y Schenk, 1984; Entwistle, 1987; Biggs, 1988; Marton, 1988; Trigwell y Prosser, 1990).

La taxonomía SOLO no mide las características psicológicas de un sujeto a nivel general, sino en relación al tipo de respuesta que da a una cuestión que se le plantea, por lo que se entiende que el nivel de una respuesta se puede modificar mediante la instrucción.

Existen estudios que avalan la validez del instrumento. Respecto a esta cuestión, Biggs y Collis han examinado la correlación con otras medidas del resultado del aprendizaje en diferentes áreas curriculares. A partir de los estudios realizados los autores concluyen que la taxonomía SOLO está relacionada con las medidas del rendimiento académico que utilizan los profesores para evaluar a los estudiantes, pero además, consideran que aporta otros datos relevantes del proceso cognitivo que las evaluaciones de los profesores no suelen incluir.

También se ha podido constatar que el resultado de la calidad del aprendizaje valorado a través de la taxonomía SOLO está relacionado con las habilidades y el proceso de estudio. Estos datos según los autores sugieren, en primer lugar, que la taxonomía está directamente relacionada con la calidad del resultado del aprendizaje y, en segundo lugar, que los niveles más elevados de la taxonomía los obtienen los estudiantes que tienen una motivación intrínseca elevada, que pretender dar significado a aquello que estudian y que evitan reproducir literalmente datos y detalles.

FIGURA 1
TAXONOMÍA SOLO (*Structure of the Observed Learning Outcome*) (Biggs y Collis, 1982)



CUADRO 1
Niveles de complejidad de la taxonomía SOLO

Niveles	Capacidad	Operaciones Implicadas	Consistencia y conclusiones
P R E E S T R U C T U R A L	Mínima: Dificultad por recordar la respuesta a la cuestión planteada. Respuestas confusas e indiferenciadas.	No hay interrelaciones lógicas. Tipos de confusiones: negación, tautología y transducción.	No hay conciencia de la necesidad de ser consistente. Se establecen conclusiones sin analizar el problema. Se repite la pregunta o se hace una transducción irrelevante.
U N I E S T R U C T U R A L	Baja: Recuerdo de un dato relevante sobre la pregunta. El estudiante debe comprender la pregunta y relacionarla con la respuesta al menos con una operación lógica.	Se puede "generalizar" pero sólo en relación a un aspecto. Inducción: requiere diseñar correctamente una conclusión general de una cuestión particular. Relacionar un aspecto particular o dato con una conclusión.	No hay conciencia de la necesidad de ser consecuente y se establecen conclusiones precipitadamente. Las conclusiones hacen referencia a un aspecto, pero al menos éste es relevante. Todas las respuestas uniestructurales pueden ser igualmente correctas pero son inconsistentes con cualquier otra. Sólo se tiene en cuenta una perspectiva.
M U L T I E S T R U C T U R A L	Media: Recuerdo de dos o más datos relevantes, aunque de forma aislada.	Se puede "generalizar" en relación a algunos aspectos limitados e independientes. Ordena adecuadamente varios datos relevantes pero falla la conexión entre ellos. Generalmente contiene: "y...y...y también...".	Las conclusiones vienen determinadas por la cantidad de aspectos que se perciben, pero si estos aspectos no están interrelacionados y la explicación se centra en un aspecto aislado de un dato, puede resultar inconsistente. Dos respuestas de este nivel pueden utilizar los mismos datos pero extraer diferentes conclusiones.

CUADRO 1 (Cont.)
Niveles de complejidad de la taxonomía SOLO

Niveles	Capacidad	Operaciones Implicadas	Consistencia y conclusiones
R E L A C I O N A L	Alta: Se interrelacionan los datos o conceptos relevantes.	Inducción. Se puede generalizar dentro de un contexto o experiencia dada utilizando los aspectos (contenidos) implicados en la situación. Da un conjunto de conceptos o principios relacionados.	El estudiante extrae conclusiones después de analizar todos los aspectos que intervienen y luego los interrelaciona de forma integrada. No hay inconsistencias dentro del contexto pero puede haber inconsistencias fuera de éste.
A B T R A C T O	Máxima: Datos relevantes interrelacionados. El estudiante necesita no únicamente relacionar la información dada, sino también comprender la relevancia de los principios abstractos primordiales de los cuales puede deducir una hipótesis y aplicarla a una situación que no ha sido dada.	Deducción e inducción. Se puede generalizar a situaciones no experimentales. Hay varias características de este nivel que debemos tener en cuenta: a) La introducción de un principio abstracto que no se ha dado directamente en los datos. b) La deducción desde este principio y la comprobación de esta deducción confrontándola con los datos. c) La introducción de una analogía que sea compatible con estos principios pero no se haya dado con los datos. d) El resultado puede ser indeterminado (los eventos pueden haber sido diferentes en distintas circunstancias).	Explica en profundidad los datos y su incidencia en la situación particular. Las inconsistencias quedan resueltas. No se plantea la necesidad de dar conclusiones cerradas sino que se mantienen abiertas de tal manera que puedan ser diferenciadas de manera lógica en posibles alternativas distintas.

Anexo 2

Relación de preguntas de examen analizadas

Exponemos a continuación las preguntas analizadas de manera que queden reflejadas su tipología y la demanda que formulan. Omitimos fragmentos del contenido para garantizar el anonimato de los profesores que nos han proporcionado el material.

Pregunta 1.

Los rasgos de.... son: a).... b).... c).... d)....

Pregunta 2.

Los.... se diferencian de.... en: a).... b).... c).... d)....

Pregunta 3.

En relación al análisis de.... uno de los siguientes no es cierto: a).... b).... c).... d)....

Pregunta 4.

Señala la respuesta que es falsa: a).... b).... c).... d)....

Pregunta 5.

Es cierto que: a).... b).... c).... d)....

Pregunta 6.

¿Qué diseño le parece menos adecuado para.....? a).... b).... c).... d)....

Pregunta 7.

En la tipología de.... la.... se asocia a cualidades como: a).... b).... c).... d)....

Pregunta 8.

Atribuirías la teoría a:

Pregunta 9.

Teniendo en cuenta que normalmente prefieres “x” que “y” de qué manera resolverías “z”:
.... b).... c).... d)....

Pregunta 10.

Escribe la regla de.....

Pregunta 11.

Los sujetos pueden utilizar distintos tipos de reglas para..... a) cierto b) falso

Pregunta 12.

Descripción de un problema: María tiene que.....

Analiza si a María le conviene más a).... b)

¿Qué grado de.... existe entre “a” y “b”?

Calcula la “x” si “z” es igual a 0,3.

Pregunta 13.

Argumenta la relación que establecerás entre estas dos afirmaciones? 1. “....” y 2. “....”.

Pregunta 14.

Indica brevemente posibles.... y sobre todo argumenta el motivo de tu respuesta.

Pregunta 15.

Teniendo en cuenta..... desarrolla una propuesta de actuación. Es muy importante que enmarques tu propuesta en función de.....

Pregunta 16.

Comenta la frase siguiente “.....”.

Pregunta 17.

Explica los criterios sobre los que te fundamentarías para teniendo en cuenta.... Justifica el modelo de... que adoptarías.

Pregunta 18.

Diseña.... y justifica.... teniendo en cuenta....

Pregunta 19.

Explica los tipos de conflictos más frecuentes en.... y posibles estrategias para solucionarlos.

Pregunta 20.

Clasifica.... según.....

Pregunta 21.

¿Qué es la

Pregunta 22.

Describe los tipos de....

Pregunta 23.

¿Qué son las.....?

Pregunta 24.

¿Qué mide la.....?

Pregunta 25.

Diferencias entre “a” y “b”

Pregunta 26.

Relaciona... y argumenta las razones para...

Pregunta 27.

Relaciona cada una de las casillas de la columna izquierda con la casilla de la columna derecha que consideres más oportuna. (columna izquierda: 1.. 2... 3... 4... /columna derecha: a... b... c... d..

Pregunta 28.

¿Qué quiere decir que.....? Analiza el papel de... en el proceso de...

Pregunta 29.

Explica la relación que existe entre "a" y "b". Relaciónalo con el objetivo final de.... que considera...

Pregunta 30.

Analiza los dos diálogos que se exponen a continuación a partir del

Pregunta 31.

En la práctica es imposible que.... La realidad muestra... ¿Cómo crees que se tiene que plantear la intervención en....?

Pregunta 32.

Ejemplifica.... teniendo en cuenta.... Explica en cada caso por qué el ejemplo escogido es representativo de....

Pregunta 33.

Comenta tu opinión respecto a las siguientes afirmaciones "....."

Pregunta 34.

En una clase de..... encontramos las siguientes condiciones a (...), b (...), c (...), d (...). ¿Qué actuaciones propondrías para intervenir en este contexto?

Pregunta 35.

Desde la perspectiva de se subraya que... Haz una propuesta de.....

Pregunta 36.

Comenta esta cita: ".....".

Pregunta 37.

¿Qué queremos decir cuando afirmamos que....? Para responder la pregunta tienes que tener en cuenta los siguientes elementos....

Pregunta 38.

Describe una actividad para realizar en..... Tienes que diferenciar en cada caso.....

Pregunta 39.

Supon que eres... y tienes que solventar la siguiente situación (...). Diseña una posible intervención teniendo en cuenta...

Pregunta 40.

¿Como se podría hacer una intervención en.... Describe brevemente los procedimientos y métodos..

Pregunta 41.

Entre los distintos ámbitos de intervención relacionados con... explica el papel de...

Pregunta 42.

Explica los factores que pueden incidir en ...

Pregunta 43.

Si se considera que....., ¿qué quiere decir...?

Pregunta 44.

(Listado de datos). Analiza el sentido de teniendo en cuenta...

Pregunta 45.

(Listado de datos). Analiza la clasificación de...

Pregunta 46.

Explica qué se consigue aumentando la precisión de... y controlando....

Pregunta 47.

La en un diseño... ¿es igual a?

Pregunta 48.

A qué corresponde la siguiente definición ".....".

Pregunta 49.

¿Son significativos los efectos de...? Analiza.... a través del método de...

Pregunta 50.

Comentar situación "....." en términos de....

Pregunta 51.

Comenta ... desde el punto de vista de....

Extended Summary

Assessment in the university environment is a broad concept, which includes the analysis and evaluation of the various bodies and elements that make up this institution with the aim of improving the educational process that takes place there. Within this context of university assessment, one of the variables which is open to assessment is students' learning. Even though this variable has been studied on a number of occasions, it is also true that today differing approaches coexist in the universities and there is little reflection or collective work on the issue. Given this situation and given that the effect of evaluation on the process and outcome of education is decisive and directly or indirectly affects its quality, our research focused on this variable.

The objectives guiding this research are as follows: 1) To investigate the effect of the type of questions and demands formulated in exams on the quality of learning, as demonstrated by students' answers to these questions. 2) To evaluate the adequacy of the SOLO taxonomy as an instrument to assess the quality of university students' learning. 3) To contrast the possible relationship between the SOLO taxonomy assessment criteria and that of university teachers. 4) To analyse the effect of teachers' assessment decisions on the quality of university students' learning.

A group of teachers from the University of Girona assisted us in this study by providing us with the assessment demands which they formulated in their exams and with the replies students wrote to these demands. The teachers taking part were also interviewed on their criteria and intentions when assessing students.

Our main findings are briefly summarized as follows: First, analysis of exam questions and subjects' answers made it clear that the same question can be interpreted from different points of view. As such, the SOLO taxonomy was seen as an effective instrument for objectively assessing questions with demands that may be tackled from different points of view, and may reveal different levels of understanding and profundity in treating a subject or issue.

Second, a question typology is usually associated with certain specific demands. This has a number of implications for the students since the skills and procedures that they have to "bring into play" in order to reply to the demands in the questions are highly diverse, and entail very different study and performance procedures during the exam. This may cause qualitative differences in students' learning outcome. This relation between question typology and the demands it makes from the teacher's point of view has important implications for teaching: A teacher who only makes one kind of demand is fomenting and evaluating certain skills and procedures in student, and ignoring others that could be equally necessary, or more so, for acquiring relevant knowledge on a subject.

Third, some demands allow only very specific aspects of learning be revealed. On the other hand, the use of the SOLO taxonomy has some limitations. It will therefore be necessary to study different alternatives to optimize the use of this tool in university studies, and examine the possibility of looking for or creating complementary or alternative tools.

Fourth, the decisions teachers make when setting an exam respond to various premises. It is possible to observed some insecurity towards the evaluation process in general and, more specifically, when a grade has to be justified.

The results show that assessing the quality of learning is a complex process. Although there is relative theoretical consensus on *what assessment means*, the

data supports teachers difficulty establishing a common generalizable criterion when undertaking a particular assessment activity. When teachers assess the quality of learning, they often focus on specific points that may cloud students' perceptions both with respect to the demand the teacher's intentions. Finally, it should not be forgotten that learning assessment has social and ethical implications that weigh heavily on teaching staff and lead to a constant worry about being objective. This may well explain why exams continue to be an indispensable and decisive instrument to classify learning outcome. These issues will act as guidelines for our ongoing research work.