

Diagnóstico de dificultades de la evaluación del aprendizaje en la universidad: un caso particular en Chile

Gloria Contreras

Doctora en Ciencias
de la Educación Académica,
Pontificia Universidad Católica
de Valparaíso, Valparaíso, Chile.
gcontrer@ucv.cl

Resumen

La evaluación del aprendizaje a nivel universitario presenta características generales asociadas a un concepto de evaluación tradicional, individualista, con énfasis en la memorización de contenidos, cuya principal función es la calificación.

En la Pontificia Universidad Católica de Valparaíso, Chile, el panorama señalado impulsó a realizar un diagnóstico respecto de la forma como los docentes evaluaban el aprendizaje de sus estudiantes en distintas carreras profesionales, mediante la aplicación de cuestionarios a estudiantes y profesores, y el desarrollo de talleres pedagógicos.

Entre las conclusiones más importantes, se puede destacar que los docentes identificaron con claridad aquellas características técnicas de los procedimientos de evaluación que los hacían válidos y confiables. Sobre los criterios de evaluación, una de las principales dificultades que tuvieron fue explicitar la calidad esperada en los trabajos de sus estudiantes. Finalmente, hay que destacar el alto nivel de concordancia entre ellos sobre dos procesos que consideraban muy importantes, pero que no implementaban debido al desconocimiento de estrategias adecuadas: retroalimentación y auto y coevaluación.

Palabras clave

Educación superior, evaluación del aprendizaje, competencias del docente, estrategias educativas, Universidad Católica de Valparaíso-Chile (fuente: Tesoro de la Unesco).

Diagnosing Learning Assessment Difficulties at the University Level: A Specific Case in Chile

Abstract

The general features of learning assessment at the university level are associated with a traditional concept of individualistic assessment focused on the memorization of contents, primarily for the purpose of grading.

The situation at the Pontificia Universidad Católica de Valparaíso (Chile) prompted a diagnosis of how teachers in different faculties assess student learning. This was done with questionnaires for students and teachers alike, and through educational workshops.

Among the most important conclusions, it was found that teachers were able to identify clearly the technical features of the procedures that make assessments valid and reliable. In terms of evaluation criteria, one of the main difficulties they faced was explaining the kind of quality they expected students to show in their work. There was consensus on two processes teachers regard as highly important: feedback and self-evaluation/co-assessment. However, those processes have not been implemented owing to lack of knowledge about the appropriate strategies.

Key words

Higher education, learning assessment, teaching skills, educational strategies, Universidad Católica de Valparaíso-Chile (Source: Unesco Thesaurus).

Diagnóstico de dificuldades na avaliação da aprendizagem na universidade: um caso específico em Chile

Resumo

A avaliação da aprendizagem no nível universitário mostra características gerais associadas a um conceito de avaliação tradicional, individualista, com ênfase em memorização de conteúdos, cuja principal função é de qualificação.

Na Universidade Católica de Valparaíso, no Chile, o quadro apontado levou a realizar um diagnóstico sobre como avaliavam os professores a aprendizagem dos alunos em diferentes carreiras, aplicando questionários para alunos e professores, assim como desenvolvendo oficinas pedagógicas.

Entre as principais conclusões, os professores identificaram claramente as características técnicas dos processos de avaliação que os faziam válidos e confiáveis. Sobre os critérios de avaliação, uma das principais dificuldades que enfrentaram foi explicitar a qualidade esperada nos trabalhos de seus alunos. Finalmente, devemos salientar seu elevado nível de concordância em dois processos muito importantes para eles, mesmo que não os implementavam devido à falta de estratégias adequadas: feedback, auto e co-avaliação.

Palavras-chave

Ensino superior, avaliação da aprendizagem, competência dos professores, estratégias de ensino, Universidade Católica de Valparaíso-Chile (fonte: Tesouro da Unesco).

Introducción

La Pontificia Universidad Católica de Valparaíso (Chile) es una institución privada de educación superior, que cuenta con subsidio del Estado. En la actualidad tiene una matrícula de aproximadamente 12.500 estudiantes, que cursan 63 carreras profesionales, y de cerca de 500 estudiantes que cursan otro tipo de programas, tales como magíster o doctorado. Sus académicos deben realizar docencia, investigación, extensión y administración, aunque la docencia es la función que más se les demanda.

No obstante su importancia, la docencia en nuestra universidad ha sido tradicionalmente subvalorada en relación con la investigación o extensión académica, situación que se repite en diversas instituciones de educación superior en otros países (Canales, 2003). Esta subvaloración se traduce, por ejemplo, en el hecho de que los docentes son profesionales a los cuales no se les exige formación pedagógica previa alguna, como tampoco se les exige como parte de su carrera académica, que se habiliten en materias de docencia universitaria.

Sin embargo, el año 2005 por primera vez la universidad cuenta con un Plan de Desarrollo Estratégico, cuyo primer lineamiento es “Formación universitaria de calidad, con sello valórico distintivo, innovadora y pertinente” y compromete institucionalmente a la universidad en una nueva mirada hacia la formación profesional y, por lo tanto, hacia la docencia (Pontificia Universidad Católica de Valparaíso, 2005).

De esta manera, se concibe ahora la docencia como una actividad profesional que implica no solo un alto grado de conocimiento de la disciplina que se enseña, sino también de fundamentos y estrategias pedagógicas, de forma que dicho saber disciplinar sea transformado en conocimiento enseñable. Por este motivo, cobra importancia que el docente conozca, desarrolle y perfeccione competencias del campo pedagógico para generar las condiciones necesarias para el aprendizaje de los estudiantes, potenciando procesos que les permitan aprender a aprender (Pontificia Universidad Católica de Valparaíso, 2008.)

A la Dirección de Desarrollo Curricular y Formativo, dependiente de la Vicerrectoría de Asuntos Docentes y Estudiantiles de la Universidad, le corresponde en parte velar por la calidad de los procesos formativos que se desarrollan

en las carreras profesionales que imparte la universidad. En este espíritu apoya la gestión de las carreras profesionales con una serie de acciones, entre las cuales se cuenta el perfeccionamiento pedagógico de sus docentes.

Como una forma de proveer una base para futuros perfeccionamientos en docencia universitaria, dicha dirección inició, a fines del año 2008, un diagnóstico acerca de las prácticas en evaluación del aprendizaje por parte de los profesores.

De acuerdo con Stake (2006), cualquier proceso que pretenda develar las fortalezas y debilidades debe considerar la participación activa de los involucrados, en este caso los docentes, y la comprensión de su quehacer. Estos postulados constituyeron los fundamentos para realizar el diagnóstico sobre las prácticas evaluativas en la universidad, que se presenta a continuación.

Evaluación del aprendizaje universitario

Concepto de evaluación

La evaluación del aprendizaje a nivel universitario constituye un aspecto central de la formación profesional, dado que mediante este proceso los docentes deben dar cuenta del desarrollo de determinadas competencias en sus estudiantes. Esto implica para las instituciones de educación superior una gran responsabilidad para con la sociedad.

Se puede señalar también que constituye un motivo de preocupación para distintos actores, como son los estudiantes, los mismos profesores y autoridades, dados los efectos de las decisiones que se toman a través de ella; por ejemplo, la promoción, reprobación e incluso pérdida de una carrera o graduación para un estudiante.

Como señala Broadfoot (2002), la evaluación es el producto de la interacción de las personas, tiempo y espacio, enlazados en una compleja red de comprensiones, motivaciones, ansiedades, expectativas, tradiciones y elecciones. Esta red conforma un marco referencial desde el cual los docentes universitarios orientan sus prácticas de evaluación.

Algunos de los elementos que influyen en este marco referencial son sus valores, sus experiencias como estudiantes, sus percepciones respecto de la disciplina que enseñan, el contexto de la institución en la que trabajan, sus ideas acerca de la enseñanza, aprendizaje y evaluación, entre otros. En la mayoría de los casos, un elemento que falta es la formación pedagógica.

Los elementos presentes en la constitución del marco referencial que orienta las prácticas de evaluación podrían explicar algunas de las características que tiene la evaluación universitaria. Barberá (2003) establece que la evaluación universitaria puede caracterizarse brevemente como un conjunto de tareas puntuales sin conexión entre ellas, que se presenta a los estudiantes en instancias separadas de la enseñanza, constituyendo más bien un resumen de lo enseñado; se trataría de situaciones descontextualizadas y estandarizadas que se aplican al finalizar un período (Herrera-Torres y Lorenzo-Quiles, 2009), que relevan el aprendizaje de contenidos y no de habilidades (Gómez Roldán, 2005), y cuya finalidad es la calificación.

El informe de la OCDE sobre el sistema de educación superior en Chile, en su apartado denominado “Calidad de la docencia, cursos y aprendizaje”, también da cuenta de procesos de enseñanza y evaluación enmarcados en un enfoque técnico y tradicionalista, con similares características a las mencionadas en el párrafo anterior. Se reconoce que si bien se ha intentado mejorar la calidad de la formación, tanto académicos como estudiantes comparten las siguientes percepciones sobre la docencia: los métodos pedagógicos usados son generalmente muy tradicionales, con énfasis en la memorización de contenidos, se fomenta el aprendizaje individual más que colaborativo y se aplican pruebas tradicionales. A esto hay que sumar la limitada participación que se otorga a los estudiantes en aspectos que después los afectarán directamente en su formación (OCDE, 2009). Esta última situación también se encuentra presente en instituciones de educación superior de otros países, tales como México (Covarrubias y Tovar, 2005).

Tras estas características se encuentra un concepto de evaluación asociado fuertemente a la calificación y posterior certificación, que desatiende los procesos de enseñanza y aprendizaje e ignora el potencial que la evaluación tiene para transformar ambos.

La evaluación en una conceptualización actualizada se comprende como un juicio de valor acerca del trabajo académico del estudiante, el que debe expresar una comparación entre su desempeño y uno o más criterios de excelencia, de manera que pueda orientar tanto el trabajo del estudiante como el del docente, evitando así que se convierta en una instancia aislada solo realizada para otorgar calificaciones (Himmel, 2001). En esta conceptualización los criterios de evaluación son un aspecto esencial, pues constituyen las orientaciones o guías desde la cual los profesores juzgarán la calidad de las tareas, procesos o productos de sus estudiantes.

Criterios de evaluación

La palabra criterio proviene del griego *kriterion*, que significa “medio para juzgar” (Sadler, 2005). Según este autor, un criterio es, en términos generales, una propiedad o característica distinguible de alguna cosa, mediante la cual su calidad puede ser juzgada o estimada, o por la cual una decisión o clasificación puede ser hecha. Entonces un criterio de evaluación constituye la orientación o guía desde la cual los profesores juzgarán la calidad de las tareas, procesos o productos de sus estudiantes.

En otras palabras, los criterios describen “qué mirar” en los trabajos de los estudiantes (Arter y McThige, 2001; Himmel, et al., 1999; Geoff, 1999). De esta manera, permiten orientar las explicaciones y decisiones, pues el juicio que se emita sobre el desempeño de los estudiantes se desprenderá de la comparación entre este y los aspectos definidos previamente como criterios.

Los criterios de evaluación pueden ser de distinta naturaleza, obedecer a distintas lógicas y, por lo tanto, generarse de variadas maneras. La forma más adecuada de establecerlos es con base en los aprendizajes que se espera que los estudiantes desarrollen a partir del proceso de enseñanza, los que deben ser comunicados con antelación al desarrollo del proceso de evaluación. En esta modalidad se pueden

establecer criterios de acuerdo con cada tarea o situación de evaluación particular o, en el otro extremo, generar criterios amplios y generales.

En el primer caso, que suele ser el más utilizado, los criterios de evaluación permanecen “anclados” a una tarea o situación de evaluación en particular, quedando prácticamente reducidos a indicadores de evaluación más que a criterios, y corriendo el riesgo de que se establezcan una serie de criterios que escapen al objetivo de aprendizaje.

En el segundo caso, los criterios de evaluación constituyen un ejercicio de abstracción y representación clara de las grandes tareas en la disciplina y sus componentes (Veslin & Veslin, 1992). Esta forma permite diseñar tareas y situaciones de evaluación diversas atendiendo a los mismos criterios; por ejemplo, explicar fenómenos en forma científica o escribir un ensayo persuasivo, bien fundamentado y claro, son criterios más generales que cualquier particular explicación o investigación. De esta manera, los criterios actúan como un puente entre el gran objetivo de aprendizaje y las idiosincrasias de una tarea en particular (Wiggins, 1998), siendo claro para los estudiantes que lo que se enfatiza son las habilidades más que los contenidos específicamente. Así, dos características importantes de los criterios de evaluación serían la capacidad de cubrir lo esencial de los aprendizajes que se desean evaluar y su grado de generalización.

Los actuales planteamientos sobre evaluación universitaria señalan que los estudiantes pueden participar activamente en los procesos de evaluación en la medida en que comprendan los criterios y para ello los docentes deben usar estrategias de comunicación, tales como respuestas modelo e indicaciones claras sobre cómo tener éxito en sus tareas (Angelo & Cross, 1993; Ramsden, 2007; Collazos, Mendoza & Ochoa, 2007).

A este respecto, es importante mencionar que tanto a nivel nacional como internacional es poca la literatura referida a aspectos teóricos y metodológicos de la comunicación de criterios de evaluación a estudiantes de nivel superior. Un autor que ha trabajado este aspecto es Royce Sadler (1998, 2005, 2009), quien propone dos maneras prácticas de comunicar criterios. Por una parte, señala que los estudiantes deben tener acceso a ejemplos concretos de trabajos de buena y de mala calidad, para que de esta

manera identifiquen qué elementos los hacen de buena o mala calidad. Por otra, establece que deben ser sometidos a procesos en que puedan juzgar sus propios trabajos y los de sus compañeros. En ambas estrategias se encuentra presente la idea de la experiencia guiada con los criterios de evaluación, es decir, el involucramiento prolongado en la actividad evaluativa compartida con y bajo el tutelaje de una persona que es un conocedor o experto, en este caso, el profesor. De esta manera los estudiantes son gradualmente expuestos al conjunto completo de criterios y a las reglas para usarlos, y pueden así construir un cuerpo de conocimiento evaluativo y acceder a grados progresivos de autonomía, lo que es vital para futuros aprendizajes.

La capacidad de los profesores de comunicar los criterios y específicamente los grados de calidad y la siguiente comprensión por parte de los estudiantes constituyen aspectos centrales de las buenas prácticas de evaluación (Shepard, 2000). Este punto es difícil para los profesores, pues en general lo que suele orientar la selección, diseño y construcción de sus procedimientos de evaluación son los contenidos, actividades o sencillamente aquello que les ha dado resultado antes y, por lo tanto, aunque es imposible evaluar sin recurrir a criterios de evaluación, estos, la mayoría de las veces, permanecen en el conocimiento tácito de los profesores y no acostumbran a expresarse en términos de criterios de evaluación con sus estudiantes.

Calidad técnica de los procedimientos de evaluación

Un procedimiento de evaluación se refiere a cualquier medio por el cual se recoge la evidencia sobre el aprendizaje de los estudiantes, la cual, una vez contrastada con los criterios y analizada, permitirá emitir un juicio con fundamento y tomar las mejores decisiones acerca del aprendizaje y la enseñanza (Himmel, et al., 1999).

La tarea de diseñar y aplicar un procedimiento de evaluación y analizar la evidencia de

aprendizaje que recoge es muy delicada, si se tienen en cuenta los profundos y duraderos efectos de las decisiones que se toman sobre los estudiantes a partir de este proceso. Para que efectivamente un procedimiento de evaluación sirva de base sólida tanto para mejorar los procesos de enseñanza y aprendizaje, como para certificar adecuadamente el aprendizaje, debe cumplir con las características de dar cuenta en forma válida y fidedigna de los conocimientos y habilidades que están presentes en los aprendizajes que los profesores desean desarrollar en forma progresiva en sus estudiantes. En otras palabras, el procedimiento de evaluación debe “alinearse” con los aprendizajes consignados en los programas, expresados ya sea como objetivos o competencias.

Los conceptos asociados a estas características son conocidos en el ámbito de la medición con los nombres de validez y confiabilidad. Para expresarlo en términos sencillos, a nivel de aula, la validez debe comprenderse como la capacidad de un procedimiento de evaluación de dar cuenta de los conocimientos, habilidades, destrezas y actitudes que realmente se quieren evaluar, y la confiabilidad como la característica que indica en qué medida se puede tener confianza en la precisión de las informaciones que recoge y que se utilizarán para emitir un juicio o tomar una decisión (Himmel, et al., 1999).

A este respecto es importante señalar que el tipo de validez que se reconoce como más importante es el denominado “de constructo”, que alude a la coherencia entre el constructo hipotético que se va a medir y las tareas de evaluación que se presentan a los estudiantes. Otro tipo de validez es el “instruccional”, que se relaciona con la coherencia entre lo enseñado y lo evaluado.

En esta última década ha cobrado fuerza un nuevo tipo de validez que se ha denominado “validez de consecuencias”, que hace relación con la intencionalidad del proceso evaluativo. Esto quiere decir que si deseamos usar la evaluación con un sentido formativo, debemos diseñar y aplicar procedimientos de evaluación, y trabajar con la evidencia proveniente de ellos, de manera coherente con dicho propósito (Stobart, 2006).

Para que las características mencionadas se encuentren presentes, entonces es necesario cuidar una serie de aspectos, tales como la cantidad de situaciones de evalua-

ción que se pueden presentar al estudiante para que, en efecto, evidencie su aprendizaje. Obviamente una muestra mayor de tareas o situaciones arrojará una información más confiable que si se le presenta solo una o dos. Si se amplía esta idea se puede pensar en la confiabilidad de un proceso de evaluación completo y no solo de un procedimiento aislado, ya que una gama amplia de fuentes de información de evaluación puede servir mejor para tomar buenas decisiones que lo arrojado por una sola prueba, por ejemplo.

La claridad de las instrucciones e indicaciones dadas a los estudiantes también es un factor importante, pues ante indicaciones poco claras es muy probable que no expliciten el conocimiento que el profesor desea, no porque no sepan, sino porque no entienden la tarea; por tanto, ¿qué grado de confianza se puede tener en una información obtenida de esta manera? Además, por la misma razón, los estudiantes pueden establecer una relación con el conocimiento que difiere de la que demanda el profesor.

Por último, para cumplir con estas condiciones es absolutamente necesario que el profesor tenga pleno dominio de su disciplina, ya que es la única forma de determinar cuáles conceptos y procesos son más o menos importantes para que reflejen adecuadamente la extensión y profundidad de la disciplina en su enseñanza y evaluación (Taylor & Bobbit, 1996).

Impacto de la evaluación en el aprendizaje de los estudiantes universitarios

El qué y cómo los docentes evalúan comunica a los estudiantes universitarios qué habilidades, actitudes y contenidos disciplinarios y profesionales es necesario relevar y cuáles no. A este respecto, Biggs (2007) señala que lo que aprendan y como aprendan los estudiantes dependerá en gran parte de cómo ellos piensen que serán evaluados. Elton y Johnston (2002) incluso señalan que si se desea cambiar el aprendizaje de los estudiantes universitarios, hay que cambiar la forma de evaluarlos.

Puesto que la evidencia proveniente de investigaciones internacionales sugiere que los estudiantes piensan que la evaluación siempre define el camino por seguir en su formación profesional y no el perfil de egreso o currículo declarado (Biggs, 2007; Brown & Glasner, 2007; Ramsden, 2007; Entwistle, 2000), un sistema universitario cuyos procesos de evaluación no reflejan los aprendizajes o competencias que deben desarrollar sus estudiantes, está promoviendo una formación profesional poco adecuada.

La forma en que la evaluación influye en los enfoques de aprendizaje, entendidos como intenciones que orientan y condicionan la actuación del estudiante durante el proceso de aprendizaje (Ramsden, 2007), ha sido estudiada internacionalmente. En términos generales, las investigaciones concluyen que métodos de evaluación tradicionales, memorísticos y descontextualizados inducen un enfoque superficial del aprendizaje; en cambio, métodos de evaluación alternativos y con mayores exigencias cognitivas inducen un enfoque profundo del aprendizaje (Gulikers, et al., 2006; Barrow, 2006; Struyven, et al., 2005; Thompson & Falchikov, 1998).

Por ejemplo, pruebas de opción múltiple con énfasis en respuestas de memorización impulsan a los estudiantes hacia un enfoque superficial, pues las perciben como de baja exigencia cognitiva, mientras que un ensayo de final abierto tiende a fomentar un enfoque profundo, pues perciben que les exige procesos cognitivos de más alto nivel (Scouller, 1998). Y aunque una vez que se ha promovido un enfoque superficial es difícil inducir un enfoque profundo, el cambio de las pruebas de opción múltiple por las de tipo ensayo ha cambiado la tendencia general de los estudiantes desde un enfoque a otro (Struyven, Dochy & Janssens, 2005).

Teniendo estos elementos presentes, conceptualización de la evaluación, calidad de los procesos de evaluación e impacto en los estudiantes, se procedió a realizar un análisis del actual estado de la evaluación del aprendizaje en la formación de pregrado en la Pontificia Universidad Católica de Valparaíso.

Metodología

El trabajo de diagnóstico se diseñó pensando en otorgar la mayor participación posible a los docentes de la universidad, por lo cual en noviembre del 2008 se en-

vió una convocatoria abierta a quienes quisieran asistir. Se tenía también la intención de que esta participación produjera aprendizajes en el área de la evaluación, de manera que los docentes pudieran implementar algunas innovaciones en el corto plazo, por ejemplo, introduciendo cambios en el siguiente semestre. De esta manera se planificaron tres talleres, cada uno con dos horarios de libre elección.

Los propósitos generales de este diagnóstico, en función de los docentes, fueron detectar fortalezas y debilidades en el área de la evaluación del aprendizaje y establecer las bases para un futuro perfeccionamiento en dicha área.

Se creó para estos talleres un cuestionario sobre necesidades en evaluación del aprendizaje, en el que los docentes debían pronunciarse individualmente ante un listado respecto de, por una parte, aquellas actividades que consideraban más o menos importantes en sus prácticas evaluativas y, por otro, la frecuencia con la que desarrollaban dichas actividades. Para ello contaban con una escala graduada del 1 al 4, en la que el 1 representaba la menor importancia o frecuencia en la práctica, y el 4, la mayor importancia o frecuencia. La discrepancia entre importancia asignada y frecuencia en las prácticas, es decir, su cercanía o lejanía, arrojaría los focos de interés para un futuro perfeccionamiento en el área.

Por otra parte, se contaba con los resultados del Cuestionario de Opinión Estudiantil del año 2008, instrumento que es parte del Sistema de Evaluación del Desempeño Docente de la universidad. Dicho cuestionario está estructurado en cinco dimensiones: compromiso (del estudiante) con la asignatura; enseñanza para el aprendizaje; evaluación para el aprendizaje; ambiente para el aprendizaje y responsabilidad formal. Las respuestas de los estudiantes van de 1 a 4, de las cuales, la opción 1 representa la percepción más deficiente, y la 4, la mejor.

El primer taller tuvo como objetivos “analizar colectivamente instrumentos de evaluación

de los profesores participantes desde la perspectiva de su calidad técnica” y “comprender las características técnicas más importantes de los procedimientos de evaluación”. Para esto se solicitó previamente a los docentes que trajeran una copia de algún procedimiento de evaluación escrito, como por ejemplo una prueba, que hubieran aplicado durante el año, junto al programa del curso correspondiente.

La segunda parte de este primer taller tuvo como objetivos “discutir las necesidades de perfeccionamiento de los/las profesores/as en cuanto a evaluación del aprendizaje a nivel universitario” y “contrastar estas necesidades con lo arrojado por la aplicación del Cuestionario de Opinión Estudiantil en la Dimensión *Evaluación para el aprendizaje*”.

El segundo taller tuvo como objetivos “comprender la importancia de los criterios de evaluación” y “comprender la relación entre objetivos de aprendizaje y situaciones de evaluación”. Para ello se les pidió realizar un trabajo manual siguiendo un número de instrucciones muy específicas. El objetivo de la segunda etapa fue “generar criterios de evaluación coherentes con los objetivos de aprendizaje de la asignatura”.

Para el tercer taller se retomó el objetivo “generar criterios de evaluación en coherencia con los objetivos de aprendizaje consignados en el programa” y se agregó el objetivo “diseñar una situación de evaluación, con su respectivo procedimiento de evaluación, en coherencia con los criterios generados anteriormente”.

En cada taller se comentaron los resultados de la aplicación del Cuestionario de Opinión Estudiantil durante el 2008 en aspectos referidos a la evaluación del aprendizaje, como una forma de ir contrastando la opinión de los estudiantes y las prácticas de sus docentes en esta área.

Aunque muchas de las actividades propuestas fueron individuales, se solicitó a los participantes que discutieran grupalmente una guía de trabajo con una serie de preguntas sobre las dificultades o facilidades que iban encontrando para desarrollar lo solicitado, se les pidió que lo escribieran y finalmente se solicitó su exposición al resto de los grupos. Con la autorización de los participantes, se grabaron sus intervenciones y posteriormente fueron transcritas.

Participaron en este diagnóstico 35 profesores, de las áreas de Agronomía, Comercio, Filosofía, Historia, Ma-

temáticas, Pedagogía, Periodismo, Psicología, Química, Religión, Ingeniería en Alimentos, Ingeniería Bioquímica, Ingeniería Industrial, Ingeniería Informática, Ingeniería Eléctrica e Ingeniería en Transporte.

Análisis

El análisis individual y conjunto, proveniente del Cuestionario sobre Necesidades en Evaluación del Aprendizaje, Cuestionario de Opinión Estudiantil y trabajo con docentes en los talleres, se desarrolló siguiendo un enfoque cualitativo, dadas las características de la información disponible. Se organizó dicha información tanto por contenidos como en función de los objetivos iniciales del diagnóstico; luego se trató de identificar frecuencias, omisiones y declaraciones que se constituyeran en categorías relevantes (LeCompte, 2000). Aquí hay que mencionar que las categorías surgieron principalmente de las temáticas propuestas para los talleres, pero también de algunas recurrencias emergentes en el trabajo con los docentes participantes.

A continuación se presenta el análisis conjunto proveniente de las tres fuentes mencionadas: Cuestionario sobre Necesidades en Evaluación del Aprendizaje, Cuestionario de Opinión Estudiantil y trabajo con docentes en los talleres. Se analizan dos grandes categorías: calidad técnica de los procedimientos de evaluación y criterios de evaluación, las que a su vez se dividen en cuatro y tres subcategorías respectivamente.

Calidad técnica de los procedimientos de evaluación

Coherencia entre aprendizajes por evaluar y aprendizajes por desarrollar presentes en el programa de la asignatura

Un programa de asignatura consiste en una descripción de la estructura del curso considerando tiempos, actividades del docente y de los estudiantes, situaciones de evaluación, bi-

bliografía, entre otros. En particular, los aprendizajes allí expresados, ya sea como objetivos o competencias, indican lo que se espera que los estudiantes aprendan como resultado de la enseñanza y, por lo tanto, bien formulados orientan tanto al docente como al estudiante, pues hacen clara la meta a la que se quiere llegar.

En el primer taller se realizó una actividad sencilla, en la que los docentes participantes debieron analizar los aprendizajes consignados en sus programas de asignaturas a partir de una serie de preguntas. Las conclusiones a esta actividad se pueden sintetizar como sigue: en algunos casos los programas simplemente no contemplaban los aprendizajes que se debían desarrollar en los estudiantes. En otros, su redacción hacía difícil que pudieran orientar la enseñanza, aprendizaje y evaluación, ya sea porque señalaban lo que el docente realizaría (por ejemplo, “Introducir al estudiante al método científico y el diseño experimental”), porque los verbos utilizados eran ambiguos o demasiado amplios (por ejemplo, “Valorar la vida de Jesús”), o porque eran más bien actividades que debían realizar los estudiantes y no aprendizajes (por ejemplo, “Escribir un informe”).

Una vez establecida la necesidad de cambiar la redacción de los aprendizajes en sus programas de asignatura, los docentes iniciaron el trabajo de análisis de sus pruebas escritas. En la primera parte, el propósito era identificar los aprendizajes que se estaban evaluando con dichos instrumentos y establecer su coherencia con lo establecido en los programas, situación que podría dar cuenta en parte de la validez de dichos instrumentos.

Para la prueba correspondiente al área de ciencias no se presentaron mayores dificultades, pues los docentes consideraron que los aprendizajes estaban presentes de manera clara y sin ambigüedades, como por ejemplo, “Relacionar conceptos”, “Plantear hipótesis” y “Plantear objetivos generales y específicos para el estudio de la hipótesis propuesta”.

En el área de Ingeniería, los docentes identificaron aprendizajes más bien procedimentales, sin entrar al contenido específico de la prueba presentada, como, por ejemplo, “Interpretación de información”, “Traducción al lenguaje matemático” y “Capacidad de resolver ecuaciones”.

Quizás la prueba que más dificultades presentó para identificar los aprendizajes fue la del ámbito de Religión, por-

que la mayoría de los docentes no eran del área, por la gran cantidad de aprendizajes que aparentemente se quería evaluar y por la naturaleza de los aprendizajes en dicha área. Ejemplos de aprendizajes identificados en el nivel cognitivo fueron: “Conocer el actuar de Jesús en una situación de angustia”; “Conocer los hechos y dichos fundamentales de la vida de Jesús”; “Vincular las ciencias religiosas con otras ciencias y hechos de la vida diaria” y “Uso o manejo de la Biblia”. Otros aprendizajes identificados más bien a nivel actitudinal fueron: “Ver la dimensión humana y divina de Jesús”; “Valorar la vida de Jesús como hombre y Dios verdadero” y “Establecer contacto personal con Jesús”.

Además de los aprendizajes mencionados, los docentes identificaron algunos más bien procedimentales, que debieran estar ya desarrollados en el estudiante, como requisito previo para presentarse a una prueba escrita de esta naturaleza. Un claro ejemplo de ello es “Comprensión de lectura”, identificado como aprendizaje para evaluar por dos grupos. Aquí es importante destacar que se debe distinguir entre aquellos aprendizajes que se desea evaluar con cualquier procedimiento y aquellos que el estudiante debe necesariamente poner en juego para desempeñarse en forma adecuada, como es comprender lo que lee.

Los profesores identificaron en todas las pruebas aprendizajes tanto memorísticos como de orden superior, pero con una clara mayoría de estos últimos. Los verbos utilizados para iniciar la descripción de estos aprendizajes fueron: clasificar, sintetizar, interpretar, relacionar, describir, comparar, concluir, analizar, reproducir.

La variedad de aprendizajes identificados por los docentes en una misma prueba mueve a una reflexión: ¿están refiriéndose todos a lo mismo, por ejemplo, con el aprendizaje “analizar”? Con las actividades posteriores quedó establecido que había una variedad de interpretaciones al mencionar aprendizajes para desarrollar o evaluar, con lo que se hacía necesario en el futuro seleccionar o construir una clasificación de aprendizajes.

Los docentes participantes indicaron que idealmente debiera evaluarse la capacidad de aplicar lo aprendido a “la práctica” o a “situaciones prácticas”. Para que la tarea sea de aplicación debe estar diseñada de tal manera que el estudiante use sus conocimientos en una situación que es nueva para él. Puesto que no hubo profundización respecto del término “a la práctica”, suponemos que los profesores se refieren a situaciones cotidianas o relacionadas con el futuro ámbito laboral de los estudiantes. Este último aspecto se relaciona con la denominada “Evaluación auténtica”, cuya principal característica es que examina el desempeño de los estudiantes en forma directa y por lo tanto posibilita realizar inferencias válidas acerca de cómo se podrían desenvolver los estudiantes en el mundo adulto o profesional (Wiggins, 1990).

Una vez identificados estos aprendizajes, los docentes participantes debían establecer su grado de coherencia con los aprendizajes consignados en los programas, pero tuvieron algunas dificultades debido a que, como se señaló, los aprendizajes en los programas eran demasiado amplios, otros muy ambiguos, y en otros casos sencillamente no estaban presentes.

Sin embargo, cuando los aprendizajes estaban bien definidos en los programas de asignaturas y en las pruebas, los participantes no tuvieron problemas en establecer su grado de correspondencia.

La coherencia entre los aprendizajes que se evalúan con los procedimientos de evaluación y los aprendizajes consignados en los programas de asignaturas fue el aspecto más mencionado por los participantes como una característica indispensable en un procedimiento de evaluación, aunque con matices. En algunos casos se mencionó que cualquier procedimiento de evaluación debía ser construido basados en objetivos de aprendizaje. Otros docentes complementaron esta idea señalando que estos objetivos debían ser los desarrollados por los estudiantes en el contexto del curso dictado, lo cual es una consideración importante, pues no todos los aprendizajes planteados al inicio del curso se logran desarrollar por diversas razones. Quizás ese es el sentido de las siguientes respuestas: “Por lo menos parte de ella (la prueba) debe corresponder a objetivos/resultados o aprendizajes esperados” y “Debe haber correspondencia con los contenidos consensuados previamente con los alumnos”.

Señalaron también que para que esto pueda realizarse de manera adecuada, es necesario que los objetivos de aprendizaje estén claramente establecidos en los programas de asignatura y que sus implicancias sean comprendidas por los/as estudiantes.

Dado que las pruebas presentadas no claraban los aprendizajes que se debían evaluar y dadas las características de los programas de asignaturas, no se pueden obtener muchas conclusiones respecto a este punto, pero sí se puede afirmar que los docentes participantes son conscientes de la importancia de este tipo de validez.

Sobre la validez instruccional, hay que destacar que la segunda pregunta del Cuestionario de Opinión Estudiantil, que es: “¿Los procedimientos de evaluación utilizados por el/la docente son coherentes con los contenidos tratados y las actividades desarrolladas durante el curso?”, obtuvo el promedio más alto en la dimensión, que fue de 3,4, y comentarios tales como “Es totalmente coherente en la manera de desarrollar los contenidos, evaluarlos y explicar cómo fueron evaluados”. Sin embargo, algunos docentes no cumplirían con esta condición, en especial en lo referente al nivel de exigencia cognitiva de sus pruebas.

Respecto de aspectos formales de las pruebas

Como se señaló, la calidad de las indicaciones e instrucciones puede afectar la validez y confiabilidad de cualquier procedimiento de evaluación. En el caso de las pruebas analizadas, se observó claridad en la redacción de las preguntas, pero no así en las instrucciones de tipo formal referidas al tiempo en que se debían desarrollar, formas de responder, materiales posibles de utilizar, entre otros.

Por ejemplo, en un caso se consideró que el espacio disponible para presentar la prueba era excesivamente reducido, pues se limitaba a media página, que era ocupada por las preguntas y sin dejar campo para las respuestas de los estudiantes. Esta situación, sumada a la ausencia de instrucciones, puede inducir a los estudiantes a

diversas interpretaciones respecto de la estructura de sus respuestas. Por ejemplo, podrían responder en la extensión que les parezca. Tal como señaló un profesor: “Alguien puede responder en una línea y otro en media página. ¿Cuál es el nivel de extensión apropiado al objetivo del curso?”.

Los docentes participantes señalaron que la formalidad de un procedimiento de evaluación era una característica importante, ya que podría influir en la calidad de las respuestas de los estudiantes, y que esta formalidad debiera estar dada por una estructura que facilitara su comprensión y que además los orientara. Para ello debiera incluir, en orden, un título, el nombre de la signatura, la identificación de la unidad académica, el nombre del estudiante; luego, las instrucciones respecto del tiempo disponible para contestar, materiales que se pueden usar y puntaje o ponderación de las distintas tareas y un enunciado explicativo de lo que la prueba pretende medir.

Por otra parte, la buena redacción de las preguntas o tareas, letra clara, interlineado adecuado y espacios para respuestas, si corresponde; también fueron considerados aspectos importantes para tener en cuenta.

Señalaron, además, que para que esto pueda cumplirse sería necesario tener un diseño previo del instrumento, tratar de responderlo y de esta manera mejorarlo antes de su aplicación definitiva; por ejemplo, adecuándola al tiempo real de los estudiantes, opinión con la que coincide Ramsden (2007), quien plantea que un docente no debiera aplicar una prueba cuyas preguntas no pudiera responder él mismo. Los profesores especificaron que idealmente el instrumento debiera ser validado por un colega o un ayudante.

No obstante las observaciones que estos docentes realizaron a sus propias pruebas, ante la pregunta del Cuestionario de Opinión Estudiantil: “¿Las instrucciones e indicaciones para desarrollar los procedimientos de evaluación aplicados por el/la docente son claras y precisas?”, los estudiantes otorgaron un promedio general de 3,3 y buenos comentarios. Los estudiantes que opinaron que las instrucciones no se otorgan o que no son claras, constituyen solo el 12% de los encuestados. Al respecto se encontraron comentarios como los siguientes: “Las pruebas que se realizan son poco claras, a menudo no se entiende lo que se pregunta”.

Sobre la asignación de puntajes o ponderaciones de las preguntas, tareas o actividades presentadas a los estudiantes

Este aspecto solo pudo ser establecido al consultar con los profesores que habían confeccionado las pruebas, ya que ninguna de ellas especificaba puntajes. Todos señalaron que cada pregunta valía lo mismo.

Este hecho fue considerado poco adecuado por los participantes del taller, ya que cada pregunta en cada prueba, como se había analizado anteriormente, respondía a distintos niveles de exigencia, apuntaba a distintos objetivos del programa y, por lo tanto, esta diversidad debiera verse reflejada en los puntajes. Por esta razón, un profesor señaló: “La ponderación es subjetiva”.

Si todas las preguntas tienen igual ponderación, el mensaje para los estudiantes es que todo es igualmente relevante como aprendizaje. Además, si hay un objetivo de aprendizaje más importante que otro en el programa de la asignatura, esto debiera reflejarse en la forma de ponderar. Una situación como esta afecta la interpretación de las calificaciones, pues estas no podrán ser comparables. Por ejemplo, el 4,0 de un estudiante no reflejará el mismo nivel de aprendizaje que el 4,0 de otro. Sadler (2009) señala que la comparabilidad en los grados o calificaciones es muy relevante en los actuales contextos sociales, pues es un aspecto que se relaciona con la justicia y la equidad, con la posibilidad de realizar predicciones válidas y, a futuro, con la igualdad de oportunidades laborales.

En esta etapa los participantes expresaron que los procedimientos de evaluación debían ser objetivos, en el sentido de ser equitativos o de no contener sesgos para los estudiantes, asegurando de esta manera un trato justo para todos. Además, que debían estar contextualizados, por una parte, a las características propias de cada carrera y, por otra, al nivel de la carrera que se encuentran cursando los estudiantes.

Otros aspectos

Aunque la pregunta estaba acotada a los procedimientos de evaluación, algunos docentes aportaron otros elementos, que se relacionan con la situación de evaluación. Uno de ellos es la adecuada comprensión por parte del alumnado acerca de cómo será evaluado, antes de que esto ocurra. Se incluyen aquí aspectos tales como el conocimiento del formato de evaluación, materiales que se van a emplear, modalidad de la evaluación y plazos, entre otros.

Finalmente, se destaca el hecho de que los docentes consideran necesario que sus estudiantes conozcan las pautas de corrección, pues “Así el alumno se retroalimenta conociendo sus errores, y se alcanza mayor transparencia de la nota obtenida”. Este elemento es importante, pues alude directamente a la evaluación como instrumento para ayudar a mejorar.

A este respecto se puede mencionar que en el Cuestionario de Opinión Estudiantil, ante la pregunta: “¿El análisis y los comentarios de los resultados de las evaluaciones me ayudan a ver mis errores y así mejorar mis aprendizajes?”, obtuvo un promedio de 3,0, siendo el más bajo en la dimensión de evaluación, y la mayor desviación estándar (0,57). Los mejores puntajes iban acompañados por comentarios tales como “Cuando nos entrega los resultados, nos hace ver los errores para así mejorar nuestro aprendizaje” y “Siempre aclara cada error que tenemos en una evaluación”. No obstante, al observar las distribuciones de frecuencias, se encontró que un 23% de los encuestados opinaron que estos comentarios no se realizaban o, si se realizaban, no les ayudaban a mejorar. Dado que una de las características de la buena retroalimentación es su pertinencia en el tiempo, quizás una explicación a la situación descrita se encuentre en las respuestas a una pregunta correspondiente a otra dimensión y se relaciona con el tiempo que demoran los docentes en entregar los resultados de las evaluaciones. Este comentario refleja lo señalado: “La entrega de los resultados de las pruebas son muy tardías, llegando a ser una hora antes de la siguiente prueba”.

Esta situación fue reconocida por los profesores participantes, pues en el Cuestionario sobre necesidades en evaluación del aprendizaje consignaron que la retroalimentación era un aspecto muy relevante pero poco

practicado por ellos, debido al desconocimiento de estrategias adecuadas para este fin.

Criterios de evaluación

Como se ha señalado, una de las primeras actividades de los talleres fue el análisis de pruebas propias para concluir respecto de cuáles debían ser las características de un buen procedimiento de evaluación. En esta etapa, aún sin que el equipo docente hubiera tratado el tema, los docentes participantes mencionaron que los criterios de evaluación debían ser coherentes con los objetivos de la evaluación, estar consignados en una pauta de evaluación y además ser claros para todos, es decir, profesor y estudiantes.

Sin embargo, a pesar de esta claridad, puesto que el conjunto de criterios de evaluación debe desprenderse, principalmente, de los aprendizajes que deben desarrollar los estudiantes y estar consignados en el programa de asignatura, una de las posibles dificultades que podrían tener estos profesores para generar y expresar sus criterios lo constituye la ausencia o ambigüedad en la redacción de los aprendizajes en sus programas, situación que fue establecida en la primera parte del taller.

Una característica importante que no fue mencionada por los participantes es el conocimiento por parte de los estudiantes de los criterios previamente a ser sometidos a cualquier situación de evaluación. No obstante esta omisión, en términos globales los profesores de la universidad parecieran cumplir con esta característica, ya que en la primera pregunta de la dimensión de evaluación en el Cuestionario de Opinión Estudiantil: “¿El/la docente comunica claramente los criterios de evaluación y calificación con los que seremos evaluados?”, se obtiene un promedio de 3,2 y algunos comentarios favorables tales como “Para las evaluaciones, el docente nos entrega una pauta que debemos seguir, tiempo para hacer las preguntas en relación con esta y nos ayuda a mejorar; en cada clase dedica un tiempo especial para ver nuestros avances”. A pesar de

este buen promedio, algunos estudiantes seleccionaron las opciones 1 y 2, lo que indica que perciben que sus docentes no comunican los criterios o que lo hacen con poca claridad. Esto los llevaría a tener dificultades para comprender los requerimientos de una situación o tarea de evaluación. Comentarios que dan cuenta de la situación: “No sabemos cómo seremos evaluados” y “Los criterios que ocupa el profesor para la corrección no son conocidos y genera una cierta desconfianza al saber la nota”.

A continuación se presentan las categorías formuladas para concluir sobre los criterios de evaluación.

Identificación de criterios de evaluación

El tema general de los criterios fue abordado principalmente en los talleres 2 y 3, iniciándose la discusión a partir de un trabajo práctico que los participantes desarrollaron en grupos. El trabajo consistió básicamente en la construcción de un aparato, para lo cual siguieron instrucciones de una hoja entregada, que además contenía una foto del aparato. Al finalizar debían señalar por escrito las razones por las cuales daban por terminado el trabajo. La idea central aquí es que sus razones constituyeron finalmente sus criterios de evaluación.

Las razones que los participantes expusieron ante esta pregunta pueden sintetizarse como sigue: **Porque se cumple con el objetivo requerido**. Todos los grupos, sin excepción, identificaron como objetivo principal de la actividad la construcción del aparato tal como se mostraba en la foto, utilizando los materiales entregados. Las razones que expusieron fueron, por ejemplo: “Se mueve de acuerdo con la figura”, “Es idéntico al modelo entregado”, “La construcción realiza los movimientos indicados y cumple con la utilización de los materiales”.

Esto ocurrió porque, aunque no se declaraba como objetivo, las instrucciones, la foto, los materiales entregados, la ausencia de instrucciones orales por parte del equipo docente y la falta de consultas por parte de los participantes, indujeron a pensar de esa manera. Por lo tanto, una vez identificado este objetivo, la tarea de cada grupo fue cumplirlo y a partir de allí generaron sus criterios de evaluación. Este razonamiento lo habían expresado anteriormente los docentes al describir las características de un buen procedimiento de evaluación.

Una vez que los participantes expusieron en grupo sus razones, el equipo docente comunicó en forma escrita sus criterios para evaluar los trabajos, que estaban relacionados con aspectos actitudinales, tales como “perseverancia” y “trabajo colaborativo”. Esta situación generó una discusión debido al alto grado de discrepancia que percibieron los participantes entre sus criterios y los del equipo docente.

Las razones de esta discrepancia radicaban justamente en la identificación del propósito de la tarea, que para ellos fue construir el aparato. Por ejemplo, dos grupos respondieron: “Nosotros nos centramos en llegar al producto final y no al proceso para llegar al resultado” y “Se consideró como criterio lograr los movimientos, o sea, obtener el resultado, y el equipo docente consideró el desarrollo en grupo”.

Además, señalaron que ninguno de los criterios explicitados al final por el equipo docente estaba señalado en las instrucciones escritas que se habían entregado. Sus comentarios fueron: “Los objetivos no estaban explicitados, se esperaba que el criterio de evaluación estuviese junto con las instrucciones. El criterio de evaluación del equipo docente no fue expuesto claramente” (en forma previa) y “El enunciado no es explícito sobre los criterios de evaluación”.

Esta reflexión es muy relevante, pues en el primer taller, al señalar características de los procedimientos de evaluación, los docentes no mencionaron que los estudiantes debían conocer y comprender los criterios en forma previa a cualquier situación de evaluación.

Otro comentario interesante provino de un grupo del área de ciencias: “La perseverancia, como elemento por evaluar, no estaba en el planteamiento, y disciplinariamente, como somos del área de las ciencias, la síntesis o la integración funcional es algo muy importante para nosotros, entonces lo identificamos como criterio”. Siguiendo la misma línea, otro grupo, en el cual había un profesor de Artes, mencionó: “Además

consideramos la creatividad como un aspecto importante de la evaluación y no solo seguir indicaciones”. Lo interesante de estos comentarios radica en la naturaleza y origen de los criterios de evaluación, ya que estos pueden provenir de la experiencia, del contexto, de la cultura u otros, pero en gran parte se van conformando por el área disciplinar que cultiva cada docente. Sadler (1983) denomina criterios constitutivos a este tipo de criterios, es decir, a aquellos cuyas bases se encuentran en un consenso de los miembros de una disciplina sobre categorías de estándares, conceptos y metodologías.

Como síntesis respecto a la identificación de criterios de evaluación se puede establecer que, en ausencia de criterios explicitados previamente por el equipo docente, los participantes identificaron aquellos que se generaban de manera casi natural desde el propósito de la tarea y desde las instrucciones otorgadas por escrito. En menor medida identificaron criterios relacionados con su disciplina.

Por otra parte, todos los grupos coincidieron en que cambiarían sus acciones si es que hubiesen conocido los criterios antes de iniciar el trabajo, aunque hubo matices en el grado de cambio y sus razones. Algunos grupos señalaron que cambiarían, poniendo más énfasis en algunas acciones que permitieran que el equipo docente notara que trabajaban de acuerdo con sus criterios de evaluación. Por ejemplo: “Mostraríamos perseverancia y trabajo colaborativo y haríamos ostentación de ello”, “Habríamos explicitado aquello que se dio de manera espontánea”, “Se habría enfatizado en aquellos aspectos por los cuales se evaluaría el trabajo” y “Algo, no mucho. Hubiésemos preguntado la opinión de cada uno (de los participantes), aunque, igual cuando hicimos el trabajo, conversamos todos”.

Un grupo comentó que: “En general habríamos cambiado las acciones con el fin de alcanzar los objetivos propuestos. No obstante, algunos plantean que existen ventajas al no saberlos, porque permite desarrollar un trabajo más libre o creativo”. A este respecto, otro grupo les dio la razón, comentando que cambiarían “Porque como alumno, el objetivo es ser bien evaluado”, es decir, sus razones habrían sido estratégicas.

Los grupos concordaron en que conocer los criterios previamente les permitiría desarrollar una planificación

adecuada, orientar las acciones y, por lo tanto, aprovechar mejor los recursos disponibles (materiales, humanos, entre otros) y ser más eficientes. Agregaron que ello permitiría vislumbrar mejor el o los objetivos de la tarea. Un grupo agregó que de esta manera se evitarían algunas frustraciones, pues los evaluados sabrían desde un principio los criterios.

Un grupo en particular resaltó que, de conocer los criterios previamente, tendrían la posibilidad de realizar una autoevaluación del desempeño individual de los integrantes, antes de dar por finalizado el trabajo. Esta es una observación muy interesante pues, si se entiende la autoevaluación como un proceso mediante el cual los estudiantes van a emitir juicios fundados sobre sus propios aprendizajes, entonces es necesario que primero conozcan y compartan los criterios de evaluación.

Para finalizar, los docentes expresaron sentirse “sorprendidos”, “confundidos”, “molestos” y “defraudados”, cuando vieron la discrepancia entre los criterios identificados por ellos y los presentados por el equipo docente. Al explayarse en esta respuesta, algunos señalaron haber sentido estar trabajando sin rumbo, aunque igual cumplieron con los criterios.

Respecto de esta etapa del taller, se puede concluir que generó una muy buena dinámica, pues, en un momento determinado, los profesores no sabían a qué atenerse debido a que no comprendían la incoherencia entre los criterios presentes en la tarea y los presentados por el equipo docente. Básicamente el propósito fue poner en discusión la importancia de la comprensión de los criterios por parte de los estudiantes y que, para este fin, deben estar claramente establecidos con ciertas características mínimas. Además, hizo que los docentes participantes se pusieran en el lugar de sus estudiantes y comprendieran sus temores y sentimientos.

Generación de criterios de evaluación

Para que los docentes participantes pudieran generar criterios con las características requeridas,

en una primera etapa se les entregó y comentó un material con definiciones y características de los criterios de evaluación, tales como la cobertura, generalidad, claridad y nivel de calidad. Posteriormente se les presentaron tres formas distintas de redactar criterios, en las que las características iban apareciendo de manera progresiva. También se realizó una presentación con diversas formas de conceptualizar los criterios de evaluación, redacciones, formas de comunicación, errores más comunes, entre otros.

La conclusión a esta etapa del taller fue que la redacción debiera realizarse con palabras claras, sin ambigüedades, tanto para el docente como para el estudiante; debieran distinguirse niveles de desempeño en la calidad del trabajo de los estudiantes y los conceptos usados debieran ser comprendidos previamente por los estudiantes.

Es importante destacar que en el Cuestionario sobre Necesidades en Evaluación del Aprendizaje, aplicado en la primera sesión con los profesores participantes, la generación de criterios de evaluación coherentes con los aprendizajes consignados en los programas ocupó el segundo lugar de importancia.

Dificultades para generar criterios de evaluación

En el trabajo de taller, los profesores participantes debieron generar criterios de evaluación y describirlos. Para ello se les hizo entrega de una pauta escrita y se les pidió específicamente que seleccionaran uno o más objetivos de aprendizajes presentes en sus programas de asignatura y desde ahí generaran criterios de evaluación, poniendo especial atención en dejar claramente establecido el nivel de calidad con que esperaban se manifestara cada aspecto o característica contenida en los criterios por parte de sus estudiantes.

Una vez que cada grupo hubo terminado su trabajo, se le solicitó que lo intercambiaran con otros, de manera que a partir de una pauta, pudieran evaluar entre ellos sus propios trabajos. Finalmente debieron contestar varias preguntas referidas a las dificultades encontradas en las distintas etapas y a lo aprendido en la actividad.

La tarea solicitada presentó algunas dificultades provenientes desde distintas fuentes. La primera provino de la redacción de los objetivos de aprendizaje en los programas de asignatura, o sencillamente la inexistencia de objetivos. Por ejemplo, los docentes comentaron que “Los objetivos

del curso están planteados de manera muy amplia, lo que dificulta la realización de la actividad”.

Un ejemplo de objetivo que podría presentar esta dificultad es “Conocer la política de la Corona sobre la administración de los territorios ultramarinos y el espíritu que la sostuvo y los cambios hacia fines del siglo XVIII” (Programa de una asignatura de Licenciatura en Historia), ya que el verbo “conocer” es excesivamente amplio y da lugar para diversas interpretaciones.

Otra dificultad, para algunos, provino de la naturaleza de la tarea encomendada, pues lo que se solicitó a los docentes fue que establecieran criterios a partir de los objetivos de aprendizaje, lo que implicaba trabajar con criterios amplios y abstractos referidos a habilidades, y no con criterios específicos referidos a contenidos o tareas particulares, lo que hubiera sido más cercano a su experiencia de generar criterios de evaluación a partir de actividades de evaluación.

Pero la mayor dificultad que reconocieron los profesores estuvo en la descripción del componente de calidad de cada criterio de evaluación. Esto se debe principalmente a que describir la calidad implica señalar con claridad el nivel con que el profesor espera que los estudiantes le presenten los distintos aspectos que componen una tarea. Algunas conclusiones de los profesores fueron: “La redacción de la componente de calidad fue difícil”; “Cuesta ser preciso en describir la calidad de la conducta”; “Definir el límite del aprendizaje esperado en los criterios de calidad”; “Sí, fue difícil en términos de criterios de calidad”.

Una fuente de dificultades en este ámbito estuvo en la tendencia de algunos profesores a describir la calidad con sinónimos, situación que un grupo planteó con la siguiente afirmación: “La dificultad en expresar de forma explícita los términos *adecuado* o *bien*”. Un ejemplo de esto es el siguiente: se solicita como criterio a los estudiantes la “Expresión de ideas propias”, y en la descripción de la calidad se señala: “Manifiesta su propio punto de vista a partir de los conceptos revisados”.

Un aspecto positivo de esta actividad estuvo en el trabajo de coevaluación entre los mismos docentes participantes. Por ejemplo, ante este tipo de redacción, un grupo que evaluó el trabajo de otro le manifestó: “El lenguaje es comprensible pero falta precisión que evite ambigüedades” y “Los criterios de calidad deben hacerse aún más explícitos, de manera que se pueda establecer cuáles conductas son las deseables, por ejemplo, el manejo de conflictos”.

Otra dificultad para la redacción de la calidad en los criterios de evaluación, aunque solo en dos casos, estuvo en que no se describió la calidad sino el conjunto de pasos secuenciado que tenía que seguir un estudiante para finalizar con éxito la etapa establecida en el criterio. Por ejemplo: a) “Lee comprensivamente el problema planteado”, b) “Identifica elementos que le permiten diseñar el proceso de resolución algorítmica”, c) “Desarrolla el algoritmo, entregando una respuesta al elemento planteado”. Junto con ello, aquí se vislumbra otro problema y es que el profesor ha instalado un criterio, del cual luego no podrá dar cuenta por tratarse de una prueba escrita: “Lee comprensivamente el problema planteado”, aunque la resolución del problema supone que el estudiante ha leído comprensivamente.

No obstante los casos mencionados, se encontraron ejemplos en que la calidad estaba bien descrita. Por ejemplo, ante el aspecto “Manejo de conceptos”, un grupo describió la calidad como: “Les entrega el significado correcto a los conceptos; los aplica adecuadamente en contextos específicos y da ejemplos fundamentados del uso de los conceptos en la vida cotidiana”.

La superposición de criterios también fue mencionada como una dificultad. Esta superposición podía estar dada por la repetición de un criterio, a veces con distintos términos, como también por la inclusión de un criterio en otro más amplio. Si esto ocurre, entonces, al aplicar el criterio a alguna respuesta de los estudiantes, se los estaría premiando o castigando dos o más veces. Estas fueron algunas respuestas de los docentes participantes: “En algunos casos hubo superposición de ellos” y “La principal fue no superponer los criterios”. Un ejemplo lo constituye el siguiente par de criterios que fue puesto por separado por una profesora: “Selecciona una o más técnicas metalúrgicas para analizar un problema” y “Descarta la técnica metalúrgica que no le aportará a su objetivo”. El grupo que

evaluó su trabajo consideró con razón que el segundo criterio estaba contenido en el primero.

Un profesor del área de Artes mencionó *que* “la excesiva definición reduce la libertad y manipula al estudiante”. Esta observación es muy importante, pues, por una parte, se refiere a la fragmentación en pequeños trozos de lo que un profesor espera ver como aprendizaje, pudiendo perderse el foco global, y, por otra, tal como lo mencionó un grupo, se corre el peligro de que el estudiante se restrinja a lo que aparece consignado como criterio, preocupándose solo por cumplir para la calificación.

En términos generales, los docentes participantes señalaron que una de las dificultades provino del hecho de trabajar en grupo con otros profesores que no eran necesariamente de la misma especialidad y que, en algunos casos, no se conocían. Este hecho dificultaba avanzar con fluidez en el trabajo grupal. No obstante, agregaron que esta misma situación muchas veces la vivían los propios estudiantes y que ellos debieran tomar en cuenta las posibles limitaciones que traen estas diferencias.

Finalmente comentaron que a veces el lenguaje utilizado por el equipo docente, proveniente del área de la pedagogía, le era un poco lejano. Por ejemplo, un profesor señaló que tuvo “dificultad para distinguir los conceptos usados (habilidades, criterios, competencias, objetivos)”.

Conclusiones

El trabajo de diagnóstico desarrollado permite establecer algunas conclusiones y trazar líneas para el futuro.

La evaluación del aprendizaje en las carreras profesionales de la Pontificia Universidad Católica de Valparaíso tiene características que la enmarcan más bien en un enfoque tradicional, de corte técnico administrativo. Es decir, se asocia la evaluación a medición y calificación, y no a una reflexión fundamentada sobre la calidad de aquello que aprenden los estudiantes.

La falta de orientaciones institucionales, traducida en ausencia de reglamentos específicos a esta área; la poca claridad o ausencia de aprendizajes para desarrollar en los programas de asignaturas; la ausencia de instrucciones formales en las pruebas escritas; las formas de asignar puntajes poco adecuadas para reflejar el aprendizaje que se evalúa y el uso de resultados, en su mayoría solo para asignar calificaciones y no para mejorar la enseñanza o el aprendizaje, son elementos que avalan esta conclusión.

Sin embargo, hay otros elementos que permiten concluir que existe una buena base para que se puedan cambiar las prácticas evaluativas.

En primer lugar, uno de los mayores logros de este trabajo ha sido su capacidad para promover la reflexión en los docentes acerca de la importancia de la evaluación como medio para mejorar la enseñanza y el aprendizaje. Los participantes tuvieron una actitud abierta al aprendizaje, a la colaboración entre pares y la reflexión. En segundo término, fueron capaces de identificar aquellas características técnicas más importantes de los procedimientos de evaluación, y otras correspondientes a la evaluación más bien como un proceso. Por último, y de acuerdo con la percepción estudiantil, los docentes tienen un amplio dominio de la disciplina que enseñan, base indispensable para llevar a cabo cualquier proceso de evaluación.

No obstante ser estas conclusiones emanadas del trabajo con un grupo de profesores en particular, la experiencia ganada en la Dirección de Desarrollo Curricular y Formativo permite establecer que el panorama a nivel de universidad es bastante similar en cuanto a fortalezas y debilidades. Por esta razón los siguientes lineamientos pueden ser acogidos por la institución completa respecto de la docencia en carreras profesionales.

Por una parte, es necesario remover este enfoque tradicional, más bien técnico administrativo, lo que implica estudio y reflexión sobre las propias concepciones y prácticas. En forma paralela debe diseñarse e implementarse un nuevo enfoque, cuidando las condiciones de trabajo de los docentes, ya que actualmente deben

realizar docencia, investigación, extensión y, en algunos casos, atender tareas administrativas.

En lo específico, la implementación de un nuevo enfoque debiera considerar la generación y socialización de alguna clasificación de aprendizajes a nivel de educación superior, que permita que tanto docentes como estudiantes orienten su quehacer en el aula con una comprensión conjunta de lo que se debe desarrollar como enseñanza y aprendizaje respectivamente. Ello permitiría, en parte, mayor claridad para los docentes acerca de cómo generar y comunicar a los estudiantes la calidad de lo que esperan, es decir, debiera facilitar las estrategias de redacción y comunicación de criterios de evaluación. Además les ayudaría a poner el acento en la calidad del aprendizaje y no en la cantidad.

En un nivel general, este trabajo requeriría de un equipo académico bien consolidado, disponibilidad de tiempo de los profesores, apoyo a cada uno de ellos en sus respectivas escuelas y facultades, y el compromiso institucional de manera que sea toda la universidad la que se involucre y no solo algunos docentes.

Si se puede perfeccionar en forma sistemática a los profesores sobre los aspectos mencionados y se establecen políticas adecuadas en la universidad, por ejemplo, traducidas en normas y procedimientos generales, es posible contar con un sistema o modelo de evaluación del aprendizaje que efectivamente contribuya a mejorar los procesos formativos y permita, a su vez, certificar adecuadamente el aprendizaje de los estudiantes.

Agradecimientos

Se agradece muy especialmente la colaboración de las siguientes profesionales de apoyo de la Unidad de Evaluación Curricular: Priscila Cárdenas, Angélica Pizarro y Carmen Gloria Zúñiga.

Bibliografía

- Angelo, Tomas & Cross, Patricia (1993). *Classroom Assessment Techniques. A Handbook for College Teachers*. 1ª ed. San Francisco: Jossey-Bass.
- Arter, Judith & McHige, Jay (2001). *Scoring Rubrics in the Classroom. Using Performance Criteria for Assessing and Improving Student Performance*. 1ª ed. California: Corwin Press.
- Barberá, Elena (2003). Estado y tendencias de la evaluación en educación superior. *Revista de la Red Estatal de Docencia Universitaria*, vol. 3 (2), 5-18.
- Barrow, Mark (2006). Assessment and student transformation: linking character and intellect. *Studies in Higher Education*, vol. 31 (3), 357-372.
- Biggs, John (2007). *Teaching for Quality Learning at University*. 2ª ed. New York: Open University Press.
- Broadfoot, Patricia (2002). Assessment for lifelong learning: Challenges and choices. *Assessment in Education*, vol. 9 (2), 5-7.
- Brown, Sally & Glasner, Ángela (2007). *Evaluar en la universidad. Problemas y nuevos enfoques*. 1ª ed. Madrid: Narcea.
- Canales, Alejandro (2003). Docencia: Tensiones estructurales en la valoración de la actividad. En Mario Rueda, Frida Díaz-Barriga, Mónica Díaz (Eds.). *Evaluar para comprender y mejorar la docencia en la educación superior* (pp. 71-76). México: Universidad Autónoma Metropolitana - Universidad Nacional Autónoma de México - Universidad Autónoma de Baja California.
- Collazos, César Alberto, Mendoza, Jair & Ochoa, Sergio F. (2007). Mejorar los esquemas de evaluación mediante procesos de colaboración. *Educación y Educadores*, vol. 10 (1), 79-88 [citado junio 25, 2010]. Disponible en: <http://educacionyeducadores.unisabana.edu.co/index.php/eye/issue/view/126/showToc>.
- Covarrubias, Patricia & Tovar, Luisa (2005). Las representaciones de los estudiantes sobre la evaluación del aprendizaje. En Miguel Ángel Campos Hernández (Coord.), *Construcción del conocimiento en el proceso educativo* (pp. 265-292). México: CESU-UNAM.
- Elton, Lewis & Johnston, Brenda (2002). *Assessment in Universities: a critical review of research* (LTSN Generic Centre) [citado marzo 17, 2009]. Disponible en: <http://eprints.soton.ac.uk/59244/01/59244.pdf>.
- Entwistle, Noel (2000). Promoting deep learning through teaching and assessment: conceptual frameworks and educational contexts. Paper presented at the TLRP Conference, Leicester [citado abril 30, 2008]. Disponible en: <http://www.etl.tla.ed.ac.uk/docs/entwistle2000.pdf>.
- Geoff, Isaacs (1999). Assessment criteria, standards and marking schemes. Teaching and Educational Development Institute. The University of Queensland, Australia [citado agosto 9, 2007]. Disponible en: <http://www.tedi.uq.edu.au/downloads/Criteria.pdf>.

- Gómez Roldán, Ignacio (2005). Competencias profesionales: una propuesta de evaluación para las facultades de ciencias administrativas. *Educación y Educadores*, vol. 8, 45-66 [citado junio 25, 2010]. Disponible en: <http://educacionyeducadores.unisabana.edu.co/index.php/eye/article/view/1621/2059>.
- Gulikers, Judith, et al. (2006). Relations Between Student Perceptions of Assessment Authenticity, Study Approaches and Learning Outcomes. *Studies in Educational Evaluation*, vol. 32, pp. 381-400.
- Herrera-Torres, Lucía & Lorenzo-Quilmes, Oswaldo (2009). Estrategias de aprendizaje en estudiantes universitarios. Un aporte a la construcción del espacio europeo de educación superior. *Educación & Educadores*, vol. 12 (3), 75-98 [citado junio 26, 2010]. Disponible en: <http://educacionyeducadores.unisabana.edu.co/index.php/eye/article/view/1532/1838>.
- Himmel, Erika, Olivares M., Angélica, & Zabalza, Javier (1999). *Hacia una evaluación educativa. Aprender para evaluar y evaluar para aprender*. Vol. I. 1ª ed. Santiago: PUC-Mineduc.
- Himmel, Erika (2001). La evaluación de los aprendizajes en la educación superior. *Revista Calidad en la Educación*, N° 15 [citado enero 27, 2009]. Disponible en: http://www.cse.cl/public/Secciones/seccionpublicaciones/doc/33/cse_articulo98.pdf.
- LeCompte, Margaret (2000). Analyzing Qualitative Data. *Theory into Practice*, vol. 39 (3), 146-154.
- OCDE (2009). *La Educación Superior en Chile*. París: OCDE-BIRD.
- Pontificia Universidad Católica de Valparaíso (2005). Desarrollo estratégico. Hacia el 2010 [citado mayo 5, 2009]. Disponible en: <http://www.planestrategico.ucv.cl/>.
- Pontificia Universidad Católica de Valparaíso (2008). Visión institucional de la formación de pregrado [citado mayo 5, 2009]. Disponible en: http://ucv.altavoz.net/prontus_unidacad/site/artic/20080110/asocfile/20080110123547/libro_formato_grande_28_.pdf
- Ramsden, Paul (2007). *Learning to Teach in Higher Education*. 2ª ed. New York: Routledge Falmer.
- Sadler, D. Royce (1983). Evaluation and the Improvement of Academic Learning. *The Journal of Higher Education*, vol. 54 (1), 60-79.
- Sadler, D. Roy ce (1989). Formative assessment and the design of instructional systems. *Instructional Science*, vol. 18, 119-144.
- Sadler, D. Royce (2005). Interpretations of criteria-based assessment and grading in higher education. *Assessment & Evaluation in Higher Education*, vol. 30 (2), 175-194.
- Sadler, D. Royce (2009). Grade integrity and the representation of academic achievement. *Studies in Higher Education*, vol. 34 (7), 807-826.

- Scouller, Karen (1998). The Influence of Assessment Method on Students' Learning Approaches: Multiple Choice Question Examination versus Assignment Essay. *Higher Education*, vol. 35 (4), 453-472.
- Shepard, Lorrie (2000). The Role of Assessment in a Learning Culture. *Educational Researcher*, vol. 29 (7), 4-14.
- Stake, Robert (2006). *Evaluación comprensiva y evaluación basada en estándares*. 1ª ed. Barcelona: Graó.
- Stiggins, Rick, et al. (2007). Classroom Assessment for Learning. *Doing it right-Using it well*. 1ª ed. New Jersey: Pearson Education.
- Stobart, Gordon (2006). The Validity of Formative Assessment. En J. Gardner (Ed.). *Assessment and Learning* (pp. 133-146). London: Sage Publications.
- Struyven, Katrien, Dochy, Filip & Janssens, Steven (2005). Students' perceptions about evaluation and assessment in higher education: a review. *Assessment & Evaluation in Higher Education*, vol. 30 (4), 325-341.
- Taylor, Catherine & Bobbit Nolen, Susan (1996). What Does the Psychometrician's Classroom Look Like? Reframing Assessment Concept in the Context of Learning. *Education Policy Analysis Archives*, vol. 4 (17) [citado julio 15, 2008]. Disponible en: <http://epaa.asu.edu/epaa/v4n17.html>.
- Thompson, Karen & Falchikov, Nancy (1998). "Full on until the sun comes out: The effects of assessment on students' approaches to studying". *Assessment and Evaluation in Higher Education*, vol. 23 (4), 379-390.
- Veslin, Jean & Veslin, Odette (1992). *Corriger des copies. Évaluer pour former*. 1ª ed. París: Hachette.
- Wiggins, Grant (1990). The case for authentic assessment. *Practical Assessment, Research & Evaluation*, vol. 2 (2) [citado septiembre 20, 2009]. Disponible en: <http://pareonline.net/getvn.asp?v=2&n=2>.
- Wiggins, Grant (1998). *Educative Assessment. Designing assessment to inform and improve student performance*. 1ª ed. San Francisco: Jossey-Bass, 1998.