



UNIVERSIDAD
COMPLUTENSE
MADRID

Proyecto de Innovación y Mejora de la Calidad Docente

Convocatoria **2014**

Nº de proyecto: **307**

Meta-evaluación del programa Docencia de la UCM

Carlos Gallego López / Nieves Rojo Mora

Facultad de Psicología

1. Objetivos propuestos en la presentación del proyecto

El principal objetivo de una universidad es alcanzar la excelencia y adaptarse a los cambios sociales para contribuir al progreso de la sociedad. Una de los factores fundamentales que determinan la excelencia de cualquier institución universitaria es la calidad docente de sus profesores. Por ello, la universidad debe llevar a cabo una evaluación mediante un proceso continuo que permita conocer su estado y corregir aquellos aspectos que precisen de mejora.

Desde finales del siglo XX, en la universidad Complutense de Madrid (UCM) ha habido varios intentos de llevar a cabo dicha evaluación, intentos que no han perdurado en el tiempo, hasta que, finalmente, en el año 2005 y en el marco del modelo DOCENTIA, propuesto por la Agencia Nacional para la Evaluación de la Calidad y la Acreditación (ANECA), desarrolló e implantó con carácter experimental el Programa DOCENTIA. Este proyecto de evaluación se planteó con un triple objetivo: conocer el grado de la calidad con que se desarrollan las actividades docentes; mejorar la calidad de las actividades docentes y obtener valoraciones individuales de la calidad docente del profesorado que presta sus servicios en la UCM. A partir del curso 2008-09, se hizo un nuevo diseño del programa que se ha aplicado de forma voluntaria hasta el curso 2014-2015. Desde que se empezó a aplicar el programa, el instrumento de evaluación utilizado ha sufrido numerosos cambios hasta llegar al que se está aplicando en la actualidad. Han sido cambios que, sin duda, lo han ido mejorando pero que, en general y según información recibida por los responsables, no se han basado en un estudio riguroso de sus características psicométricas.

En este proyecto nos hemos planteado, precisamente, el análisis de las propiedades psicométricas de los instrumentos del Programa DOCENTIA, a partir de los datos existentes, con un procedimiento doble ciego que garantice el anonimato de centros, departamentos y profesores para detectar sus puntos fuertes y débiles, y elaborar propuestas de mejora para estos últimos.

El objetivo principal del presente proyecto ha sido determinar si el Programa DOCENTIA es adecuado para conseguir los objetivos que se pretenden con su aplicación (conocer la calidad, mejorarla y evaluar de forma fiable, válida y discriminativa al profesorado), determinando sus puntos fuertes y los aspectos susceptibles de mejora y, en el caso de estos últimos, establecer recomendaciones para aumentar su eficacia y utilidad.

Para la consecución del objetivo general se proponen los siguientes objetivos específicos:

1. Estudiar la dimensionalidad de los instrumentos (evidencias tanto de validez de contenido como de estructura interna).
2. Comparar las dimensiones extraídas con las que se proponen desde la literatura como dimensiones básicas de la calidad docente (evidencias de validez de contenido).
3. Examinar el acuerdo con los requisitos y dimensiones establecidos por la ANECA para la acreditación de las titulaciones.
4. Analizar la fiabilidad como consistencia interna de las dimensiones encontradas

y sus bloques.

5. Estudiar la validez referida a criterios externos, con datos objetivos relativos al rendimiento de sus estudiantes y otros datos objetivos relacionados con conductas del profesor.
6. Determinar si las puntuaciones generadas por bloques son útiles para diferenciar entre docentes eficaces e ineficaces.
7. Analizar las preguntas de los diferentes instrumentos y su capacidad discriminativa para diferenciar entre docentes eficaces e ineficaces.
8. Estudiar la capacidad de discriminación de los diferentes procedimientos para la evaluación válida de los profesores.
9. Determinar el grado de acuerdo entre valoraciones de la calidad del profesor con los diferentes sistemas.
10. Estudiar los sistemas de ponderación o asignación de pesos a los diferentes aspectos, determinando si son óptimas para la finalidad pretendida.
11. Evaluar si el sistema proporciona retroalimentación adecuada al profesorado para su mejora, analizando los cambios experimentados en sucesivas evaluaciones.
12. Proponer un sistema adecuado de elaboración de las puntuaciones (global, por dimensiones,...) que proporcione una información más clara al profesorado.

2. Objetivos alcanzados

A continuación, se indica en qué medida se han alcanzado los objetivos propuestos. En el documento: *INFORME META_EVALUACIÓN DOCENTIA UCM*, se explican de forma detallada los análisis realizados para alcanzar los objetivos y los resultados obtenidos.

Objetivo 1. Alcanzado

Estudiar la dimensionalidad de los instrumentos (evidencias tanto de validez como de estructura interna).

Objetivo 2. Alcanzado

Comparar las dimensiones extraídas con las que se proponen desde la literatura como dimensiones básicas de la calidad docente (evidencias de validez de contenido).

Objetivo 3. Alcanzado

Examinar el acuerdo con los requisitos y dimensiones establecidos por la ANECA para la acreditación de las titulaciones.

Objetivo 4. Alcanzado

Analizar la fiabilidad como consistencia interna de las dimensiones encontradas y sus bloques.

Objetivo 5. No alcanzado

Estudiar la validez referida a criterios externos, con datos objetivos relativos al rendimiento de sus estudiantes y otros datos objetivos relacionados con conductas del profesor.

Este objetivo no ha sido alcanzado ya que no ha sido posible obtener información sobre el rendimiento de los alumnos ni sobre otras conductas académicas de los profesores que pudieran servir como criterio externo. El procedimiento que se ha seguido ha sido el de analizar (o reanalizar) los datos de alumnos y profesores obtenidos en años anteriores, 2013-2014 y 2014-2015. Aunque en el caso de los profesores podría ser posible, aunque realmente costoso, recopilar nuevos datos y fundirlos con los ya existentes, en el caso de los alumnos no sería posible. Uno de los procedimientos de recogida de información, el mando, no identifica al alumno, más allá del código que le asigna; no sería, por tanto, posible recoger nueva información en un momento posterior al de la misma contestación de la encuesta.

Objetivo 6. Alcanzado

Determinar si las puntuaciones generadas por bloques son útiles para diferenciar entre docentes eficaces e ineficaces.

Objetivo 7. Alcanzado

Analizar las preguntas de los diferentes instrumentos y su capacidad discriminativa para diferenciar entre docentes eficaces e ineficaces.

Objetivo 8. Alcanzado

Estudiar la capacidad de discriminación de los diferentes procedimientos para la evaluación válida de los profesores.

Objetivo 9. Alcanzado

Determinar el grado de acuerdo entre valoraciones de la calidad del profesor con los diferentes sistemas.

Objetivo 10. Parcialmente alcanzado

Estudiar los sistemas de ponderación o asignación de pesos a los diferentes aspectos, determinando si son óptimas para la finalidad pretendida.

Este objetivo ha sido parcialmente alcanzado ya que, con los datos disponibles, no fue posible determinar los pesos óptimos, dado que no existen variables criterio. No obstante, los análisis psicométricos realizados nos permiten afirmar que los pesos aplicados para calcular los distintos aspectos y las dimensiones de ellos derivadas no son los más adecuados. Se podría utilizar otro sistema distinto de ponderación de los datos de la encuesta utilizando la información de las saturaciones de los análisis llevados a cabo. También, se podrían eliminar de la fórmula para obtener los aspectos de las distintas dimensiones algunas variables de autoinforme, departamento y centro que no discriminan.

Objetivo 11. No alcanzado

Evaluar si el sistema proporciona retroalimentación adecuada al profesorado para su mejora, analizando los cambios experimentados en sucesivas evaluaciones.

Aunque inicialmente nos propusimos analizar y comparar los datos de sucesivas evaluaciones para estudiar su evolución, finalmente, hemos analizado sólo los de los años 2012-2013 y 2013-2014, y nos hemos centrado fundamentalmente en este último. La razón ha sido que el instrumento de evaluación ha ido cambiando de una evaluación a otra, siendo muy grande el cambio sufrido entre el curso 2012-2013 y el curso 2013-2014. Ello ha hecho casi imposible la comparación, inicialmente pensada, de sucesivas evaluaciones.

Objetivo 12. Parcialmente alcanzado

Proponer un sistema adecuado de elaboración de las puntuaciones (global, por dimensiones,...) que proporcione una información más clara al profesorado.

Los resultados ponen de relieve que las dimensiones a priori establecidas para cumplir con las tres propuestas por ANECA no coinciden con las derivadas de los resultados empíricos. Los resultados apuntan hacia dos dimensiones: cumplimiento formal y calidad docente.

3. Metodología empleada en el proyecto

Se han analizado los datos que la UCM ha recogido durante los cursos 2012-2013 y 2013-2014 sobre la evaluación de la actividad docente del profesorado (Programa DOCENTIA) y que nos han facilitado los Servicios Informáticos de la Unidad de Calidad y Datawarehouse del Vicerrectorado de Calidad de la UCM. Estos datos proceden de distintas fuentes de información: de la encuesta contestada por los estudiantes, del autoinforme del profesor, del informe de la institución (centro y departamento) y los procedentes de las bases de datos de la Universidad.

En primer lugar, ha sido necesaria lectura de los datos y conversión de los ficheros proporcionados por el Rectorado al formato en que han sido analizados. A continuación, se ha llevado a cabo la depuración de los datos: corrección de errores, eliminación de información inservible, transformación de variables, etc.

Hemos utilizado un procedimiento que podríamos considerar de doble ciego en el sentido de que los datos relativos a profesores, departamentos y centros han sido codificados por el servicio informático del Vicerrectorado de Calidad de la UCM de forma opaca, de manera que, en ningún momento del análisis, hemos tenido información sobre a quién correspondían los datos que estábamos analizando. Tampoco hemos tenido información sobre qué planes de estudio, áreas, lasignaturas, etc. eran evaluadas en cada momento. En el informe final, por tanto, centros, departamentos, profesores, asignaturas, etc. aparecerán identificados con el código previamente asignado, que deberá ser decodificado por los servicios informáticos.

Siguiendo el procedimiento de evaluación utilizado por la UCM, los datos fueron agregados por grupo evaluador, es decir, no se obtuvo una única evaluación por profesor, sino que cada profesor tenía varias puntuaciones en función del número de grupos distintos que le habían evaluado.

En cuanto a los análisis descriptivos, se han analizado los resultados de la evaluación, tanto la puntuación final como las dimensiones y los aspectos que constituyen cada una de ellas, en función de diferentes variables de interés: centro, área, tipo de asignatura, tipo de profesor, tipo de docencia y sexo. También, se ha correlacionado la puntuación obtenida en determinados ítems con las dimensiones que constituyen el cuestionario.

Para el análisis psicométrico se realizaron análisis factoriales exploratorios y confirmatorios de la encuesta del alumnado y de los aspectos y dimensiones utilizadas para la obtención de la puntuación final. Se calcularon varios coeficientes de fiabilidad como consistencia interna.

Se construyeron tipologías de profesores con los resultados de la encuesta y con los aspectos, mediante análisis de conglomerados bietápico. Ambas tipologías son similares con elevado coeficiente de acuerdo (.80). Con estas tipologías se determinaron los tamaños de efecto de las variables para diferenciar entre profesores de alta y media calidad docente (no existe un grupo de profesores de baja calidad docente estrictamente hablando).

Como resultado de los análisis llevados a cabo, hemos destacado los puntos fuertes y débiles del procedimiento de evaluación utilizado hasta el momento por la UCM (Programa DOCENTIA) y hemos hecho las recomendaciones pertinentes.

4. Recursos humanos (Máximo 1 folio)

En la tabla que se presenta a continuación, se incluyen, por orden alfabético, las personas que han participado en el desarrollo del proyecto y las tareas realizadas por cada una de ellas. Como se observa, unos de los participantes, Elisa Pérez Moreno, no estaba incluido en la memoria inicial. Se solicitó su inclusión al Vicerrectorado de Calidad ya que se consideró importante su participación. En el curso del desarrollo del proyecto se ha producido la reasignación de algunas tareas, reasignación que queda recogida en la tabla.

Participantes	Dirección del proyecto	Preparación de las bases de datos	Revisión teórica y análisis de los modelos de la evaluación de la calidad docente del profesorado (*)	Análisis de las propiedades psicométricas de los instrumentos de evaluación y análisis de las tipologías de profesores	Establecimiento de puntos fuertes, puntos débiles del sistema y recomendaciones	Elaboración del informe final
R. Alcalá Quintana		X				X
M. Antino		X				
A. Calles Doñate			X			X
G. Castaño Collado			X			X
C. Gallego López	X	Solicitante del proyecto				
M.J. Hernández Lloreda				X	X	X
M.V. Hernández Lloreda		X	X			
P. Herreros de Tejada	X	X				
R. Martínez Arias	X			X	X	X
E. Pérez Moreno		X	X			X
T. Ruiz Gallego-Largo	X	X			X	X

(*) Incluye las tareas:

- Estudio de la utilidad de la evaluación de la calidad docente para la mejora continua del profesorado
- Análisis de la utilidad para la institución

5. Desarrollo de las actividades

Las actividades se organizaron a través de tareas que se han ido realizando a lo largo del proyecto:

Preparación de las bases de datos. La tarea de preparación de la base de datos incluye la siguientes actividades:

1. Obtención de los datos. Los datos que nos facilitaron estaban almacenados en formato EXCEL y en diferentes ficheros, uno al menos por cada una de las fuentes de información: alumnos, profesores, departamentos y centros. Cada uno de ellos con diferentes hojas en que aparecían respuestas, preguntas, códigos, etc. Dado que el programa de análisis de datos utilizado en el estudio fue el SPSS, tuvimos que leerlos, reorganizarlos y almacenarlos de nuevo en el formato adecuado.

Inicialmente se nos proporcionaron los datos de las evaluaciones del curso 2011-1012 y 2012-2013. La información relativa al curso 2013-2014 no se nos pudo proporcionar hasta que no estuvo concluida la evaluación correspondiente. Hasta finales del mes de marzo del 2014 no tuvimos todos los datos de este curso.

2. Gestión y transformación de los archivos de SPSS. Dado el sistema de almacenamiento, fueron varias y lentas las tareas que se tuvieron que realizar para preparar el fichero de datos: fusión en un solo archivo de todos los datos relativos a la evaluación de los alumnos y a la información del profesor y de la institución; reestructuración de los datos; agregación de los datos por profesor, transformación de variables; etc.
3. Depuración de la base de datos. Esta actividad, imprescindible para garantizar la calidad y consistencia de los datos, fue realizada en dos fases. En una primera fase, hubo que realizar una depuración a fondo de la base de datos proporcionada por los servicios informáticos ya que, en ocasiones, detectamos algunos errores: codificación errónea de variables, categorización incorrecta, valores fuera de rango, etc. Estos errores se le comunicaron a los servicios informáticos que procedieron a aclararlos y a proporcionarlos la información correcta cuando fue necesario.

En una segunda fase, tuvimos que depurar los errores que pudieron producirse al pasar los datos de formato Excel a formato SPSS y de la manipulación del archivo de datos efectuada en el punto 2.

Revisión teórica y análisis de los modelos de la evaluación de la calidad docente del profesorado. Esta tarea que tenía como propósito fundamental la definición del constructo "calidad docente" y el estudio y comparación de diferentes modelos de evaluación de la calidad del profesorado universitario, incluye las siguientes actividades:

1. Revisión y análisis en la literatura especializada del concepto de calidad docente del profesorado en el ámbito de la institución universitaria.
2. Análisis de los modelos de evaluación de la calidad docente del profesorado implementados en distintas universidades.

3. Análisis en profundidad del programa de apoyo para la evaluación de la actividad docente del profesorado universitario (Programa DOCENTIA) desarrollado por la ANECA y estudio comparativo de las dimensiones del modelo propuesto por la ANECA y del modelo DOCENTIA actualmente implementado en la UCM.

Análisis de las propiedades psicométricas de los instrumentos de evaluación. Se trata del núcleo central del proyecto. Previo al análisis de las propiedades psicométricas, propiamente dicho, se ha llevado a cabo un análisis descriptivo de las variables de interés.

1. Análisis descriptivo de ítems de la encuesta, aspectos, dimensiones y puntuación total: globales, por centros, áreas, tipo de asignatura, categoría, sexo y antigüedad del profesor.
2. Análisis de las propiedades psicométricas de la encuesta y dimensiones del modelo
 - 2.1. Se llevaron a cabo análisis factoriales exploratorios y confirmatorios de los ítems de la encuesta del alumnado, utilizando para cada uno de ellos la mitad de los grupos evaluadores. Los resultados confirman dos factores, muy correlacionados, referidos a cumplimiento formal y calidad docente. Se obtuvieron varios coeficientes de fiabilidad como consistencia interna.
 - 2.2. Prueba del modelo con las dimensiones de ANECA y modelos alternativos. Se realizaron análisis factoriales exploratorios y confirmatorios de los nueve aspectos que constituyen las dimensiones. Se puso a prueba un primer modelo con las tres dimensiones teóricas que proporcionó mal ajuste. Se pusieron a prueba modelos alternativos con dos y una dimensión, siendo la mejor la de dos dimensiones que también podrían identificarse con las obtenidas en el análisis de la encuesta de los estudiantes.
 - 2.3. Se realizaron análisis de conglomerados en un intento de construir una tipología del profesorado basado en las evaluaciones recibidas,

Establecimiento de puntos fuertes, puntos débiles del sistema y recomendaciones. El estudio de los puntos fuerte y débiles, se ha llevado a cabo a lo largo de todo el desarrollo del proyecto. Se detallarán, junto a las recomendaciones, en el informe final que se presenta junto a esta Memoria.

Puntos fuertes

Hacen referencia, sobre todo, a la importante mejora del instrumento de evaluación aplicado hasta el curso 2012-2013 y el aplicado en el curso 2013-14. Se ha simplificado tanto la encuesta a los estudiantes como el autoinforme del profesor y el informe de centro.

Ha habido un incremento de la participación de los profesores en un procedimiento de evaluación que, hasta el momento, es voluntario.

Se observa que el instrumento de evaluación no establece diferencias por sexo y establece diferencias mínimas por antigüedad, por lo que no parece que discrimine a los profesores mayores.

La encuesta del alumnado en su última versión 2013-2014 produce factores similares a los encontrados en la literatura (aunque se podrían eliminar algunos ítems).

Algunos de los ítems a los que se les atribuye menos pesos coinciden con los que después del análisis sugerimos que se pueden eliminar.

Puntos débiles

Han sido agrupados en los siguientes apartados:

- A. Modelo teórico (p.e., los pesos de los ítems y de las dimensiones se supone que son arbitrarios).
- B. Recogida de datos (p.e., no han sido considerados los datos recogidos por una de las dos fuentes de recogida de datos, mandos ó página web).
- C. Base de datos (p.e., no se ha depurado la base de datos).
- D. Instrumento de evaluación
 - D.I. Diseño de la encuesta
 - 1. Encuesta del alumno (p.e., no existe a categoría de NS/NC).
 - 2. Autoinforme de profesor e informe de la institución (p.e., algunos ítems incluidos en los informes correspondientes no tienen relación con la calidad docente del profesor).
 - D.II. Cálculo de la puntuación final (p.e., la inclusión de los ítems en las distintas dimensiones no están basados en un modelo teórico).
- E. Limitaciones del análisis (p.e., sesgos debidos a la voluntariedad de la participación de los profesores).
- F. Otros aspectos (p.e., devolución de la información al profesor poco clara sobre el significado de las dimensiones).

Recomendaciones

Están ligadas a los puntos débiles encontrados y se centran, principalmente, en la necesidad de contar con un modelo teórico claro de la evaluación de la calidad docente del profesorado y en determinados aspectos que pueden mejorar la calidad del instrumento de evaluación. En el informe se detallan estas recomendaciones pero, como ejemplo, señalamos alguna de ellas: debería existir la categoría de respuesta NS/NC; algunos ítems, como asistencia a clase o asistencia a tutorías, deberían funcionar como llaves para ponderar, modular o eliminar la encuesta del alumno; la unidad de análisis debería ser el profesor en lugar del grupo, etc.