

# PROPIEDADES PSICOMÉTRICAS DEL CUESTIONARIO ActEval SOBRE LA ACTIVIDAD EVALUADORA DEL PROFESORADO UNIVERSITARIO

*[Psychometric properties of the ActEval questionnaire on university  
teachers' assessment activity]*

by/por

[Article record](#)

[About authors](#)

[HTML format](#)

**Biencinto, Chantal** ([alameda@edu.ucm.es](mailto:alameda@edu.ucm.es))

**Carpintero, Elvira** ([ecarpintero@edu.ucm.es](mailto:ecarpintero@edu.ucm.es))

**García-García, Mercedes** ([mergaci@edu.ucm.es](mailto:mergaci@edu.ucm.es))

[Ficha del artículo](#)

[Sobre los autores](#)

[Formato HTML](#)

## Abstract

This article presents the results of the empirical validation of the questionnaire ActEval (Self-Report on the Evaluation Activity of university teachers). ActEval attempts to identify how important teachers consider a variety of assessment tasks, whether they feel competent, and to what extent they use them in their daily practice, in line with the new competences under the European Higher Education Area. Validation is studied using Non-Metric Multidimensional Scaling (PROXSCAL) and the reliability results show high levels of internal consistency. The data suggest that some items in the questionnaire should be reviewed in order to obtain an effective tool for the analysis of evaluation practice.

## Keywords

Competence assessment at University; perception of the teacher; construct validity; internal consistency.

## Resumen

El presente artículo tiene como objetivo presentar los resultados de la validación empírica del cuestionario ActEval (*Autoinforme sobre la Actividad Evaluadora del profesorado universitario*), que permite conocer en qué medida el profesorado universitario considera importante, se siente competente y utiliza una serie de tareas de *evaluación orientada al aprendizaje* en su práctica diaria, acordes con las nuevas competencias derivadas del Espacio Europeo de Educación Superior. Los resultados de fiabilidad muestran elevados índices de consistencia interna y la validación empírica mediante el procedimiento de Escalamiento Multidimensional No-Métrico (PROXSCAL) sugiere la revisión de algunos ítems del cuestionario para obtener un instrumento eficaz para el análisis de la práctica evaluadora.

## Descriptores

Evaluación de competencias en la universidad, percepción del profesor, validación de constructo, consistencia interna.

El perfil del profesorado universitario está cambiando y requiere, para asegurar una docencia de calidad y adecuada a la nueva situación que demanda la convergencia del Espacio Europeo de Educación Superior (EEES), un perfil transferencial, flexible y polivalente, capaz de adecuarse a la diversidad y responder ante los continuos cambios (Bozu y Canto, 2009). Las tareas asociadas a este nuevo perfil exigen un rol del profesorado con implicaciones y competencias docentes, algunas de ellas ya presentes y otras que deben ser repensadas (Butcher *et al.*, 2006; Cano, 2005, 2008; Valcárcel, 2003; Zabalza, 2003). Entre estas últimas se encuentra la evaluación, que deja de ser únicamente un proceso de calificación y control para constituirse en un proceso optimizador de los aprendizajes (Allen, 2000; Bain, 2006; Dochy *et al.*, 2000; McDonald *et al.*, 2000; Padilla *et al.*, 2010), de modo que pone su énfasis en la denominada *evaluación orientada al aprendizaje*.

Carless, Joughin y Mok (2006) consideran que para que la *evaluación orientada al aprendizaje* suceda, deben darse tres condiciones esenciales: a) tareas de evaluación auténticas que demanden la aplicación de competencias “reales”, cercanas al ámbito de desarrollo profesional de los estudiantes, b) retroalimentación efectiva, con oportunidades reales de mejora y cambio en el futuro (*feedforward*), lo cual redundará no solo en la propia ejecución de las tareas sino en la satisfacción del alumnado; y c) implicación y participación de los estudiantes en su propio proceso de evaluación (Simon, 2010). Quizás este último sea uno de los cambios más importantes en el proceso de evaluación, valorando sus contribuciones tanto en el diseño de indicadores (compartiendo con los estudiantes sus ideas o, al menos, recogiendo sus criterios) como en el propio proceso evaluativo, tomando parte en el mismo y, por tanto, teniendo como referencia su autoevaluación y la evaluación entre compañeros (coevaluación). En este sentido, distintos estudios recomiendan la participación de los estudiantes en la evaluación (Boud y Falchikov, 2007; Boud, 2010; Chen, 2008; Ljungman y

Silén, 2008; López Pastor, 2008; Rodríguez, Ibarra y Gómez, 2011), destacando las ventajas para el alumnado relacionadas con una mayor implicación, un aprendizaje más autónomo, y un aumento de la confianza y de la responsabilidad (Falchikov, 2005; MacDonald, 2011; Pastor, 2008, 2011).

Aunque cada vez es más frecuente la inclusión de tareas de evaluación formativa en el ámbito universitario, todavía sigue siendo el examen final la técnica primordialmente utilizada en las universidades españolas (Ibarra y Rodríguez, 2010). El EEES supone un cambio de concepción educativa que debe repercutir en la modificación de los procedimientos y la actualización de los instrumentos de evaluación, considerando la evaluación como una actividad de aprendizaje vinculada con la mejora.

Sin embargo, del mismo modo que cualquier estrategia educativa debe ser enseñada y aprendida, los procedimientos de autoevaluación y evaluación compartida o coevaluación también requieren su tiempo y su proceso de aprendizaje, por lo que la práctica de una *evaluación orientada al aprendizaje* no puede esperarse se incorpore de modo inmediato y automático a las prácticas docentes. Este entretencimiento del proceso y las dificultades asociadas son, en ocasiones, esgrimidas por los propios docentes como argumentos para evitar dichas prácticas.

La realidad es que cualquier cambio que se introduzca en el sistema educativo, y más concretamente en la formación superior, requiere la participación efectiva y la disposición del profesorado. Para Prieto y Contreras (2008), son los docentes que entienden la evaluación como un proceso de mejora de su propia práctica, los que buscan la participación de los estudiantes, frente a aquellos docentes preocupados únicamente por la corrección de las respuestas, para quienes la evaluación se convierte en un proceso de control y no tanto de mejora. En este sentido, es preciso conocer, en primer lugar, las creencias que los docentes tie-

nen sobre la evaluación, su función y la intencionalidad del proceso.

Sin duda, estrechamente vinculadas a las creencias sobre evaluación se encuentran las prácticas y actuaciones llevadas a cabo por el profesorado, no solo de evaluación, sino en general de la enseñanza. Citando a Sanmartí (2007): “dime qué y cómo evalúas y te diré qué y cómo enseñas y qué y cómo tus alumnos aprenden” (pág. 19). Por ello, junto a las concepciones sobre evaluación, parece importante conocer las prácticas que se desarrollan en el aula con respecto a la evaluación para poder modificar, constatar o incorporar nuevas prácticas evaluativas.

En consecuencia, parece necesario conocer la percepción del profesorado sobre su propia actividad evaluadora. El autoinforme es un instrumento de reflexión que permite al docente universitario contrastar sus prácticas habituales de evaluación y valorar en qué medida considera importante o relevante la incorporación de procedimientos alternativos vinculados con la *evaluación orientada al aprendizaje* en la dinámica diaria del aula. El objetivo de este artículo es presentar los resultados de la validación empírica de un instrumento, *Autoinforme sobre la Actividad Evaluadora del profesorado universitario* (cuestionario ActEval), diseñado para conocer en qué medida el profesorado universitario considera importante, se siente competente y utiliza una serie de tareas de evaluación en su práctica diaria, acordes con las nuevas competencias derivadas del EEES.

## Método

### Participantes

La totalidad de la muestra que cumplimenta el instrumento se compone de 427 docentes

universitarios de 18 universidades públicas españolas. La distribución en función del sexo es casi homogénea, pues el 49,4% son hombres y el 50,6% mujeres.

Respecto a rama de conocimiento a la que pertenecen los participantes, el mayor porcentaje son del área de *Ciencias Sociales y Jurídicas* (44,3%), seguido de *Ciencias de la Salud* (15,2%), *Arte y Humanidades* (15,2%) y en menor porcentaje *Ciencias* (13,1%) e *Ingeniería y Arquitectura* (12,2%).

En cuanto a la categoría laboral del profesorado, la muestra es equilibrada pues 49,6% del profesorado es personal contratado, mientras que el 50,3% restante funcionario.

Por último, cabe señalar que el 49,4% del profesorado disfruta de más de 15 años de experiencia laboral, y únicamente el 7,7% tiene una experiencia inferior a 3 años, por lo que se trata de una muestra con elevado conocimiento del ámbito universitario.

### Instrumento

La versión actual del *Autoinforme sobre la actividad evaluadora del profesorado universitario* (*Cuestionario ActEval*) ha sido validado teóricamente y por expertos en una fase anterior (Quesada, Rodríguez e Ibarra, 2013).

El cuestionario ActEval es un autoinforme que permite recoger información sobre las prácticas de evaluación de los docentes universitarios, como parte de sus tareas profesionales. Está compuesto por un total de 31 ítems organizados en cuatro dimensiones, correspondientes a actividades fundamentales de evaluación, tal y como se describe en la tabla 1.

Tabla 1. Dimensiones e ítems del cuestionario ActEval

DIMENSIÓN	DESCRIPCIÓN	ÍTEMS
Diseño y planificación de la evaluación	Procedimientos, instrumentos y adaptación de los criterios a las competencias de la materia.	1, 9, 18, 20, 25, 30, 31
Seguimiento de los estudiantes	Transmisión de información a los estudiantes sobre su aprendizaje, al inicio, durante y al final.	2, 3, 19, 21, 22, 23, 27, 28
Participación de los estudiantes en la evaluación	Consenso y negociación de criterios, autoevaluación, evaluación entre iguales, y coevaluación.	4, 10, 11, 12, 13, 14, 15, 16, 17, 26, 29
Seguimiento, mejora y adaptación de la evaluación	Actualizaciones e innovaciones en el proceso, tras su reflexión.	5, 6, 7, 8, 24

Cada uno de los ítems del cuestionario se valora a través de una escala Lickert comprendida entre 1 (nada) y 6 (totalmente) en los siguientes criterios:

- importancia*, definida como el grado de interés y relevancia que tiene para los docentes universitarios la actuación concreta en el proceso de evaluación del aprendizaje de los estudiantes universitarios;
- competencia*, definida como grado en el que se consideran preparados, con destreza, para realizar o llevar a cabo la actuación; y
- utilización*, determinado como el grado en el que se acostumbra a ejecutar o realizar la actuación en el desempeño como docente universitario.

### Procedimiento

Para la validación empírica del cuestionario ActEval, se aplica de forma virtual a través de un formulario construido con la herramienta *SurveyMonkey* y se almacenan los datos en Excel para posteriormente construir la base de datos con SPSS.

### Análisis de datos

El análisis estadístico se realiza con el paquete SPSS (versión 19.0). Para analizar la consistencia interna del instrumento se utiliza el estadístico Alpha de Cronbach y para la validez de constructo se emplea el método de

Escalamiento Multidimensional No-métrico (PROXSCAL). Dadas las particularidades de la escala de los datos el procedimiento llevado a cabo para el cálculo de las disimilaridades se describe con más detalle en el apartado de resultados.

## Resultados

### Consistencia interna

La consistencia interna del instrumento completo analiza los 31 ítems que lo componen, considerando los tres criterios (importancia, competencia y utilización) se analiza a través del estadístico Alpha de Cronbach atendiendo para su interpretación al baremo clásico establecido por Nunnaly (1978). En todos los casos la interpretación indica una fiabilidad excelente, superior a 0,93. Así, para el criterio de importancia la fiabilidad es de 0,936, para el criterio de utilización de 0,935 y para el criterio de competencia, 0,961.

El estudio de fiabilidad por dimensiones del instrumento indica una consistencia interna muy buena en tres de ellas, siendo excelente para la dimensión relativa a la participación de los estudiantes en la evaluación, tal y como se observa en la tabla 2 (para el cálculo de la fiabilidad se toman en conjunto los datos de los tres criterios de análisis).

Tabla 2. Fiabilidad del instrumento por dimensiones

DIMENSIÓN	Alpha de Cronbach	En cada dimensión	Nº de elementos	Interpretación
Planificación y diseño de la evaluación	,890	7	21	Muy buena
Seguimiento de los estudiantes	,896	8	24	Muy buena
Participación de los estudiantes en la evaluación	,936	11	33	Excelente
Seguimiento, mejora y adaptación de la evaluación	,891	5	15	Muy buena

### Validez de constructo

Para analizar la estructura interna del instrumento, hemos optado por el Escalamiento Multidimensional No-Métrico (PROXSCAL) como alternativa al Análisis Factorial Clásico (AFC). La razón principal de esta opción radica en la no adecuación del AFC a los datos recogidos. Recordemos que la escala utilizada es tipo Likert con 6 categorías ordenadas, por tanto se incumplen de partida los supuestos del AFC (López, Pérez y Ramos, 2011; Rivas y Martínez Arias, 1991).

El Escalamiento Multidimensional No-Métrico es el término que se utiliza para describir procedimientos en los que, a partir de las “*distancias*” (disimilaridades) entre un conjunto de puntos, el objetivo es encontrar “*configuraciones*” de los puntos, preferiblemente en un número pequeño de dimensiones. Si denotamos con  $d_{rs}$  a la medida asociada de los individuos  $r$  y  $s$ , se dice que se trata de una disimilaridad si:

- i)  $d_{ij} \geq 0$  para todo  $i, j$
- ii)  $d_{ii} = 0$  para todo  $i$
- iii)  $d_{ij} = d_{ji}$  para todo  $i, j$ .

Por configuraciones nos referimos a un conjunto de coordenadas que se pueden representar en un “*mapa*”. En ese sentido, puede considerarse una generalización de la idea de *componentes principales*. De hecho, debe tenerse cuidado con la notación. Las coordenadas procedentes del escalamiento multidimensional a veces reciben el nombre de *coordenadas principales*.

El procedimiento más utilizado para llevar a cabo un escalamiento multidimensional es probablemente el conocido como ALSCAL, desarrollado en los años 70 por Takane, Young y De Leeuw (1979) e implementado en diferentes programas estadísticos, entre ellos SPSS.

Sin embargo, este método presentaba problemas de *suboptimalidad*, es decir, frecuentemente el algoritmo finalizaba con la detección de un mínimo local, no global, de la función. Investigaciones posteriores llevaron al desarrollo de un procedimiento alternativo: PROXSCAL. Este algoritmo fue presentado por Busing en la conferencia SOFSTAT’97 (Busing, 1997) como una alternativa más eficiente frente al ALSCAL. El método utiliza el algoritmo de De Leeuw y Heiser (1980).

En nuestro caso, la dificultad principal reside en la construcción de la matriz de disimilaridades entre ítems a partir de los datos originales. Se decide seguir el procedimiento reseñado en Cuadras (1996) basado en los resultados obtenidos por Gower (1971) cuando los datos originales no son de escala.

En concreto, sea  $R_{ij}$  la correlación de Spearman entre los ítems  $i$  y  $j$  calculada a partir de los datos recogidos en los 427 sujetos participantes. Es claro que esta correlación puede entenderse como un coeficiente de similaridad entre los ítems, con valores que van desde el -1 (mínima similaridad) al 1 (máxima similaridad). Como es conocido,  $R_{ii} = 1$ . El procedimiento propuesto por Cuadras construye una disimilaridad definida por:

$$d_{ij} = \text{SQRT}(2 * (1 - R_{ij}))$$

Es inmediato comprobar que se trata efectivamente de una disimilaridad (se cumplen las propiedades i, ii y iii mencionadas anteriormente). Este método de construcción de disimilaridades se ha llevado a la práctica en numerosas aplicaciones entre las que cabe destacar Arenas y Cuadras (2002) y Pavoine, Vallet, Dufour, Gachet y Daniel (2009).

Una vez definida la disimilaridad se procede a aplicar el escalamiento multidimensional con el procedimiento PROXSCAL.

Para la presentación de los resultados se analiza la estructura interna del instrumento com-

pleto por criterio de valoración, comenzando con el criterio definido como *importancia*; a continuación, el de *competencia*; y por último, el criterio basado en la *utilización*.

### Escalamiento multidimensional: criterio de IMPORTANCIA

El *stress* y las medidas de ajuste indican la eficacia con la que se aproximan las distancias de la solución a las distancias originales. En los resultados obtenidos tras el análisis encontramos indicadores muy buenos del ajuste de la solución (ver tabla 3).

Tabla 3. Medidas de ajuste y stress del criterio importancia

Stress bruto normalizado(Sbn)	,01140
Stress-I	,10675(a)
Stress-II	,29835(a)
S-Stress	,02671(b)
Dispersión explicada (D.A.F.)	,98860
Coefficiente de congruencia de Tucker (CCT)	,99429

PROXSCAL minimiza el stress bruto normalizado.

a Factor para escalamiento óptimo = 1,012.

b Factor para escalamiento óptimo = ,994.

Cada uno de los cuatro estadísticos de *stress* mide el desajuste de los datos, mientras que la dispersión explicada y el Coeficiente de Congruencia de Tucker miden el ajuste. Las medidas de *stress* cercanas a 0 (Sbn=0,01140) y las medidas de ajuste que se aproximan a 1 (CCT=0,99429), indican las mejores soluciones posibles, por tanto estamos ante soluciones excelentes.

El gráfico 1 muestra las dos primeras coordenadas principales asociadas a los 31 ítems del cuestionario. Para facilitar la interpretación, se han marcado con colores diferentes los ítems pertenecientes a cada uno de los cuatro bloques. Como puede apreciarse en el mapa se agrupan de manera muy clara los ítems correspondientes al bloque tercero (participación de

los estudiantes en la evaluación, en color rojo). Esto se corresponde con la alta consistencia interna que se había encontrado al estudiar este bloque. Frente a esta tendencia general existen dos excepciones: el ítem 15, del bloque tercero, aparece separado, mientras que el ítem 21, del segundo bloque (seguimiento de los estudiantes, en color verde) se posiciona con los del tercero. Los ítems del segundo bloque también aparecen agrupados con la excepción de los ítems 19 y 28. Los otros dos bloques (primero y cuarto) no presentan una separación tan acusada, por lo que es necesario pasar al estudio de las siguientes coordenadas principales, tercera y cuarta. En el gráfico 2 se puede apreciar la concentración de los ítems de estos dos bloques, atendiendo sobre todo a sus valores en el eje 4.

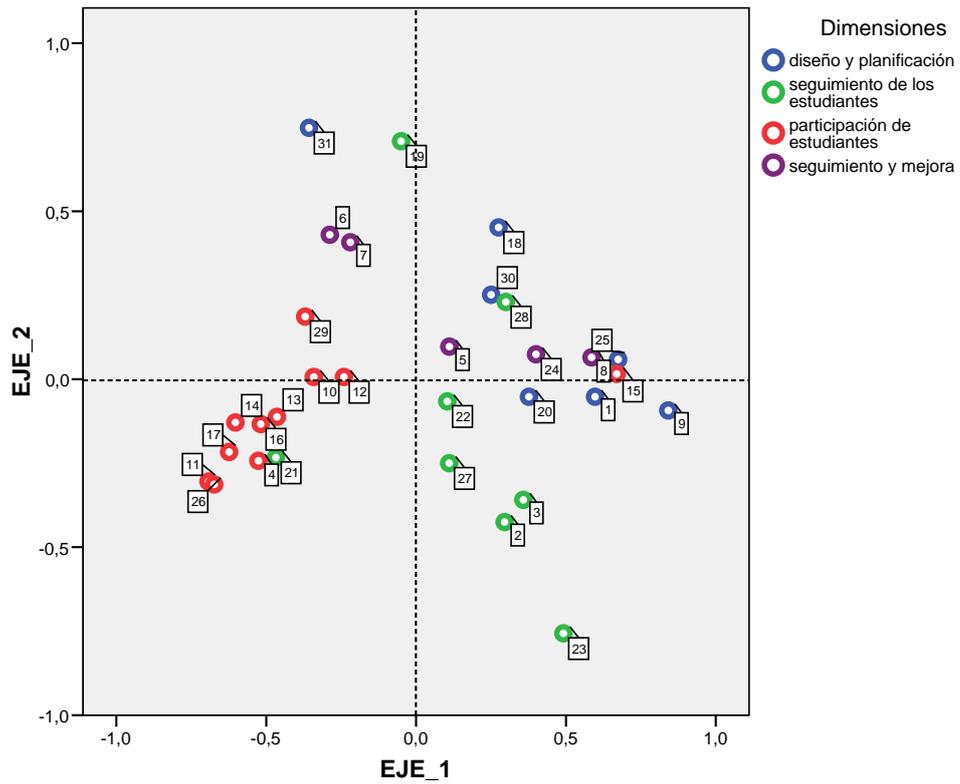


Gráfico 1. Coordenadas 1 y 2

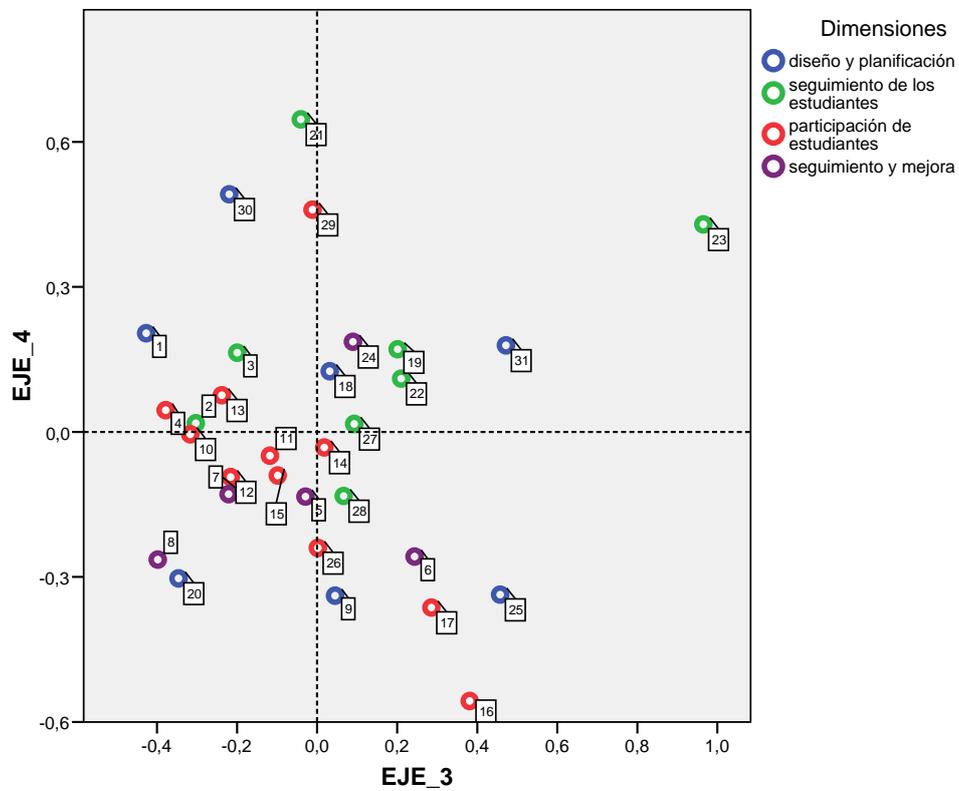


Gráfico 2. Coordenadas 3 y 4

### Escalamiento multidimensional: criterio de COMPETENCIA

Los principales indicadores a interpretar en este segundo criterio nos dan información acerca de la elevada calidad de los resultados.

El índice de *stress* tiene un valor excelente ( $S_{bn}=0,010$ ) y la proporción de varianza de las disparidades iniciales explicada por las distancias entre los ítems asciende a un 99,46% (CCT) (ver tabla 4).

Tabla 4. Medidas de ajuste y stress del criterio competencia

Stress bruto normalizado (Sbn)	,01062
Stress-I	,10307(a)
Stress-II	,30264(a)
S-Stress	,02564(b)
Dispersión explicada (D.A.F.)	,98938
Coefficiente de congruencia de Tucker (CCT)	,99467

PROXSCAL minimiza el stress bruto normalizado.

a Factor para escalamiento óptimo = 1,011.

b Factor para escalamiento óptimo = ,994.

En el Gráfico 3, que representa las dos primeras coordenadas principales, se aprecia, nuevamente, de manera clara la agrupación de todos los ítems del bloque 3 salvo el ítem 15, y todos los del bloque 2 salvo los ítems 19 y 28

(en color verde). En este caso aparece más clara la agrupación de los ítems del resto de grupos salvo, quizás, los ítems 18 y 20, del bloque primero, que tienden a aparecer mezclados con los del bloque cuarto.

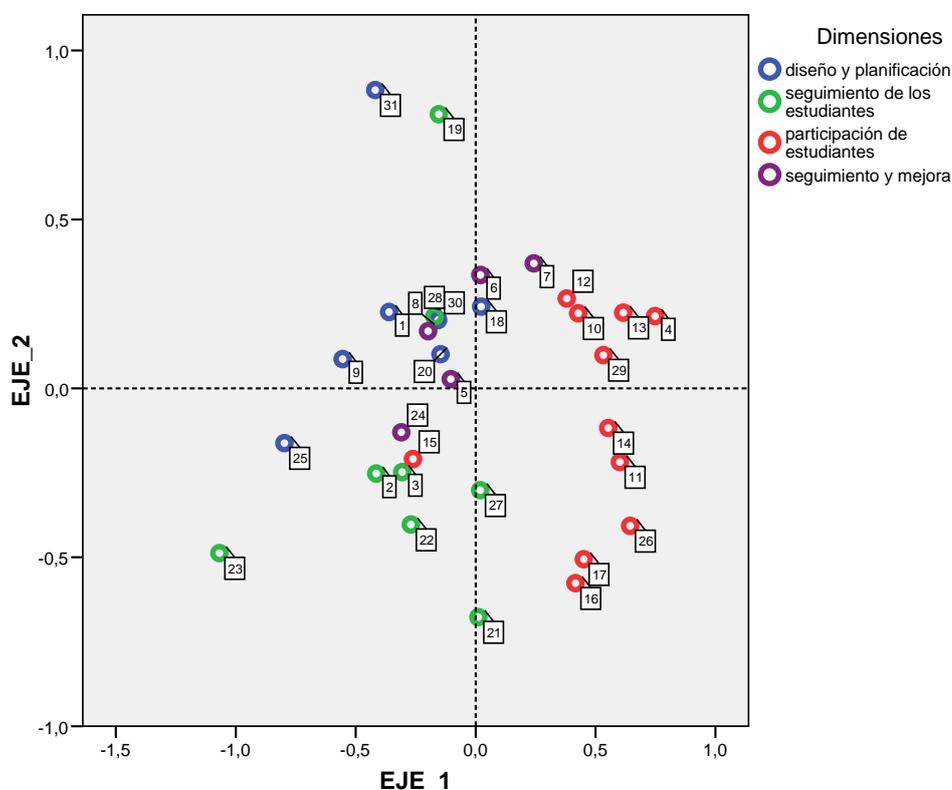


Gráfico 3. Coordenadas 1 y 2

## Escalamiento multidimensional: criterio de UTILIZACIÓN

Para este último criterio, relacionado con la ejecución de la tarea evaluadora, el indicador del stress se sitúa en el 0,010 y el Coeficiente

de Congruencia de Tucker (CCT) 0,994, lo que indica que, al igual que en los casos anteriores, se obtienen valores de calidad y ajuste que podemos calificar de excelentes (ver tabla 5).

Tabla 5. Medidas de ajuste y stress en el criterio utilización

Stress bruto normalizado	,01101
Stress-I	,10494(a)
Stress-II	,28931(a)
S-Stress	,02881(b)
Dispersión explicada (D.A.F.)	,98899
Coeficiente de congruencia de Tucker	,99448

PROXSCAL minimiza el stress bruto normalizado.

a Factor para escalamiento óptimo = 1,011.

b Factor para escalamiento óptimo = ,993.

En el Gráfico 4, que representa las dos primeras coordenadas principales, se repiten los patrones detectados en el gráfico 1. Así, se agrupan de manera muy clara los ítems correspondientes al bloque tercero salvo el ítem 15, mientras que el ítem 21, del segundo bloque se posiciona con los del tercero. Los ítems del

segundo bloque también aparecen agrupados con la excepción de los ítems 19 y 28. A diferencia del Gráfico 1, ahora sí que aparecen agrupados de manera clara los ítems del primer bloque, a excepción del ítem 20, y los del bloque cuarto, a excepción del 24.

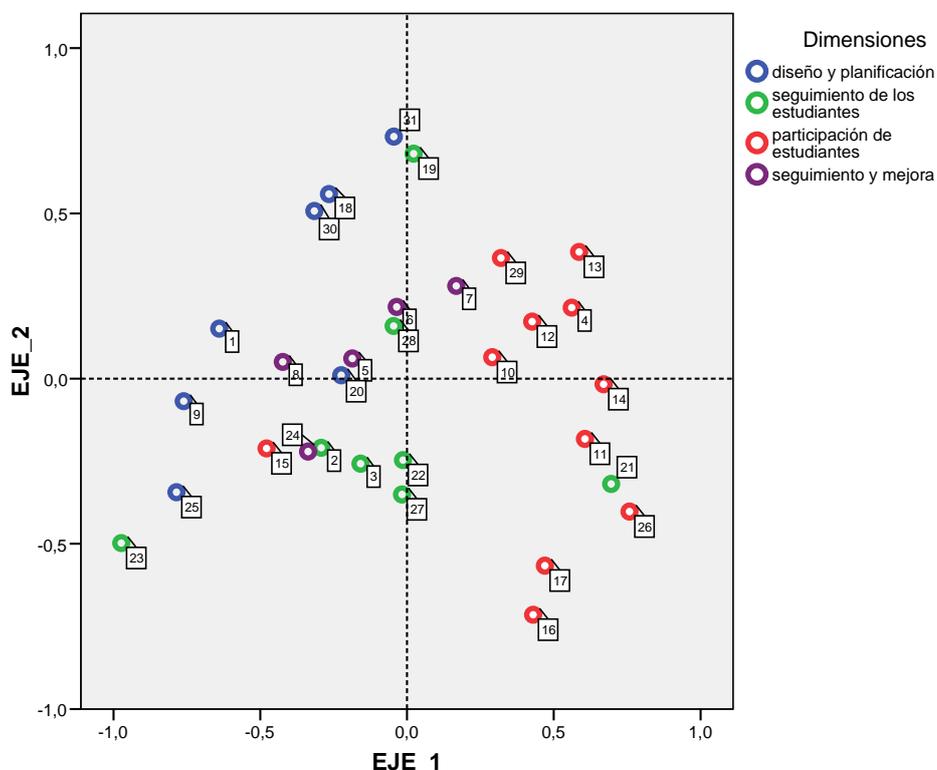


Gráfico 4. Coordenadas 1 y 2

Así pues, revisando de forma conjunta los resultados obtenidos en los tres aspectos evaluados referidos a los criterios de importancia, competencia y utilidad del cuestionario, se ratifica la existencia de los bloques propuestos en la estructura de ítems. Sin embargo, debe destacarse la mayor concentración en alguno de los bloques (siempre el segundo y tercero, y en menor medida el primero y el cuarto), así como la continuada presencia de alguno de los ítems alejado de los que teóricamente serían sus ítems afines, esto es, pertenecientes al mismo bloque (ítems 15, 21, 19 y 28).

### Conclusiones y discusión

El objetivo de este estudio es analizar las propiedades psicométricas del *Autoinforme sobre la actividad evaluadora del profesorado universitario (Cuestionario ActEval)* para confirmar su utilidad como medio para conocer la importancia que los docentes conceden a cada una de las actuaciones y tareas de evaluación, y en qué medida se consideran preparados para abordarlas, así como en qué grado las utilizan en su actividad cotidiana. La realización individual del *Autoinforme*, un instrumento que destaca por sus buenas cualidades psicométricas, permite conocer tanto la concepción como la actuación del profesorado universitario con respecto a la evaluación realizada durante el proceso educativo.

La consistencia interna de la escala es elevada, tanto en el análisis por criterios de valoración como en sus dimensiones teóricas. Aunque el instrumento ya se había validado teóricamente, mediante juicio de expertos (Quesada, Rodríguez e Ibarra, 2013), los resultados empíricos encontrados mediante el Escalamiento Multidimensional No Métrico indican la necesaria revisión de algunos ítems, principalmente de la dimensión *Seguimiento de los estudiantes* (19, 21 y 28) y *Participación* (ítem 15), pues se sitúan alejados de sus ítems teóricamente afines.

En cualquier caso, los resultados sobre las características técnicas del instrumento sugieren aplicaciones prácticas de utilidad para la

mejora universitaria y su adecuación al EEES. La disposición de un instrumento que recoja, no solo la percepción del profesorado universitario sobre las actividades de evaluación, sino la competencia percibida en dichas tareas y la utilidad real de las mismas permite conocer desde una perspectiva global el pensamiento y la actuación de los docentes sobre una de sus funciones esenciales. Este conocimiento permitirá analizar si las prácticas de evaluación actuales en la universidad son coherentes, o al menos indican una tendencia, con los nuevos retos que alguno expertos plantean al profesorado universitario (Boud, 2010; JISC, 2010).

La *evaluación orientada al aprendizaje*, coherente con los nuevos planteamientos educativos, plantea que el proceso de evaluación será efectivo en la medida en que se encuentre situado en el eje central del proceso de aprendizaje, ofreciendo a los estudiantes la oportunidad de participar en el mismo y de recibir una adecuada información sobre sus ejecuciones, de modo que les permita aprender y mejorar, además de intentar recabar información *objetiva* sobre sus errores, omisiones o equivocaciones (Boud *et al.*, 2010).

En cualquier caso, y más en el ámbito docente, es esencial realizar una reflexión crítica sobre las actuaciones profesionales. Frecuentemente, los análisis se refieren a los procesos de enseñanza, siendo menos frecuente la mejora de los procedimientos de evaluación, cuando son éstos los que actúan como dinamizadores del proceso de aprendizaje de los estudiantes. Analizar la práctica educativa, a través de los procedimientos reales llevados a cabo en el aula, permite enriquecer, por un lado, la idea del profesorado sobre la misma y por otro lado, mejorar o favorecer la función docente.

Aunque, sin duda, realizar un cambio en los procedimientos, instrumentos y concepciones de evaluación no solo requiere la modificación de las creencias docentes, sino que necesita, de forma paralela, un cambio en las consideraciones que los estudiantes tienen. Las posibles resistencias, las posiciones pasivas o una me-

nor implicación y responsabilidad, podría cercenar los procesos de *evaluación centrada en el aprendizaje*.

Cabría, por tanto, dirigir las acciones hacia un enfoque de evaluación profunda que, aunque no exento de dificultades y obstáculos, derivados principalmente de la carga de trabajo, la falta de experiencia y las percepciones y creencias de la comunidad educativa, puede ser implementado estableciendo pilares de actuación coherentes, coordinados y responsables (Sánchez, 2011) que permitan mejorar los procesos educativos.

## Referencias

- Adam, C. & King, K. (1995). Towards a framework for student self-assessment. *Innovations in Education and Training International*, 32, 336–343.
- Álvarez-Rojo, Víctor; Asensio-Muñoz Inmaculada; Clares, José; del-Frago, Rakel; García-Lupión, Beatriz; García-Nieto, Narciso; García-García, Mercedes; Gil, Javier; González-González, Daniel; Guardia, Soledad; Ibarra, Marisol; López-Fuentes, Rafael; Rodríguez-Diéguez, Antonio; Rodríguez-Gómez, Gregorio; Rodríguez-Santero, Javier; Romero, Soledad; Salmerón, Purificación (2009). Perfiles docentes para el espacio europeo de educación superior (EEES) en el ámbito universitario español. *RELIEVE*, v. 15, n. 1. [http://www.uv.es/RELIEVE/v15n1/RELIEVEv15n1\\_1.htm](http://www.uv.es/RELIEVE/v15n1/RELIEVEv15n1_1.htm).
- Andreu-Andrés, M<sup>a</sup>A. & García-Casas, M. (2006). Evaluación, coevaluación y autoevaluación del trabajo en grupo en la lectura de mapas topográficos. En: F. Watts & A. García-Carbonell (Eds.): *La evaluación compartida: investigación multidisciplinar*. Valencia, Universidad Politécnica de Valencia, pp. 69-90.
- Arenas, C. & Cuadras, C.M. (2002). Recent statistical methods based on distances. *Contributions to Science*, 2 (2), 183-191.
- Biggs, J. y Tang, C. (2009). *Teaching for Quality Learning at University*. Maidenhead: Open University Press, McGraw-Hill.
- Bloxham, S. y Boyd, P. (2007). *Developing Effective Assessment in Higher Education. A Practical Guide*. Londres: Open University Press, McGraww-Hill.
- Boud, D. (2006). Foreword. En: C. Bryan & K. Clegg. *Innovative Assessment in Higher Education* (pp. xvii-xix). New York: Routledge.
- Boud, D. (2010). *Assessment 2020: Seven propositions for assessment reform in higher education*. Sydney: Australian Learning and Teaching Council.
- Bozu, Z. & Canto, P.J. (2009). El profesorado universitario en la sociedad del conocimiento: competencias profesionales docentes. *Revista de Formación e Innovación Educativa Universitaria*, 2, 87-97.
- Briggs, C.L., Stark, J.S. & Rowland-Poplawsky, J. (2003). How Do We Know a “Continuous Planning” Academic Program When We See One? *The Journal of Higher Education*, 74 (4), 361-385.
- Busing, F. (1997). *PROXSCAL: A multidimensional scaling program*. En: A. Bandilla & F. Faulbaum. *Softstat 97. Advances in Statistical Software. The 9<sup>th</sup> Conference on the Scientific Use of Statistical Software, 1997*, Heidelberg.
- Butcher, C., Davies, C. & Highton, M. (2006). *Designing learning from module outline to effective teaching*. Abindgon: Routledge.
- Carballo, J.L., Pérez-Jover, M<sup>a</sup>V., Espada, J.P., Orgilés, M. & Piqueras, J.A. (2012). Propiedades psicométricas de la escala multidimensional de expresión social para la evaluación de habilidades sociales en el contexto de internet. *Psicothema*, 24, 121-126.
- Carless, D. (2003). *Learning-oriented assessment. Comunicación presentada en la Evaluation and Assessment. Conference, University of South Australia, Adelaide, 25 de noviembre de 2003*.
- Carless, D. (2011). *Learning-oriented assessment and the development of student learning capacities*. Ponencia presentada en el Congreso Internacional EVALtrends. Cádiz, 9 al 11 de marzo.

- Carless, D., Joughin, G. & Liu, N.F. (2006). *How Assessment supports learning: learning-oriented assessment in action*. Hong Kong University Press.
- Chen, Y.M. (2008). Learning to Self-assess Oral Performance in English: A Longitudinal Case Study. *Language Teaching Research*, 12 (2), 235-262.
- Cheng, W. & Warren, M. (1997). Having second thoughts: student perceptions before and after a peer assessment exercise. *Studies in Higher Education*, 22, 233-239.
- Cuadras, C.M. (1996). *Métodos de análisis multivariante*. Barcelona: EUB.
- De Leeuw, J., & Heiser, W.J. (1980). Multidimensional Scaling with Restrictions on the Configuration. En P.R. Krishnaiah (Ed.), *Multivariate Analysis*, Vol. 5, Amsterdam, Holland: North Holland Publishing Company.
- Dochy, F., Segers, M. & Sluijsmans, D. (1999). The use of self-, peer and co-assessment in higher education: a review. *Studies in Higher Education*, 24 (3), 331-350.
- Falchikov, N. (2005). *Improving Assessment through Student Involvement: Practical solutions for aiding learning in higher and further education*. Abingdon: Routledge.
- Gibbs, G. (2006). How assessment frames student learning. En Cordelia Bryan & Karen Clegg, *Innovative Assessment in higher education* (23-36). Abingdon: Routledge.
- Gibbs, G. & Simpson, C. (2004). Conditions under which assessment supports students' learning. *Learning and Teaching in Higher Education*, Issue 1.
- Gower, J.C. (1971). A general coefficient of similarity and some of its properties. *Biometrics*, 27, 857-874.
- Hanrahan, S. & Isaacs, G. (2001). Assessing self- and peer assessment: the students' views. *Higher Education Research & Development*, 20 (1), 53-70.
- Ibarra, M.S. (Dir.) (2008). *EvalCOMIX: Evaluación de competencias en un contexto de aprendizaje mixto*. Cádiz: Servicio de Publicaciones de la Universidad de Cádiz.
- Ibarra, M.S. & Rodríguez, G. (2010). Aproximación al discurso dominante sobre la evaluación del aprendizaje en la universidad. *Revista de Educación*, 351, 385-407.
- Ibarra, M.S., Rodríguez, G. & Gómez, M.A. (2012). La evaluación entre iguales: beneficios y estrategias para su práctica en la universidad. *Revista de Educación*, 359.
- Ibarra Sáiz, María Soledad; Rodríguez-Gómez, Gregorio & Gómez-Ruiz, Miguel-Ángel (2010). La planificación basada en competencias en los másteres oficiales: un reto para el profesorado universitario. *RELIEVE*, v. 16, n. 1.  
[http://www.uv.es/RELIEVE/v16n1/RELIEVEv16n1\\_6.htm](http://www.uv.es/RELIEVE/v16n1/RELIEVEv16n1_6.htm)
- Knight, P. (1995). *Assessment for Learning in Higher Education*. London: Kogan Page.
- Kruskal, J.B. (1964). Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis. *Psychometrika*, 29, 1-28.
- Kruskal, J.B. (1964). Nonmetric multidimensional scaling: a numerical method. *Psychometrika*, 29, 115-129.
- Kruskal, J.B. & Wish, M. (1978). *Multidimensional Scaling*. London: Sage.
- Litwin, M. (1995). *How to measure survey reliability and validity*. London: Sage.
- Ljungman, A.G. & Silén, C. (2008). Examination involving students as peer-examiners. *Assessment & Evaluation in Higher Education*, 33 (3), 289-300.
- López Pastor, V.M. (Coord.) (2009). *La evaluación formativa y compartida en docencia universitaria: propuestas, técnicas, instrumentos y experiencias*. Madrid: Narcea.
- Meulman, J. & Heiser, W.J. (2005). *SPSS Categories 14*. Chicago: SPSS Inc.
- Monereo, C. (2009). La enseñanza de tareas de autorregulación a través de tareas auténticas. En P. Díez (Coord.), *Nuevas funciones de la evaluación*. Madrid: Ministerio de Educación.
- Nunnally, J.C. (1978). *Psychometric theory*. New York: McGraw-Hill Book.
- Olivia, P. & Henson, K. (1980). What are the Essential Generic Teaching Competencies?. *Theory into Practice*, 9 (2), 117-121.

- Osterlind, S.J. (1989). *Constructing test items*. Boston: Kluwer.
- Pavoine, S., Vallet, J., Dufour, A.B., Gachet, S. & Daniel, H. (2009). On the challenge of treating various types of variables: application for improving the measurement of functional diversity. *Oikos*, 118, 391-402.
- Peña, D. (2002). *Análisis de datos multivariantes*. Madrid: McGraw-Hill.
- Pope, N. (2005). The impact of stress in self and peer assessment. *Assessment & Evaluation in Higher Education*, 30 (3), 51-63.
- Quesada, V., Rodríguez, G. & Ibarra, M.S. (2013, en prensa). ActEval: un instrumento para el análisis y reflexión sobre la actividad evaluadora del profesorado universitario. *Revista de Educación*, 362, Septiembre-diciembre.
- Rodríguez, G. (Dir.) (2009). *EvalHIDA: Evaluación de Competencias con Herramientas de Interacción Dialógica Asíncronas (foros, blogs y wikis)*. Cádiz: Servicio de Publicaciones de la Universidad de Cádiz.
- Rodríguez, G., Ibarra, M.S. & Gómez, M.A. (2011). e-Autoevaluación en la universidad: Un reto para profesores/as y estudiantes. *Revista de Educación*, 356, 401-430.
- Sánchez, J. (2011). Evaluación de los aprendizajes universitarios: una comparación sobre sus posibilidades y limitaciones en el Espacio Europeo de Educación Superior. *Revista de Formación e Innovación Educativa Universitaria*, 4 (1), 40-54.
- Takane, Y., Young, F.W. & De Leeuw, J. (1977). Nonmetric individual differences multidimensional scaling: an alternating least squares method with optimal scaling features. *Psychometrika*, 42, 593-600.
- Villardón, L. (2006). Evaluación del aprendizaje para promover el desarrollo de competencias. *Educación Siglo XXI*, 24, 57-76.
- Yalcin, I. & Amemiya, Y. (2001). Nonlinear Factor Analysis as a Statistical Method. *Statistical Science*, 16, 275-294.

---

### NOTAS

Resultados procedentes del Proyecto “Re-Evalúa: Reingeniería de la e-Evaluación, tecnologías y desarrollo de competencias en profesores y estudiantes universitarios” (Ref. P08-SEJ-03502). Proyecto financiado parcialmente por la Consejería de Innovación, Ciencia y Empresa de la Junta de Andalucía en su convocatoria de proyectos de excelencia.

---

## ABOUT THE AUTHORS / SOBRE LOS AUTORES

**Biencinto, Chantal** ([alameda@edu.ucm.es](mailto:alameda@edu.ucm.es)). Doctora en Filosofía y Ciencias de la Educación por la Universidad Complutense de Madrid. Actualmente es Profesora en el Departamento de Métodos de Investigación y Diagnóstico en Educación de la Facultad de Educación. Pertenece al grupo de investigación Pedagogía Adaptativa. Entre las líneas de investigación destacan la e-evaluación de competencias interpersonales en el contexto universitario y la Metodología de Investigación. Es la autora de contacto para este artículo. Su dirección postal es: Facultad de Educación UCM. C/ Rector Royo Villanova, s/n. 28040 Madrid (España). [Buscar otros artículos de](#)

[esta autora en Google Académico / Find other articles by this author in Scholar Google](#)



**Carpintero, Elvira** ([ecarpintero@edu.ucm.es](mailto:ecarpintero@edu.ucm.es)). Doctora en Psicopedagogía por la Universidad Complutense de Madrid. Actualmente es profesora en el Departamento de Métodos de Investigación y Diagnóstico en Educación de la Facultad de Educación de la Universidad Complutense de Madrid. Sus principales líneas de investigación se centran en las estrategias de aprendizaje y el transfer de conocimientos, la educación adaptativa y la evaluación de competencias interpersonales a través de plataformas virtuales. Pertenece al grupo de investigación Pedagogía Adaptativa y al grupo de Medida y Evaluación de Sistemas Educativos de la Universidad Complutense de Madrid. Es coordinadora del proyecto de mentoría entre compañeros SOU-estuTutor. Su dirección postal es: Facultad de Educación UCM. C/ Rector Royo Villanova, s/n. 28040 Madrid (España). [Buscar otros](#)

[artículos de esta autora en Scholar Google/ Find other articles by this author in Scholar Google](#)



**García-García, Mercedes** ([mergaci@edu.ucm.es](mailto:mergaci@edu.ucm.es)). Doctora en Ciencias de la Educación. Profesora Titular en el Departamento de Métodos de Investigación y Diagnóstico en Educación de la Universidad Complutense de Madrid. Directora del Servicio de Orientación Educativa de la Facultad de Educación. Su docencia se centra principalmente en las áreas de Pedagogía Diferencial y Orientación Educativa. Codirectora del grupo de investigación 940424 Pedagogía Adaptativa y miembro del grupo HUM-688 EVALFOR, sus líneas de investigación en los últimos años se centran en Evaluación de Estrategias de Adaptación Educativa en ESO, Tutoría universitaria en el EEES, y e-Evaluación de competencias interpersonales en contextos universitarios.. Su dirección postal es: Facultad de Educación UCM. C/ Rector Royo Villanova, s/n. 28040 Madrid (España). [Buscar otros](#)

[artículos de esta autora en Scholar Google/ Find other articles by this author in Scholar Google](#)



Biencinto, Chantal; Carpintero, Elvira & García-García, Mercedes (2013). Propiedades psicométricas del cuestionario ActEval sobre la actividad evaluadora del profesorado universitario. *RELIEVE*, v. 19, n. 1, art. 2.  
[DOI:10.7203/relieve.19.1.2611](https://doi.org/10.7203/relieve.19.1.2611).

## ARTICLE RECORD / FICHA DEL ARTÍCULO

<b>Reference / Referencia</b>	Biencinto, Chantal; Carpintero, Elvira & García-García, Mercedes (2013). Propiedades psicométricas del cuestionario ActEval sobre la actividad evaluadora del profesorado universitario. <i>RELIEVE</i> , v. 19, n. 1, art. 2. <a href="https://doi.org/10.7203/relieve.19.1.2611">DOI:10.7203/relieve.19.1.2611</a>
<b>Title / Título</b>	Propiedades psicométricas del cuestionario ActEval sobre la actividad evaluadora del profesorado universitario. <i>[Psychometric properties of the ActEval questionnaire on university teachers' assessment activity]</i> .
<b>Authors / Autores</b>	Biencinto, Chantal; Carpintero, Elvira & García-García, Mercedes
<b>Review / Revista</b>	RELIEVE (Revista ELectrónica de Investigación y EValuación Educativa), v. 19, n. 1
<b>ISSN</b>	1134-4032
<b>Publication date / Fecha de publicación</b>	2012 ( <b>Reception Date</b> : 2012 October 30; <b>Approval Date</b> : 2013 March 3. <b>Publication Date</b> : 2013 March 21)
<b>Abstract / Resumen</b>	<i>This article presents the results of the empirical validation of the questionnaire ActEval (Self-Report on the Evaluation Activity of university teachers). ActEval attempts to identify how important teachers consider a variety of assessment tasks, whether they feel competent, and to what extent they use them in their daily practice, in line with the new competences under the European Higher Education Area. Validation is studied using Non-Metric Multidimensional Scaling (PROXSCAL) and the reliability results show high levels of internal consistency. The data suggest that some items in the questionnaire should be reviewed in order to obtain an effective tool for the analysis of evaluation practice.</i> El presente artículo tiene como objetivo presentar los resultados de la validación empírica del cuestionario ActEval ( <i>Autoinforme sobre la Actividad Evaluadora del profesorado universitario</i> ), que permite conocer en qué medida el profesorado universitario considera importante, se siente competente y utiliza una serie de tareas de <i>evaluación orientada al aprendizaje</i> en su práctica diaria, acordes con las nuevas competencias derivadas del Espacio Europeo de Educación Superior. Los resultados de fiabilidad muestran elevados índices de consistencia interna y la validación empírica mediante el procedimiento de Escalamiento Multidimensional No-Métrico (PROXSCAL) sugiere la revisión de algunos ítems del cuestionario para obtener un instrumento eficaz para el análisis de la práctica evaluadora.
<b>Keywords / Descriptores</b>	<i>Competence assessment at University; perception of the teacher; construct validity; internal consistency.</i> Evaluación de competencias en la universidad, percepción del profesor, validación de constructo, consistencia interna.
<b>Institution / Institución</b>	Facultad de Educación. Universidad Complutense de Madrid (España)..
<b>Publication site / Dirección</b>	<a href="http://www.uv.es/RELIEVE">http://www.uv.es/RELIEVE</a>
<b>Language / Idioma</b>	Español & English version (Title, abstract and keywords in English & Spanish)

## RELIEVE

**Revista ELectrónica de Investigación y EValuación Educativa**  
*E-Journal of Educational Research, Assessment and Evaluation*

[ISSN: 1134-4032]

© Copyright, RELIEVE. Reproduction and distribution of this articles it is authorized if the content is no modified and their origin is indicated (RELIEVE Journal, volume, number and electronic address of the document).

© Copyright, RELIEVE. Se autoriza la reproducción y distribución de este artículo siempre que no se modifique el contenido y se indique su origen (RELIEVE, volumen, número y dirección electrónica del documento).