

## **DIFERENCIAS EN LAS PUNTUACIONES DE LAS ENCUESTAS DE VALORACIÓN DEL PROFESORADO EN FUNCIÓN DEL TIPO DE CUESTIONARIO: COMPARATIVA CUESTIONARIOS LIKERT VS CUESTIONARIOS BARS**

Differences in the scores of teacher evaluation surveys depending on the type of questionnaire: Comparative Likert questionnaires vs BARS questionnaires

Diferenças nas pontuações dos questionários de avaliação dos docentes, de acordo com o tipo de questionário: comparação de questionários Likert vs questionários BARS

**Luis Matosas-López (1)**

(1) Universidad Rey Juan Carlos, España. Teléfono: +34 658621389. Correo electrónico: [luis.matosas@urjc.es](mailto:luis.matosas@urjc.es)

### **Resumen**

Cuando de la evaluación del profesorado universitario se trata, la mayor parte de instituciones optan por el uso de encuestas de satisfacción con cuestionarios Likert, no obstante, existen otras alternativas. Este trabajo explora las posibles diferencias en las puntuaciones arrojadas por las encuestas de valoración del profesorado en función del tipo de cuestionario empleado. En la investigación toman parte 126 estudiantes. Los participantes valoran el desempeño de un único docente y asignatura utilizando, por una parte, un cuestionario Likert y, por otra, un cuestionario con episodios de comportamiento BARS (Behaviorally Anchored Rating Scales). Se realizan análisis descriptivos y de pruebas paramétricas (prueba t-Student). Los hallazgos muestran la existencia de diferencias significativas en los resultados de las encuestas en función de la tipología de cuestionario empleada. Mientras que los cuestionarios Likert parecen incrementar las valoraciones, los cuestionarios BARS arrojan puntuaciones más modestas, pero, probablemente, más ajustadas al desempeño real del docente.

**Palabras clave:** *Calidad de la educación; eficiencia del profesorado, evaluación del profesorado; cuestionario; universidad*

### **Abstract**

When it comes to the evaluation of the university teaching staff, most institutions opt for the use of satisfaction surveys with Likert questionnaires, however, there are other alternatives. This work explores the possible differences in the scores thrown by the teachers' assessment surveys depending on the type of questionnaire used. 126 students participate in the investigation. Participants rate the performance of a single teacher and subject using, on the one hand, a Likert questionnaire and, on the other, a questionnaire with BARS (Behaviorally Anchored Rating Scales). Descriptive and parametric analysis (t-Student test) are carried out. The findings show the existence of significant differences in the surveys' ratings depending on the type of questionnaire used. While the Likert questionnaires seem to increase the ratings, the BARS questionnaires show more modest scores, but, probably, more adjusted to teacher's real performance.

**Keywords:** *Quality of education; teacher efficiency; teacher assessment; questionnaire; university*

### **Resumo**

A maioria das instituições optam pelo uso de questionários de satisfação de Likert quando têm de avaliar os professores universitários, apesar de existirem outras alternativas. Este trabalho explora as possíveis diferenças nas pontuações obtidas pelos questionários de avaliação dos docentes, de acordo com o tipo de questionário usado. Participaram nesta investigação 126 alunos. Os participantes valoram o desempenho de um único professor e respetiva unidade curricular, usando por um lado, um questionário de Likert e, por outro, um questionário com episódios de comportamento BARS (Behaviorally Anchored Rating Scales). Realizam-se análises descritivas e testes paramétricos (teste t de Student). Os resultados mostram a existência de diferenças significativas nos resultados dos questionários, em função do tipo de questionário utilizado. Enquanto que os questionários de Likert parecem aumentar os valores das avaliações, os questionários BARS obtêm pontuações mais modestas, mas, provavelmente, mais ajustadas ao desempenho real do docente.

**Palavras-chave:** *Qualidade da educação; eficiência dos professores, avaliação dos docentes; questionário; Universidade*

## **1. Introducción**

La adaptación del sistema universitario a las directrices del EEES ha llevado a la implementación de diferentes procesos de fiscalización de la calidad. Una de las líneas de actuación es la del control de la calidad en la actividad docente. Desde que en 2007 la Agencia Nacional de Evaluación de la Calidad y Acreditación (ANECA) pusiera en marcha el proyecto DOCENTIA, este se ha convertido en el principal mecanismo de gestión de la calidad del desempeño del profesorado universitario en España (ANECA, 2015). Este programa, al igual que ocurre con los sistemas de gestión de la calidad docente en otros países, persigue dos grandes objetivos. Por un lado, la mejora de la labor docente (Villa, 2008) y, por otro, la promoción profesional del profesorado (Subbaye & Vithal, 2017).

La satisfacción de ambos objetivos requiere de la existencia de indicadores, y estos generalmente se constituyen a partir de encuestas de satisfacción cumplimentadas por el alumnado. A la hora de desarrollar estas encuestas la mayor parte de las instituciones optan por cuestionarios Likert (Lizasoain-Hernández, Etxeberria-Murgiondo, & Lukas-Mujika, 2017; Molero & Ruiz, 2005), no obstante, para tal propósito, también existen cuestionarios con episodios de comportamiento o BARS (Behaviorally Anchored Rating Scales) (Dickinson & Zellinger, 1980; Matosas-López & Leguey-Galán, 2018), e incluso cuestionarios que emplean preguntas de tipo abierto (Nasser-Abu Alhija & Fresko, 2009). El presente trabajo explora las posibles diferencias en las puntuaciones arrojadas por las encuestas de valoración del profesorado en función del tipo de cuestionario empleado. La investigación se centra en la comparativa entre cuestionarios Likert y cuestionarios BARS.

## **2. Método**

El estudio se desarrolla en la Facultad de Ciencias Jurídicas y Sociales de la Universidad Rey Juan Carlos en Madrid. La investigación, con el propósito de mantener el rigor comparativo de los hallazgos, se concentra en la valoración del desempeño de un mismo docente, en diferentes grupos en los que imparte idéntica asignatura. En el estudio participan 126 estudiantes seleccionados por muestreo incidental (De-Juanas & Beltrán, 2013). La edad promedio de los participantes se sitúa en 20.03 años ( $DT = 1.48$ ), siendo el 53.71% de sexo femenino y el 46.29% de sexo masculino.

El investigador emplea, por una parte, un cuestionario con escalas de Likert (figura 1) y, por otra, un cuestionario con episodios de comportamiento BARS (figura 2). Ambos cuestionarios constan de diez preguntas con cinco opciones de respuesta en cada caso.

**SEÑALA TU NIVEL DE ACUERDO SOBRE LA SIGUIENTE AFIRMACIÓN RELATIVA AL DESEMPEÑO DE TU PROFESORA. MARCA SOLO UNA OPCIÓN. (1=Muy en desacuerdo, 2=En desacuerdo, 3=Ni de acuerdo ni en desacuerdo, 4=De acuerdo, 5=Muy de acuerdo)**

1.- Introducción a la asignatura

El profesor/a explica detalladamente en sus primeros días de clase la guía docente de la asignatura, o la parte del temario que imparte.  1  2  3  4  5

**Figura 1.** Ejemplo de pregunta del cuestionario Likert

En el primer cuestionario las opciones de respuesta se representan en forma de niveles de acuerdo, en el segundo en forma de episodios de comportamiento representativos de la labor docente. Con el propósito de garantizar una correcta comprensión de las preguntas, el contenido de ambos cuestionarios es validado recurriendo a la técnica del juicio de expertos (Lacave, Molina, Fernández & Redondo, 2016).

**SEÑALA EL CONJUNTO DE EPISODIOS DE COMPORTAMIENTO QUE MEJOR DEFINE EL DESEMPEÑO DE TU PROFESORA. MARCA SOLO UNA OPCIÓN.**

1.- Introducción a la asignatura

El profesor/a NO presenta todos los puntos clave de la guía docente al inicio del curso (programa de estudios, competencias, objetivos, metodología de trabajo, ECTS, recursos en Aula Virtual ...); NO expone la importancia de la asignatura en términos académicos / profesionales; NO describe el plan cronológico de contenidos de la asignatura ni la inversión de tiempo requerida en cada parte (clases, exámenes, tareas ...) y NO proporciona una descripción detallada de la bibliografía / materiales de apoyo, ni de la forma de usarlos

El profesor/a expone la importancia de la asignatura en términos académicos / profesionales

El profesor/a presenta todos los puntos clave de la guía docente al inicio del curso (programa de estudios, competencias, objetivos, metodología de trabajo, ECTS, recursos en Aula Virtual ...) y proporciona una descripción detallada de la bibliografía / materiales de apoyo, además de la forma de usarlos

El profesor/a presenta todos los puntos clave de la guía docente al inicio del curso (programa de estudios, competencias, objetivos, metodología de trabajo, ECTS, recursos en Aula Virtual ...); expone la importancia de la asignatura en términos académicos / profesionales y proporciona una descripción detallada de la bibliografía / materiales de apoyo, además de la forma de usarlos

El profesor/a presenta todos los puntos clave de la guía docente al inicio del curso (programa de estudios, competencias, objetivos, metodología de trabajo, ECTS, recursos en Aula Virtual ...); expone la importancia de la asignatura en términos académicos / profesionales; describe el plan cronológico de contenidos de la asignatura y la inversión de tiempo requerida en cada parte (clases, exámenes, tareas ...) y proporciona una descripción detallada de la bibliografía / materiales de apoyo, además de la forma de usarlos

**Figura 2.** Ejemplo de pregunta del cuestionario BARS

Ambos cuestionarios son administrados a los participantes durante la última semana del curso. Tras recabar y depurar las respuestas, el investigador realiza, por un lado, un análisis descriptivo de los datos y, por otro, un análisis de pruebas paramétricas para muestras relacionadas empleando la prueba t-Student.

No obstante, antes de abordar el análisis de pruebas paramétricas, el investigador extrae el estadístico Shapiro-Wilk con el propósito de comprobar que los promedios de los ítems en ambos cuestionarios siguen una distribución normal (*p-valor* Likert = .465 y *p-valor* BARS = .378). Todos los análisis se desarrollan empleando el software SPSS V25.

### 3. Resultados

Los resultados de los análisis descriptivos, recogidos en la tabla 1, muestran como las valoraciones promedio en el cuestionario Likert son superiores a las alcanzadas en el cuestionario BARS en ocho de las diez preguntas.

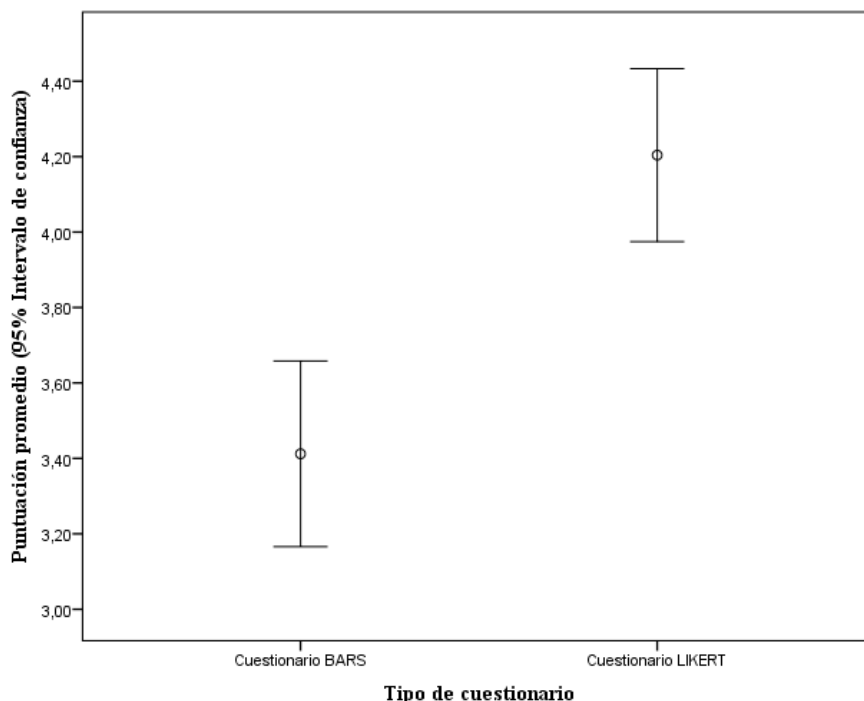
De igual forma, se observa también como las puntuaciones obtenidas en el cuestionario BARS reflejan un mayor grado de acuerdo entre los participantes. No obstante, el valor de *DT* en el segundo cuestionario es inferior al del primero en siete de los diez ítems.

**Tabla 1.**

*Resultados descriptivos*

	Cuestionario Likert		Cuestionario BARS	
	Promedio	DT	Promedio	DT
1. Introducción a la asignatura	4.33	1.049	3.33	1.029
2. Descripción del sistema de evaluación	4.50	.856	3.47	.761
3. Gestión del tiempo	4.04	1.235	3.39	.921
4. Coherencia organizativa	3.74	1.297	3.90	.990
5. Implementación del sistema de evaluación	4.60	.706	3.79	.816
6. Resolución de dudas	4.59	.860	4.06	.861
7. Capacidad explicativa	3.69	1.236	3.94	1.115
8. Disponibilidad general	4.25	.977	3.34	.921
9. Facilidad de seguimiento	4.17	.939	3.44	.944
10. Satisfacción general	4.13	1.124	3.46	.985

El análisis de pruebas paramétricas muestra, considerando las puntuaciones promedio obtenidas en cada uno de los diez ítems, unos resultados de la prueba t-Student que revelan como la puntuación promedio difiere de forma significativa en ambos instrumentos. Para el umbral de significatividad  $\pm = .05$  adoptado, el estadístico  $t = 5.325$  para 17.909 grados de libertad muestra un *p-valor* < .001.



**Figura 3.** Gráfico de barras de error simple

En línea con lo anterior, el gráfico de barras de error simple (figura 3) muestra la ausencia de solapamiento entre las puntuaciones medias de los cuestionarios, corroborando así la presencia de diferencias significativas.

#### **4. Conclusiones**

Los hallazgos del presente trabajo muestran la existencia de diferencias significativas en las puntuaciones de las encuestas de valoración del profesorado, en función de la tipología de cuestionario utilizada. Los resultados revelan que mientras los cuestionarios Likert parecen incrementar las valoraciones, los cuestionarios BARS arrojan puntuaciones más modestas.

Estudios previos revelan que el empleo de episodios de comportamiento en la construcción de cuestionarios BARS permite aumentar la objetividad de la valoración y reducir la influencia de potenciales variables de sesgo (Bernardin & Beatty, 1984; Harari & Zedeck, 1973). Esto, en opinión del autor, indica que las valoraciones del cuestionario BARS reflejan, probablemente, de manera más fiel la labor realizada por el profesor durante el curso. En línea con Matosas-López y Leguey-Galán (2019), los resultados de la presente investigación proporcionan indicios que hacen pensar que las valoraciones del profesorado realizadas empleando cuestionarios BARS permiten al alumnado realizar una valoración más ajustada al desempeño real del docente.

## Referencias

- ANECA. (2015). *Programa de apoyo para la evaluación de la actividad docente del profesorado universitario (Docentia)*. Madrid. Recuperado a partir de [http://www.aneca.es/content/download/13305/164819/file/DOCENTIA\\_nuevadoc\\_v1\\_final.pdf](http://www.aneca.es/content/download/13305/164819/file/DOCENTIA_nuevadoc_v1_final.pdf)
- Bernardin, H. J., & Beatty, R. W. (1984). *Performance appraisal: Assessing human behavior at work*. (PWS, Ed.). Boston: Kent Pub. Co.
- De-Juanas Oliva, Á., & Beltrán Llera, J. A. (2013). Valoraciones de los estudiantes de ciencias de la educación sobre la calidad de la docencia universitaria. *Educación XXI*, 17(1), 59-82. <https://doi.org/10.5944/educxx1.17.1.10705>
- Dickinson, T. L., & Zellinger, P. M. (1980). A comparison of the behaviorally anchored rating and mixed standard scale formats. *Journal of Applied Psychology*, 65(2), 147-154. <https://doi.org/10.1037//0021-9010.65.2.147>
- Harari, O., & Zedeck, S. (1973). Development of Behaviorally Anchored Scales for the Evaluation of Faculty Teaching. *Journal of Applied Psychology*, 58(2), 261-265. <https://doi.org/10.1037/h0035633>
- Lacave Rodero, C., Molina Díaz, A. I., Fernández Guerrero, M. M., & Redondo Duque, M. A. (2016). Análisis de la fiabilidad y validez de un cuestionario docente. *Revista de Investigación en Docencia Universitaria de la Informática*, 9(1), 2.
- Lizasoain-Hernández, L., Etxeberria-Murgiondo, J., & Lukas-Mujika, J. F. (2017). Propuesta de un nuevo cuestionario de evaluación de los profesores de la Universidad del País Vasco. Estudio psicométrico, dimensional y diferencial. *RELIEVE - Revista Electrónica de Investigación y Evaluación Educativa*, 23(1). <https://doi.org/10.7203/relieve.23.2.10436>
- Matosas-López, L., & Leguey-Galán, S. (2018). Implementación de Behavioral Anchored Rating Scales (BARS) para la evaluación del profesorado universitario en asignaturas de modalidad Online. En P. G. H. y R. H. M. Carlos Monge López (Ed.), *I Congreso Virtual Internacional y III Congreso Virtual Iberoamericano sobre Recursos Educativos Innovadores CIREI* (p. 204). Madrid: Fundación General de la Universidad de Alcalá.
- Matosas-López, L., & Leguey-Galán, S. (2019). Actitud del alumnado hacia el uso de instrumentos BARS (Behavioral Anchored Rating Scales) para la valoración del

profesorado universitario. En AIDIPE (Ed.), *XIX Congreso Internacional de Investigación Educativa: Investigación Comprometida para la Transformación Social*. Madrid: Asociación Interuniversitaria de Investigación Pedagógica.

Molero López-Barajas, D. M., & Ruiz Carrascosa, J. (2005). La evaluación de la docencia universitaria. Dimensiones y variables más relevantes. *Revista de Investigación Educativa*, 23(1), 57-84. Recuperado a partir de <http://revistas.um.es/rie/article/view/98341>

Nasser-Abu Alhija, F., & Fresko, B. (2009). Student evaluation of instruction: What can be learned from students' written comments? *Studies in Educational Evaluation*, 35(1), 37-44. <https://doi.org/10.1016/j.stueduc.2009.01.002>

Subbaye, R., & Vithal, R. (2017). Teaching criteria that matter in university academic promotions. *Assessment & Evaluation in Higher Education*, 42(1), 37-60. <https://doi.org/10.1080/02602938.2015.1082533>

Villa, A. (2008). La excelencia docente. *Revista de Educación, Número Ext(1)*, 177-212. Recuperado a partir de [http://www.revistaeducacion.mec.es/re2008\\_08.htm](http://www.revistaeducacion.mec.es/re2008_08.htm)