Revista Iberoamericana para la Investigación y el Desarrollo Educativo
ISSN 2007 - 7467

https://doi.org/10.23913/ride.v15i29.2100

Artículos científicos

Evaluación del desempeño docente por estudiantes de una universidad pública del noroeste de México: análisis descriptivos y comparativos

Evaluation of teaching performance by students of a public university in northwest Mexico: descriptive and comparative analyzes

Avaliação do desempenho docente de estudantes de uma universidade pública do noroeste do México: análises descritivas e comparativas

Patricio Henríquez Ritchie
Universidad Autónoma de Baja California, México
phenriquez@uabc.edu.mx
https://orcid.org/0000-0002-1026-3379

Juan Carlos Pérez Morán
Universidad Autónoma de Baja California, México
jucarpint@gmail.com
https://orcid.org/0000-0001-8849-2326

Sofía Contreras Roldán
Universidad Autónoma de Baja California, México
sofiacontreras@gmail.com
https://orcid.org/0000-0003-1987-3445

Carlos Javier del Cid
Universidad Autónoma de Baja California, México
carlos.cid@uabc.edu.mx
https://orcid.org/0000-0001-7579-7331

Revista Iberoamericana para la Investigación y el Desarrollo Educativo
ISSN 2007 - 7467

Resumen

La FCAyS de la UABC (México) implementa semestralmente la evaluación del desempeño docente desde la opinión de los estudiantes a través de la Escala de Evaluación del Desempeño Docente (EEDDocente), compuesta por tres subescalas: i) planeación y organización de la enseñanza, ii) calidad de la enseñanza y iii) evaluación y retroalimentación del aprendizaje. Se reportan los resultados del período 2023-1. El objetivo es analizar el desempeño docente desde la mirada de los estudiantes a partir de variables personales (sexo, edad) y académicas (áreas de conocimiento, etapas de formación). Se realizó un estudio cuantitativo con alcance descriptivo y comparativo, con base en análisis estadísticos inferenciales paramétricos (t-student para muestras independientes, ANOVA). Los resultados muestran que las variables sexo, edad y áreas de conocimiento impactan significativamente en la opinión de los estudiantes respecto al desempeño de sus profesores. Las comparaciones realizadas arrojaron diferencias significativas (95% de confianza) a favor de la opinión de hombres, estudiantes del cuartil mayor de edad (Q3), que se encontraban cursando la etapa terminal de sus estudios y aquellos del área de conocimiento de ciencias jurídicas. Sin embargo, estos resultados deben interpretarse con cautela y considerarse como un punto inicial para la realización de otras investigaciones enfocadas en el análisis y evaluación del desempeño docente en el nivel educativo superior, abarcando otras áreas de conocimiento y diversificando las perspectivas de estudio, de manera que se incorporen otros agentes educativos (docentes, directivos, expertos externos) y otros acercamientos metodológicos (autoevaluación, coevaluación, heteroevaluación externa).

Palabras claves: evaluación docente; evaluación formativa; educación superior; áreas de conocimiento.

Abstract

The FCAyS of UABC (Mexico) implements the evaluation of teaching performance every semester from the opinion of the students through the application of the Teaching Performance Evaluation Scale (EEDDocente), which is made up of three subscales: i) planning and organization of teaching, ii) quality of teaching and iii) evaluation and feedback of learning. This article reports the results for the period 2023-1. The objective is to analize teaching performance from the perspective of the FCAyS students of the UABC based on personal variables (sex, age) and academic variables (areas of knowledge and stages of training). A quantitative study with descriptive and comparative scope was carried out, based

on parametric inferential statistical analyzes (t-student for independent samples, ANOVA). The results show that the variables sex, age and areas of knowledge significantly impact the students' opinions regarding the performance of their teachers. The comparisons made showed significant differences (95% confidence) in favor of the opinion of men, students of the oldest quartile (Q3), who were in the final stage of their studies and those in the area of knowledge of legal sciences. However, these results should be interpreted with caution and considered as a starting point that encourages the conduct of other research focused on the analysis and evaluation of teaching performance at the higher education level, covering other areas of knowledge and diversifying the study perspectives, so that other educational agents (teachers, managers, external experts) and other methodological approaches (self-assessment, co-assessment, external hetero-assessment).

Keywords: teaching evaluation; formative evaluation; higher education; areas of knowledge.

Resumo

A FCAyS da UABC (México) implementa semestralmente a avaliação do desempenho docente a partir da opinião dos alunos por meio da aplicação da Escala de Avaliação de Desempenho Docente (EEDDocente), que é composta por três subescalas: i) planeamento e organização do ensino, ii) qualidade do ensino e iii) avaliação e feedback da aprendizagem. Este artigo relata os resultados para o período 2023-1. O objetivo é descrever o desempenho docente na perspectiva dos alunos do FCAyS da UABC a partir de variáveis pessoais (sexo, idade) e variáveis acadêmicas (áreas de conhecimento e etapas de formação). Foi realizado um estudo quantitativo com escopo descritivo e comparativo, baseado em análises estatísticas inferenciais paramétricas (t-student para amostras independentes, ANOVA). Os resultados mostram que as variáveis sexo, idade e áreas de conhecimento impactam significativamente a opinião dos alunos quanto ao desempenho de seus professores dentro das dimensões medidas pelo instrumento EEDDocent. As comparações efetuadas evidenciaram diferenças significativas (95% de confiança) a favor da opinião dos homens, dos alunos do quartil mais antigo (Q3), que se encontravam na fase final dos estudos, e dos da área do conhecimento. das ciências jurídicas. Contudo, estes resultados devem ser interpretados com cautela e considerados como um ponto de partida que incentiva a realização de outras pesquisas focadas na análise e avaliação do desempenho docente no nível superior, abrangendo outras áreas do conhecimento e diversificando as perspectivas de estudo, para que são incorporados

outros agentes educativos (professores, gestores, especialistas externos) e outras abordagens metodológicas (autoavaliação, coavaliação, heteroavaliação externa.

Palavras-chave: avaliação docente; avaliação formativa; ensino superior; áreas do conhecimento.

Fecha Recepción: Abril 2024 **Fecha Aceptación:** Septiembre 2024

Introduction

Research on the evaluation of teacher performance in higher education began to develop during the first half of the 20th century in the United States, based on the measurement of student learning; that is, learning was used as a reference basis to evaluate the effectiveness of teaching. Later, in this same context, the mechanism for evaluating teacher performance from the students' perspective began to become widespread, based on the argument that it is the main users of the educational service who must evaluate the quality of the teaching received. Following the changes in higher education financing policies towards the 1980s, the evaluation of teacher performance became a fundamental component of accountability in universities (Alcaraz-Salarirche, 2015; Cisneros-Cohernour and Stake, 2010; García-Olalla et al., 2022; Zhao et al., 2022; Zamora Serrano, 2021). By 1990, in the context of globalization and the privatization of knowledge, the concepts of quality and educational equity became part of the dominant discourse of educational policy in universities. With this, the evaluation of all its components, including teaching, was extended to most higher education institutions (Galaz Ruiz et al., 2019).

The evaluation of teaching can be understood as a systematic process of collecting information about the performance of teachers, which can seek both summative and formative purposes (Fink, 2008; Liebowitz, 2021; UNESCO, 2006). On the one hand, the evaluation of teaching performance from a summative and high-consequence approach is associated with the processes of hiring, promotion, delivery of incentives and granting of recognition to teachers, as well as with the accountability and decision-making of educational authorities. On the other hand, the evaluation of teaching performance from a formative approach provides information to design training or updating programs and thus provide feedback and improve those aspects of teaching that are identified as areas of opportunity, and thus, promote the professional development of teachers (García Cabrero et al., 2008; Gómez and Valdéz, 2019; Liebowitz, 2021; Silva Huaman et al., 2022).

Currently, the use of scales and questionnaires for the evaluation of teaching performance from the perspective of students is a common practice in higher education institutions (Gómez and Valdés, 2019; Wang and Guan, 2017; Zamora, 2021). In this sense, some authors (García, 2014; Mohammadi, 2021) have highlighted that the use of scales and questionnaires as instruments for evaluating teacher performance are ideal tools for measuring the effectiveness of teaching at the higher education level. In addition to this, Bazán-Ramírez et al. (2021) mention that resorting to the use of self-report scales to evaluate teachers' performance from the students' perspective has associated administrative and academic purposes that can contribute to the understanding of the teaching-learning process both in the classroom and at the school level.

However, the use of self-report scales applied to students as the only method to evaluate teaching performance is a recurring practice in universities that is strongly questioned due to social desirability and its effects on the objectivity of the responses, as well as obtaining evidence of validity and reliability of the instruments (Benton and Young, 2018; Boysen, 2016; Gómez and Valdés, 2019; Newton et al., 2019; Zamora Serrano, 2021). In this way, in recent years there has been a growing interest in diversifying the methods, techniques and instruments for teacher performance evaluation processes, as well as the sources of information and participating subjects (students, teachers, managers, external experts, among others (Bleiberg et al., 2023; Romero and Martínez, 2017; Zamora Serrano, 2021). For his part, Kikut Valverde (2018) points out that, beyond the limitations of scale methods to evaluate performance from the students' perspective, they provide a variety of evidence that helps to better understand teaching in the classroom. Molero and Carrascosa (2005) argue that students are the main observers of their teachers' performance and their assessment, so under adequate representative samples, the reliability of the information collected is not put at risk.

In Mexico, the application of scales to evaluate teacher performance from the students' perspective began to proliferate since the 1990s. The process of evaluation of teaching performance has been in use since 1960 in various higher education institutions and its proliferation occurred from the 1990s onwards (Canales et al., 2004; Gómez and Valdés, 2019). Currently and at a general level, this process continues to be carried out under the same methodological guidelines. However, some authors (García Garduño and Medécigo Shej, 2014; Rueda et al., 2010) assert that this process is used with mainly summative or administrative functions, counting that more than 80% of higher education institutions use

the results as a basis for promotion programs or salary incentives. Therefore, it is of utmost importance to plan and implement teacher performance evaluation processes from the students' perspective that allow obtaining valuable information to enrich, improve, provide feedback and perfect the teaching practices of university professors, based on reliable, valid and relevant instruments around this object of study. The purpose of this document is to present the results of the evaluation of teaching performance from the students' perspective, which was implemented in the period 2023-1 at the Faculty of Administrative and Social Sciences (FCAyS) of the Autonomous University of Baja California (UABC), Mexico. In particular, the results of the evaluation of teaching performance by subscale and dimensions are described and characterized, based on personal variables (sex, age) and academic variables (area of knowledge and stages of training) of the students. Likewise, the results of the comparative analysis of the dimensions of teaching performance by sex, age, area of knowledge and training stage are shown.

Method

Spatio-temporal context and participants

The FCAyS is an academic unit of the UABC that concentrates eight undergraduate programs (Law, Business Administration, Accounting, Computer Science, Education Sciences, Communication Sciences, Psychology and Sociology), in addition to two common cores (TC_administration and TC_social) grouped into three areas of knowledge: Legal Sciences, Administrative Sciences and Social Sciences. In particular, in the first two semesters, the subjects of the undergraduate programs in the areas of Administrative Sciences and Social Sciences are integrated into a common core, except for the Bachelor of Law. During the period 2023-1, according to official figures from the UABC, the total enrollment of FCAyS students was 4,180 students. For the purposes of the study, a random sample of 1,480 students was considered, representing 35.4% of the total population for this period. 47.8% of the participating students were enrolled in Bachelor's programs in the area of Administrative Sciences, and 32.1% in Bachelor's programs in the area of Social Sciences (see Table 1):

Table 1. *Participants by educational program and areas of knowledge.*

Degree	n	%	Areas	n	%
Law	298	20.1%	Legal Sciences	298	20.1%

Business Administration	305	20.6%	Administrative Sciences	708	47.8%
Computing	3	0.2%			
Accountancy	226	15.3%			
TC_administration	174	11.8%			
Comunication	102	6.9%	Social Sciences	474	32.1%
Psychology	215	14.5%			
Education	80	5.4%			
Sociology	18	1.2%			
TC_social	59	4.0%			
n total=	1,480	100%	n total=	1,480	100%

Source: own elaboration

Regarding the distribution of participants by gender, as shown in Table 2, between six and seven out of 10 students who participated in the study were women (64.3%, n=951). In particular, in the area of knowledge of Social Sciences, the participation of women was greater. (70.7%, *n*=335).

Table 2. *Percentage of participants by gender.*

Areas	Gender	n	%
Legal Sciences	Women	183	12.4%
	Men	115	7.7%
Administrative Sciences	Women	433	29.3%
	Men	275	18.6%
Social Sciences	Women	335	22.6%
	Men	139	9.4%
	N=	1,480	100%

Source: own elaboration

Similarly, Table 3 shows the percentage of participants by age quartile. The average age of the sample was 22.2 years (S.D.=5.2 years), with a minimum of 17 years and a maximum of 67 years. Young people under 19 years old belong to the first quartile (Q1), students between 20 and 22 years old belong to the second quartile (Q2), and those over 23 years old belong to the third quartile (Q3). On the other hand, nearly half of the participants (47.3%) were located in the second quartile (Q2), while 26.9% and 25.8% of the participants are within the Q1 and Q3 quartiles, respectively.

Tabla 3. *Percentage of participants by quartile according to age range.*

Age Range	Quartile	n	%
Under 19 years old	Q1	398	26.9%
Between 20 to 22 years old	Q2	700	47.3%
Over 23 years old	Q3	382	25.8%
	N=	1,480	100%

Source: own elaboration

Table 4 presents the percentage of participating students by quartile according to age range and area of knowledge. The highest percentage of participants in the three areas of knowledge belongs to the second quartile (between 20 and 22 years old). In particular, the highest percentage of participants by quartile according to age range (22.3%) is concentrated in students between 20 and 22 years old in the area of Administrative Sciences. In contrast, the lowest percentage of participants (5.9%) is concentrated in young students under 19 years old in the area of Social Sciences.

Table 4. *Percentage of participants by quartiles according to age ranges and areas of knowledge***.**

Age Quartile		Legal Sciences	Administrative Sciences	Social Sciences
Q1	n	91	220	87
	%	6.1%	14.9%	5.9%
Q2	n	117	330	253
	%	7.9%	22.3%	17.1%
Q3	n	90	158	134
	%	6%	10.7%	9.1%
Subtotal	n	298	708	474
	%	20.1%	47.8%	32.1%

Source: own elaboration

Instrument

The Teaching Performance Evaluation Scale (EEDDocente), designed by researchers from the FCAyS (Henríquez et al., 2023), was applied with the purpose of measuring teaching performance from the students' perspective, based on three main subscales: (1) Planning and organization of teaching, (2) Quality of teaching, and (3) Evaluation and feedback on learning.

The Teaching Planning and Organization subscale measures, from the student's perspective, the teacher's ability to explain the subject matter in a clear and organized manner,

as well as the execution of different classroom organization activities. This subscale is made up of two dimensions: Class planning, which is made up of 10 items on a four-point scale (1 = Did not explain, 2 = The explanation was confusing, 3 = Partially clear explanation, 4 = Clearly explained); and Classroom Organization, which is made up of 16 items on a four-point Likert-type scale (1 = Strongly disagree, 2 = Disagree, 3 = Agree, and 4 = Strongly agree).

The Teaching Quality subscale, on the other hand, consists of 16 items on a four-point Likert-type scale (1 = Strongly disagree, 2 = Disagree, 3 = Agree, and 4 = Strongly agree). This measures the teacher's mastery in: the application of teaching techniques, the ability to link the subject content with other subjects in the academic program, the promotion of student participation in class, the establishment of rules for coexistence, adjustments to the activity plan based on student suggestions, among other teaching activities.

Likewise, the Learning Assessment and Feedback subscale measures, from the student's perspective, the teacher's ability to apply learning assessment strategies, processes, methods, and techniques, as well as the actions for monitoring, feedback, and improving learning based on the results of the assessments. This subscale is made up of two dimensions: Quality of assessment strategies, composed of 14 items on a four-point Likert-type scale (1 = Strongly disagree, 2 = Disagree, 3 = Agree, and 4 = Strongly agree); and Use of learning assessment strategies, composed of 11 items on a four-point scale (1 = Never, 2 = Sometimes, 3 = Almost always, 4 = Always) that measure the frequency of use by teachers of different learning assessment strategies.

Regarding validity and reliability, the EEDDocente has evidence of construct validity of internal structure and factorial invariance (Henríquez et al., 2023). The scale presents a moderate correlation index between the items (0.64) and an overall Cronbach's Alpha reliability coefficient of 0.92. The model contains three factors (F1 = Teaching planning, F2 = Teaching quality, and F3 = Learning assessment and feedback) whose Confirmatory Factor Analysis shows acceptable fit indices ($\chi 2 = 251.21$; $df = 87$, $p = 0.000$; $CFI = 0.868$; $TLI = 0.841$; $GFI = 0.936$; $NNFI = 0.814$; $RMSEA = 0.034$; $SRMR = 0.057$). In addition, there are factorial invariance studies by area of knowledge and stage of training in samples of university students that reveal an adequate fit of the Configurational model, and significant differences in the parameters of the Weak, Strong and Strict models. Thus, it can be said that the EEDDocente presents solid evidence for carrying out comparative studies.

Procedure and data analysis

The EEDDocente was administered online using a Google Form in the middle of the 2023-1 period. Authorization and support were requested from the FCAyS management to enter the classrooms with the electronic link to access the instrument and a QR code. The students answered immediately, using their own electronic equipment (smartphones, laptops, and tablets). The average response time was approximately 20 minutes and the questionnaire was available for two weeks to be answered by students who did not have access to the application at the time.

Once the data was collected, the file was configured in the SPSS statistical package, version 26.0, the responses were emptied, the database was purified (deletion of lost values or cases), and the statistical analysis began. Subsequently, the basic descriptives of each variable were obtained (percentages, frequency distributions, central tendency indicators, and dispersion indices). Likewise, parametric inferential statistical analyses (Student t-test for independent samples and ANOVA) were performed to compare statistical means between contrast groups based on contextual variables (sex, age range, area of knowledge and stage of training).

Results

This section shows the results of each subscale and dimension of the EEDDocente by sex, age range, and area of knowledge. Table 5 shows the central tendency indicators (mean and standard deviation) of the EEDDocente subscales and dimensions at a general level. It should be noted that all of them were estimated by means of an ordinal scale of four categories associated with attributes of quality, frequency, and agreement. In this sense, the average of each subscale and dimension allows us to obtain a general overview of the students' opinion on the performance of their teachers, which reflects favorable scores on the Teaching Quality subscale (X = 3.5), and the dimensions of Class Planning (X = 3.4), Class Organization (X = 3.2), and Quality of Educational Assessment Activities (X = 3.2). It is worth mentioning that three reagents were eliminated from subscale 2. Quality of teaching that collected information in a negative sense about the opinion of the students in the teaching activities (The teacher is confusing when explaining complex topics, X = 2.7; The teacher reinforces the memorization and retention of the subject content, X = 3.2; and The teacher usually generates a threatening and distrustful environment in the class, X = 3.4). On the other hand, it is striking that the dimension Use of learning assessment strategies, which

refers to the frequency with which the teacher uses certain instruments and strategies for assessing learning, reports a lower average ($\underline{X}$ = 2.4).

Table 5. *Indicators of central tendency of subscales and dimensions of the EEDDocent.*

Subescales	Dimensions	Mean	D.E.
Planning and organization of teaching	Class planning	3.4	0.7
	Class organization	3.2	0.8
Teaching quality	---	3.5	0.6
Learning evaluation and feedback	Quality of educational evaluation activities	3.2	0.8
	Use of learning assessment strategies	2.4	0.8

Source: own elaboration

Table 6 shows the results of a comparison of means for independent samples (t-student) carried out from the variable sex regarding the opinion of students regarding the use of learning assessment strategies by their teachers. As can be seen, the results showed a significant difference in favor of men ((X)= 2.5; sig. = 0.000, 95% confidence), who indicate that their teachers more frequently use assessment instruments and strategies, such as research papers, presentations, questionnaires, portfolios, multiple choice exams and individual essays, mainly.

Table 6. *T-student for independent samples of the Use of learning evaluation strategies according to the variable sex of the students.*

Dimension	Descriptives	Gender		T-Student (Sig.)
		Women	Men	
Use of learning assessment strategies	Mean	2.34	2.50	0.000
	D.E.	0.8	0.9	

Source: own elaboration

Table 7 shows the results of the analysis of contrast of means (ANOVA) by subscales and dimensions of the EEDDocente and the quartiles by age range of the students. Significant differences were observed between the contrast groups (95% confidence) in subscale 2. Quality of teaching (sig. = 0.000) and in the dimensions Class planning (sig. = 0.001), Class organization (sig. = 0.000), Quality of educational assessment activities (sig. = 0.000) and Use of learning assessment strategies (sig. = 0.000). It should be noted that, in all the significant differences observed, the group of students over 23 years old (Q3) obtained the highest statistical means, while the group of young people under 19 years old (Q1) obtained the lowest means.

Table 7. *Comparison of means (ANOVA) by subscales and dimensions of the EEDDocent according to the quartile of the age range of the students.*

Dimensions and subdimensions	Descriptives	Q1	Q2	Q3	ANOVA (Sig.)
Class planning	Mean	3.32	3.36	3.50	0.001
	D.E.	0.8	0.8	0.7	
Class organization	Mean	3.14	3.23	3.36	0.000
	D.E.	0.8	0.7	0.7	
Teaching quality	Mean	3.35	3.50	3.63	0.000
	D.E.	0.7	0.6	0.6	
Quality of educational evaluation activities	Mean	3.09	3.21	3.33	0.000
	D.E.	0.8	0.8	0.8	
Use of learning assessment strategies	Mean	2.21	2.41	2.58	0.000
	D.E.	0.8	0.8	0.9	

Source: own elaboration

In turn, Table 8 shows the results of a comparison of means analysis (ANOVA) by areas of knowledge, based on the subscales and dimensions of the EEDDocente. As can be seen, the students of the knowledge area of Legal Sciences obtained the highest mean with significant differences (95% confidence) in the dimensions Class organization (X = 3.3, sig. = 0.011) and Use of learning assessment strategies (X = 2.52, sig. = 0.019). On the other hand, the students of the knowledge area of Social Sciences obtained the highest mean in subscale 2. Quality of teaching (X = 3.57, sig. = 0.000). It is striking that the students of the knowledge area of Administrative Sciences showed the lowest means in all the contrasts carried out around the subscales and dimensions of the EEDDocente.

Table 8. *Comparison of means (ANOVA) by subscales and dimensions of the EEDDocent according to the area of knowledge to which the students are assigned.*

Subescales and dimensions	Descriptives	Area of knowledge			ANOVA (Sig.)
		Legal	Administrative	Social	
Class organization	Media	3.33	3.18	3.27	0.011
	D.E.	0.7	0.8	0.8	
Teaching quality	Media	3.56	3.42	3.57	0.000
	D.E.	0.6	0.7	0.6	
Use of learning assessment strategies	Media	2.52	2.37	2.37	0.019
	D.E.	0.9	0.8	0.8	

Source: own elaboration

Finally, Table 9 shows the results of the comparative analysis between the means (ANOVA) according to the stage of training that the students were in at the time of answering

the scale. Again, significant differences were observed (95% confidence) between the contrast groups and subscale 2. Quality of teaching (sig. = 0.000), as well as in the dimensions Class planning (sig. = 0.001), Class organization (sig. = 0.001), Quality of educational assessment activities (sig. = 0.000) and Use of learning assessment strategies (sig. = 0.000). It is striking that, in all the significant differences observed, the group of students who were in the Terminal training stage of their study program showed the highest statistical means, while the lowest means were observed in students in the Basic and Disciplinary training stages.

Table 9. *Comparison of means (ANOVA) by subscales and dimensions of the EEDDocent according to the training stage of the students***.**

Subescales and dimensions	Descriptives	Training Stages			ANOVA (Sig.)
		Basic	Discipline	Terminal	
Class planning	Media	3.26	3.41	3.45	0.001
	D.E.	0.8	0.7	0.7	
Class organization	Media	3.11	3.27	3.30	0.001
	D.E.	0.8	0.7	0.8	
Teaching quality	Media	3.30	3.55	3.56	0.000
	D.E.	0.8	0.6	0.6	
Quality of educational evaluation activities	Media	3.05	3.23	3.32	0.000
	D.E.	0.8	0.8	0.7	
Use of learning assessment strategies	Media	2.29	2.39	2.54	0.000
	D.E.	0.8	0.8	0.9	

Source: own elaboration

Discussion

The findings reported in this study coincide with the contributions made by some authors in research carried out in other contexts regarding the evaluation of teacher performance within the higher education level. In this regard, García Garduño and Medécigo Shej (2014) highlighted that students from a public university in the central region of Mexico in the area of Social Sciences and Humanities established the following priority criteria for evaluating an effective teacher: the teaching method/didactics used, the knowledge/mastery of the subject matter, as well as the teacher's punctuality and attendance. In turn, the ineffectiveness criteria established by the students were: the teacher's attitudes, values, and personality, such as interaction with the group, treatment, and interest in teaching. In this sense, EEDDocente attempts to investigate these same aspects through three dimensions and two subdimensions:

i) class organization (teaching planning and teaching activities), ii) teaching quality, and iii) feedback and learning assessment (assessment strategy and frequency). The results provided attempt to analyze and estimate the level of student satisfaction, although the challenge remains of determining to what extent the evaluation that students make about their teachers' performance could represent the degree of learning achieved by students in the courses taught by teachers.

For its part, in the study reported by Kikut Valverde (2018) it is highlighted that women showed a greater tendency to participate and respond to teacher performance evaluation questionnaires at the higher education level, while those students who had a better academic performance, reflected in the passing levels of the subjects, evaluated their teachers better. Within this academic variable, a positive relationship stands out between those who obtain the best grades and a better evaluation of their teachers, in aspects related to the professor's work so that they better understand the content and the clarity with which it is presented. In the case of the present study, no evidence was found that allows establishing an incidence of the academic performance of the students in the results of the evaluation of the performance of the teachers of the FCAyS of the UABC. In this sense, an important area of opportunity for future lines of research is seen within the evaluation of teacher performance: the analysis of the impact of the academic variables of university students in the evaluation of the performance of their teachers, considering aspects of class organization, quality of teaching and evaluation-feedback of learning.

Conclusions

The analysis of the results of the evaluation of teaching performance from the students' perspective, which is applied in the FCAyS of the UABC, allowed us to identify strengths and weaknesses in the subscales and dimensions considered in relation to teaching. In particular, it was found that male students, compared to female students, are more willing to indicate that their teachers more frequently use learning assessment strategies related to research papers, presentations, questionnaires, portfolios, multiple choice exams and individual essays. This is interesting if one considers that nearly two thirds of the sample of participating students (64.3%) were women. In addition, the students within the oldest quartile (Q3) showed a significantly more favorable opinion regarding all the subscales and dimensions considered in the EEDDocente, compared to the students belonging to the youngest quartiles (Q1 and Q2). At the academic level, the training stage that students were

in was also a variable that influenced the evaluation of teachers' performance: those in the terminal stages (last semesters) showed significantly higher evaluation rates in each of the dimensions and subdimensions considered in the instrument, compared to students in the initial stages of their degrees. In addition, students in the area of Legal Sciences showed significantly higher evaluation levels for teaching activities (clarity and sequentiality in the presentation of content, justification and association of prior knowledge, promotion of meaningful and collaborative learning, among others), quality of teaching (promotion of collaborative, situated learning, case studies, debates) and frequency of use of learning assessment strategies and instruments by teachers, while those in the area of Administrative Sciences showed the lowest rates.

Future lines of research

The findings reported in this study confirm the contributions and benefits of using scales to assess teacher performance from the perspective of students for training and educational improvement purposes. Among the limitations of the study, it is considered relevant to mention the lack of other sources that provide convergent evidence of teacher performance from other perspectives. It is recommended that future studies include other instruments that expand the exploration of relevant constructs to provide feedback on teacher performance improvement, and implement a teaching evaluation strategy with a formative approach that integrates other evaluation methods (for example, self-assessments, peer assessment, classroom performance observation, to mention a few) that provide valuable information from different sources for decision-making to improve teaching and the design of training and continuing education policies and programs for FCAyS teachers.

References

Alcaraz-Salarirche, N. (2015). Aproximación histórica a la evaluación educativa: de la generación de la medición a la generación ecléctica. *Revista Iberoamericana de Evaluación Educativa*, *8*(1). https://dialnet.unirioja.es/servlet/articulo?codigo=5134142

Bazán-Ramírez, A., Pérez-Morán, J. C. y Bernal-Baldenebro, B. (2021). Criteria for teaching performance in psychology: invariance according to age, sex, and academic Stage of Peruvian students. *Front. Psychol*, *12*. doi:10.3389/fpsyg.2021.764081

Benton S. y Young, S. (2018). Best Practices in the Evaluation of Teaching. *IDEA Center, paper 69. https://files.eric.ed.gov/fulltext/ED588352.pdf*

Bleiberg, J., Brunner, E., Harbatkin, E., Kraft, M.A, y Springer, M.G. (2023). *The effect of teacher evaluation on achievement and attainment: evidence from statewide reforms* [Working Paper]. Annenberg Brown University. https://scholar.harvard.edu/sites/scholar.harvard.edu/files/mkraft/files/bleiberg_et_al._2021_teacher_eval.pdf

Boysen, G. A. (2016). Using student evaluations to improve teaching: evidence-based recommendations. *Scholar. Teach. Learn. Psychol. 2*, 273-284. doi: 10.1037/stl0000069

Canales, A., Luna, E., Díaz Barriga, F., Monroy, M., Díaz, M. y García, J.M. (2004). Aproximaciones metodológicas al análisis y la evaluación de la docencia. En Rueda, M. y Díaz Barriga, F. (Coord.). *La evaluación de la docencia en la universidad. Perspectivas desde la investigación y la intervención profesional*. CESU-UNAM.

Cisneros-Cohernour, E., y Stake, R. (2010). La evaluación de la docencia en educación superior: de evaluaciones basadas en opiniones de estudiantes a modelos por competencias. *Rev. Iberoam. Eval. Educ*., *3*, 218-231.

Fink, D. (2008). *Evaluating teaching: a new approach to an old problem in to improve the academy: resources for faculty, instructional, and organizational development.* Eds. S. Chadwick-Blossey and D. R. Robertson. Jossey-Bass.

Galaz Ruiz, A., Jiménez-Vázquez, M.S. y Díaz Barriga, A. (2019). Evaluación del desempeño docente en Chile y México. Antecedentes, convergencias y consecuencias de una política global de estandarización. *Perfiles Educativos*, *41*(163), 156-176. https://perfileseducativos.unam.mx/iisue_pe/index.php/perfiles/article/view/58935/52097

García, J. M. (2014). ¿Para qué sirve la evaluación de la docencia? Un estudio exploratorio de las creencias de los estudiantes. *Education Policy Analysis Archives, 22,* 1-20. https://www.redalyc.org/pdf/2750/275031898010.pdf

García Cabrero, B., Loredo Enríquez, J. y Carranza Peña, G. (2008). Análisis de la práctica educativa de los docentes: pensamiento, interacción y reflexión. *Revista Electrónica de Investigación Educativa, Especial.* http://redie.uabc.mx/redie/article/viewFile/200/345

García Garduño, J. M. y Medécigo Shej, A. (2014). Los criterios que emplean los estudiantes universitarios para evaluar la ineficacia docente de sus profesores. *Perfiles Educativos, 36*(143), 124-139. https://www.scielo.org.mx/scielo.php?pid=S0185-26982014000100008&script=sci_abstract

García-Olalla, A., Villa-Sánchez, A., Aláez, M. y Romero-Yesa, S. (2022). Aplicación y resultados de un sistema para evaluar la calidad de la docencia universitaria en una década de experimentación. *Rev. Investig. Educ., 40,* 51-68. doi: 10.6018/rie.401221

Gómez, L. F. y Valdés, M. G. (2019). La evaluación del desempeño docente en la educación superior. *Revista Psicológica Educativa Propósitos y Representaciones, 7*(2), 479-515. https://revistas.usil.edu.pe/index.php/pyr/article/view/255/633

Henríquez, P.S., Morán-Pérez, J.C., Del Cid García C.J. y Zamora, J.E. (2023). Factor structure and invariance of the scale to measure teaching performance in the area of social sciences. *Frontiers in Education, 8*. doi:10.3389/feduc.2023.1229129

Kikut Valverde, L. (2018). *Influencia de algunas variables sobre los resultados de la evaluación del desempeño docente.* Centro de Evaluación Académica: Universidad de Costa Rica. https://kerwa.ucr.ac.cr/handle/10669/81591

Liebowitz, D. (2021). Teacher evaluation for accountability and growth: should policy treat them as complements or substitutes? *Labour Econ., 71*. doi: 10.1016/j. labeco.2021.102024

Mohammadi, M. (2021). Dimensions of teacher performance evaluation by students in higher education. *Shanlax Int. J. Educ*., *9*, 18-25. doi: 10.34293/education.v9i2.3673

Molero, D. y Carrascosa, J. (2005). La evaluación de la docencia universitaria. Dimensiones y variables más relevantes. *Revista de Investigación Educativa, 23*(1), 57-84. https://revistas.um.es/rie/article/view/98341/94041

Newton, G., Poung, K., Laila, A., Bye, Z., Bettger, W. y Cottenie, K. (2019). Perception of biology instructors on using student evaluations to inform their teaching. International Journal of Higher Education, *81(*1), 133-147. doi: 10.5430/ijhe.v8n1p133

Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura (UNESCO) (2006). Evaluación del desempeño y carrera profesional docente. Un estudio comparado entre 50 países de América y Europa. Oficina Regional de Educación para América Latina y el Caribe. http://unesdoc.unesco.org/images/0015/001529/152934s.pdf

Romero, T. y Martínez, A. (2017). Construcción de instrumentos de evaluación del desempeño docente universitario desde una perspectiva cualitativa. *Revista Universitaria del Caribe*, *18*(1), 34-42. https:// www.researchgate.net/publication/320210350_ Construccion_de_instrumentos_de_evaluacion_del_ desempeno_docente_universitario_desde_una_perspectiva_cualitativa

Rueda, M., Luna, E., García Cabrero, B. y Loredo, J. (2010). La evaluación de la docencia en las universidades públicas mexicanas: un diagnóstico para su comprensión y mejora. *Revista Iberoamericana de Evaluación Educativa*, *3*(1), 77-92. https://revistas.uam.es/riee/article/view/4506

Silva Huaman, J., Solís Trujillo, B.P., Huaman Vargas, J. y Quispe Merma, F. (2022). La evaluación formativa en el desempeño docente universitario: Revisión sistemática de literatura. *TecnoHumanismo Revista Científica*, *2*(4) 1-14. doi: 10.53673/th.v2i4.177

Wang, D. F. y Guan, L. (2017). Higher education quality evaluation from the perspective of students: theoretical construction and reflection. *J. Nat. Inst. Educ. Admin*., *5*(75). doi: 10.3969/j.issn.1672-4038.2017.05.005

Zamora Serrano, E. (2021). La evaluación del desempeño docente mediante cuestionarios en la universidad: su legitimidad según la literatura y los requerimientos para que sea efectiva. *Revista Actualidades Investigativas en Educación, 21*(3), 1-23. doi: 10.15517/aie.v21i3.46221

Zhao, L., Xu, P., Chen, Y., y Yan, S. (2022). A literature review of the research on students' evaluation of teaching in higher education. *Frontiers in Psychology*, *13*. doi: 10.3389/fpsyg.2022.1004487

Rol de Contribución	Autor (es)
Conceptualización	Patricio Henriquez Ritchie (principal), Juan Carlos Pérez Morán (igual)
Metodología	Patricio Henriquez Ritchie (principal), Carlos Javier del Cid (que apoya), Sofia Contreras Roldán (que apoya)
Software	
Validación	Patricio Henriquez Ritchie (principal), Sofia Contreras Roldán (que apoya)
Análisis Formal	Juan Carlos Pérez Morán (principal), Carlos Javier del Cid (que apoya)
Investigación	Patricio Henriquez Ritchie (princiàl), Sofia Contreras Roldán (que apoya), Juan Carlos Pérez Morán (que apoya), Carlos Javier del Cid (que apoya)
Recursos	Patricio Henriquez Ritchie (principal), Juan Carlos Pérez Morán (igual)
Curación de datos	Patricio Henriquez Ritchie (principal), Juan Carlos Pérez Morán (igual)
Escritura - Preparación del borrador original	Patricio Henriquez Ritchie (principal), Sofia Contreras Roldán (que apoya), Juan Carlos Pérez Morán (que apoya), Carlos Javier del Cid (que apoya)
Escritura - Revisión y edición	Patricio Henriquez Ritchie (princiàl), Juan Carlos Pérez Morán (que apoya)
Visualización	Patricio Henriquez Ritchie
Supervisión	Patricio Henriquez Ritchie (principal), Sofia Contreras Roldán (que apoya), Juan Carlos Pérez Morán (que apoya), Carlos Javier del Cid (que apoya)
Administración de Proyectos	Patricio Henriquez Ritchie
Adquisición de fondos	Patricio Henriquez Ritchie