EDUCATIONIS
MOMENTUM

ARTÍCULO ORIGINAL

Comentarios abiertos en evaluación docente desde la opinión estudiantil: análisis por género del docente

Open-ended Comments in Teaching Evaluation from the Student Perspective: Analysis by Instructor Gender

Erika Paola Reyes Piñuelas [1]*

[1] Universidad Autónoma de Baja California, Mexicali, México

* Autor de correspondencia: erikareyes@uabc.edu.mx

Resumen

Contexto: la evaluación docente desde la perspectiva estudiantil constituye un elemento para la mejora educativa; sin embargo, los comentarios abiertos suelen ser infravalorados frente a los indicadores numéricos. **Objetivo**: analizar la relación entre comentarios abiertos estudiantiles y valoración global docente, identificando diferencias por género, en la licenciatura de Odontología. Estudio cuantitativo con 50 docentes de una universidad del noroeste mexicano y 655 comentarios estudiantiles categorizados en positivos, negativos, neutros y mixtos. **Metodología**: se calculan medidas de tendencia central, comparan medias de porcentajes por género y estiman correlaciones de Pearson. **Resultados**: indican que las docentes mujeres recibieron mayor proporción de comentarios negativos. La correlación entre comentarios negativos y valoración global fue casi perfecta en mujeres ($r = -0.921$; $r^2 = 84.8$ %) y fuerte en hombres ($r = -0.774$; $r^2 = 59.9$ %), ambas significativas ($p < 0.001$). **Conclusión**: los comentarios negativos presentan una asociación más intensa con la valoración global en las docentes mujeres, lo que plantea la necesidad de considerar posibles sesgos de género en la evaluación estudiantil e integrar el análisis de comentarios abiertos en procesos de evaluación formativa y equitativa.

Palabras clave: retroalimentación negativa, sesgo de género, educación superior, evaluación docente, percepciones de estudiantes, educación odontológica

Abstract

Context: teaching evaluation from the student perspective constitutes a key element for educational improvement; however, open-ended comments are often undervalued compared to numerical indicators. **Aim**: analyze the relationship between student open-ended comments and overall teaching evaluations in Dentistry, identifying potential differences by gender. **Methodology**: a quantitative study was conducted involving fifty faculty members from a Mexican university and 655 student comments, which were categorized as positive, negative, neutral, or mixed. Measures of central tendency were calculated, mean percentages were compared by gender, and Pearson correlation coefficients were estimated. **Results**: indicate that female faculty received a higher proportion of negative comments. Furthermore, the correlation between negative comments and overall evaluation scores was nearly perfect among women ($r = -0.921$; $r^2 = 84.8\%$) and strong among men ($r = -0.774$; $r^2 = 59.9\%$), both statistically significant ($p < 0.001$). **Conclusion**: negative comments showed a stronger association with overall evaluation scores

OPEN ACCESS

Como citar (APA 7.ª) Reyes Piñuelas, E. P. (2026). Comentarios abiertos en evaluación docente desde la opinión estudiantil: análisis por género del docente. *Educationis Momentum, 12*(1), 1-16. https://doi.org/10.36901/em.v12i1.1879

Historia del artículo

Recibido: 2026.03.30

Aprobado: 2026.05.15

Publicado: 2026.XX.XX

Copyright: © 2026 *Educationis Momentum.* Este es un artículo de acceso abierto distribuido bajo los términos de Creative Commons Atribution 4.0 (CC BY 4.0) que permite el uso, la distribución y la reproducción sin restricciones en cualquier medio, siempre que se otorguen los créditos correspondientes al autor original y a la fuente.

Disponibilidad de los datos: Los autores declaran que los datos, materiales, instrumentos y demás recursos utilizados en esta investigación no se encuentran disponibles para su consulta debido a que no se obtuvo el consentimiento para que la información se socializara y se compartiera en bases de datos.

Contribución de los autores

EPRP: conceptualización, análisis formal, administración de proyecto, investigación, metodología, tratamiento de datos, recursos, software, supervisión, validación, visualización de resultados, redacción: borrador original, redacción: revisión y edición.

Declaración de conflicto de intereses

La autora declara que no existe ningún conflicto de intereses financiero, profesional o personal que pudiera haber influido en los resultados presentados en este estudio.

Financiamiento

Investigación sin financiamiento específico:

Esta investigación no recibió financiamiento específico de agencias del sector público, comercial o sin fines de lucro.

Declaración ética

Cuando corresponda aprobación de un comité de ética:

La autora declara que la investigación se realizó conforme a los principios éticos aplicables. El estudio fue aprobado por el Comité de Ética de y Evaluación de Investigación y Posgrado, de la Facultad de Ciencias Humanas, con número de aprobación 101/2787. Así mismo, se utilizaron reportes institucionales del SED, cuyos registros analizados no contenían información que permitiera identificar individualmente a los estudiantes evaluadores. El uso de esta información se realizó conforme a las disposiciones institucionales aplicables y bajo criterios de confidencialidad y protección de los participantes.

Objetos de ciencia abierta

Los reportes institucionales utilizados como fuente de datos, provenientes del SED en su modalidad de Opinión de los Alumnos, no se encuentran disponibles para consulta pública debido a las disposiciones institucionales de confidencialidad y protección de la información. Por esta razón, los datos utilizados en el análisis no se encuentran depositados en un repositorio de acceso abierto. Los materiales o instrumentos desarrollados específicamente para esta investigación podrán ponerse a disposición previa solicitud, cuando su divulgación no implique restricciones institucionales o comprometa la confidencialidad de la información.

Agradecimientos

La autora agradece a la Universidad Autónoma de Baja california por el apoyo brindado durante el desarrollo de la investigación.

for female faculty, highlighting the need to consider potential gender biases in student evaluations and to integrate the analysis of open-ended comments into formative and equitable evaluation processes.

Keywords: negative feedback, gender bias, higher education, teacher evaluation, student perceptions, dental education

En el contexto de la educación superior en México, la institucionalización de la evaluación de la docencia ha dejado de ser un ejercicio administrativo aislado para consolidarse como uno de los mecanismos de aseguramiento de la calidad y de las políticas de mejora continua. Aunque sus orígenes respondieron a lógicas de rendición de cuentas y estímulos económicos, las instituciones enfrentan hoy el desafío de diseñar sistemas que reconozcan la complejidad de la práctica docente y promuevan su transformación pedagógica (Hernández-Villafaña & Luna, 2023), que debe superar enfoques fiscalizadores o punitivos (Ochoa & Moya, 2019) para convertirse en un dispositivo reflexivo que impulse mejoras reales y sostenibles en los procesos de enseñanza y aprendizaje.

En el ámbito de las ciencias de la salud, donde la experticia disciplinar y la utilidad práctica son fundamentales, el proceso evaluativo suele reducirse a métricas cuantitativas que ignoran la riqueza de significado contenida en los comentarios abiertos de los estudiantes (Ripoll-Nuñez & Arrieta-Caycedo, 2022).

Zheng et al. (2023) señalan diversas investigaciones que advierten que estos instrumentos no son neutrales, ya que están permeados por sesgos de género que imponen estándares evaluativos fluidos y prejuicios que afectan de manera desproporcionada a las docentes mujeres, perpetuando las desigualdades estructurales e institucionales en la educación superior. Mientras que a los hombres se les suele valorar por su carisma o conocimientos técnicos, roles tradicionalmente asociados a la autoridad y la competencia intelectual, las mujeres enfrentan expectativas de rol vinculadas al cuidado y la afectividad, lo que resulta en valoraciones cualitativas más polarizadas y críticas que impactan su imagen profesional (Gelber et al., 2022).

Estas asimetrías se materializan en una brecha de género persistente en la educación superior, donde las mujeres enfrentan mayores obstáculos estructurales para alcanzar rangos de profesorado titular y puestos de dirección universitaria (Sigurdardottir et al., 2023). En las áreas de Ciencias de la Salud, esta problemática se agudiza al operar bajo jerarquías históricamente masculinizadas que restringen el acceso de las académicas a las esferas de toma de decisiones, consolidando entornos institucionales complejos que condicionan de origen el desarrollo profesional femenino.

Particularmente en el ámbito de la Odontología, ejercer en estos contextos implica para las docentes una constante sobrecarga de labor emocional y una marcada segregación jerárquica, las académicas de esta disciplina continúan subrepresentadas, ocupando significativamente

menos jefaturas de departamento, coordinaciones clínicas y decanatos en comparación con sus pares masculinos (Madhu et al., 2026). Esta falta de representatividad en las cúpulas de poder, sumada a los sesgos que enfrentan en las evaluaciones estudiantiles, termina frenando su promoción y perpetuando el techo de cristal institucional (Gelber et al., 2022; Weber et al., 2025).

La importancia de este trabajo radica en la necesidad de transitar hacia una cultura de la evaluación que privilegie la evaluación formativa y el diálogo reflexivo, proporcionando evidencias que permitan transformar las prácticas institucionales hacia modelos más equitativos y centrados en la calidad del aprendizaje, al visibilizar cómo el género actúa como un filtro en la percepción estudiantil, particularmente en las ciencias de la salud, donde la formación clínica ha estado históricamente marcada por dinámicas de poder y asimetrías de género en el ejercicio profesional y la persistencia de sesgos en los procesos de evaluación docente (Weber et al., 2025). De modo que este estudio justifica la relevancia de diseñar sistemas evaluativos integrales que protejan la integridad del profesorado y fomenten el desarrollo profesional continuo en las facultades de Odontología.

Ante este panorama, el presente estudio tiene como objetivo analizar la relación entre los comentarios abiertos estudiantiles y la valoración global docente en el contexto de la Odontología, identificando patrones diferenciales según el género del profesorado. Así mismo, la pregunta de investigación que guía este análisis es: ¿cómo influye el género del docente en la relación entre los comentarios abiertos estudiantiles y la evaluación global de la enseñanza?

Caracterización de los comentarios abiertos

Diversos estudios han documentado la variabilidad en la participación estudiantil en la evaluación docente y en la proporción de estudiantes que emiten comentarios abiertos. Por ejemplo, Vargas-Madriz y Nocente (2023) señalan que la tasa de respuesta puede alcanzar alrededor del 60 % cuando se implementan estrategias como recordatorios por correo electrónico. En contraste, Padgett (2021) reporta niveles particularmente altos en un curso de enfermería comunitaria, con tasas de participación entre el 94 % y el 98 %; de estos, entre el 77 % y el 90 % incluyeron comentarios escritos, lo que representa una proporción sustancial del total del estudiantado. Por su parte, en un estudio longitudinal en Australia, Gelber et al. (2022) identificaron una tasa de respuesta global del 38.6 %, con diferencias por sexo (42 % en mujeres frente a 33 % en hombres).

En contextos europeo y estadounidense se reportan tasas de participación similares en la sección cualitativa, donde el 70 % (aproximadamente) de los estudiantes opta por emitir comentarios abiertos en las evaluaciones docentes (Brockx et al., 2012; Zimmaro et al., 2006). Estos autores también señalan que la probabilidad de emitir comentarios varía según la disciplina y el tamaño del grupo, siendo mayor en cursos de menor tamaño y en áreas con mayor interacción docente-estudiante.

En cuanto a algunos hallazgos reportados por la literatura que estudia los comentarios de los estudiantes, se confirma que las docentes mujeres suelen enfrentar una mayor carga de valoraciones negativas y una menor proporción de comentarios positivos en comparación con los hombres. Zheng et al. (2023) confirman que las mujeres reciben un porcentaje significativamente superior de comentarios negativos en todos los campos académicos, existe evidencia de que los estudiantes son más generosos al calificar a los hombres, estos concentran más comentarios positivos y mixtos, incluso cuando su desempeño es cuestionado en el texto; y los comentarios dirigidos a las profesoras tienden a ser de una polaridad sentimental más alta y extrema (más negativos en las críticas y más positivos en los elogios), mientras que las valoraciones para los hombres suelen ser más neutrales o estables.

Por su parte, Binderkrantz y Bisgaard (2024) señalan que los estudiantes hombres son más propensos a emitir comentarios negativos hacia las docentes mujeres, lo que contribuye a reducir su proporción global de positividad en comparación con los docentes varones. Así García (2000) identifica patrones diferenciados por área disciplinar en la evaluación docente. En ciencias sociales los profesores superan a las profesoras en entusiasmo, estimulación y retroalimentación. En humanidades la tendencia se invierte en la dimensión de retroalimentación, donde las mujeres reciben mejores calificaciones. En ciencias, aunque la evaluación global favorece a los profesores, no se evidencia que el género del estudiante modere dicha valoración.

Sin embargo, un hallazgo divergente se presenta en el estudio de Lakeman et al. (2022), el cual no identificó una relación dependiente entre el género del docente y la recepción de comentarios ofensivos o angustiantes en su muestra específica, sugiriendo que en ciertos contextos de *feedback* abusivo, la agresión puede ser generalizada, aunque las mujeres reporten críticas más específicas sobre su apariencia o vestimenta.

Sesgos de género en la evaluación docente: aportes de la teoría de la congruencia de rol y los estándares cambiantes

La teoría de los estándares cambiantes (*shifting standards* en su traducción en inglés) ofrece un sustento relevante para comprender cómo operan los sesgos en la evaluación. Esta perspectiva plantea que los juicios hacia las personas no se realizan con base en criterios universales, sino en estándares que se ajustan implícitamente según el grupo social de pertenencia, como el género. En este sentido, Zheng et al. (2023) documentan que las evaluaciones en contextos educativos pueden estar mediadas por expectativas diferenciadas, lo que influye en la forma en que se interpretan y ponderan los comentarios estudiantiles.

Los trabajos de Biernat (2003) profundizan en esta lógica al señalar que los estándares de evaluación pueden ser más flexibles o estrictos dependiendo del grupo evaluado, generando juicios que aparentan ser equivalentes pero que en realidad reflejan criterios distintos. Estos

planteamientos permiten interpretar que el impacto de los comentarios negativos puede ser mayor en las docentes debido a la aplicación de estándares más exigentes. Así, esta teoría contribuye a explicar la presencia de sesgos sutiles en la evaluación docente, los cuales no siempre son evidentes en los análisis cuantitativos tradicionales.

Por otro lado, la teoría de la congruencia de rol propuesta por Eagly y Karau (2002) explica el origen de los prejuicios hacia las mujeres en posiciones de liderazgo a partir de la discrepancia entre los estereotipos de género y las expectativas asociadas a dichos roles. Según este enfoque, los roles femeninos se asocian con características comunales (empatía, apoyo), mientras que los roles de autoridad —como la docencia— implican atributos de agencia como la dirección y el control. Esta incongruencia puede traducirse en dos formas de sesgo: percibir a las mujeres como menos adecuadas para ciertos roles o evaluar más críticamente su desempeño cuando los ejercen.

En el ámbito de la evaluación docente, diversos estudios han documentado patrones consistentes con esta teoría; por ejemplo, Boring (2017) y MacNell et al. (2015) evidencian que las evaluaciones estudiantiles tienden a estar influidas por el género del profesorado, afectando las calificaciones y percepciones de efectividad. Es decir, las valoraciones no son completamente objetivas, sino que están mediadas por expectativas sociales que pueden derivar en evaluaciones más exigentes o desfavorables hacia las mujeres, incluso en contextos donde el desempeño es comparable.

Metodología

El estudio se enmarca en un enfoque cuantitativo, con un diseño no experimental de alcance correlacional y corte transversal (Hernández-Sampieri & Mendoza, 2018). Se analizó la relación entre el porcentaje de comentarios negativos emitidos por estudiantes y la valoración global docente, estableciendo comparaciones por género. Dado que no se manipuló ninguna variable y los datos fueron recolectados en un único momento temporal, el diseño responde a las características de un estudio observacional transversal *ex post facto* (Kerlinger & Lee, 2002).

Muestra

Se incluyen cincuenta reportes de unidad de aprendizaje del profesorado provenientes del Sistema de Evaluación Docente (SED) del segundo periodo del año 2024. Como criterios de inclusión se consideraron aquellos reportes que contaran con al menos tres comentarios abiertos y con un mínimo de diez estudiantes participantes en la evaluación. La muestra presentó una distribución relativamente equilibrada entre asignaturas prácticas (n = 22; 44 %) y teóricas (n = 28; 56 %). Los reportes corresponden a tres unidades académicas del área de Odontología, con un total de cincuenta docentes, 1680 estudiantes inscritos, de los cuales 1346 participaron en la evaluación y 655 emitieron comentarios abiertos, esta diferencia se debe a que la sección cualitativa del instrumento es de carácter estrictamente voluntario,

por lo que no todos los estudiantes deciden expresarse por escrito (ver tabla 1). En conjunto, los datos permiten analizar la participación estudiantil tanto en la evaluación cuantitativa como en la emisión de comentarios abiertos, de tipo cualitativo.

Tabla 1

Características de los reportes por unidad de aprendizaje del docente

Unidad académica	Docentes	Estudiantes		
		Inscritos	Participan en la evaluación	Participan en comentarios
Facultad de Odontología, Mexicali	16	463	383	195
Facultad de Odontología, Tijuana	20	790	675	298
Centro de Ciencias de la Salud, Valle de las Palmas	14	427	288	162
Total	50	1680	1346	655

La tabla 2 muestra diferencias relevantes en la distribución del profesorado por unidad académica, género y nivel de estudios. En cuanto al género, en Mexicali predomina el profesorado femenino, mientras que en Tijuana y Valle de las Palmas predomina el masculino, en conjunto, existe ligera mayoría de hombres. Respecto al nivel de estudios, la especialidad es el grado más frecuente en las tres sedes, seguida de licenciatura.
La maestría se concentra principalmente en Tijuana, y el doctorado se mantiene constante en las tres unidades académicas.

Tabla 2

Características de los docentes

Variable		Unidad académica de docentes		
		Facultad de Odontología, Mexicali	Facultad de Odontología, Tijuana	Centro de Ciencias de la Salud, Valle de las Palmas
Género	Femenino	11	6	6
	Masculino	5	14	8
Total		16	20	14
Nivel de estudios	Licenciatura	4	4	4
	Especialidad	6	6	5
	Maestría	3	7	2
	Doctorado	3	3	3
Total		16	20	14

Instrumentos y materiales

Los datos del estudio provienen del Sistema de Evaluación Docente (SED) de la Universidad Autónoma de Baja California (UABC). Este sistema, implementado en 1988 y automatizado en el año 2000, opera al cierre de cada semestre, y sus resultados impactan tanto en los programas de estímulos académicos como en los indicadores de acreditación institucional (Arámburo & Luna, 2013).

La información contenida en los reportes por unidad de aprendizaje del SED constituyó el insumo principal para el análisis. Estos reportes incluyen: datos de identificación de la asignatura (carrera, unidad de aprendizaje y tipo de clase), del profesorado (género y grado académico) y de los estudiantes (número total de encuestados y tamaño del grupo). Así mismo integran la evaluación sumativa, que comprende la valoración global del docente y la calificación general del cuestionario.

En cuanto al Cuestionario de Evaluación Docente con base en la Opinión de los Alumnos que cuenta con evidencias de validez y confiabilidad (Reyes et al., 2020) se reportan, en primer lugar, los resultados por dimensión (planeación y gestión del proceso de enseñanza-aprendizaje y conducción y valoración del proceso de enseñanza-aprendizaje), mediante frecuencias y porcentajes; y, en segundo lugar, los resultados por reactivo, también expresados en frecuencias y porcentajes. Finalmente, en un listado se enumeran los comentarios abiertos que de manera opcional los estudiantes tienen para expresar algún aspecto de la práctica de su docente.

Procedimiento y análisis de datos

Los reportes por unidad de aprendizaje del profesorado provienen del SED en opinión de los alumnos. La captura de la información se realizó por docente en el programa SPSS. Para la categorización de los comentarios abiertos emitidos por los estudiantes se empleó una estrategia de análisis de contenido cualitativo basada en un enfoque convencional, en la cual las categorías emergen directamente del análisis de los datos (Hsieh & Shannon, 2005). Esta estrategia implicó la definición de categorías en comentarios positivos, negativos, neutros y mixtos, los cuales fueron establecidos a partir de su revisión por parte de dos evaluadores. Estos realizaron una primera lectura individual de un subconjunto de datos para familiarizarse con el contenido, seguida de acuerdos sobre los criterios para cada categoría:

- Positivos: referencias que expresan satisfacción, apoyo, entusiasmo o sentimientos favorables hacia el docente o la materia (ejemplo: Me gustaban sus clases, son muy tranquilas y se aprende mucho. Además de que evalúa de manera justa, con empatía y explica muy bien [comentario 539]).
- Negativos: comentarios que reflejan críticas, aversión, insatisfacción emocional o en la enseñanza (ejemplo: No le gusta aclarar dudas, cuando uno se acerca a preguntarle algo siempre se escuda en que eso ya deberíamos saberlo, y cuando le pedimos ayuda para revisar a

un paciente siempre va con la mentalidad de regañarnos, en lugar de explicarnos bien la técnica adecuada [comentario 190]).

- Neutros: declaraciones que carecen de una carga afectiva clara, respuestas indecisas o puramente informativas que no expresan ni aprobación ni desaprobación (ejemplo: La clase fue a las ocho de la mañana [comentario 27]).
- Mixtos: Textos que integran tanto elementos positivos como negativos (ejemplo: Buena maestra capacitada e informada pero algo rígida con sus métodos y su manera de enseñar [comentario 409]).

Posteriormente, con el fin de garantizar la confiabilidad del proceso, se llevó a cabo un proceso de calibración en el que ambos evaluadores revisaron conjuntamente un mismo conjunto de comentarios, discutieron las discrepancias y alcanzaron consenso sobre los criterios de clasificación.

Este método de revisión y discusión en equipo, conocido como codificación con acuerdo, refuerza la credibilidad y confiabilidad del análisis, aspectos esenciales en los enfoques cualitativos convencionales que se alinean con las recomendaciones de Hsieh y Shannon (2005) para asegurar la confiabilidad del proceso analítico. Se preservó el anonimato, dado que los reportes no contenían nombres del profesorado ni información que permitiera identificar a los estudiantes.

Para llevar a cabo un análisis de datos se combinaron estrategias descriptivas y correlacionales, con un enfoque particular en la comparación por género del docente, se organizaron en tres fases:

1. Se calcularon medidas de tendencia central (medias, desviaciones estándar, frecuencias) para caracterizar la distribución de los 655 comentarios emitidos por los estudiantes, clasificados en cuatro categorías: positivos, negativos, neutros y mixtos.
2. Se estimaron las medias de porcentajes de comentarios (positivos, negativos, neutros y mixtos) desagregadas por grupo.
3. Se examinó la relación entre la evaluación cualitativa (porcentaje de comentarios negativos) y la evaluación cuantitativa (valoración global asignada por los estudiantes), para ello se evaluó el supuesto de normalidad de las variables mediante la prueba de Shapiro-Wilk, adecuada para muestras menores a cincuenta casos. Los resultados indicaron que las variables no seguían una distribución normal ($p < 0.05$). En función de lo anterior, para comparar el número de comentarios negativos según el género del profesorado, se utilizó la prueba no paramétrica U de Mann-Whitney, además se realizaron análisis de correlación de Pearson de manera diferenciada por género, con el propósito de examinar la relación entre el porcentaje de comentarios negativos y la valoración global docente, para estimar la fuerza y dirección de la asociación lineal entre ambas variables dentro de cada grupo

Resultados

Se procedió a caracterizar la distribución de los 655 comentarios emitidos por los estudiantes, clasificados en cuatro categorías: positivos, negativos, neutros y mixtos. Para ello, se calcularon medidas de tendencia central para el porcentaje de comentarios recibidos por cada docente en cada categoría. Cabe destacar que ocho de cada diez estudiantes participaron en la evaluación docente, de los cuales una proporción menor decidió emitir comentarios abiertos a sus docentes.

En términos generales, el porcentaje total de comentarios recibidos por docente varió considerablemente, con un mínimo de 18 % y un máximo de 86 % (M = 50.4 %). La tabla 3 presenta las medidas de tendencia central para cada categoría. Los resultados muestran que los comentarios positivos constituyen la categoría con mayor presencia (M = 63.5 %, Mediana = 75.0 %), aunque con una alta variabilidad (DE = 35.2). Por su parte, los comentarios negativos presentan una media de 25.3 % y una mediana de 15.5 %, evidenciando una distribución asimétrica con tendencia hacia valores bajos, aunque con un rango que alcanza el 100 % en algunos casos. Las categorías neutros y mixtos registran las proporciones más bajas. Los comentarios neutros presentan una media de 3.2 % y una mediana de 0 %, indicando que más de la mitad de los docentes no recibieron comentarios de este tipo. Los comentarios mixtos alcanzan una media de 8.0 % (Mediana = 2.0 %), con una variabilidad moderada (DE = 13.3).

Tabla 3

Medidas de tendencia central de los comentarios emitidos por estudiantes

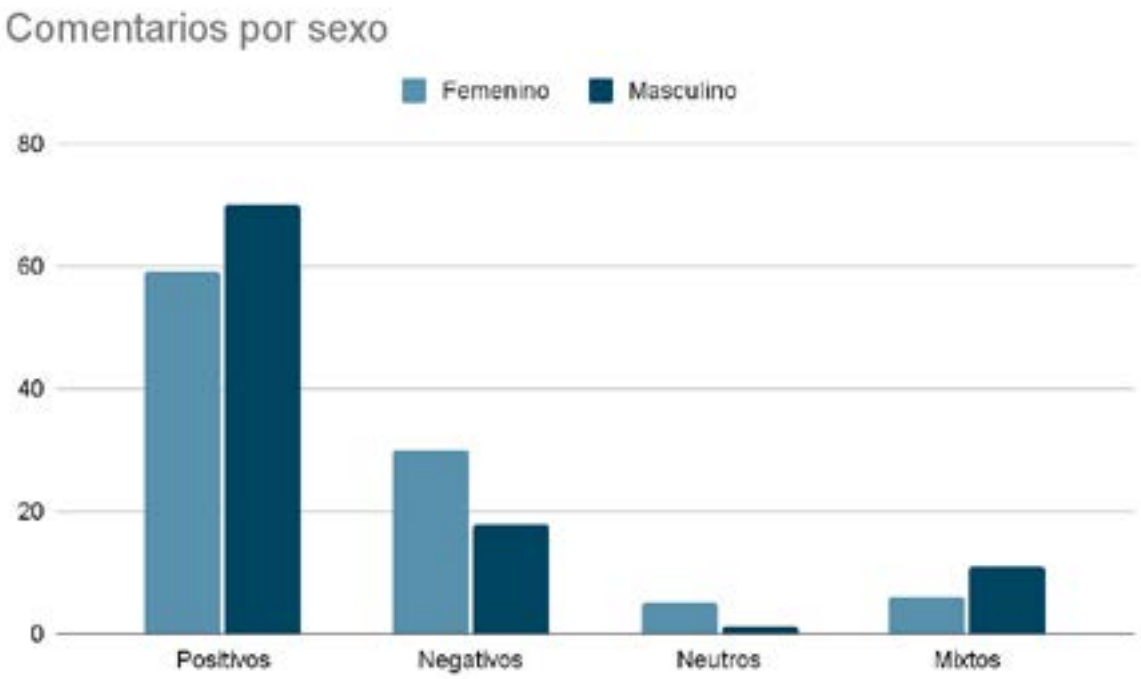
Porcentaje de comentarios	Rango		Media	Mediana	Desviación estándar
	Mínimo	Máximo			
Positivos	0	100	63.5	75.0	35.2
Negativos	0	100	25.3	15.5	30.2
Neutros	0	33	3.2	.0	6.4
Mixtos	0	66	8	2.0	13.3

Nota. Los valores corresponden a porcentajes de comentarios por docente en cada categoría.

En la figura 1 se muestran diferencias descriptivas por género, las docentes mujeres presentan menor proporción de comentarios positivos (59 %) en contraste con los hombres (70 %) y mayor de negativos (30 % y 18 % respectivamente), además de más comentarios neutros. En contraste, los hombres concentran más comentarios positivos y mixtos.

Figura 1

Media de porcentajes de comentarios por género



Con el propósito de analizar las diferencias en el número de comentarios negativos, según el género del profesorado, se aplicó la prueba U de Mann-Whitney. Los resultados indicaron que no existen diferencias estadísticamente significativas entre hombres (n = 18; rango promedio = 22.64) y mujeres (n = 31; rango promedio = 26.37) en el número de comentarios negativos (U = 236.50, Z = -0.906, p = 0.365). Aunque las docentes mujeres presentaron un rango promedio ligeramente mayor, esta diferencia no fue significativa y mostró un tamaño de efecto pequeño (r = -0.13).

Tabla 4

Comparación del número de comentarios negativos según género del profesorado

Género	*n*	Rango promedio	U de Mann–Whitney	Z	*p*	*r*
Hombre	18	22.64				
Mujer	31	26.37	236.50	−0.906	.365	−.13

Nota. *r* = tamaño de efecto calculado como Z/√N.

En el caso del análisis de correlación de Pearson para examinar la relación entre el porcentaje de comentarios negativos y la valoración global en el subgrupo de docentes mujeres (n = 31), los resultados mostraron una correlación negativa fuerte y estadísticamente significativa (r = 0.921, p < 0.001), indicando que, a

mayor proporción de comentarios negativos, menor es la valoración numérica asignada por los estudiantes.

Tabla 5

Correlación de Pearson entre comentarios negativos y valoración global, por género del docente

Género	*n*	*r* de Pearson	r^2	*p*	IC 95 %
Mujeres	31	- .921	.848	< .001	[-0.96, -0.85]
Hombres	19	- .774	.599	< .001	[-0.90, -0.55]

Nota. IC = Intervalo de confianza bootstrap para la correlación.

Estos resultados indican que, si bien no hay diferencias en la frecuencia de comentarios negativos, su impacto sobre la valoración global es considerablemente mayor en las docentes mujeres. En este sentido, los hallazgos sugieren un efecto diferencial asociado al género del profesorado, donde los comentarios negativos adquieren un peso mayor en la evaluación de las mujeres.

Finalmente, se incorporan algunas expresiones del alumnado para robustecer la comprensión de los aspectos que subyacen en la evaluación docente, las cuales revelan un marcado contraste: mientras que las quejas hacia los profesores hombres se limitan a aspectos operativos y de planeación didáctica, las dirigidas a las docentes mujeres trascienden lo pedagógico para evaluar, de manera punitiva, su personalidad, temperamento e interacción socioafectiva. La tabla 5 clasifica estos núcleos temáticos y recupera ejemplos textuales representativos según su registro original.

Tabla 5

Clasificación de comentarios negativos

Núcleo temático	Comentarios dirigidos a docentes mujeres	Comentarios dirigidos a docentes hombres
Interacción y clima de aula	Ataques a la personalidad y hostilidad. Son descritas como déspotas, groseras y prepotentes, se denuncia humillación pública. Ej.: «Fue muy prepotente tanto con alumnos y pacientes, solo juzgaba con comentarios groseros» [comentario 79].	Desviación del tiempo. Se critica que se distraen conversando o usando tecnicismos, sin reportar hostilidad verbal. Ej.: «Se desvía mucho del tema por querer ser buena onda, pero en realidad solo nos atrasamos» [comentario 499].
Gestión didáctica y evaluación	Rigidez y opacidad. Se critica la severidad injustificada al calificar y el infundir miedo para resolver dudas. Ej.: «Baja demasiado la calificación... hace que le tengas miedo, a veces trata de humillarte» [comentario 289].	Abandono y falta de rigor, las quejas por delegar la clase a exposiciones de alumnos, videos o pasantes. Ej.: «Prácticamente nosotros damos las clases... no nos resolvía dudas, la mesa clínica la dio un pasante». [comentario 550].
Cumplimiento operativo	Impuntualidad. Constantes quejas por llegar tarde a las clínicas, restando tiempo de práctica clínica. Ej.: «Siempre estuvo llegando tarde a la clínica... de 20 a 40 minutos tarde, lo cual nos quita tiempo» [comentario 367].	Ausentismo e improvisación. Señalan ausencia de las aulas, falta de planeación y cancelación de laboratorios. Ej.: «Siempre desaparece a mitad del semestre y deja de dar las clases y sube tareas sin haber explicado» [comentario 20].

Discusión

Los resultados del presente estudio muestran una tasa de participación estudiantil del 80 % en el SED, con un porcentaje total de comentarios recibidos por docente que varió considerablemente entre un mínimo de 18 % y un máximo de 86 % (M = 50.4 %). Este hallazgo resulta congruente con lo reportado por Padgett (2021), quien documentó niveles particularmente altos de participación en cursos de enfermería comunitaria (94 % - 98 %), así como con Brockx et al. (2012) y Zimmaro et al. (2006), quienes encontraron que entre el 67 % y el 70 % de los estudiantes emiten comentarios escritos.

La alta tasa de participación (80 %) se ubica en el rango superior de lo reportado por la literatura, lo que podría explicarse por el contexto disciplinar de Odontología, caracterizado por una alta interacción docente-estudiante en entornos clínicos, tal como sugieren Zimmaro et al. (2006) al señalar que la probabilidad de emitir comentarios es mayor en áreas con mayor interacción. Así mismo, la amplia variabilidad encontrada en el porcentaje de comentarios recibidos por docente (18 % - 86 %) es consistente con la literatura que señala que la participación no es homogénea entre docentes, pudiendo estar influida por factores como el tamaño del grupo, la disciplina y las características del curso (Vargas-Madriz & Nocente, 2023; Zimmaro et al., 2006).

Los resultados descriptivos evidencian que las docentes mujeres reciben una menor proporción de comentarios positivos y una mayor de negativos en comparación con sus pares hombres, lo cual es consistente con la literatura especializada. En particular, Zheng et al. (2023) documentan que las mujeres tienden a recibir más comentarios negativos y menos positivos en distintos campos académicos, además de que sus evaluaciones suelen presentar mayor polarización, con críticas más severas y elogios más intensos. El análisis de la relación entre los comentarios abiertos estudiantiles y la valoración global docente en el contexto de la formación en Odontología, incorporando la comparación por género del profesorado, permitió conocer que si bien no se identificaron diferencias estadísticamente significativas en la cantidad de comentarios negativos entre hombres y mujeres, los análisis evidenciaron un patrón diferencial en su impacto: los comentarios negativos se asociaron de manera más intensa con la valoración global en las docentes mujeres que en los hombres. Estos hallazgos confirman la relevancia de considerar no solo la frecuencia, sino también el peso que adquieren los comentarios en la evaluación docente, particularmente desde una perspectiva de género.

Estas asimetrías reflejan la brecha de género persistente en la academia, donde las mujeres enfrentan obstáculos estructurales, jerarquías masculinizadas y sesgos en la evaluación que limitan su promoción y visibilidad (Sigurdardottir et al., 2023; Weber et al., 2025). En Odontología, esto se traduce en menor representación en cargos de liderazgo y en un impacto diferencial en evaluaciones, reforzando el techo de cristal y perpetuando desigualdades, especialmente en áreas con dinámicas de poder históricamente masculinas y en contextos clínicos que afectan la percepción del desempeño profesional femenino

Los hallazgos del presente estudio son consistentes con la teoría de los estándares cambiantes *(shifting standards)*, la cual plantea que las evaluaciones no se realizan bajo criterios uniformes, sino que se ajustan en función del grupo social del evaluado. En esta línea, las investigaciones de Zheng et al. (2023) evidencian que el género influye en la forma en que se interpretan y ponderan los juicios evaluativos, activando expectativas diferenciadas hacia hombres y mujeres. En congruencia con estos planteamientos los resultados obtenidos muestran que, aunque no existen diferencias significativas en la cantidad de comentarios negativos recibidos, estos tienen un impacto considerablemente mayor en la valoración global de las docentes mujeres. Este patrón sugiere que los mismos indicadores cualitativos no son valorados de manera equivalente, sino que adquieren un peso diferencial dependiendo del género del profesorado, lo que refuerza la hipótesis de la presencia de sesgos implícitos en la evaluación estudiantil.

Los resultados del presente estudio también son congruentes con la teoría de la congruencia de rol propuesta por Eagly y Karau (2002), en tanto evidencian que las evaluaciones hacia las docentes mujeres pueden estar mediadas por expectativas de género que influyen en la forma en que se juzga su desempeño. Aunque no se encontraron diferencias en la cantidad

de comentarios negativos, estos tuvieron un impacto significativamente mayor en la valoración global de las mujeres, lo que sugiere una evaluación más exigente o penalizadora ante posibles incongruencias con el rol esperado.

La persistencia del sesgo de género que enfrentan las docentes mujeres en ámbitos académicos continúa afectando no solo su percepción profesional, sino también los procesos de evaluación en los que participan. A pesar de los avances, las expectativas sociales y los estereotipos de género suelen traducirse en una mayor exigencia y una menor tolerancia a las críticas dirigidas a las mujeres, lo que se refleja en la forma en que sus evaluaciones son percibidas y ponderadas. Este estigma puede generar efectos negativos en su carrera profesional, limitando oportunidades de promoción y reconocimiento, y alimentando desigualdades estructurales que permanecen arraigadas en el sistema educativo.

Este hallazgo se alinea también con lo reportado por Gelber et al. (2022), quienes documentan la persistencia de sesgos de género en contextos educativos, donde las mujeres suelen ser evaluadas bajo estándares más estrictos. Estos autores señalan que los estudiantes suelen utilizar criterios diferentes para juzgar a hombres y mujeres: mientras a ellos se les premia por su experiencia técnica o humor, a ellas se las penaliza más severamente por aspectos interpersonales o fallos percibidos en la organización. Estos marcos teóricos y empíricos respaldan la interpretación de que las diferencias observadas no radican en la frecuencia de los juicios negativos, sino en el peso diferencial que estos adquieren en función del género del profesorado.

Lo anterior puede traducirse en consecuencias profesionales, considerando que estas valoraciones suelen emplearse en decisiones de permanencia, promoción y asignación de incentivos; el efecto diferencial identificado contribuye a profundizar desigualdades de género en el ámbito académico, al penalizar de manera más severa a las mujeres bajo criterios que no operan de forma equivalente entre grupos (Vargas-Madriz & Nocente, 2023).

Se subraya que no basta con contar cuántos comentarios negativos recibe un docente, es imperativo considerar el peso y la intensidad que el estudiantado asigna a esos comentarios, ya que el género opera como un filtro que hace que la misma crítica tenga un impacto mucho mayor en la reputación y calificación global de una mujer que en la de un hombre

Entre las limitaciones del estudio se encuentra que se desarrolló en una institución y disciplina específicas, por lo que los resultados deben interpretarse con cautela al considerar su generalización a otros contextos. Así mismo, el uso de datos agregados por docente ofrece una visión global del fenómeno, aunque limita el análisis de la variabilidad a nivel individual. Por otro lado, si bien la categorización de los comentarios se realizó mediante consenso entre evaluadores, no se descarta la posibilidad de matices interpretativos propios del análisis cualitativo. Se recomienda ampliar el análisis a otros contextos y disciplinas, emplear diseños multinivel

con datos individuales, así como la incorporación de algunas variables contextuales puede ser relevante en futuras investigaciones para una comprensión más integral del fenómeno.

Finalmente, con este estudio se busca poner acento en la necesidad de transitar a una cultura de la evaluación que privilegie la evaluación formativa y el diálogo reflexivo, proporcionando evidencias que permitan transformar las prácticas institucionales hacia modelos más equitativos y centrados en la calidad real del aprendizaje. Al visibilizar cómo el género actúa como un filtro en la percepción estudiantil, este estudio justifica la relevancia de diseñar sistemas evaluativos integrales que protejan la integridad del profesorado y fomenten el desarrollo profesional continuo en las facultades de Odontología, así también invita a las instituciones a complementar los sistemas cuantitativos con análisis cualitativos sistemáticos que incorporen la perspectiva de género, fortaleciendo una cultura evaluativa que no solo mida, sino que comprenda y apoye el desarrollo docente en su diversidad.

Referencias

Arámburo, V., & Luna, E. (2013). La influencia de las características, profesor y del curso en los puntajes de evaluación docente. *Revista Mexicana de Investigación Educativa, 18*(58), 949-968. https://ojs.rmie.mx/index.php/rmie/article/view/780

Biernat, M. (2003). Toward a Broader View of Social Stereo Typing. *American Psychologist, 58*(12), 1019-1027. https://doi.org/10.1037/0003-066X.58.12.1019

Binderkrantz, A. S., & Bisgaard, M. (2024). A gender affinity effect: the role of gender in teaching evaluations at a Danish university. *High Education 87*, 591-610. https://doi.org/10.1007/s10734-023-01025-9

Boring, A. (2017). Gender biases in student evaluations of teaching. *Journal of Public Economics, 145*, 27-41. https://doi.org/10.1016/j.jpubeco.2016.11.006

Brockx, V., & Mortelmans, D. (2012). The Student as a Commentator: Students' Comments in Student Evaluations of Teaching. *Procedia - Social and Behavioral Sciences, 69*, 1122-1133. https://doi.org/10.1016/j.sbspro.2012.12.042

Eagly, A. H., & Karau, S. J. (2002). Role congruity theory of prejudice toward female leaders. *Psychological Review, 109*(3), 573-598. https://doi.org/10.1037/0033-295X.109.3.573

García, J. M. (2000). ¿Qué factores extraclase o sesgos afectan la evaluación docente en la educación superior? *Revista Mexicana de Investigación Educativa, 5*(10), 303-325. https://ojs.rmie.mx/index.php/rmie/article/view/1576

Gelber, K., Brennan, K., Duriesmith, D., & Fenton, E. (2022). Gendered mundanities: gender bias in student evaluations of teaching in political science. *Australian Journal of Political Science, 57*(2), 199-220. https://doi.org/10.1080/10361146.2022.2043241

Hernández-Sampieri, R., & Mendoza, C. (2018). *Metodología de la investigación: Rutas cuantitativa, cualitativa y mixta.* McGraw-Hill.

Hernández-Villafaña, A. D., & Luna, E. (2023). Metaevaluación del sistema de evaluación docente en una universidad pública mexicana. *Archivos Analíticos de Políticas Educativas, 31*(11). https://doi.org/10.14507/epaa.31.7501

Hsieh, H., & Shannon, S. (2005). Three Approaches to Qualitative Content Analysis. *Qualitative Health Research, 15*(9), 1277-1288. https://doi.org/10.1177/1049732305276687

Kerlinger, F., & Lee, H. (2002). *Investigación del comportamiento: Métodos de investigación en ciencias sociales.* McGraw-Hill.

Lakeman, R., Coutts, R., Hutchinson, M., Lee, M., Massey, D., Nasrawi, D., & Fielden, J. (2022). Appearance, insults, allegations, blame and threats: an analysis of anonymous non-constructive student evaluation of teaching in Australia. *Assessment & Evaluation in Higher Education, 47*(8), 1245-1258. https://doi.org/10.1080/02602938.2021.2012643

MacNell, L., Driscoll, A., & Hunt, A. N. (2015). What's in a name: Exposing gender bias in student ratings of teaching. *Innovative Higher Education, 40*(4), 291-303. https://doi.org/10.1007/s10755-014-9313-4

Madhu N., Kohli D., & Sangalli, L. (2026). Gender and Leadership Trajectories in Academic Dentistry. *European Journal of Dental Education, 0*, 1-11. https://doi.org/10.1111/eje.70109

Ochoa, L., & Moya, C. (2019). La evaluación docente universitaria: retos y posibilidades. *Folios, 49*, 41-60. https://doi.org/10.17227/folios.49-9390

Padgett, S. (2021). «He just teaches whatever he thinks is important»: Analysis of comments in student evaluations of teaching. *Nursing Inquiry, 28*, e12411, 1-11. https://doi.org/10.1111/nin.12411

Reyes, E., Luna, E., & Caso, J. (2020). Evidencias de validez del Cuestionario de Evaluación de la Competencia Docente Universitaria. *Perfiles Educativos, 42*(169), 106-122. https://doi.org/10.22201/iisue.24486167e.2020.169.58931

Ripoll-Nuñez, K., & Arrieta-Caycedo, C. A. (2022). ¿Qué es un docente de calidad? Perspectivas de docentes y estudiantes de una institución de educación superior en Colombia. *Revista Colombiana de Educación*, (85), 9-29. https://doi.org/10.17227/rce.num85-11690

Sigurdardottir, M., Rafnsdottir, G., Jónsdóttir, A., & Kristofersson, D. (2023). Student evaluation of teaching: gender bias in a country at the forefront of gender equality. *Higher Education Research & Development, 42*(4), 954-967. https://doi.org/10.1080/07294360.2022.2087604

Vargas-Madriz, L., & Nocente, N. (2023). Exploring students' willingness to provide feedback: A mixed methods research on end-of-term student evaluations of teaching. *Social Sciences & Humanities Open, 8*(1). https://doi.org/10.1016/j.ssaho.2023.100525

Webber, S., Babal, J., Allen A., Chen, L., Nackers, K., Allen, B., Jewell, D., Kieren, M., & Kelly, M. (2025). Unveiling Gender Bias: An Update on Faculty Teaching Evaluations in Academic Medicine. *Academic Pediatrics, 25*(1), 102562. https://doi.org/10.1016/j.acap.2024.08.007

Zheng, X., Vastrad, S., He, J., & Ni, C. (2023). Contextualizing gender disparities in online teaching evaluations for professors. *PLoS ONE, 18*(3): e0282704. https://doi.org/10.1371/journal.pone.0282704

Zimmaro, D., Gaede, C., Heikes, E., Shim, M., & Lewis, K. (2006). A Study of Students' Written Course Evaluation Comments at a Public University. Division of Instructional Innovation and Assessment. Universidad de Texas. https://testingservices.utexas.edu/sites/default/files/A-Study-of-Students-Written-Course-Evaluation-Comments-at-a-Public-University-2006.pdf