



**VOL. 29, Nº 2 (Julio, 2025)**

ISSN 1138-414X, ISSNe 1989-6395

DOI: 10.30827/profesorado.v29i2.32229

Fecha de recepción: 15/12/2024

Fecha de aceptación: 07/05/2024

## **PROFESORES EN EVALUACIÓN: EL IMPACTO DEL GÉNERO EN LOS RESULTADOS DE LA EVALUACIÓN DOCENTE UNIVERSITARIA**

*Faculty members in evaluation: the impact of gender on the results of student evaluations of teaching*



**Rosario Undurraga,  
José Tomás Montalva &  
Emilio Guerrero Barrera**  
*Universidad Finis Terrae*

**E-mail:**

[mrundurraga@uft.cl](mailto:mrundurraga@uft.cl)

[jmontalva@uft.cl](mailto:jmontalva@uft.cl)

[eguerrero@uft.cl](mailto:eguerrero@uft.cl)

**ORCID:**

<https://orcid.org/0000-0001-5801-2125>

<https://orcid.org/0009-0009-5436-783X>

<https://orcid.org/0009-0004-3781-1234>

### **Resumen:**

La evaluación docente es la herramienta más utilizada para evaluar la calidad de la enseñanza y el desempeño docente en las universidades, con efectos tanto en la evaluación de la docencia como en las posibilidades de promoción y estabilidad laboral del profesorado. A pesar de su extenso uso, su credibilidad y objetividad han sido cuestionadas debido a las diversas fuentes de error identificadas. Este estudio cuantitativo analiza los resultados de la evaluación docente de una universidad chilena en su totalidad y en dos carreras típicamente feminizadas y masculinizadas (enfermería e ingeniería) durante el periodo 2022-2024. Se considera el sexo del estudiantado y del profesorado planteando ocho hipótesis. Los resultados muestran que las docentes mujeres son

mejor evaluadas que sus pares varones en el promedio de las cuatro dimensiones de la práctica docente, sin embargo, a nivel institucional, el estudiantado recomienda de igual manera a profesores varones que a profesoras mujeres. En contraste, en unidades generizadas, la evaluación de la práctica docente es consistente con la recomendación, siendo las docentes mujeres mejor evaluadas y más recomendadas que los docentes varones. Se concluye instando a reflexionar sobre los sesgos de género y sus consecuencias en la academia en pro de alcanzar una mayor equidad de género en las instituciones de educación superior.

*Palabras clave:* enseñanza superior; evaluación del profesor; evaluación docente; sesgo de género; universidad.

### Abstract:

Student evaluation of teaching is the most widely used tool to assess the quality of teaching and teaching performance in universities, with effects both on the evaluation of teaching and on the promotion and job stability possibilities of faculty members. Despite its extensive use, its credibility and objectivity have been questioned due to various sources of error detected. This quantitative study analyzes the results of teacher evaluations at a Chilean university, both institutionally and within two careers typically gendered (nursing and engineering) during the 2022-2024 period. The study considers the gender of both students and faculty, proposing eight hypotheses. The results show that female faculty members are rated higher than their male counterparts on average across all four dimensions of teaching practice. However, at the institutional level, students recommend male faculty just as frequently as female faculty. In contrast, within gendered units, the evaluation of teaching practices aligns with student recommendations, with female faculty being better evaluated and more frequently recommended than male faculty. The article concludes by urging reflection on gender biases and their impact on academia to achieve greater gender equity in higher education institutions.

*Key Words:* academic performance; gender bias; higher education; student evaluation of teaching; university.

## 1. Evaluación de la docencia

La evaluación de la docencia en las universidades es clave para garantizar la calidad del proceso educativo. A nivel internacional, la evaluación docente o *Student Evaluations of Teaching* (SETs) es la forma más común de medir la calidad de la enseñanza y el desempeño docente en las instituciones de educación superior (Cook et al., 2024; Gourley & Madonia, 2020; Peterson et al., 2019). A su vez, la evaluación docente tiene un impacto significativo en las posibilidades de promoción, carrera académica y estabilidad laboral del profesorado (Linse, 2017; Peterson et al., 2019).

Para asegurar un instrumento estandarizado y objetivo que mida la calidad de la enseñanza y el desempeño docente en las universidades, se han analizado los indicadores de la evaluación de la práctica docente en la educación superior (Fernández et al., 2021), una buena docencia universitaria (García et al., 2021), los sistemas de evaluación del profesorado universitario (Gómez Sánchez & Rumbo Arcas, 2023), así como las fortalezas y debilidades del modelo de evaluación del profesorado universitario según las entidades certificadoras (Aravena Gaete & Gairán Sallán, 2021), entre otros aspectos. Sin embargo, la credibilidad y validez de los resultados de la evaluación docente han sido cuestionadas porque se han encontrado

diversas fuentes de error (Quansah et al., 2024). Se ha documentado que la evaluación docente presenta sesgos, entre ellos, anomalías estadísticas, factores temporales y sesgos cognitivos que inciden en los resultados de la evaluación docente (Cook et al., 2024). Se ha mostrado que las características del estudiantado, como el género y la edad, afectan en la evaluación que realizan de sus profesores (Hejase et al., 2014; Worthington, 2002). Froment et al. (2021) advierte que la credibilidad del docente y la motivación del estudiantado universitario tienen un efecto positivo en la evaluación del desempeño docente. Además de las características del estudiantado, las del profesorado, como el género y la raza/etnia del docente, afectan las evaluaciones del estudiante sobre la evaluación docente (Wallace et al., 2018). Por ende, los sesgos de género parecieran tener un rol importante en los resultados de la evaluación docente en las instituciones de educación superior.

El sesgo de género se presenta cuando docentes hombres y mujeres reciben evaluaciones diferentes que no se explican por diferencias objetivas en la calidad de la enseñanza (Mengel et al., 2019). En el ámbito de las ciencias sociales, Felkey & Batz-Barbarich (2021) demostraron que los instrumentos de evaluación favorecen a los profesores hombres, particularmente, en economía. En los Países Bajos, Mengel et al. (2019) encontraron que las mujeres son evaluadas sistemáticamente peor que los hombres, especialmente por estudiantes hombres, quienes las califican un 21% peor en comparación con sus colegas varones, sesgo que se intensifica en cursos con contenidos de matemáticas y en docentes jóvenes. En Francia, Boring (2017) mostró que los estudiantes evalúan a las profesoras de acuerdo con estereotipos de género, favoreciendo a los docentes hombres sin evidencia de que sean mejores. En México, Arceo-Gomez & Campos-Vazquez (2019) observaron que los estudiantes tienden a enfocarse más en la apariencia y personalidad de las profesoras, usando términos asociados a cualidades de servicio, mientras que los términos masculinos impactan negativamente las evaluaciones de las profesoras mujeres. En Estados Unidos, Korte et al. (2013) mostraron diferencias sistemáticas en la percepción de estudiantes mujeres y hombres respecto a sus docentes y concluyeron que las mujeres tienden a calificar mejor a sus docentes en la calidad de la enseñanza. Sin embargo, un experimento realizado por Peterson et al. (2019) en la Universidad Estatal de Iowa, en el que se sensibilizó al estudiantado sobre los sesgos inconscientes, mostró una reducción significativa en el sesgo de género, mejorando las evaluaciones hacia las profesoras mujeres.

En Chile, a pesar de la relevancia de la evaluación docente para la evaluación de la enseñanza, la promoción y la estabilidad laboral del profesorado universitario como para el aseguramiento de la calidad y el proceso de acreditación institucional, no hay literatura abundante sobre sesgos en la evaluación docente en instituciones de educación superior. De los pocos estudios en la materia, Medel & Asún (2014) analizaron patrones de género en la evaluación del estudiantado universitario respecto a sus docentes en la Facultad de Ciencias Sociales de la Universidad de Chile, encontrando diferencias sustantivas en los resultados de los Cuestionarios de Evaluación Docente entre académicos hombres y mujeres. El estudiantado evalúa mejor a las académicas en responsabilidad y habilidades docentes, mientras que a los

académicos varones en dominio disciplinario. Las profesoras mujeres son significativamente mejor evaluadas por estudiantes de primer ciclo que por estudiantes de segundo ciclo de formación y son los estudiantes hombres quienes peor evalúan a sus profesoras. Recientemente, el estudio de Paredes et al. (2024) en la Facultad de Economía y Negocios de la Universidad de Chile sugiere que las características faciales, particularmente, la relación ancho-altura de la cara como indicador de la dominancia percibida, tiene impacto en la evaluación docente, especialmente en el caso de las mujeres. Evidencian un marcado sesgo de género, donde las mismas características que benefician a los profesores hombres son perjudiciales para las profesoras. Adicionalmente, existe literatura sobre la evaluación docente a nivel escolar en Chile (Fardella, 2012; Galaz-Ruiz & Anaya-Torres, 2023; Lara-Subiabre & Angulo-Brunet, 2020; Manzi et al., 2011), pero no hay mucha más evidencia sobre la evaluación docente y sesgos en otras universidades del país durante la última década.

Este artículo aporta al campo de conocimiento de la evaluación de la calidad docente en la educación superior y los posibles sesgos de género que pueden incidir en los resultados de la evaluación docente. Para ello, se analizan los resultados de la evaluación docente del periodo 2022-2024 en una universidad chilena de la región Metropolitana y examina las diferencias entre distintos aspectos evaluados. ¿En qué dimensiones son mejor evaluadas las profesoras universitarias mujeres y hombres? ¿Existen diferencias en los resultados de la evaluación docente según el sexo del estudiantado? ¿Cómo es la evaluación docente según carreras feminizadas y masculinizadas? ¿Pueden las diferencias en los resultados de la evaluación docente de profesores hombres y mujeres atribuirse a sesgos de género?

### **1.1. Evaluación docente en una universidad chilena**

En la Universidad Finis Terrae (UFT), la Encuesta de Evaluación Docente ha desempeñado un rol central en la recolección de percepciones sobre el desempeño docente por parte del estudiantado. Sin embargo, hasta el año 2022, el proceso de implementación de esta encuesta enfrentaba limitaciones que comprometían la calidad y utilidad de la información obtenida. La encuesta tenía una tasa de respuesta inferior al 20%, lo que contribuía al cuestionamiento de la validez y de la utilidad de los resultados obtenidos.

Para abordar estas deficiencias, la Universidad emprendió un proceso de rediseño del instrumento, inspirado en los modelos de evaluación docente utilizados en universidades de renombre internacional y nacional. Este rediseño incluyó la colaboración activa de representantes de cada facultad y la consulta directa al estudiantado, contribuyendo a que el nuevo instrumento respondiera a las necesidades específicas de la comunidad universitaria. Esta nueva versión de la encuesta de evaluación docente fue implementada por primera vez en el segundo semestre del año 2022, logrando avances significativos en términos de validez, confiabilidad y tasa de respuesta.

En coherencia con su importancia institucional, la evaluación docente en la Universidad Finis Terrae no solo cumple una función formativa, orientada a la mejora continua, sino que también tiene implicancias contractuales relevantes. De acuerdo con el reglamento académico vigente, los resultados obtenidos en este instrumento pueden incidir en procesos de calificación, promoción e incluso en la pérdida de la calidad de académico/a si las evaluaciones son consistentemente deficientes. Por ello, el instrumento se diseñó para equilibrar criterios de objetividad y pertinencia, resguardando tanto su utilidad para el desarrollo profesional como su validez en contextos decisionales de la gestión académica.

Por otro lado, uno de los cambios más relevantes corresponde a la aplicación de la encuesta durante el último mes de clases, asegurando que la evaluación de la docencia no se viera influenciada por las calificaciones finales de la asignatura. De hecho, los análisis estadísticos confirman que no existe correlación significativa entre el rendimiento académico de los estudiantes y la valoración de la docencia (Universidad Finis Terrae, 2023), lo que respalda la independencia de la evaluación docente con los resultados obtenidos en ese curso. El proceso, además, comprende la generación de reportes de forma inmediata, lo que permite al profesorado obtener una retroalimentación oportuna.

Para garantizar una alta tasa de participación, se implementaron incentivos que otorgan prioridad en la toma de ramos a los estudiantes que responden la encuesta. Estos estudiantes pueden inscribir sus cursos antes que aquellos que no completaron la totalidad de sus encuestas asignadas, lo que representa una ventaja importante para programar su agenda, horarios y elegir profesores. Así, aunque no existe una obligación formal de responder la evaluación docente, el incentivo fue clave para lograr tasas de acceso y respuesta superiores al 80%.

La escala del instrumento se ajustó para medir la frecuencia de prácticas docentes observadas, en lugar de basarse en criterios de acuerdo o de logro, lo que facilita la interpretación de los resultados y promueve una mejora continua en áreas específicas. La encuesta de evaluación docente incluye 25 preguntas agrupadas en seis dimensiones diseñadas para evaluar integralmente la percepción estudiantil sobre la calidad docente y la experiencia educativa. Dos dimensiones abordan aspectos generales de la asignatura y el/la docente (dimensión global) y el logro de resultados de aprendizaje (dimensión diseño y planificación). Las otras cuatro dimensiones evalúan la práctica docente basándose en buenas prácticas de enseñanza: organización de la enseñanza (preparación, dominio y planificación), metodologías de aprendizaje (estrategias activas y participativas), clima (respeto y apoyo en el aula) y evaluación y retroalimentación (claridad y utilidad de las evaluaciones). Además, la encuesta incluye una sección sobre ayudantías con dos preguntas específicas sobre los o las ayudantes de cátedra.

Este estudio considera las cuatro dimensiones relacionadas con la práctica docente y el ítem “¿Recomendarías a este profesor a un amigo o amiga?”, para analizar la percepción del estudiantado sobre la docencia y explorar posibles sesgos de género en las evaluaciones.

## 2. Metodología

Este estudio cuantitativo analiza los resultados de la evaluación docente de la Universidad Finis Terrae a partir del segundo semestre de 2022 hasta el primer semestre de 2024, periodo que coincide con la implementación de la nueva encuesta docente. La muestra incluye 146.925 encuestas realizadas por 10.751 estudiantes, abarcando 25 escuelas distribuidas en 9 facultades. Los datos corresponden a cuatro periodos de evaluación docente realizados en los semestres mencionados.

Para analizar los resultados de la evaluación docente y posibles sesgos de género, este artículo plantea ocho hipótesis a contrastar:

1. Las docentes mujeres son mejor evaluadas que sus pares hombres en la dimensión organización de la enseñanza.
2. Los docentes hombres son mejor evaluados que sus pares mujeres en el ítem dominio y preparación del contenido de la dimensión metodología.
3. Las docentes mujeres son mejor evaluadas que sus pares hombres en la dimensión clima.
4. Las docentes mujeres son mejor evaluadas que sus pares hombres en la dimensión evaluación y retroalimentación.
5. Las docentes mujeres son mejor evaluadas que sus pares hombres considerando el promedio de las cuatro dimensiones que evalúan la práctica docente.
6. Las docentes mujeres son mejor evaluadas que sus pares hombres en la pregunta “¿Recomendarías este docente a un amigo o amiga?”.
7. El estudiantado evalúa mejor a docentes de su mismo género a nivel universidad.
8. El estudiantado evalúa mejor a los docentes de su sexo opuesto en unidades típicamente generizadas (enfermería e ingeniería).

Para el análisis, se consolidaron los datos de evaluación docente de los cuatro procesos mencionados, integrando las 146.925 encuestas. Se generaron cinco bases de datos: cuatro correspondientes a cada dimensión evaluada (organización de la enseñanza, metodología, clima, y evaluación y retroalimentación), y una quinta para la pregunta “¿Recomendarías este docente a un amigo o amiga?”.

Se realizaron tres tipos de análisis. Primero, un análisis ANOVA para evaluar la varianza en las respuestas a las 17 preguntas de las dimensiones y de la pregunta referida a la recomendación docente, considerando nivel universidad (UFT), enfermería e ingeniería. A nivel institucional, el valor p fue significativo ( $p < 0.005$ ) en todas las preguntas con la excepción de P15 y P5. A nivel enfermería, el valor p

fue significativo en todas las preguntas con la excepción de P14. A nivel ingeniería, el valor p fue significativo en todas las preguntas con la excepción de P9 y P5.

Con estos resultados, se consideró apropiado realizar una prueba t de Welch aplicada a las 17 preguntas de las dimensiones y de la pregunta referida a la recomendación docente, para comparar las medias entre docentes hombres y mujeres, verificando la significancia estadística de las diferencias. Finalmente, se efectuó un análisis de regresión múltiple para determinar el impacto de las variables género docente y género estudiante en los puntajes obtenidos por los docentes.

Las preguntas asociadas a las cuatro dimensiones pedagógicas utilizan una escala Likert de cuatro niveles: (4) Siempre o casi siempre, (3) Muchas veces, (2) Pocas veces, (1) Nunca o casi nunca. Para la pregunta de recomendación de ese docente, las opciones fueron: (4) Me encantó, lo recomiendo, (3) Lo recomiendo, (2) No lo recomiendo, (1) Definitivamente no lo recomiendo. Los resultados se presentan siguiendo el orden de las hipótesis planteadas.

### 3. Resultados

El cuerpo docente evaluado se conforma de 742 hombres (53,4%) y 648 mujeres (46,6%) (tabla 1). A nivel universidad, el estudiantado que respondió la evaluación docente se compone de 6.448 mujeres (46,6%), 4.287 hombres (39,9%) y 16 estudiantes cuya información respecto de su género figura como “No Disponible” (0,1%) (tabla 2).<sup>1</sup>

En la carrera de enfermería, respecto al género del cuerpo docente evaluado, 25 docentes son mujeres (67,6%) y 12 profesores son hombres (32,4%) (tabla 1). El género del estudiantado de la carrera de enfermería que respondió la encuesta corresponde a 379 mujeres (80,8%) y 90 hombres (19,2%) (tabla 2).

En la escuela de ingeniería (ingeniería civil industrial, ingeniería civil informática y telecomunicaciones, ingeniería civil en realidad virtual y videojuegos), la distribución de género del cuerpo docente evaluado corresponde a 34 profesoras (35,1%) y 63 profesores (64,9%) (tabla 1). El género del estudiantado de la escuela de ingeniería que respondió la encuesta corresponde a 237 mujeres (25,7%), 682 hombres (74%) y 0,2% (2 casos) “No Disponible” (tabla 2).

Tabla 1  
*Distribución por género de docentes evaluados: UFT, enfermería e ingeniería*

UFT		Enfermería		Ingeniería	
N	%	N	%	N	%

<sup>1</sup> La variable género considera las opciones femenino, masculino y no disponible, según la información de registro del estudiante.

Docentes mujeres	648	46.6%	25	67.6%	34	35.1%
Docentes hombres	742	53.4%	12	32.4%	63	64.9%

Fuente: Elaboración propia.

Tabla 2

*Distribución por género de estudiantes que respondieron encuesta docente: UFT, enfermería e ingeniería*

Estudiantes	UFT		Enfermería		Ingeniería	
	N	%	N	%	N	%
Mujeres	6448	60.0%	379	80.8%	237	25.7%
Hombres	4287	39.9%	90	19.2%	682	74.0%
Género no disponible	16	0.1%	0	0.0%	2	0.2%

Fuente: Elaboración propia.

Las tablas 1 y 2 muestran que el cuerpo docente a nivel universidad es relativamente parejo entre hombres y mujeres (53,4% y 46,6%, respectivamente), mientras que la carrera de enfermería y la escuela de ingeniería están mayoritariamente feminizada y masculinizada, respectivamente. A continuación, se presenta el análisis de los resultados de la evaluación docente según el orden de las hipótesis planteadas.

### 3.1 Hipótesis 1: Las docentes mujeres son mejor evaluadas que sus pares hombres en la dimensión organización de la enseñanza.

El análisis muestra que, a nivel universidad, las docentes mujeres obtienen calificaciones superiores a los docentes hombres en la dimensión organización de la enseñanza (tabla 3).<sup>2</sup>

Tabla 3

*Evaluación dimensión organización de la enseñanza según género del docente, UFT*

Docentes	N	Organización de la enseñanza	
		Media	DE
Mujeres	648	3.59	.73
Hombres	742	3.56	.76

Fuente: Elaboración propia.

<sup>2</sup> Preguntas dimensión organización de la enseñanza: (P9) Explicitó el aporte de la asignatura al desarrollo profesional y perfil de egreso; (P10) Presentó los tópicos, estructura o los objetivos de aprendizaje en cada clase; (P11) Cumplió con los contenidos declarados en el programa de asignatura; (P12) Fue puntual en el inicio y término de cada sesión de clases; (P13) Utilizó bibliografía actualizada y/o pertinente a los contenidos del curso.

Tabla 4  
*t de Welch preguntas organización de la enseñanza*

Dimensión	UFT			Enfermería			Ingeniería		
	t	df	P	t	df	P	t	df	P
Organización de la enseñanza									
P9	-3.3	159348	.001*	-5.116	4130	.001*	-1.821	7521	.069
P10	-12.41	162225	.001*	-5.944	3995	.001*	-7.51	8059	.001*
P11	-10.65	162166	.001*	-1.525	4587	.127	-7.8	8262	.001*
P12	-5.122	160532	.001*	-3.01	4493	.001*	-7.052	8099	.001*
P13	-8.652	161562	.001*	-2.953	4250	.001*	-5.427	7821	.001*

\* P < .05

Fuente: Elaboración propia.

La prueba t de Welch indica que las diferencias en las medias no son producto del azar. Las cinco preguntas analizadas a nivel universidad mostraron diferencias significativas en cuanto a la media obtenida por docentes hombres y mujeres. El análisis confirma la hipótesis 1 que las docentes mujeres son mejor evaluadas que los docentes varones en la dimensión organización de la enseñanza (tabla 4).

### 3.2 Hipótesis 2: Los docentes hombres son mejor evaluados que sus pares mujeres en el ítem dominio y preparación del contenido.

En la dimensión metodología, se analizó la percepción del estudiantado respecto al dominio y preparación del contenido de su asignatura de docentes hombres y mujeres.<sup>3</sup>

Tabla 5  
*Evaluación pregunta dominio de contenido de la dimensión metodología según género del docente, UFT*

	Dominio contenido UFT			Dominio contenido Enfermería			Dominio contenido Ingeniería		
	N	Media	DE	N	Media	DE	N	Media	DE
Mujeres	648	3.68	.66	25	3.79	.54	34	3.59	.74
Hombres	742	3.69	.66	12	3.81	.50	63	3.52	.80

Fuente: Elaboración propia.

A partir de los resultados, a nivel universidad, se confirma la hipótesis de que los docentes hombres son mejor evaluados que sus pares mujeres en el dominio de

<sup>3</sup> El ítem analizado de la dimensión metodología fue: (P14) Demostró un dominio y preparación del contenido que imparte en su asignatura. No fueron consideradas las preguntas: Ofreció ejemplos clarificadores, aplicados al contexto profesional y/o cotidianidad; Promovió actividades para intencionar la interacción con y entre las y los estudiantes; Utilizó y aprovechó las distintas funciones de eFinis como apoyo a la realización de las clases; Promovió la resolución de problemas de manera crítica y autónoma por parte de las y los estudiantes.

contenidos. Esto sucede tanto a nivel institucional como en enfermería, no así en ingeniería, donde las docentes mujeres son mejor evaluadas que sus pares varones en el dominio y preparación del contenido (tabla 5). Esto llama la atención considerando que las disciplinas STEM son tradicionalmente masculinas.

Tabla 6  
*t de Welch pregunta dominio y preparación del contenido*

Dimensión	UFT			Enfermería			Ingeniería		
	t	df	P	t	df	P	t	df	P
Metodología									
P14	3.5479	158445	.001*	1.6759	4772	.094	-4.471	8066	.001*

\* P < .05

Fuente: Elaboración propia.

La prueba t de Welch para la pregunta 14 de la dimensión metodología mostró diferencias significativas en las medias entre docentes hombres y mujeres a nivel universidad. En enfermería no se encontraron diferencias significativas, mientras que en ingeniería las docentes mujeres fueron mejor evaluadas que los hombres (tabla 6). Estos resultados confirman parcialmente la hipótesis 2, indicando que los hombres son mejor evaluados en el dominio de contenido a nivel universidad, pero no en ingeniería, donde las mujeres muestran una puntuación más alta.

### 3.3 Hipótesis 3: Las docentes mujeres son mejor evaluadas que sus pares hombres en la dimensión clima.

Al analizar la dimensión clima del aula<sup>4</sup>, se observa que las docentes mujeres son mejor evaluadas que los profesores hombres tanto a nivel universitario como en unidades generizadas (tabla 7).

Tabla 7  
*Evaluación dimensión clima según género del docente UFT, enfermería e ingeniería*

	Dimensión clima UFT			Dimensión clima Enfermería			Dimensión clima Ingeniería		
	N	Media	DE	N	Media	DE	N	Media	DE
Docentes									
Mujeres	648	3.63	.71	25	3.75	.60	34	3.53	.80
Hombres	742	3.61	.73	12	3.66	.69	63	3.45	.85

Fuente: Elaboración propia.

<sup>4</sup> Preguntas dimensión clima: (P19) Propició un ambiente para que las y los estudiantes hicieran preguntas durante las clases; (P20) Realizó las clases en un ambiente de respeto y cordialidad; (P21) Se vio accesible para consultas y responder inquietudes (presencialmente, eFinis, correos, etc.).

Tabla 8  
*t de Welch dimensión clima*

Dimensión	UFT			Enfermería			Ingeniería		
	Clima	t	df	P	t	df	P	t	df
P19	-4.099	160341	.001*	-4.803	3924	.001*	-4.595	7823	.001*
P20	-5.403	61183	.001*	-3.501	4029	.001*	-4.965	7996	.001*
P21	-6.474	160682	.001*	-5.039	4115	.001*	-5.041	7868	.001*

\* P < .05

Fuente: Elaboración propia.

Los resultados de la prueba t de Welch para las tres preguntas de la dimensión clima mostraron diferencias significativas en cuanto a la media obtenida por docentes hombres y mujeres a nivel universidad, enfermería e ingeniería (tabla 8).

En síntesis, los resultados confirman la hipótesis 3: las docentes mujeres son mejor evaluadas que sus pares hombres en la dimensión clima.

### 3.4 Hipótesis 4: Las docentes mujeres son mejor evaluadas que sus pares hombres en la dimensión evaluación y retroalimentación.

Se analizó la dimensión evaluación y retroalimentación, que incluye preguntas sobre la entrega oportuna de las evaluaciones, criterios explícitos de evaluación, posibilidades de retroalimentación, así como formas variadas de evaluación.<sup>5</sup>

Tabla 9  
*Dimensión evaluación y retroalimentación según género del docente: UFT, enfermería e ingeniería*

	Dimensión evaluación y retroalimentación UFT			Dimensión evaluación y retroalimentación Enfermería			Dimensión evaluación y retroalimentación Ingeniería		
	N	Media	DE	N	Media	DE	N	Media	DE
Mujeres	648	3.49	.82	25	3.59	.74	34	3.41	.88
Hombres	742	3.44	.86	12	3.45	.83	63	3.30	.93

Fuente: Elaboración propia.

A partir de los resultados, se observa que las docentes mujeres son mejor evaluadas que sus pares masculinos en la dimensión evaluación y retroalimentación. Esto sucede tanto a nivel universidad como en unidades generizadas. La diferencia a favor de las profesoras en la dimensión evaluación y retroalimentación es particularmente mayor en enfermería (tabla 9).

<sup>5</sup> Preguntas dimensión evaluación y retroalimentación: (P22) Entregó oportunamente las calificaciones durante y hacia el cierre de la asignatura; (P23) Entregó retroalimentación (*feedback*) con respecto al logro de los aprendizajes de las y los estudiantes durante el desarrollo de las clases; (P24) Realizó evaluaciones con criterios explícitos y conocidos por las y los estudiantes; (P25) Utilizó variadas formas de evaluación a lo largo de la asignatura.

Tabla 10  
*t de Welch dimensión evaluación y retroalimentación*

Dimensión	UFT			Enfermería			Ingeniería		
	t	df	P	t	df	P	t	df	P
Evaluación y Retroalimentación									
P22	-8.675	161465	.001*	-4.95	4189	.001*	-4.191	7915	.001*
P23	-8.633	160964	.001*	-7.759	3982	.001*	-5.891	7888	.001*
P24	-6.774	160482	.001*	-4.353	4397	.001*	-6.231	7905	.001*
P25	-17.6	162643	.001*	-7.924	4122	.001*	-7.419	7925	.001*

\* P < .05

Fuente: Elaboración propia.

La prueba t de Welch mostró diferencias significativas en las medias obtenidas por docentes hombres y mujeres en las cuatro preguntas de la dimensión evaluación y retroalimentación a nivel universidad. Estas diferencias también fueron significativas en enfermería e ingeniería para todas las preguntas analizadas, lo que indica una percepción consistentemente favorable hacia las docentes mujeres en esta dimensión (tabla 10).

En síntesis, los resultados confirman la hipótesis 4: las profesoras mujeres son mejor evaluadas que los docentes hombres en la dimensión evaluación y retroalimentación.

### 3.5 Hipótesis 5: Las docentes mujeres son mejor evaluadas que sus pares hombres considerando el promedio de las cuatro dimensiones que evalúan la práctica docente.

La evaluación de las cuatro dimensiones de la práctica docente (organización de la enseñanza, metodología, clima, y evaluación y retroalimentación) es superior en las docentes mujeres que en los profesores varones.

Tabla 11  
*Evaluación por dimensión según género del docente a nivel universidad*

	N	Organización de la enseñanza		Metodología		Clima		Evaluación y Retroalimentación	
		Media	DE	Media	DE	Media	DE	Media	DE
Docentes									
Mujeres	648	3.59	.73	3.52	.81	3.63	.71	3.49	.82
Hombres	742	3.56	.76	3.49	.84	3.61	.73	3.44	.86

Fuente: Elaboración propia.

A nivel universidad, la diferencia a favor de las profesoras es similar en las cuatro dimensiones evaluadas (tabla 11).

Tabla 12

*Promedio evaluación según género del docente: universidad, enfermería e ingeniería*

Docentes	Promedio UFT			Promedio enfermería			Promedio ingeniería		
	N	Media	DE	N	Media	DE	N	Media	DE
Mujeres	648	3.56	.77	25	3.68	.65	34	3.52	.84
Hombres	742	3.53	.80	12	3.59	.73	63	3.43	.88

Fuente: Elaboración propia.

Al comparar el promedio de la evaluación docente de las cuatro dimensiones analizadas a nivel universidad, enfermería e ingeniería, es posible observar que en los tres niveles la evaluación de las docentes mujeres es superior a la de los docentes hombres (tabla 12).

Tabla 13

*Promedio evaluación cuatro dimensiones según género del estudiante: enfermería e ingeniería*

Docentes	Evaluación estudiantes mujeres enfermería		Evaluación estudiantes hombres enfermería		Evaluación estudiantes mujeres ingeniería		Evaluación estudiantes hombres ingeniería	
	Media	DE	Media	DE	Media	DE	Media	DE
Mujeres	3.69	.64	3.64	.70	3.49	.84	3.46	.83
Hombres	3.60	.72	3.55	.75	3.40	.88	3.37	.89

Fuente: Elaboración propia.

Los resultados muestran consistencia en las evaluaciones al considerar el género del estudiantado. Tanto hombres como mujeres en enfermería e ingeniería evalúan mejor a las docentes mujeres que a los hombres en el promedio de las cuatro dimensiones de la evaluación docente (tabla 13). Estos hallazgos coinciden con estudios previos, como el de Medel & Asún (2014), que señalan que los estudiantes hombres tienden a evaluar más bajo en general a sus docentes. En particular, los estudiantes hombres de ingeniería destacan como el grupo con las evaluaciones más bajas hacia docentes de ambos géneros.

En síntesis, se confirma la hipótesis 5: las docentes mujeres son mejor evaluadas que los hombres en el promedio de las cuatro dimensiones de la práctica docente, tanto a nivel universidad como en enfermería e ingeniería.

### **3.6 Hipótesis 6: Las docentes mujeres son mejor evaluadas que sus pares hombres en la pregunta “¿Recomendarías este docente a un amigo o amiga?”.**

Esta hipótesis analiza la recomendación del estudiantado acerca del docente en evaluación.

Tabla 14  
*Recomendación de ese docente por género del docente: universidad, enfermería e ingeniería*

Docentes	Recomienda docente UFT			Recomienda docente Enfermería			Recomienda docente Ingeniería		
	N	Media	DE	N	Media	DE	N	Media	DE
Mujeres	648	3.31	.84	25	3.42	.74	34	3.29	.90
Hombres	742	3.31	.84	12	3.34	.77	63	3.14	.92

Fuente: Elaboración propia.

A nivel universidad, a pesar de que el promedio de las dimensiones evaluadas es mejor para las docentes mujeres que para los docentes varones, al momento de recomendar un docente a un amigo o amiga, el estudiantado tiende a recomendar de igual manera a un docente varón que a una docente mujer. Sin embargo, en el caso de la carrera de enfermería, típicamente feminizada, el estudiantado recomienda más a las docentes mujeres que a los profesores hombres. De igual forma, a nivel ingeniería, facultad masculinizada, las profesoras mujeres son más recomendadas que los docentes hombres por sus estudiantes (tabla 14).

Tabla 15  
*t de Welch pregunta recomendación a docente*

Dimensión	UFT			Enfermería			Ingeniería		
	t	df	P	t	df	P	t	df	P
P5 Recomendación	0.7933	159291	.427	3.2218	5400	.001*	-0.398	5495	.691

\* P < .05

Fuente: Elaboración propia.

Los resultados de la prueba t de Welch para la pregunta “¿Recomendarías este docente a un amigo o amiga?”, demuestra que no se observan diferencias significativas en cuanto a la media obtenida por docentes hombres y mujeres a nivel universidad. A nivel enfermería, sí se encontraron diferencias significativas. A nivel ingeniería, no se encontraron diferencias significativas en las evaluaciones de docentes hombres y docentes mujeres (tabla 15).

En síntesis, la hipótesis 6, que las docentes mujeres son mejor evaluadas que los profesores hombres en la pregunta “¿Recomendarías este docente a un amigo o amiga?”, se confirma en unidades generizadas, pero se rechaza a nivel universidad.

### **3.7 Hipótesis 7: El estudiantado evalúa mejor a docentes de su mismo género a nivel universidad.**

A nivel universidad, analizando el promedio de la evaluación de las cuatro dimensiones se observa que los estudiantes hombres evalúan de manera más crítica a sus docentes (hombres y mujeres) que las estudiantes mujeres (tabla 16).

Tabla 16  
*Promedio evaluación cuatro dimensiones según género del estudiante, UFT*

Docentes	N	Evaluación estudiantes mujeres		Evaluación estudiantes hombres	
		Media	DE	Media	DE
Mujeres	648	3.57	.76	3.53	.78
Hombres	742	3.55	.79	3.50	.81

Fuente: Elaboración propia.

Se realizó un análisis de regresión múltiple para evaluar el impacto de las variables género docente, género estudiante y su interacción en los puntajes de evaluación docente, utilizando estos puntajes como variable dependiente. Las variables independientes fueron codificadas como 1 para masculino y 0 para femenino, incluyendo una interacción entre ambas para analizar posibles efectos combinados.

A nivel universidad, se analizó cada pregunta de las cuatro dimensiones evaluadas. En todos los modelos, el valor de  $R^2$  ajustado fue cercano a cero, indicando que estas variables no explican significativamente los resultados de las evaluaciones. Además, el coeficiente de interacción fue estadísticamente significativo en solo dos preguntas.<sup>6</sup>

En síntesis, la hipótesis 7, que plantea que los estudiantes evalúan mejor a docentes de su mismo género, se rechaza para estudiantes hombres y se confirma para estudiantes mujeres.

### **3.8 Hipótesis 8: El estudiantado evalúa mejor a los docentes de su sexo opuesto en unidades típicamente generizadas (enfermería e ingeniería).**

Los estudiantes hombres y mujeres de enfermería e ingeniería evalúan mejor a sus docentes mujeres que a sus profesores hombres al considerar el promedio de las cuatro dimensiones sobre práctica pedagógica (tabla 13), por tanto, no se cumple que se valore más a los profesores varones en enfermería.

Se replicó el análisis de regresión múltiple en enfermería e ingeniería para evaluar el impacto de las variables género docente, género estudiante y su interacción en los puntajes de evaluación docente. En los 17 modelos analizados, correspondientes a cada pregunta de las cuatro dimensiones, el valor de  $R^2$  ajustado fue cercano a cero en ambas unidades, indicando que estas variables no explican significativamente los resultados obtenidos. Además, el coeficiente de interacción no fue estadísticamente significativo en ninguna pregunta.

<sup>6</sup> (P12) Fue puntual en el inicio y término de cada sesión de clases; (P18) Promovió la resolución de problemas de manera crítica y autónoma por parte de las y los estudiantes.

En síntesis, la hipótesis 8, que plantea que el estudiantado evalúa mejor a docentes de su sexo opuesto en unidades generizadas, se rechaza para enfermería y se confirma para ingeniería.

#### 4. Discusión

Antes de discutir los resultados obtenidos, resulta pertinente detenerse brevemente en la conceptualización de la evaluación docente y su función institucional, ya que no toda evaluación cumple los mismos fines ni se fundamenta en los mismos supuestos. En la Universidad Finis Terrae, el instrumento aplicado se enmarca en un enfoque formativo y comprensivo, orientado a visibilizar buenas prácticas docentes desde la percepción estudiantil. Sin embargo, también posee consecuencias contractuales: sus resultados se utilizan como insumo en procesos de calificación, promoción y, eventualmente, pérdida de calidad académica según el reglamento interno. Esta doble función refleja la tensión presente en los paradigmas evaluativos contemporáneos: desde visiones centradas en el control y la rendición de cuentas hasta perspectivas críticas que promueven la mejora de los procesos de enseñanza-aprendizaje (Linse, 2017).

Los resultados de este estudio permiten discutir sobre el impacto del género en la evaluación docente en una universidad chilena. Las docentes mujeres reciben mejores evaluaciones que sus pares hombres en las dimensiones que evalúan la práctica docente, incluyendo organización de la enseñanza, clima del aula, metodología, y evaluación y retroalimentación. Sin embargo, esta ventaja no se traduce en una mayor recomendación como docentes, pues el estudiantado recomienda en la misma medida a docentes varones que a académicas mujeres. Esto no se explica por las diferencias en la evaluación en la calidad de la enseñanza (Mengel et al., 2019), lo que abre un espacio crítico para analizar cómo las evaluaciones docentes pueden albergar sesgos de género.

La evaluación docente es un mecanismo central para garantizar la calidad educativa, pero también es un insumo crucial en procesos de promoción y asignación de incentivos para el cuerpo docente. En este sentido, los resultados del presente estudio plantean interrogantes sobre cómo estas evaluaciones pueden contribuir o, por el contrario, obstaculizar el avance hacia la equidad de género en la academia. Si bien las mujeres destacan en todas las dimensiones de la práctica pedagógica, la evaluación global parece estar más influida por juicios implícitos que por la calidad docente. Este patrón, documentado en investigaciones internacionales como las de Boring (2017) y Mengel et al. (2019), sugiere que los sesgos de género persisten incluso en sistemas diseñados para demostrar objetividad en la evaluación.

A nivel disciplinar, opuestas en su composición de género, los resultados muestran similitudes. Tanto en la carrera de enfermería como en la facultad de ingeniería, las docentes mujeres obtienen evaluaciones superiores a los profesores hombres. Esto derriba creencias que enclaves masculinizados y feminizados tienden a

sobrevalorar a los varones. Por el contrario, este estudio muestra que las profesoras universitarias mujeres son mejor evaluadas que los docentes hombres independiente de si la carrera es generizada.

Otro hallazgo relevante es que el género del estudiante no actúa como un predictor claro de las evaluaciones. Si bien los estudiantes hombres tienden a ser más críticos, evaluando más bajo a sus docentes varones y mujeres en comparación a las estudiantes mujeres, tanto estudiantes hombres como mujeres valoran favorablemente a las docentes mujeres en las dimensiones de la práctica docente. Este resultado es consistente con Peterson et al. (2019), en cuanto a que las percepciones de género no dependen únicamente del evaluador y sus características, sino también de factores culturales que pueden aumentar o mitigar los sesgos de género.

Nuestros resultados muestran sesgos de género en la evaluación docente, aunque con incertidumbres respecto a si pueden atribuirse al proceso de evaluación, al instrumento o bien a factores culturales.

## 5. Conclusiones

Este artículo ha evidenciado que las docentes mujeres en una universidad chilena obtienen evaluaciones superiores que los profesores varones en las dimensiones que evalúan la práctica docente: organización de la enseñanza, clima del aula, metodología, y evaluación y retroalimentación. Estos resultados reafirman su capacidad para sobresalir en habilidades pedagógicas clave. Sin embargo, a pesar de estas calificaciones positivas, el estudiantado tiende a recomendar de igual manera a los docentes hombres que a las docentes mujeres, revelando un sesgo de género persistente en la evaluación global. Este hallazgo y la relevancia de los sesgos coincide con la literatura internacional, como los estudios de Boring (2017) y Mengel et al. (2019) en otros contextos educativos. Sin embargo, nuestro estudio distingue por la evaluación favorable de las docentes mujeres en todas las dimensiones de la práctica docente.

El análisis destaca que las académicas mujeres son evaluadas favorablemente tanto en unidades feminizadas (enfermería) como masculinizadas (ingeniería), logrando mejores resultados a nivel global en las cuatro dimensiones. Además, el género del estudiante no parece ser un factor determinante, ya que estudiantes hombres y mujeres califican mejor a las docentes mujeres en dimensiones específicas. Esto sugiere transversalidad al contexto disciplinar, en particular, considerando que el estudiantado y profesorado de ingeniería es masculinizado, mientras que la carrera de enfermería feminizada, y que los estudiantes hombres evalúan más bajo al profesorado que las estudiantes mujeres. A pesar de estas diferencias, las docentes mujeres son mejor evaluadas que sus pares varones en carreras con distinta composición de género.

En suma, este estudio aporta un análisis detallado y actualizado sobre las relaciones de género en la evaluación docente, examinando la totalidad de respuestas de estudiantes evaluadores durante dos años en una universidad chilena. Este enfoque permite develar la persistencia de sesgos de género en un contexto cultural y disciplinar específico, destacándose por su alcance y profundidad en comparación con investigaciones previas. No obstante, el trabajo presenta ciertas limitaciones a considerar en futuras investigaciones. Este estudio se centra en una universidad en particular, por lo que sus resultados y análisis son aplicables a esa institución en específico. Sin embargo, la sistematización en el uso del instrumento de evaluación docente y la identificación de sesgos de género pueden aportar a otras instituciones educativas y al debate nacional e internacional en estas materias. Para fortalecer la aplicabilidad de este estudio a otros contextos sería valioso contrastar estas dinámicas de género en universidades públicas y privadas, así como en diferentes regiones. Además, aunque el instrumento mejora en confiabilidad, sigue siendo insuficiente para capturar integralmente la calidad de la enseñanza. Advertir sobre sesgos inconscientes al momento de responder la evaluación docente, complementar la evaluación docente con observaciones directas, análisis cualitativos y autoevaluaciones enriquecería la comprensión de estas dinámicas de género y la evaluación de la calidad de la docencia.

Otra limitación de este estudio es la falta de datos cualitativos que permitan comprender mejor las percepciones del estudiantado y las experiencias de los docentes, lo que podría ayudar a identificar las causas subyacentes de las diferencias observadas en el caso de la recomendación docente. Asimismo, no se consideraron factores contextuales como la carga laboral, el tipo de asignatura o el nivel académico de los docentes, que podrían influir en las evaluaciones. Estas limitaciones relevan la necesidad de adoptar un enfoque más integral y diverso en futuras investigaciones, que profundice en las interacciones entre género, evaluación docente y los contextos institucionales, para sentar bases sólidas hacia políticas que promuevan la equidad de género en la academia.

Nuestros hallazgos plantean desafíos relevantes tanto para mitigar los sesgos de género en la evaluación docente como para recompensar los resultados favorables de las profesoras mujeres en su evaluación. Esto se adiciona a sesgos de género encontrados en otros momentos del proceso de enseñanza-aprendizaje (Resa Ocio, 2023; Simbürger & Undurraga, 2013), a la intersección del género con otras categorías sociales como raza/etnia, edad, orientación sexual, estado civil, que (re)producen desigualdades sociales en el mercado laboral (Crenshaw, 1991; McCall, 2005; Undurraga, 2019; Undurraga & Bazoret, 2015) y a otras inequidades de género experimentadas en la academia (Guizardi et al., 2024, 2022), como el reparto desigual del poder en puestos de toma de decisiones, los sesgos de género en la investigación y la docencia, la ausencia de políticas de género en universidades, entre otros aspectos (Mora et al., 2023; Resa Ocio & Rabazas Romero, 2020; Undurraga et al., 2021). Eliminar los sesgos de género en procesos tan relevantes en la trayectoria académica como lo es la evaluación docente es un pendiente para avanzar hacia la equidad de género en las universidades, tanto para afirmar la

evaluación en base a la calidad como para garantizar la promoción en base al desempeño.

Este artículo plantea nuevas interrogantes sobre la relación entre los resultados de la evaluación docente y las estrategias institucionales para asegurar que las mujeres reciban el merecido reconocimiento en términos de promoción e incentivos por su labor docente, en particular, con efectos en su jerarquización académica y remuneraciones. Resolver estas cuestiones es esencial para avanzar hacia una equidad de género en las instituciones de educación superior.

### Referencias bibliográficas

- Aravena Gaete, M.E., & Gairín Sallán, J. (2021). Evaluación del desempeño docente: Una mirada desde las agencias certificadoras. *Profesorado, Revista de Currículum y Formación del Profesorado*, 25(1), 297-317. <https://doi.org/10.30827/profesorado.v25i1.8302>
- Arceo-Gomez, E., & Campos-Vazquez, R. (2019). Gender stereotypes: The case of MisProfesores.com in Mexico. *Economics of Education Review*, 72, 55-65. <https://doi.org/10.1016/j.econedurev.2019.05.007>
- 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>
- Crenshaw, K. (1991). Mapping the Margins: Intersectionality, Identity Politics, and Violence against Women of Color. *Stanford Law Review*, 43(6), 1241-1299.
- Fardella, C. (2012). Verdades sobre la docencia, efectos y consecuencias subjetivas de la evaluación docente en Chile. *Revista De Psicología*, 21(1), 209-227. <https://doi.org/10.5354/0719-0581.2012.19996>
- Felkey, A., & Batz-Barbarich, C. (2021). Can Women Teach Math (and Be Promoted)? A Meta-analysis of Gender Differences across Student Evaluations of Teaching. *AEA Papers and Proceedings*, 111, 184-189. <https://doi.org/10.1257/pandp.20211125>
- Fernández, M.Á, Jiménez, M.I., & García, F.J. (2021). La evaluación de la práctica docente en la educación superior: indicadores. En M.Á. Fernández y A. Matas (Eds.), *Evaluación de la actividad docente como mecanismo transformador de*

- la Universidad del Siglo XXI* (33-42). Dykinson, S.L.  
<https://doi.org/10.2307/j.ctv2s0j cq0.6>
- Froment, F., Bohórquez, M.R., & García González, A.J. (2021). El impacto de la credibilidad docente y la motivación del estudiante en la evaluación de la docencia. *Revista Española de Pedagogía*, 79(280), 413-436.  
<https://www.jstor.org/stable/48621959>
- Galaz-Ruiz, A., & Anaya-Torres, E.D. (2023). Repensando la evaluación docente: Chile-México. *Estudios pedagógicos (Valdivia)*, 49(1), 203-218.  
<https://dx.doi.org/10.4067/s0718-07052023000100203>
- García, F.J., Fernández, M.Á., & Jiménez, M.I. (2021). Dimensiones de la evaluación para la buena docencia universitaria. En M.Á. Fernández y A. Matas (Eds.), *Evaluación de la actividad docente como mecanismo transformador de la Universidad del Siglo XXI* (53-62). Dykinson, S.L.  
<https://doi.org/10.2307/j.ctv2s0j cq0.8>
- Gómez Sánchez, T.F., & Rumbo Arcas, B. (2023). Estudio en clave comparada de la evaluación del profesorado universitario. *Profesorado, Revista de Currículum y Formación del Profesorado*, 27(1), 373-397.  
<https://doi.org/10.30827/profesorado.v27i1.21576>
- Gourley, P., & Madonia, G. (2020). The impact of tenure on faculty course evaluations. *Education Economics*, 29(1), 73-104.  
<https://doi.org/10.1080/09645292.2020.1852391>
- Guizardi, M., González Torralbo, H., & Stefoni, C. (2024). Women in Latin American Academia: Gender Inequalities in the Social Sciences of Chile, Mexico, Brazil, and Argentina. *Estudios Ibero Americanos*, 50(1), 1-18.  
<https://doi.org/10.15448/1980-864X.2024.1.44895>
- Guizardi, M., González Torralbo, H., & Stefoni, C. (2022). The Shoemaker and Her Barefooted Daughter: Power relations and gender violence in university contexts. *Frontiers: A Journal of Women Studies*, 43(1), 32-67.  
<https://doi.org/10.1353/fro.2022.0001>
- Hejase, A.J., Hejase, H.J., & Kaakour, R.S.A. (2014). The impact of students' characteristics on their perceptions of the evaluation of teaching process. *International Journal of Management Sciences*, 4(2), 90-105.
- Korte, L., Lavin, A., & Davies, T. (2013). Does gender impact business students' perceptions of teaching effectiveness? *Journal of College Teaching & Learning (TLC)*, 10(3), 167-178. <https://doi.org/10.19030/tlc.v10i3.7933>
- Lara-Subiabre, B., & Angulo-Brunet, A. (2020). Reflexiones de profesores chilenos en contextos de evaluación docente. *Perspectiva Educacional*, 59(3), 24-44.  
<https://dx.doi.org/10.4151/07189729-vol.59-iss.3-art.1083>

- Linse, A.R. (2017). Interpreting and using student ratings data: Guidance for faculty serving as administrators and on evaluation committees. *Studies in Educational Evaluation* 54, 94-106. <https://doi.org/10.1016/j.stueduc.2016.12.004>
- Manzi, J., González, R., & Sun, Y. (2011). *La evaluación docente en Chile*. MIDE UC. [https://www.mideuc.cl/libroed/pdf/La\\_Evaluacion\\_Docente\\_en\\_Chile.pdf](https://www.mideuc.cl/libroed/pdf/La_Evaluacion_Docente_en_Chile.pdf)
- McCall, L. (2005). The Complexity of Intersectionality. *Signs: Journal of Women in Culture and Society*, 30(3), 1771-1800.
- Medel, R., & Asún, R. (2014). Surveys on teacher evaluation and gender biases: an exploratory study. *Calidad en la educación*, (40), 171-199. <https://dx.doi.org/10.4067/S0718-45652014000100006>
- Mengel, F., Sauermann, J., & Zölitz, U. (2019). Gender Bias in Teaching Evaluations. *Journal of the European Economic Association*, 17(2), 535-566. <https://doi.org/10.1093/jeea/jvx057>
- Mora, C., Undurraga, R., & Simbürger, E. (2023). The multidimensionality of care in remote work: women academics in Chile during the COVID-19 pandemic. *Culture and Organization*, 30(5), 463-484. <https://doi.org/10.1080/14759551.2023.2294999>
- Paredes, V., Pino, F., & Díaz, D. (2024). Does facial structure explain differences in student evaluations of teaching? The role of fWHR as a proxy for perceived dominance. *Economics & Human Biology*, 54. <https://doi.org/10.1016/j.ehb.2024.101381>
- Peterson, D.A.M., Biederman, L.A., Andersen, D., Ditonto, T.M., & Roe, K. (2019). Mitigating gender bias in student evaluations of teaching. *PLoS ONE* 14(5): e0216241. <https://doi.org/10.1371/journal.pone.0216241>
- Quansah, F., Cobbinah, A., Asamoah-Gyimah, K., & Hagan, J. (2024). Validity of student evaluation of teaching in higher education: a systematic review. *Front. Educ.* 9: 1329734. <https://doi.org/10.3389/feduc.2024.1329734>
- Resa Ocio, A. (2023). Igualdad de género y formación inicial del profesorado en España: entre la utopía y la realidad. *Profesorado, Revista de Currículum y Formación del Profesorado*, 27(1), 255-275. <https://doi.org/10.30827/profesorado.v27i1.21192>
- Resa Ocio, A., & Rabazas Romero, T. (2020). Organizaciones y políticas supranacionales: una mirada desde la educación superior y la igualdad de género. *Revista Española de Educación Comparada*, (37), 281-298. <https://doi.org/10.5944/reec.37.2021.27793>
- Simbürger, E., & Undurraga, R. (2013). Jerarquías epistemológicas: formas de invisibilizar el género en las carreras de sociología en Chile. En C. Mora (Ed.),

*Desigualdad en Chile: La continua relevancia del género* (171-195). Ediciones Universidad Alberto Hurtado.

Undurraga, R. (2019). Who will get the job? Hiring practices and inequalities in the Chilean labour market. *Bulletin of Latin American Research*, 38(5), 575-590. <https://doi.org/10.1111/blar.12888>

Undurraga, R., & Barozet, E. (2015). Pratiques de recrutement et formes de discrimination des femmes diplômées - le cas du Chili. *L'Ordinaire des Amériques*, 219. <https://doi.org/10.4000/orda.2357>

Undurraga, R., Simbürger, E., & Mora, C. (2021). Desborde y desazón versus flexibilidad y concentración: Teletrabajo académico y género en tiempos de pandemia. *Revista Polis Latinoamericana* 20(59), 12-38. <http://dx.doi.org/10.32735/S0718-6568/2021-N59-1594>

Universidad Finis Terrae (2023). Documento interno “Dimensiones y estándares para la medición de la calidad de la docencia en la Universidad Finis Terrae”, diciembre de 2023.

Wallace, S.L., Lewis, A.K., & Allen, M.D. (2018). The State of the Literature on Student Evaluations of Teaching and an Exploratory Analysis of Written Comments: Who Benefits Most? *College Teaching*, 67(1), 1-14. <https://doi.org/10.1080/87567555.2018.1483317>

Worthington, A.C. (2002). The Impact of Student Perceptions and Characteristics on Teaching Evaluations: A case study in finance education. *Assessment & Evaluation in Higher Education*, 27(1), 49-64. <https://doi.org/10.1080/02602930120105054>

**Contribuciones del autor:** R.U.: concepción del artículo, introducción, revisión de la literatura, metodología, hipótesis, análisis, discusión, conclusiones, referencias, redacción, revisión, formato. J.T.M.: antecedentes evaluación docente en UFT, hipótesis, discusión, conclusiones, redacción. E.G.B.: metodología, hipótesis, análisis estadístico, resultados.

**Financiación:** Esta investigación no recibió financiación externa.

**Conflicto de intereses:** Los autores declaran que no existen conflictos de intereses para la publicación de este manuscrito.

**Declaración ética:** La presente investigación se realizó conforme a los principios éticos establecidos por la comunidad científica.

**Cómo citar este artículo:**

Undurraga, R., Montalva, J.T., & Guerrero Barrera, E. (2025). Profesores en evaluación: El impacto del género en los resultados de la evaluación docente universitaria. *Profesorado. Revista de Currículum y Formación de Profesorado*, 29(2), 1-23. DOI: <https://doi.org/10.30827/profesorado.v29i2.32229>