

## ENCUESTAS DE EVALUACIÓN DOCENTE Y SESGOS DE GÉNERO: UN ESTUDIO EXPLORATORIO

### Introducción

En el contexto de las demandas a las que actualmente son sometidas las universidades y que tensionan su rol tradicional como productoras autónomas de conocimiento, exigiéndoles también generar y difundir saberes prácticos beneficiosos para sus estudiantes, o bien, el desarrollo nacional (Van Vucht Tijssen & Weert, 2005), se ha tendido a generalizar en ellas la implementación de diversos mecanismos de rendición de cuentas y evaluación de la calidad de sus productos (Aedo, 2010). Dentro de estos nuevos instrumentos de gestión, uno de los que ha alcanzado mayor extensión en los últimos años ha sido el desarrollo y aplicación regular de Cuestionarios de Evaluación de la Docencia (CED), dirigidos a los estudiantes (Zabaleta, 2007).

Las universidades chilenas no han quedado al margen de este proceso. Por el contrario, en los últimos años se ha difundido la aplicación de diversos mecanismos de evaluación y control de calidad (OECD, 2013), dentro de los que destaca la implementación de variados modelos para evaluar la calidad de la docencia impartida (Contreras, Faúndez, Gutiérrez, Jiménez, Ponce y Silva, 2007). Pese al interés puesto en desarrollar sistemas multidimensionales de valoración de la enseñanza, en la práctica, los CED dirigidos a los estudiantes representan el tipo de instrumento más empleado y muchas veces el único (Salazar, 2008).

Más allá de los debates que se han generado respecto del significado y potenciales consecuencias ideológicas de la implantación de estos instrumentos (Haskell, 1997; Valsan & Sproule, 2008), la ubicuidad alcanzada por los CED hace necesario asegurar la máxima validez a sus resultados. Lamentablemente, al contrario de la fuerte tradición de investigación de factores que invalidan estas mediciones

en la literatura anglosajona, en América Latina y Chile existe muy poca evidencia sistemática al respecto (Garduño, 2000).

Esta carencia de información puede no ser problemática en temas independientes de la cultura, pues en ese caso sería legítimo extrapolar a Chile los hallazgos encontrados en la literatura extranjera, pero podría ser decisiva frente a factores de invalidez que sean función de la cultura local. En este contexto, debido al carácter “masculino” que incluso hoy en día parece tener la institución universitaria en América Latina (Bonder, 1998; Sagot, 2002) y el mayor nivel de desigualdad de género que existe en estas sociedades respecto de los países industrializados (Kliksberg, 2003), creemos que ignorar si en nuestros países operan sesgos de género en la evaluación que hacen los estudiantes universitarios de sus académicos y académicas es una carencia crucial, pues esa variable podría introducir un importante elemento de invalidez y discriminación en este difundido sistema de evaluación.

Con el objetivo de comenzar a cerrar esa brecha de información, en este artículo se presenta un estudio que tiene por objetivo determinar y cuantificar la presencia de diferencias de género en la evaluación que hacen los estudiantes universitarios de sus profesores y profesoras, empleando datos provenientes de una encuesta de evaluación de la docencia aplicada a todos los estudiantes de la Facultad de Ciencias Sociales de la Universidad de Chile en el 2007.

## 1. Diferencias de género y educación superior

Aunque la situación actual está lejos de ser la ideal, es reconocido que en muchos países del mundo las discriminaciones legales o fácticas asociadas al género de las personas han tendido a disminuir en los últimos años (Social Watch, 2012). Pese a que estos positivos cambios también han ocurrido en América Latina, esos mismos informes internacionales señalan igualmente que nuestro subcontinente continúa estando por debajo de Europa y de Norteamérica en términos de equidad de género.

Debido a que una posible causa de este retraso son los valores, creencias y prácticas con los que las personas operan en su cotidianidad, resulta de suma importancia estudiar en profundidad las discriminaciones de género existentes dentro de las instituciones de educación superior latinoamericanas, ya que ellas no solo transmiten conocimientos, sino que son en sí mismas mecanismos de reproducción cultural, social y valórica (Bourdieu & Passeron, 1990), por lo que su funcionamiento puede tender a reforzar y actualizar los valores e ideas dominantes en su sociedad.

A continuación presentaremos investigaciones que evidencian la existencia de sesgos de género en las instituciones de educación superior, tanto en la experiencia de las académicas universitarias, como en la interacción profesor(a)/alumno(a).

### 1.1. Diferencias de género en la experiencia de las académicas universitarias

Los estudios acerca de la situación de las académicas en la educación superior generalmente ponen el foco en la desigualdad de salarios. Sin embargo, la socióloga estadounidense Sandra Acker (1983) afirma que para entender las diferencias de género en el mundo de la academia hay que considerar tres dimensiones o problemas generales: las demandas de compatibilizar la familia y una carrera académica; el dominio de los hombres sobre la producción y administración del conocimiento; y la situación de poder de una mayoría masculina sobre una minoría femenina.

Estas tres dimensiones estarían operando de manera simultánea y su importancia individual se daría en mayor o menor proporción, dependiendo de variables contextuales específicas de cada institución y localidad. En Chile los principales estudios en torno a esta materia se han centrado en los últimos dos puntos mencionados por la autora, es decir, el dominio de los hombres sobre la producción y administración del conocimiento y el control que tiene una mayoría masculina respecto de las posiciones de poder.

Acerca del último punto, parece relativamente claro que, pese a que en los últimos años la universidad chilena se ha abierto

progresivamente a contratar académicas, aún se está en presencia de un mundo principalmente masculino. Ejemplo de ello es que entre los (las) académicos(as) menores de 35 años las mujeres representan el 47%, mientras que en el grupo de 65 o más años solo constituyen el 22% (Olavarría, 2011). Sin embargo, respecto de los cargos directivos, Saracostti (2006) señala que las diferencias son mucho más claras: en el 2005 las mujeres representaban solo el 8% del total de rectores de universidades, mientras que el 67% de las jefaturas de carrera se encontraba en manos de varones.

En relación con el control sobre las oportunidades de producción de conocimiento, Berríos (2005) señala que las diferencias de género son también notables. Por ejemplo, por cada proyecto Fondecyt aprobado que tiene a una mujer como investigadora responsable, se aprueban 3,7 proyectos cuyo investigador responsable es un hombre.

Es interesante notar que este predominio masculino no solo se observa en indicadores cuantitativos, sino que también es posible detectarlo en las relaciones cotidianas dentro del sistema. Por ejemplo, Reyes (2006), estudiando las interacciones entre académicos de una universidad chilena, encontró que el cuerpo docente de la universidad actúa como agente reproductor de estructuras simbólicas que asignan el predominio a los varones, aspecto que se puede notar, entre otros espacios, en las reuniones de Consejo de carrera, donde se tiende a valorar y prestar más atención a las reflexiones masculinas por sobre las femeninas, quienes además acostumbran emitir menos comentarios y opiniones.

La consecuencia de todo lo anterior es que el principal capital por el que luchan los(las) académicos(as), es decir, la “buena reputación” o el “prestigio” como potenciales productores y difusores de conocimiento (Bourdieu, 2003), no solo está distribuido en forma desigual por género, sino que el acceso a él es diferencialmente dificultoso: las académicas estarían estigmatizadas por una imagen simbólica negativa, debiendo mostrar estándares de desempeño mayores que los hombres para neutralizar dicha marca y obtener iguales niveles de recompensa (Berríos, 2005).

## 1.2. La importancia del género en la interacción profesor-estudiante

Existe amplia evidencia internacional acerca de la relevancia que tiene el género en la interacción profesor-estudiante, sobre todo a nivel de educación escolar. La mayoría de los estudios indica que dicha interacción tiene efectos estadísticamente significativos y diferenciados por género, en una diversa gama de resultados como los puntajes en las pruebas, las percepciones de los profesores respecto de los alumnos y el entusiasmo de estos por las materias impartidas (Dee, 2007). Otros estudios resaltan el hecho de que cuando el alumno y el profesor son del mismo sexo, los resultados en el rendimiento escolar del alumno tienden a mejorar (Ammermueller & Dolton, 2006; Holmlund & Sund, 2008).

A nivel nacional, Flores (2007) identificó que los profesores y profesoras escolares de asignaturas científicas tienden a interactuar más con los alumnos varones, estableciendo contacto con preguntas que, en su mayoría, van dirigidas a ellos. En ese sentido, los niños tendrían un efecto de imán sobre los profesores, mientras que las niñas asumen posiciones más de espectadoras que de partícipes de la clase. Más recientemente, Cabezas (2009) estudió el efecto del género de los profesores y profesoras en los resultados en las pruebas Simce. El autor encontró que, en general, tener un profesor del mismo género que el del alumno tiene un impacto positivo en las pruebas de matemática, tanto en colegios de un solo sexo, como mixtos. En tanto, en las pruebas de lenguaje también habría un impacto positivo, pero solamente para colegios de un solo sexo.

Pese a que no se han encontrado estudios que confirmen estas tendencias en las universidades chilenas, sí hay investigaciones internacionales que han corroborado que estos resultados se replican a nivel universitario. El primero de todos es el artículo de Hall y Sandler (1982), donde los autores concluyen que el género es un factor determinante en las interacciones que se dan entre profesor y alumno en la educación superior. Es así como, por medio de distintas mediciones, se documenta que la participación de las estudiantes es cuantitativa y cualitativamente distinta de la de los hombres: sus

intervenciones en clases se ven sofocadas por el uso de un lenguaje sexista, patrones diferenciados de interacción con el profesor y un tipo de comunicación que en forma sistemática las deja en desventaja y silencio. Más recientemente, otros estudios (Auster & MacRone, 1994; Canada & Pringle, 1995) han corroborado que la participación en la sala de clases se ve afectada por patrones de género.

En suma, es posible afirmar que hombres y mujeres se enfrentan a realidades distintas al ingresar a la educación superior: las académicas se sitúan en una situación de poder y prestigio desmejorado frente a sus colegas varones, mientras que las estudiantes pueden verse sometidas a interacciones sociales y pedagógicas diferenciadas respecto de sus compañeros masculinos.

Nuestra hipótesis es que estas diferencias pueden tener efecto en las evaluaciones que hacen los y las estudiantes respecto de la calidad pedagógica de sus profesores y profesoras.

## 2. Sesgos de género en las encuestas de evaluación docente

Pese a lo difundido que es evaluar la calidad de la docencia universitaria a través de CED, la antigüedad que tiene esta práctica y la investigación acumulada, la comunidad científica no tiene un juicio unánime respecto de su validez.

En efecto, si bien hay destacados investigadores que consideran demostrada su validez (Aleamoni, 1999; Marsh & Roche, 2000), otro grupo no menos persistente de autores señala la existencia de factores de invalidez, dentro de los cuales se mencionan la indulgencia de los profesores (Greenwald & Gillmore, 1997), los rasgos de personalidad de los docentes (Clayson & Sheffet, 2006), el prestigio previo de los académicos (Griffin, 2001) y el género del docente (Basow, Phelan & Capotosto, 2006), entre otros. Estos factores son los que la literatura ha denominado “sesgos” o “factores extraclase” (Marsh, 1984), refiriéndose con ello a todas las variables que se encuentran más allá de la calidad de la instrucción y que afectan la valoración que hacen los estudiantes de sus profesores.

Pese a que parece claro que los factores extraclase tienen la potencialidad de distorsionar la evaluación que hacen los estudiantes de la calidad docente de los académicos (Onwuegbuzie, Witcher, Collins, Filer, Wiedmaier & Moore 2007; Garduño, 2008; Basow, 2000), la discusión está centrada en torno a si esta distorsión tiene un efecto realmente significativo como para invalidar los resultados obtenidos (Marsh, 1984), tema que debería evaluarse en cada situación y respecto de cada factor extraclase.

En relación con el género, las investigaciones acerca de los factores extraclase que se centran en este aspecto se han ido incrementando en el mundo anglosajón a medida que ha ido aumentando la proporción de docentes mujeres en la educación superior (Basow, 1995).

En términos generales, gran parte de los estudios ha identificado que si bien ambos sexos son usualmente evaluados de forma distinta, los sesgos producto del género son menores y dependen de otros factores como la raza de los docentes o incluso a veces favorecen a las mujeres, quienes recibirían, en conjunto, mejores evaluaciones que los hombres (Tatro, 1995; Garduño, 2000; Smith, Yoo, Farr, Salmon & Miller, 2007; Smith, 2009).

Uno de los investigadores que más ha contribuido a establecer esta conclusión es Kenneth Feldman, quien en 1992 publicó un metaanálisis en torno a 17 estudios experimentales en los que no se observaban diferencias apreciables en la evaluación de los docentes según su sexo. Más aún, en un metaanálisis posterior, Feldman (1993) encontró que las profesoras eran ligeramente mejor evaluadas que sus colegas en 28 de los 39 estudios analizados. En términos de dimensiones, las profesoras fueron mejor evaluadas en sensibilidad, preocupación, disponibilidad y ayuda.

Las conclusiones de estos estudios han sido contestadas desde dos argumentos: en primer lugar, se ha señalado que estudiar la discriminación de género empleando instrumentos cuantitativos tiene una serie de limitaciones que dificulta su observación y, en segundo lugar, se ha cuestionado la tendencia a interpretar estos resultados a partir de puntajes agregados.

Respecto del primer punto, Sprague y Massoni (2005) argumentaron que si bien en una escala numérica no observan mayores diferencias de género en las puntuaciones medias otorgadas a profesores hombres y mujeres, estos estudios suponen la existencia de una métrica universal y que la puntuación que se les asigna a los profesores corresponde a conductas equivalentes. Los autores afirman que esos supuestos serían falsos en el caso del género, pues se ha podido observar que los estudiantes utilizan “líneas base” diferentes para evaluar a profesores hombres y mujeres. Esto solo se percibiría cuando se estudian cualitativamente los estereotipos de género subyacentes a los adjetivos con los cuales los estudiantes caracterizan a sus buenos y malos profesores.

Sobre el segundo argumento, la socióloga estadounidense Susan Basow (2000) señala que, usualmente, cuando los ítems se agrupan en puntajes agregados, no aparecen efectos significativos de género, lo que sugiere que estos pueden operar a nivel micro, por lo que la tendencia a utilizar puntajes agregados sería la razón por la cual los estudios cuantitativos no encuentran diferencias entre las puntuaciones de profesores hombres y mujeres.

La misma autora señala, además, que no se deben esperar grandes efectos en las diferencias por género en la evaluación docente, sino que lo importante es analizar los resultados de interacción entre el género y otras variables contextuales, ya que allí es donde se actualizan los patrones de género y donde se observarán efectos significativos y consistentes, indicando relaciones teóricamente relevantes desde la perspectiva de género (Basow, 1998).

Siguiendo estos argumentos, Basow (1995) estudió la interacción entre el género del profesor y el género del alumno en evaluaciones docentes, encontrando que si bien las calificaciones de los profesores no parecen verse afectadas por el sexo del estudiante, las académicas tendían a recibir calificaciones más altas por parte de las estudiantes y calificaciones más bajas de los varones. En un estudio más reciente (Basow, Phelan & Capotosto, 2006) se corrobora la existencia de esta interacción: los estudiantes calificaron como “buenas docentes” a sus profesoras en proporción significativamente

inferior a lo esperado en relación con el porcentaje de profesoras que tuvieron en clases. Las estudiantes, por su parte, evaluaron a sus profesores y profesoras como buenos y malos en proporción directa a su frecuencia en el aula. Por lo tanto, las estudiantes no parecen estar tan determinadas por el género, mientras que, en general, la pareja profesor femenino-estudiante masculino parecía estar más afectada por las dinámicas de género.

En síntesis: el estudio del género como posible factor extraclase en los CED exige no solo estudiar si hay diferencias significativas a favor de los profesores de algún sexo, sino también incluir en el análisis el sexo del estudiante que emite la evaluación.

A lo anterior se debe agregar que se ha argumentado (Vásquez y Gabalán, 2006) que el origen de los factores extraclase radica en las interacciones entre profesores y estudiantes en un determinado contexto sociocultural caracterizado por diferentes expectativas y percepciones de comportamiento, las cuales mediarían en la evaluación que el estudiante hace del desempeño docente de su profesor.

Ello es relevante en el caso del género, pues diversos estudios (Hills, Naegle & Bartkus, 2009; Smith & Cranton, 1992) han mostrado que las expectativas de los estudiantes universitarios cambian a lo largo de su carrera, estando los alumnos de primeros años más interesados en aspectos relacionados con la claridad expositiva del profesor, recibir apoyo para su aprendizaje o en la buena estructuración del curso, mientras que los estudiantes de los últimos años valoran más la posibilidad de discutir e integrar conocimientos, entre otros temas. Debido a que las habilidades para cumplir esas diversas expectativas pueden ser atribuidas en grados diferenciales a profesores hombres y mujeres, es posible que los sesgos de género actúen de manera distinta según el curso en el que se encuentren los estudiantes, transformándose esta variable en una moderadora de la relación entre el género y la evaluación docente.

Finalmente, es interesante notar que en América Latina también existe mucha polémica tanto respecto de las consecuencias

político-ideológicas que se derivan de la implementación de los CED (García y Beltrán, 2003; y Montes, 2009), como del grado de validez de esta evaluación, observándose muchos temores y críticas desde los distintos actores del sistema educativo: dirección, profesorado y estudiantes (Báez, Cazenave y Lagos, 2007; Aguilar, Ampuero y Loncomilla, 2008). Sorprendentemente, y al contrario de lo que ocurre en el mundo anglosajón, en nuestro subcontinente el debate respecto de la validez de los CED y la influencia del género en la evaluación docente no ha generado un volumen apreciable de investigación empírica (como excepción véanse Acevedo y Rodríguez, 2006 y Garduño, 2003) y, hasta donde alcanza el conocimiento de los autores de este artículo, actualmente constituye un campo de investigación inexistente en Chile.

El objetivo de esta investigación es contribuir a cerrar esta brecha, estudiando un tema –diferencias de género– frente al cual la distancia cultural entre Chile y aquellos países donde se ha realizado la mayor parte de los estudios es suficientemente amplia como para no poder fiarnos de los resultados disponibles en la literatura internacional.

### 3. Metodología

En los siguientes párrafos se resume la metodología utilizada para esta investigación.

#### 3.1. Características y objetivo general de la investigación

La presente investigación constituye un estudio no experimental, cuantitativo, censal y basado en datos secundarios, pues emplea como fuente de información las encuestas de evaluación docente aplicadas en el año 2007 a los estudiantes de pregrado de la Facultad de Ciencias Sociales (FACSO) de la Universidad de Chile, para evaluar la calidad de la docencia de la totalidad de sus profesores.

El objetivo general del estudio fue determinar la influencia individual y conjunta del sexo del profesor evaluado, el sexo del estudiante evaluador y el ciclo en el que se ubica el curso dictado

(es decir, formación básica, lo que implica primer o segundo año de cada carrera o formación especializada, lo que corresponde al tercer, cuarto o quinto año de la respectiva carrera), sobre las puntuaciones obtenidas por los docentes.

Con esta investigación se espera contribuir a la detección de sesgos de género en la evaluación docente en Chile. Creemos que la inexistencia en nuestro país de investigaciones previas acerca de esta temática convierte a este esfuerzo en una base para futuras pesquisas, así como en fuente de recomendaciones para la práctica de la evaluación docente con CED.

### 3.2. Participantes

Se dispuso para esta investigación de 9.132 cuestionarios de evaluación docente respondidos por 760 estudiantes de pregrado de la FACSÓ (85% de la matrícula total) durante los dos semestres del año 2007, a la totalidad de los cursos que matricularon y profesores que los dictaron, lo que implica que cada estudiante tuvo que responder un cuestionario por cada curso y profesor que tuvo durante ese año.

Se evaluaron en total 238 cursos pertenecientes a cuatro carreras (Psicología, Educación Básica Inicial, Antropología y Sociología), los que fueron dictados por 202 profesores (57,4%) y 149 profesoras (42,3%).

Producto del carácter marcadamente femenino de dos de las carreras de la Facultad (Psicología y Educación Básica Inicial), el 66,6% de los CED fueron respondidos por estudiantes mujeres. Pese a ello, se debe notar que esto no constituye un sesgo de muestreo, pues esa es la distribución por sexos en la facultad estudiada. Además, los estadísticos de significación y tamaño del efecto empleados consideran los tamaños de muestra de los subgrupos comparados, por lo que dicho factor tampoco constituye un elemento distorsionador de las asociaciones encontradas.

### 3.3. Instrumento

Para producir los datos se utilizó un cuestionario de 16 preguntas (ver Anexo 1), creado en función de preguntas pertenecientes a un instrumento anterior; otras que se estaban elaborando para diseñar un instrumento estándar para toda la universidad; también preguntas propuestas por el Centro de Estudiantes y preguntas discutidas en el Consejo de Escuela.

Como formato se optó por una estructura tipo Likert, con afirmaciones que permitían cuatro alternativas de respuesta (“Muy en desacuerdo”, “Muy de acuerdo”, “De acuerdo” y “Muy de acuerdo”). La utilización de cuatro opciones se decidió considerando que la alternativa intermedia tiene contraindicaciones, pues puede servir para ocultar la opinión real de algunos estudiantes, o bien, ser empleada para fines distintos a sostener una opinión equidistante de ambos polos (Kulas, Stachowski & Haynes, 2008).

Respecto de la dirección de las afirmaciones, se optó porque todas ellas fueran directas, es decir, que responder a la máxima categoría de acuerdo siempre indicara una percepción positiva del docente. Esto implica estar expuesto al sesgo de aquiescencia, pero se estimó que este riesgo sería menor que los errores de respuesta que aparecerían producto de incluir afirmaciones negativas.

A estas afirmaciones se agregó una pregunta final que medía la opinión global que el(la) estudiante tenía respecto del(a) profesor(a) (“En términos generales, ¿cómo calificarías al docente?”), en cinco categorías de respuesta desde “muy malo” a “muy bueno”, y que sirvió para poner a prueba la validez convergente del CED.

### 3.4. Análisis

La primera fase del análisis consistió en el estudio de la estructura, fiabilidad y validez del CED, para lo cual se realizó:

- un Análisis Factorial Exploratorio (AFE) y otro Confirmatorio (AFC) de la estructura interna del instrumento (empleando la mitad de la muestra para cada uno);
- un análisis de la fiabilidad del instrumento; y

- un análisis de la validez convergente de las puntuaciones globales del CED.

Finalmente, se construyeron las puntuaciones globales del CED, estandarizándolas luego en una métrica del 0 al 100 para que los resultados fueran interpretables como cambios porcentuales en la puntuación obtenida por profesores hombres y mujeres.

La segunda fase del estudio consistió en realizar análisis de varianza multivariados que permitieran estimar en forma conjunta la influencia del sexo de los profesores evaluados, el sexo del estudiante evaluador y el nivel en el que se dicta cada curso, sobre las puntuaciones otorgadas por los estudiantes. Este análisis fue realizado por medio del *software* SPSS (SPSS, 2010). Dado el alto tamaño de muestra disponible, no solo se estudió la existencia de relaciones significativas entre las variables, sino que también se determinó la magnitud de la fuerza de las relaciones encontradas, empleando el estadístico  $\eta^2$  cuadrado.

## 4. Resultados

Los resultados obtenidos tras la aplicación de la metodología descrita se desglosan en los siguientes puntos.

### 4.1. Validación del instrumento

Dado que se sospechaba de la presencia de una estructura multidimensional en el instrumento, la validación comenzó con un análisis factorial de su estructura interna. Con una selección aleatoria de la mitad de la muestra se realizó un AFE (sobre las correlaciones tetracóricas entre los ítems, extracción por mínimos cuadrados diagonalmente ponderados –DWLS– y rotación Oblimin) que mostró adecuados índices de ajuste (RMSEA=0.07; SRMR=0.02; CFI=0.99) a una configuración de tres factores significativamente asociados entre sí. Para confirmar el ajuste de dicha estructura se realizó un AFC con la otra mitad de la muestra, confirmándose el muy buen ajuste de una interpretación tridimensional del instrumento (RMSEA=0.056; SRMR=0.053; CFI=0.99).

Todos los ítems tuvieron marcadas y significativas cargas factoriales con su respectivo factor (mayores a 0.7), por lo que pueden ser considerados indicadores adecuados de estos, y los factores tienen altas correlaciones entre sí (entre 0.69 y 0.90), lo que autoriza a pensar que está midiendo un mismo constructo fragmentado en tres dimensiones. El análisis del contenido de los ítems que cargan en cada factor nos indica que el primero dice relación con las habilidades pedagógicas del profesor (preguntas 8 a 16 del cuestionario, ver Anexo 1); el segundo factor se refiere al dominio disciplinario del docente (preguntas 1, 2, 3 y 15 del cuestionario, ver Anexo 1); mientras que el tercer factor describe la responsabilidad que tiene el académico con el ejercicio de su docencia (preguntas 4 a 7 del cuestionario, ver Anexo 1).

En relación con la fiabilidad del instrumento y respetando la ordinalidad de los ítems, se estimó el Alfa de Cronbach Ordinal (Elosua y Zumbo, 2008), que alcanzó valores muy adecuados para el instrumento total ( $\alpha=0.957$ ) y para cada una de sus dimensiones (0.948, 0.900 y 0.868, para el primer, segundo y tercer factor respectivamente).

Por último, la correlación de Spearman entre la pregunta general: “¿Cómo calificarías al docente?” y cada una de estas puntuaciones fue muy alta y significativa (alcanzó un valor de 0.79, 0.75 y 0.57 para el primer, segundo y tercer factor respectivamente, y 0.82 para la puntuación global).

Debido a estos buenos indicadores de fiabilidad y validez del instrumento se construyó una puntuación agregada (ponderada por la carga factorial de cada ítem) y estandarizada en una escala del 0 al 100, para la escala total y cada dimensión de la escala, la que fue relacionada con el sexo del estudiante, el sexo del profesor y el ciclo de estudios como variable interviniente.

#### 4.2. Sesgos de género en la evaluación docente: efectos principales

Como se puede observar en la Tabla 1, la única variable aislada que tuvo relación significativa con la evaluación global de los profesores fue el sexo del estudiante que los evaluaba, en el sentido de que las

mujeres mostraron una tendencia a evaluar en forma un poco más positiva a sus profesores y profesoras que los estudiantes hombres (71,2 versus 68 puntos). Pese a que la fuerza de esta relación es baja ( $\eta^2 = 0.1$ ), es consistente, pues se mantuvo para todas las dimensiones del instrumento: las estudiantes evaluaron mejor a sus profesores y profesoras en habilidades pedagógicas, dominio disciplinario y responsabilidad.

Respecto de cuál género de profesores fue mejor evaluado, podemos apreciar en la Tabla 1 que, en conjunto, ambos sexos obtuvieron puntuaciones equivalentes en la escala total. No obstante, dicha paridad se obtuvo a partir de la compensación de evaluaciones significativamente diferentes en las distintas dimensiones del CED. De este modo, las profesoras fueron evaluadas mejor que sus colegas masculinos en sus habilidades pedagógicas (69,4 versus 68 puntos) y en su responsabilidad (74,9 versus 71,8 puntos), pero peor en su dominio disciplinario (70,4 versus 73,8 puntos).

Por último, observamos en la Tabla 1 que, en conjunto, no hay diferencias en la forma de evaluar de los estudiantes de primer ciclo respecto de los de segundo, aunque estos últimos tienden a evaluar leve, aunque significativamente mejor, a sus profesores y profesoras en responsabilidad (71,9 versus 70,8 puntos).

### 4.3. Sesgos de género en la evaluación docente: efectos de interacción

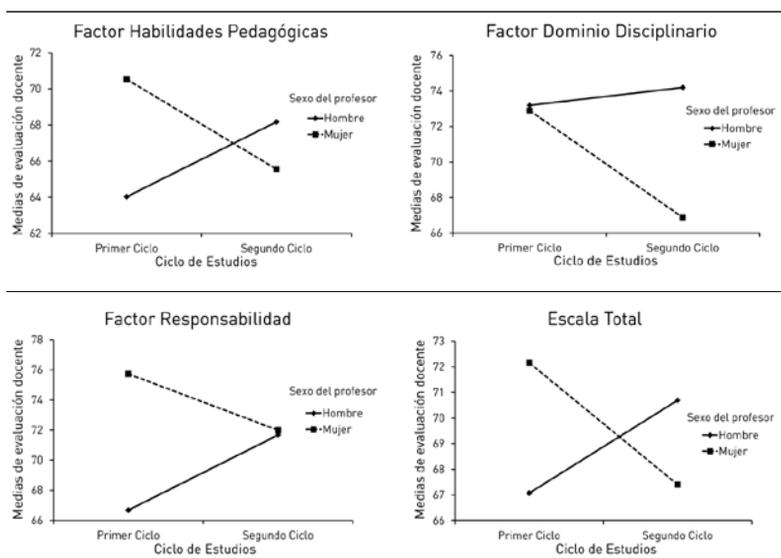
Como también podemos observar en la Tabla 1, frente a la escala total y cada una de sus dimensiones, siempre hay una relación significativa entre las puntuaciones obtenidas en la evaluación docente y la interacción entre el sexo del profesor y el ciclo de estudios del estudiante que evalúa, y esa interacción es invariablemente el efecto de orden superior estudiado que más explica las puntuaciones obtenidas por los docentes.

Tabla 1: Análisis de varianza multivariado entre la evaluación docente y el sexo del estudiante, el sexo del profesor y el ciclo de estudios del estudiante, según factores y escala total

		Factor habilidades pedagógicas		Factor dominio disciplinario		Factor responsabilidad		Escala total	
Tipo de efecto	Variables	F (sig)	Eta <sup>2</sup>	F (sig)	Eta <sup>2</sup>	F (sig)	Eta <sup>2</sup>	F (sig)	Eta <sup>2</sup>
Efectos principales	Sexo del estudiante	.000	.01	.007	.00	.000	.01	.000	.01
	Sexo del profesor	.001	.01	.000	.01	.000	.01	.113	.00
	Ciclo de estudios	.493	.00	.288	.00	.000	.00	.319	.00
Efectos de interacción	Sex estud * sex prof	.880	.00	.274	.00	.015	.00	.557	.00
	Sex estud * ciclo est	.393	.00	.255	.00	.007	.00	.339	.00
	Sex prof * ciclo est	.000	.01	.000	.01	.000	.01	.000	.01
	Sex estud * sex prof * ciclo est	.003	.00	.082	.00	.001	.00	.004	.00

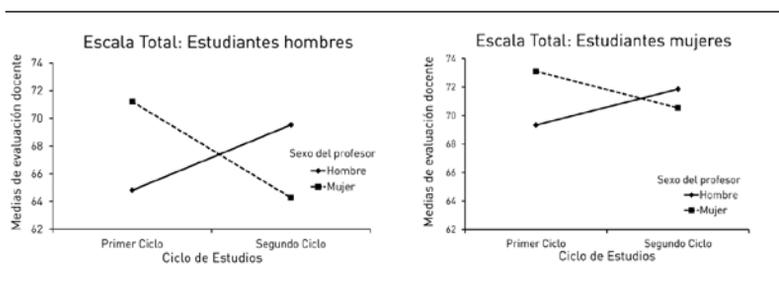
Pese a que dicho efecto es de magnitud menor ( $\eta^2 = 0.1$ ), puede implicar notorias pérdidas de puntos en la evaluación docente total. Para valorar este efecto y su dirección se puede observar la Figura 1. En ella podemos apreciar que el efecto es similar en las tres dimensiones y el instrumento total: las profesoras tienden a ser peor evaluadas por estudiantes de segundo ciclo y mejor por estudiantes del primero, al contrario de lo que ocurre con sus colegas masculinos. Ese efecto puede producir inversiones respecto de qué tipo de docente es mejor evaluado (como sucede en las habilidades pedagógicas o en la escala total), la generación de una brecha de género (como ocurre en el dominio disciplinario), o bien, igualar a los docentes de ambos sexos, como pasa en la dimensión responsabilidad. Es relevante notar que este efecto de interacción no es irrelevante en el conjunto de las evaluaciones obtenidas por hombres y mujeres, pues genera que ellas pierdan aproximadamente cuatro puntos porcentuales de evaluación, mientras los hombres ganan más o menos la misma cantidad, produciendo un cambio neto relativo al género de unos ocho puntos.

Figura 1: Puntuación promedio de evaluación docente según sexo del(a) profesor(a) y ciclo de estudios por factores y escala total



Finalmente, en relación con el efecto conjunto de las tres variables independientes, la Figura 2 muestra que, si bien tanto los hombres como las mujeres siguen el mismo patrón de cambio entre el primer y el segundo ciclo, este es más acentuado en los varones que en las estudiantes. Por ejemplo, la caída en la evaluación de las profesoras al pasar del primer a segundo ciclo es de menos de tres puntos porcentuales en el caso de que las evalúen las estudiantes, mientras que alcanza los seis puntos en el caso de que las califiquen los hombres.

Figura 2: Puntuación promedio de evaluación docente según sexo del(a) estudiante, sexo del(a) profesor(a) y ciclo de estudios por escala total



## Conclusiones

El principal hallazgo de nuestro estudio fue constatar que existen patrones de respuesta diferenciados por género en las evaluaciones docentes estudiadas, los que se manifiestan tanto en evaluadores como en evaluados y, en conjunto, favorecen a las académicas en los primeros años de enseñanza, pero las penalizan en los últimos. Además, en esta penalización son los varones quienes juegan un rol más decisivo.

Es interesante también notar que este patrón de género solo fue observable con claridad al implementar una estrategia de análisis multivariable, lo que es coincidente con lo afirmado por Basow (1998), respecto de que si no se introducen variables contextuales (en nuestro caso, el ciclo de estudios), los patrones de género pueden pasar desapercibidos. Por ello, el estudio de los efectos aislados de nuestras variables indicó simplemente la ya reportada tendencia de las mujeres a evaluar un poco mejor que los hombres –quizá por un sesgo de aquiescencia o por tener expectativas distintas que sus compañeros– (Aleamoni & Hexner, 1980) y una cierta propensión a que las académicas sean mejor evaluadas en responsabilidad y habilidades pedagógicas, mientras que sus colegas masculinos reciben mejores retroalimentaciones en dominio disciplinario, lo que coincide también con lo reportado por otros estudios realizados en el ámbito de las humanidades, las ciencias sociales y las comunicaciones (Tatro, 1995; Basow, 1998; Smith, et al., 2007).

Otra potencialidad de nuestra investigación fue la construcción de un instrumento para la evaluación de la docencia con muy buenas propiedades métricas y de estructura tridimensional, lo que nos permitió detectar patrones de respuesta de género diferenciados para cada factor. De esta forma, constatamos que mientras en las habilidades pedagógicas y en la escala total se produce una inversión de las evaluaciones de académicos y académicas según el ciclo de estudios, en responsabilidad se produce una tendencia hacia la igualdad y en dominio disciplinario se amplía la ventaja que los profesores ya tenían en el primer ciclo.

¿Cómo podemos interpretar estos hallazgos? Creemos que lo que hemos descrito efectivamente constituye evidencia de un sesgo de género en las evaluaciones docentes. No obstante, es importante reconocer que solo un estudio experimental podría descartar completamente la hipótesis alternativa que señala que la razón de los cambios en la evaluación de académicos y académicas según su ciclo de estudios se debe a diferencias reales en la calidad docente de quienes se desempeñan en ambos ciclos. Sin embargo, esta explicación alternativa resulta poco verosímil, ya que no parece fácil comprender cómo se habría producido esta desigual distribución en los profesores en una facultad completa.

Desde la teoría del género, por el contrario, parece coherente suponer que lo reportado es solo otra manifestación del dominio de los hombres sobre la producción y administración del conocimiento universitario (Berríos, 2005; Saracostti, 2006; Reyes, 2006; Olavarría, 2011), lo que se reflejaría en que las profesoras tienden a ser mejor evaluadas en temas asociados a la responsabilidad (considerado un ámbito “menor” por académicos y estudiantes), mientras que los profesores tienden a ser mejor evaluados en áreas relativas al conocimiento disciplinario, es decir, en temas mucho más cercanos a las fuentes del prestigio y autoridad académica.

En relación con la trascendencia que tuvo el ciclo de estudios en los patrones de respuesta de género, creemos que ello puede ser explicado de dos formas complementarias. En primer lugar, como efecto de un cambio de las demandas que los estudiantes de primeros

y últimos años hacen a sus docentes: se pasaría de una demanda de apoyo afectivo y efectivo para procesos de aprendizaje complejos, a solicitar espacios de discusión cuasidisciplinarios al interior de las aulas (Asún, Ruiz, Retamal, Peralta, Esquivel, Vargas y Martínez, 2013). Frente a la primera demanda, el estereotipo de las profesoras parece encajar muy bien, mientras que para la segunda, el estereotipo tradicional puede establecer el dominio masculino.

En segundo lugar, postulamos que los estudiantes de segundo ciclo, a diferencia de los recién llegados, ya se han socializado en los patrones tradicionales de prestigio académico en un sistema universitario tradicional como el chileno, introyectando la carga simbólica negativa atribuida a las académicas (Berríos, 2005), por lo que se encontrarían dispuestos a valorar más positivamente los esfuerzos de sus profesores, que los de sus profesoras.

En conjunto, nuestros resultados nos permiten afirmar que la relación entre el género y la evaluación docente no parece ser inocua, al menos en el contexto de las ciencias sociales donde la hemos estudiado. Además, se corrobora la importancia de las variables contextuales cuando se trata de develar el impacto que tienen las variables de género en estudios cuantitativos.

¿Pone todo esto en duda la validez de los cuestionarios de evaluación de la docencia en el caso chileno? No lo creemos así, pues si bien al igual que otros estudios (Onwuegbuzie, et al., 2007; Basow, 1995, 2000), hemos constatado la presencia significativa de factores extraclase en la valoración de los profesores, también hemos podido apreciar que ellos tienen un bajo impacto, por lo que si bien pueden producir distorsiones en las puntuaciones obtenidas, no impiden que, en conjunto, ellas sean representativas gruesamente de los desempeños docentes de los académicos, tal y como afirma Marsh (1984).

En cualquier caso, creemos que si se espera que estos instrumentos permitan tomar adecuadas decisiones de administración y política universitaria, se hace necesario interpretar sus resultados considerando las variables contextuales que inciden en mayor o menor

medida en los puntajes obtenidos por los profesores. Dentro de esos factores extraclase sin duda que asumen una importancia protagónica las variables de género, tanto del estudiante como del profesor, no solo por la importancia demostrada de estos factores a nivel internacional, sino porque la realidad regional en términos de equidad de género (Kliksberg, 2003) nos hace pensar que estas diferencias pueden ser aún más importantes en universidades latinoamericanas. En otras palabras, hasta que más estudios den cuenta de cómo operan los patrones de género en la evaluación docente, o de otros factores exógenos, se debe ser muy prudente a la hora de interpretar y ponderar los resultados de los CED. Incluso no se puede descartar la necesidad de generar sistemas de ponderación de sus resultados que reequilibren los desajustes producidos por esos extrafactores, una vez que se conozca mejor su impacto.

Entre las principales limitaciones de nuestra investigación, en primer lugar podemos señalar que pese a nuestra convicción respecto de haber encontrado sesgos de género en las evaluaciones docentes estudiadas, nuestro estudio no puede confirmar plenamente este aserto debido a su carácter no experimental. En segundo lugar, se estudió la presencia de patrones de género en las evaluaciones docentes de una sola facultad (la de Ciencias Sociales) en una única universidad, por lo que es posible que alguna particularidad institucional haya incidido en nuestros resultados (por ejemplo, realizar el estudio en el contexto de una población estudiantil principalmente femenina, pero con un plantel académico mayoritariamente masculino). En tercer lugar, se tomaron en cuenta solamente tres variables independientes en los modelos, por lo que los sesgos encontrados tienen un carácter exploratorio y faltaría profundizar en su estudio con modelos más complejos. Finalmente, nuestro estudio fue transversal por lo que, en rigor, no podemos asegurar que las diferencias en la evaluación de los estudiantes de primer y segundo ciclo no se deban a cambios generacionales, en vez de la explicación temporal que hemos propuesto nosotros.

De estas limitaciones y de la ausencia de otros trabajos en este campo en Chile surgen nuestras principales sugerencias de futuros estudios, pudiéndose realizar:

- investigaciones experimentales que permitan poner a prueba de manera más precisa la presencia de sesgos de género (o de otros factores extraclase) en la aplicación de CED;
- investigaciones en otros contextos institucionales y disciplinarios que permitan generalizar los resultados;
- investigaciones que incluyan más variables explicativas, tales como el tamaño del curso, el tipo de curso o nivel de especialización de los profesores;
- investigaciones longitudinales que profundicen en el impacto de la socialización universitaria por sobre los estereotipos de género de los estudiantes, entre otras posibilidades.

Esperamos que la realización del presente estudio sirva de inspiración para la concreción de este programa de investigación.

## Referencias

- Acker, S. (1983). Women, the other academics. *British Journal of Sociology of Education*, 1(1), 191-201.
- Aedo, A. (2010). Acreditación y aseguramiento de la calidad de los egresados. *CNED, Calidad de los egresados, responsabilidad institucional ineludible*, 145-151.
- Aguilar, A., Ampuero, N. y Loncomilla, L. (2008). Evaluación del desempeño docente: diagnóstico desde la perspectiva de los estudiantes. En *Cinda-Mineduc, Evaluación del desempeño docente y calidad de la docencia universitaria*, (pp.71-80). Santiago de Chile: Cinda-Mineduc.
- Acevedo, A.R. y Rodríguez, N.M. (2006). Factores de sesgo asociados a la validez de la evaluación docente universitaria: un modelo jerárquico lineal. *Archivos Analíticos de Políticas Educativas*, 14(34), 1-22.
- Aleamoni, L.M., & Hexner, P.Z. (1980). A review of the research on student evaluation and a report on the effect of different sets of instructions on student course and instructor evaluation. *Instructional Science*, 9(1), 67-84.
- Aleamoni, L.M. (1999). Student rating myths versus research facts from 1924 to 1998. *Journal of Personnel Evaluation in Education*, 13(2), 153-166.
- Ammermueller, A., & Dolton, P. (2006). Pupil-teacher gender interaction effects on scholastic outcomes in England and the USA. *Center for European Economic Research (ZEW)*. Recuperado el día 5 de julio de 2013, desde <http://ssrn.com/abstract=927689>

- Auster, C., & MacRone, M. (1994). The classroom as a negotiated social setting: an empirical study of the effects of faculty members' behaviour on students' participation. *Teaching Sociology*, 22(4), 289-300.
- Asún, R.A., Ruiz, S., Retamal, H., Peralta, M., Esquivel, A., Vargas, L., y Martínez, F. (2013). El desafío de educar a inicios del siglo XXI. Demandas de los estudiantes de ciencias sociales a la docencia universitaria. *Estudios Pedagógicos*, 39(1), 45-62.
- Báez, M., Cazenave, M. y Lagos, J. (2007) Percepción de la evaluación del desempeño docente: diagnóstico desde la perspectiva del docente. En *Cinda-Mineduc, Evaluación del desempeño docente y calidad de la docencia universitaria*, (pp.47-69). Santiago de Chile: Cinda-Mineduc.
- Basow, S.A. (1995). Student evaluations of college professors: When gender matters. *Journal of Educational Psychology*, 87(4), 656 - 665.
- Basow, S.A. (1998). Student evaluations: gender bias and teaching styles. En L.H. Collins, Chrisler, J.C., y Quina, K. (Eds), *Career strategies for women in academe: Arming Athena* (pp.135-156). Thousand Oaks: CA.
- Basow, S.A. (2000). Best and worst professors: gender patterns in students' choices. *Sex Roles*, 43(5-6), 407-417.
- Basow, S.A., Phelan J.E., & Capotosto, L. (2006). Gender patterns in college student's choices of their best and worst professors. *Psychology of Women Quarterly*, 30(1), 25-35.
- Berrios, P. (2005). El sistema de prestigio en las universidades y el rol que ocupan las mujeres en el mundo académico. *Revista Calidad en la Educación*, 23, 349-361.
- Bourdieu P. (2003). *El oficio de científico: ciencia de la ciencia y reflexividad*. Barcelona: Anagrama.
- Bourdieu, P., & Passeron, J. (1990). *Reproduction in education, society and culture*. Segunda Edición. Londres: Sage Publications.
- Bonder, G. (1998). *Estudios de la mujer en América Latina*. Washington: Organización de Estados Americanos.
- Cabezas, G. (2009). ¿Afecta el género de los profesores a los resultados académicos de los estudiantes? Diferencias de género en matemática y lenguaje. *Documento de Trabajo CPCE*. Recuperado el día 5 de septiembre de 2013, desde <http://www.cpce.cl/>
- Canada, K., & Pringle, R. (1995). The role of gender in college classroom interactions: A social context approach. *Sociology of Education*, 68(3), 161-186.

- Clayson, D.E., & Sheffet, M.J. (2006). Personality and the student evaluation of teaching. *Journal of Marketing Education*, 28(2), 149-160.
- Contreras, G. F., Faúndez, A., Gutiérrez, G., Jiménez, M., Ponce, F., y Silva, E. (2007). Diagnóstico del proceso e instrumentos usados actualmente para la evaluación del desempeño docente. En *Cinda-Mineduc, Evaluación del desempeño docente y calidad de la docencia universitaria* (pp. 81-98). Santiago de Chile: Cinda-Mineduc.
- Dee, T.S. (2007). Teachers and the gender gaps in student achievement. *Journal of Human Resources*, 42(3), 528-554.
- Elosua O.P., y Zumbo, B.D. (2008). Coeficientes de fiabilidad para escalas de respuesta categórica ordenada. *Psicothema*, 20(4), 896-901.
- Feldman, K.A. (1992). College students' views of male and female college teachers: Part I—Evidence from the social laboratory and experiments. *Research in Higher Education*, 33(3), 317-375.
- Feldman, K.A. (1993). College students' views of male and female teachers: Part II—evidence from students' evaluations of their classroom teachers. *Research in Higher Education*, 34(2), 151-211.
- Flores, R. (2007). Representaciones de género de profesores y profesoras de Matemática, y su incidencia en los resultados académicos de alumnos y alumnas. *Revista Iberoamericana de educación*, 43, 103-118.
- Garduño, J.M.G. (2000). ¿Qué factores extraclase o sesgos afectan la evaluación docente en la educación superior? *Revista Mexicana de Investigación Educativa*, 5(10), 303-325.
- Garduño, J.M.G. (2003). Profesores universitarios y su efectividad docente. Un estudio comparativo entre México y Estados Unidos. *Perfiles Educativos*, 25(100), 42-55.
- Garduño, J.M.G. (2008). El proceso perverso de la evaluación de la docencia en las universidades: un balance inicial y apuntes para mejorarlo. *Reencuentro*, 53, 9-19.
- García, M.I.A., y Beltrán, M.R. (2003). La evaluación de la docencia desde la perspectiva del propio docente. *Reencuentro*, 36, 56-64.
- Greenwald, A.G., & Gillmore, G.M. (1997). No pain, no gain? The importance of measuring course workload in student ratings of instruction. *Journal of Educational Psychology*, 89(4), 743-751.
- Griffin, B.W. (2001). Instructor reputation and student ratings of instruction. *Contemporary Educational Psychology*, 26(4), 534-552.

- Hall, R.M., & Sandler, B.R. (1982). The classroom climate: a chilly one for women? *Project on the status and education of women*. Recuperado el día 12 de junio de 2013 desde [http://www.hws.edu/offices/provost/pdf/classroom\\_climate.pdf](http://www.hws.edu/offices/provost/pdf/classroom_climate.pdf)
- Haskell, R.E. (1997). Academic freedom, tenure, and student evaluation of faculty: galloping polls in the 21st century. *Education Policy Analysis Archives*, 5(6), 1-36.
- Hills, S.B., Naegle, N., & Bartkus, K.R. (2009). How important are items on a student evaluation? A study of item salience. *Journal of Education for Business*, 84(5), 297-303.
- Holmlund, H., & Sund, K. (2008). Is the gender gap in school performance affected by the sex of the teacher? *Labour Economics*, 15(1), 37-53.
- Kliksberg, B. (2003). La discriminación de la mujer en el mundo globalizado y en América Latina. Un tema crucial para las políticas públicas. *Economía y Desarrollo*, 2(1), 6-34.
- Kulas, J.T., Stachowski, A.A., & Haynes, B.A. (2008). Middle response functioning in Likert-responses to personality items. *Journal of Business and Psychology*, 22(3), 251-259.
- Marsh, H.W., & Roche, L.A. (2000). Effects of grading leniency and low workload on students' evaluations of teaching: popular myth, bias, validity, or innocent bystanders? *Journal of Educational Psychology*, 92(1), 202-228.
- Marsh, H.W. (1984). Students' evaluation of university teaching: dimensionality, reliability, validity, potential biases, and utility. *Journal of Educational Psychology*, 76(5), 707-754.
- Montes, C.S. (2009). Las encuestas de opinión en la Universidad Autónoma de Ciudad Juárez: ¿un caso de exclusión del profesorado? *Education Policy Analysis Archives*, 17(24), 1-31.
- OECD (2013). *Revisión de políticas nacionales de educación: el aseguramiento de la calidad en la educación superior en Chile 2013*. Santiago de Chile: OECD Publishing.
- Olavarría J. (2011). La participación y el liderazgo de las mujeres que participan en la educación superior (IES) en las Américas. En *Informe final Red interamericana de formación en mujeres, género y desarrollo con equidad*. Recuperado el día 2 de diciembre del 2012, desde <http://www.ouiiohe.org/webcolam/rifged/iu/search.php?q=olavarr%C3%ADa>
- Onwuegbuzie, A.J., Witcher, A.E., Collins, K.M., Filer, J.D., Wiedmaier, C.D., & Moore, C.W. (2007). Students' perceptions of characteristics of effective college teachers: a validity study of a teaching evaluation

- form using a mixed-methods analysis. *American Educational Research Journal*, 44(1), 113-160.
- Reyes, J. (2006). Trabajadoras (es) de la educación superior y reproducciones de género. *Revista Calidad en la Educación*, 24, 397-422.
- Salazar, J. (2008). Diagnóstico preliminar sobre evaluación de la docencia universitaria. Una aproximación a la realidad en las universidades públicas y/o estatales de Chile. *Revista Iberoamericana de Evaluación Educativa*, 1(3), 67-84.
- Sagot, M. (2002). Los estudios feministas en Centroamérica: contándonos a nosotras mismas desde la academia. En UNAM-PUEG, *Seminario: Feminismos Latinoamericanos Retos y Perspectivas*. Ciudad de México: UNAM-PUEG.
- Saracostti, M. (2006). Mujeres en la alta dirección de educación superior: posibilidades, tensiones y nuevas interrogantes. *Revista Calidad en la Educación*, 25, 243-259.
- Smith R.A., & Cranton, P.A. (1992). Students' perceptions of teaching skills and overall effectiveness across instructional settings. *Research in Higher Education*, 33(6), 747-764.
- Smith, S.W., Yoo, J.H., Farr, A.C., Salmon, C.T., & Miller, V.D. (2007). The influence of student sex and instructor sex on student ratings of instructors: results from a college of communication. *Women's Studies in Communication*, 30(1), 64-77.
- Smith, B.P. (2009). Student rating of teaching effectiveness for faculty groups based on race and gender. *Education*, 129(4), 615-624.
- Social Watch. (2012). *Social Watch's Gender Equity Index (GEI)*. Recuperado el día 6 de septiembre del 2013 desde <http://www.socialwatch.org/es/node/14380>
- Sprague, J., & Massoni, K. (2005). Student evaluations and gendered expectations: what we can't count can hurt us. *Sex Roles*, 53(11-12), 779-793.
- SPSS Inc., (2010). *Spss para Windows, versión 19*. Chicago: SPSS Inc.
- Tatro, C.N. (1995). Gender effects on student evaluations of faculty. *Journal of Research and Development in Education*, 28(3), 169-173.
- Valsan, C., & Sproule, R. (2008). The invisible hands behind the student evaluation of teaching: the rise of the new managerial elite in the governance of higher education. *Journal of Economic Issues*, 42(4), 939-958.

- Van Vucht Tijssen, L., & De Weert, E. (2005). From erudition to academic competence. *Revista Española de Pedagogía*, 63(230), 123-145.
- Vásquez R.F. y Gabalán, C.J. (2006). Percepciones estudiantiles y su influencia en la evaluación del profesorado. Un caso en la Universidad Autónoma de Occidente. *Revista RELIEVE*, 12(2), 219-245.
- Zabaleta, F. (2007). The use and misuse of student evaluation of teaching. *Teaching in Higher Education*, 12(1), 55-76.

Recibido: 18/10/2013

Aceptado: 25/04/2014

## Anexo 1

## Instrumento de evaluación docente

	Muy de acuerdo	De acuerdo	En desacuerdo	Muy en desacuerdo	No aplicable
1. El profesor inició y terminó puntualmente su clase.					
2. El profesor fue puntual en la entrega de evaluaciones.					
3. El profesor asistió regularmente a clases.					
4. El profesor domina los contenidos de su clase.					
5. El profesor se comunica con claridad (al responder preguntas, explicar contenidos, etc.).					
6. Los contenidos expuestos por el profesor están actualizados y tienen relación con la realidad actual.					
7. El profesor presentó variados enfoques teóricos referentes a la materia.					
8. El profesor utiliza estrategias de enseñanza que estimulan la participación de los estudiantes.					
9. El profesor demuestra compromiso con el desarrollo del aprendizaje de los estudiantes.					
10. El profesor muestra disposición a enseñar y aclarar dudas, tanto dentro como fuera de la hora de clases.					
11. El profesor acerca al estudiante al conocimiento de la disciplina a través de situaciones concretas o ejemplos.					
12. El profesor logra un trato equitativo y respetuoso en su relación con los estudiantes.					
13. El profesor es consistente entre lo enseñado y lo exigido en las evaluaciones.					
14. El profesor prepara materiales útiles y pertinentes para el aprendizaje.					
15. El profesor entrega el programa oportunamente [al inicio del semestre].					
16. El profesor entregó pautas claras de corrección de pruebas y trabajos.					

17. En síntesis y en términos generales, ¿Cómo calificarías al docente?:

Muy bueno	Bueno	Regular	Malo	Muy malo