

¿QUÉ VALORAN MÁS LOS ESTUDIANTES DE SUS PROFESORES? UN ANÁLISIS A PARTIR DE CUESTIONARIOS DE OPINIÓN DE EVALUACIÓN DE LA DOCENCIA¹

WHAT MATTERS THE MOST TO STUDENTS? AN ANALYSIS OF OPINION QUESTIONNAIRES FOR TEACHING EVALUATION

Luis Medina Gual

Daniela Alvarado García

Luis
Medina Gual

Doctorando en Evaluación Educativa, Universidad Anáhuac. International Center for Integral Formation (ICIF). Facultad de Educación, Universidad Anáhuac y Universidad Complutense. Estudios de maestría en Investigación y Desarrollo de la Educación, Universidad Iberoamericana, becario CONACyT. Licenciado en Pedagogía, Universidad Anáhuac. Actualmente consultor responsable de Bachillerato Anáhuac en México y El Salvador para el ICIF. Su principal línea de investigación es la Evaluación educativa y del aprendizaje. Funge como evaluador de revistas arbitradas.

Correo electrónico: [medinagual@gmail.com].

Daniela
Alvarado
García

Estudiante de Licenciatura, Facultad de Educación, Universidad Anáhuac.

Correo electrónico: [dany_dag_94@hotmail.com].

¹ Se agradece el apoyo de la Red de Colegios Semper Altius para la realización del estudio.

RESUMEN

El presente estudio busca realizar un aporte metodológico en la construcción y el análisis de cuestionarios de opinión de estudiantes para la evaluación de la docencia. Se realizó un ejercicio de elaboración y validación de un instrumento que, además de los procedimientos habituales de validación, contempló al estudiante a través de entrevistas cognitivas. La aplicación final en n=31136 ocasiones en secundarias y bachilleratos de una red de colegios particulares en México y El Salvador, permitió realizar un análisis de modelación supervisada a través de árboles de decisión que mostró una alternativa estadística para el análisis de la importancia de reactivos en este tipo de instrumentos. En esta tónica, se observaron diferencias en la importancia de los ítems, según si el docente fue evaluado con puntajes altos o bajos. En los docentes evaluados con puntajes altos, los ítems de mayor importancia se refirieron a temáticas sobre didáctica y actitudes y clima del aula, mientras que en los docentes evaluados con puntajes bajos los ítems de mayor importancia correspondían a reactivos del área de actitudes y clima del aula.

Palabras clave: evaluación de profesores; competencias docentes; concepciones del profesor; perfil del profesor; percepción del estudiante.

ABSTRACT

The main objective of this study was to make a methodological contribution to the construction and analysis of opinion questionnaires for teaching evaluation. In this sense, an assessment instrument was developed and validated. This instrument used traditional validation procedures and the use of cognitive interviews. The final sample was composed of n=31136 assessments in middle and high schools from a network of private schools in Mexico and El Salvador. It was possible to assess the importance of items through the use of supervise modeling techniques (specifically by the use of decision trees analysis). Differences were observed in the item im-

portance depending to the overall assessment of the teachers. When a teacher was assessed with high scores the items that appear to be more important where related to issues concerning didactics, teacher attitudes and classroom climate. In contrast, the items about teacher attitudes and classroom climate were more important when the teachers were evaluated with overall low scores.

Keywords: teacher effectiveness; teacher evaluation; teaching skills; student's perceptions.

INTRODUCCIÓN

Dentro de un modelo educativo, un componente esencial es la definición del perfil docente. En este sentido, mucho ha sido escrito sobre el perfil docente y las competencias docentes que identifican a un profesor «de calidad» con base en características propias de su propia labor (OCEANO, 1999). Ejemplo clásico de lo anterior es el texto de Perrenoud (2007) que define diez familias de competencias para la enseñanza. De manera similar, organizaciones internacionales han seguido el mismo ejercicio (ISTE, 2004) y gobiernos de diferentes países han definido perfiles docentes con esta misma lógica (Ministerio de Educación Chileno, 2012; Secretaría de Educación Pública, 2013).

En esta tónica, lo que se «desea» de un docente se ha conceptualizado a partir de estudios empíricos y disertaciones teóricas sobre la profesión. Es decir, la definición del perfil y las competencias parte de académicos, autoridades educativas e inclusive de consultas a los mismos docentes. Valdría la pena examinar lo que lo estudiantes consideran como un docente «de calidad». Coherente con lo anterior, desde 1976, Feldman habría discutido la dicotomía de cómo las competencias o los rasgos docentes definidos pueden o no ser de relevancia para los estudiantes.

Diferentes estudios de habla inglesa reflexionan sobre ello. Un primer ejemplo sería el trabajo de Jackson (1999) que replica el trabajo

de Burdsal y Bardo (1986 en Jackson, 1999) al aplicar un cuestionario a estudiantes que permite definir seis factores de importancia para los alumnos: el *rapport* para con los alumnos; la valía o el sentido de valía del curso para el estudiante; la organización y el diseño del curso; la justicia en la evaluación y las calificaciones; la dificultad percibida y la carga de trabajo.

Catano y Harvey (2011) realizaron un trabajo posterior, a través de incidentes críticos de los estudiantes, deduciendo nueve competencias que, desde la perspectiva de los alumnos, definen la eficacia docente: comunicación; disponibilidad; creatividad; consideración sobre o al estudiante; conciencia social; habilidades de retroalimentación; profesionalismo; empatía y habilidades de resolución de problemas. Así, llegan a la conclusión que el profesor eficaz, a la luz de sus alumnos, es quien puede considerarse como un líder «transformacional».

En Nigeria, Aregbeyen (2010) identifica, a través de una encuesta a estudiantes, que les importan en primer lugar las «explicaciones claras», seguido de «el interés y la preocupación del docente por la calidad de su enseñanza», «la capacidad para ver si el alumno ha comprendido la lección», que el docente «discuta temáticas vigentes del campo disciplinar», el «respeto por los estudiantes como personas». Por otra parte, en un estudio cualitativo, Delaney *et al.* (2010) identifican rasgos del profesor eficaz: respetuoso, con conocimientos o formación, cercano, motivante, comunicativo, organizado, proactivo, profesional y con sentido del humor.

En un esfuerzo de síntesis, sería posible agrupar en tres áreas los rasgos o las características identificadas en los estudios: características sobre la didáctica de la clase, sobre la evaluación y sobre las actitudes del docente y el clima del aula. A continuación se muestra un cuadro con lo antes descrito.

**CUADRO 1. RASGOS DE IMPORTANCIA PERCIBIDOS
POR LOS ESTUDIANTES EN CUATRO ESTUDIOS**

Área	Jackson (1999)	Catano y Harvey (2011)	Aregbeyen (2010)	Delaney <i>et al.</i> (2010)
Didáctica:	Organización y diseño del curso.		Explicaciones claras.	Conocimiento disciplinar.
	Dificultad percibida.		Discusión de temáticas vigentes.	Organización.
	Carga de trabajo.			Capacidad de motivar. Comunicación.
Evaluación:	Justicia en la evaluación.	Habilidades de retroalimentación.	Capacidad para corroborar que el alumno comprenda.	
Actitudes y clima del aula:	<i>Rapport.</i>	Comunicación.	Interés por la calidad de la enseñanza.	Respeto.
	Valor del curso.	Disponibilidad.	Respeto por el estudiante.	Cercanía.
		Consideración al estudiante.		Proactividad.
		Conciencia social.		Profesionalidad.
		Profesionalismo.		Humor.
		Empatía.		
		Habilidades de resolución de problemas.		

Fuente: Elaboración propia.

Finalmente, para la educación media superior, el estudio de Sutcliff (2011) tenía a bien discutir la existencia de diferencias en la percepción de la eficacia docente con base en características étnicas de

los estudiantes. En el trabajo no se encontraron mayores diferencias según la etnicidad de los alumnos, e identifican la claridad de las expectativas para el logro y éxito en la clase como el rasgo de mayor importancia.

Por otra parte y como en cualquier proceso educativo, la definición de una serie de competencias, permite también el diseño de un mecanismo de seguimiento y mejora continua que posibilita, en este caso, la evaluación de la docencia. Coherente con lo anterior, el desarrollo de propuestas de evaluación de la docencia ha evolucionado con la misma teoría evaluativa (*cf.* Alkin, 2004). Diversas experiencias de evaluación han sido registradas. Quizá el instrumento más difundido es el cuestionario de opinión de estudiantes (*cf.* Arregui, Chaparro y Díaz, 2015), debido a la influencia de la psicometría de inicios y mediados del siglo XX, y a la relativa facilidad de recopilación y análisis de datos.

Al día de hoy se reconoce como necesaria la existencia de multiplicidad de evidencias que resulten en «modelos» de evaluación de la docencia (Schmelkes, 2015). Empero, el cuestionario de opinión sigue siendo el instrumento más común y su construcción generalmente implica el uso de procesos de validación clásicos (validación de jueces, validación estadística, confiabilidad clásica, etcétera). En todos ellos, los actores que participan son quienes definen las competencias del perfil docente ideal: investigadores y académicos, autoridades educativas y docentes.

Quizá el gran ausente es, nuevamente, el mismo actor que responde al instrumento y es centro del proceso educativo: el estudiante mismo (Catano y Harvey, 2011). La «entrevista cognitiva» es el método de validación que ha sido poco explotado en el desarrollo de cuestionarios de opinión (en comparación a otros métodos) y que contempla a este actor. Dicho método consiste en:

[...] una serie de entrevistas individuales semiestructuradas en ambiente controlado con una muestra pequeña de la población meta. Durante las entrevistas, las personas participantes completan el cuestionario en estudio y realizan una serie de pruebas para detectar problemas a la hora de contestarlo (Willis, 2005, en Molina y Smith, 2011).

El desarrollo de este método funda su práctica en la necesidad evidente de que las personas comprendan el significado o la intención de los reactivos dentro del instrumento, para dar paso a la recuperación de la información solicitada de manera autobiográfica, a la estimación de la respuesta adecuada y, finalmente, a la emisión de la respuesta elegida. Este modelo general se conoce como el Modelo Tourangeau y consta de cuatro etapas: a) comprensión; b) recuperación; c) estimación, y d) ejecución de la respuesta (Molina y Smith, 2011).

Es precisamente en esta tónica, que el presente artículo busca realizar un aporte, colocando al estudiante como centro del proceso evaluativo de los docentes. Para ello se desarrollará un instrumento que, si bien se basa en un perfil institucional de competencias docentes (ICIF, 2013), se realizará un esfuerzo de validación que permita la aportación de los estudiantes a través de entrevistas cognitivas y cuyos resultados identifiquen el tipo de temáticas a las que ellos conceden mayor importancia al momento de evaluar a sus profesores. En este sentido la pregunta central a responder será: ¿cuáles son las temáticas que «importan» más a los estudiantes sobre sus profesores?

METODOLOGÍA

El presente estudio es de corte cuantitativo, no experimental, descriptivo y correlacional. La población meta consta de estudiantes pertenecientes a una red de colegios particulares confesionales, con presencia en México y un colegio en El Salvador. A continuación se describe el proceso de elaboración y validación del instrumento empleado para el presente estudio.

Proceso de Validación del Instrumento

El instrumento diseñado parte del perfil docente definido por el ICIF (2013). Esta organización atiende a una red de colegios y

define diferentes documentos de tipo curricular. En este sentido, el perfil docente señala la existencia de diez competencias que se desean fomentar en los profesores de la red. A partir de ellas se definieron una serie de 18 reactivos para conformar el pilotaje del instrumento de evaluación de la docencia, a través de las opiniones de los estudiantes. Coherente con el referente teórico mostrado en la sección anterior, los ítems se agruparon en tres dimensiones de rasgos que la literatura menciona como de importancia para los estudiantes (cuadro 1): didáctica, evaluación y actitudes del docente, y clima del aula.

El proceso de validación del instrumento se realizó durante tres etapas. Una inicial, donde se envió la primera propuesta de reactivos a un panel de tres expertos (investigadores educativos) y de tres coordinadores de la misma red de colegios donde se realizó el proyecto. Los seis jueces contaban con un instrumento guía que facilitó la recolección de sus opiniones.

Posteriormente, la segunda versión fue aplicada en $n=25945$ ocasiones (un estudiante podía evaluar a diferente cantidad de profesores). De las aplicaciones mencionadas $n=7303$ fueron realizadas en el segundo semestre del ciclo escolar 2014-2015 y $n=18642$ en el primer semestre del ciclo escolar 2015-2016. En estas aplicaciones se realizaron análisis psicométricos de confiabilidad clásica y de validez.

En cuanto a la confiabilidad se obtuvo una Alfa de Cronbach de 0.879 para los 17 ítems del instrumento. Exceptuando dos reactivos, se observan discriminaciones ítem-total mayores a 0.30. Por otra parte, la validez se corroboró a partir de los análisis factoriales exploratorios. En éstos se obtuvieron comunalidades mayores a 0.50 con tres factores extraídos a partir de autovalores mayores a 1.0 y una varianza total explicada del 61.17%. La rotación por componentes principales devela que los dos primeros factores traslapan la totalidad de los ítems.

Con base en estos análisis se realizó un tercer esfuerzo de validación a través de entrevistas cognitivas. Este ejercicio se llevó a cabo en uno de los colegios de la misma red y contó con la participación de $n=12$ estudiantes de secundaria y $n=12$ estudiantes de bachillerato. Con la información recolectada en este ejercicio se concretó una tercera versión del instrumento, incrementando dos reactivos y afinando

términos sustentados en las recomendaciones de los alumnos. A partir de las entrevistas también fue posible corroborar que los reactivos del instrumento medían rasgos fundamentales para los estudiantes meta.

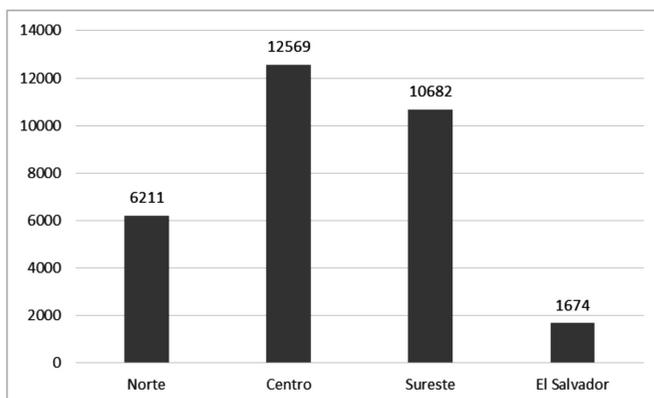
El presente estudio parte de la aplicación de la versión antes mencionada en $n=31136$ ocasiones durante el segundo semestre del ciclo 2015-2016. Esta vez se obtuvo un Alfa de Cronbach de 0.936 con $n=20$ reactivos y discriminaciones mayores a 0.30, exceptuando a cuatro reactivos. A través de un análisis factorial exploratorio se observaron comunalidades mayores a 0.50 con dos factores creados de autovalores mayores a 1.0 y una varianza total explicada del 64.38%. El instrumento final puede consultarse en los anexos.

RESULTADOS

Descripción de la muestra

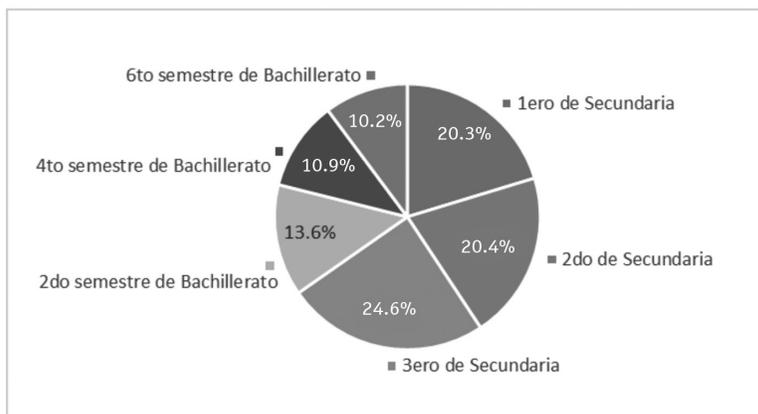
Como se mencionó con anterioridad, la muestra final se compuso por $n=31136$ aplicaciones del instrumento (un alumno podía evaluar a más de un docente). Dicha muestra se conformó por 28 diferentes colegios de la República Mexicana y un colegio de El Salvador. La siguiente gráfica representa la distribución de las $n=31136$ aplicaciones según la región de procedencia.

**DISTRIBUCIÓN DE LAS N=31136 APLICACIONES
POR REGIÓN DE MÉXICO Y EL SALVADOR**

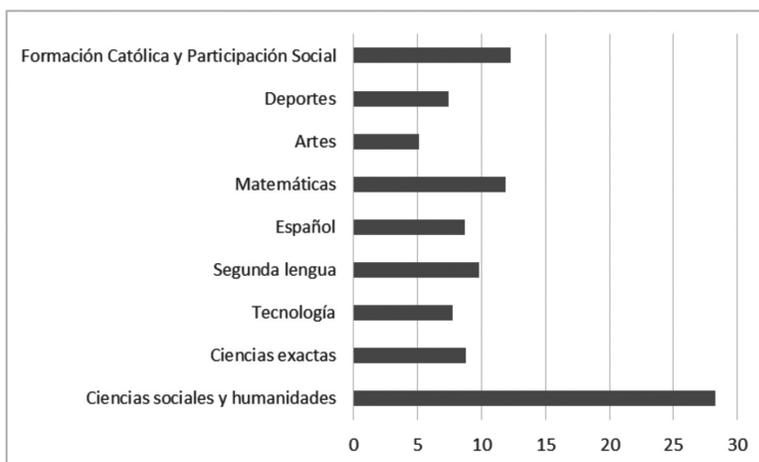


En cuanto al sexo de los respondientes, el 52.7% fueron hombres y el 47.3% mujeres. Finalmente, las siguientes gráficas muestran la distribución de las aplicaciones, dependiendo del grado del estudiante y del área disciplinar de la asignatura.

DISTRIBUCIÓN DEL GRADO DE LOS ESTUDIANTES (N=31136)

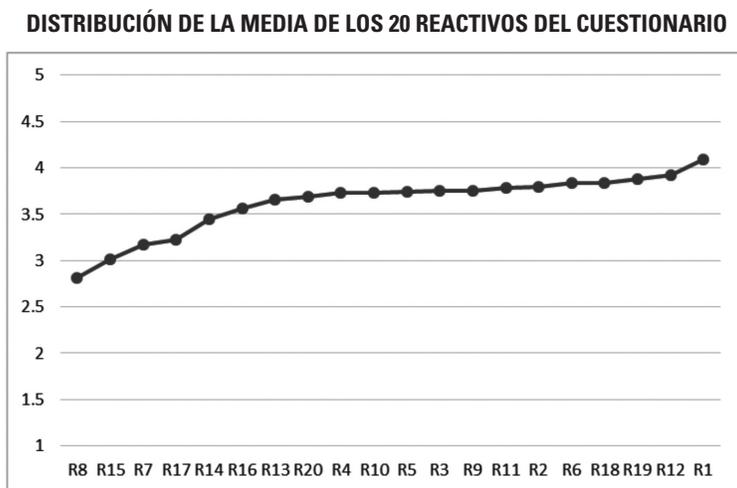


DISTRIBUCIÓN DE LAS APLICACIONES SEGÚN EL ÁREA DE CONOCIMIENTO DEL DOCENTE EVALUADO (N=31136)



Estadística Descriptiva

La escala final se compuso por 20 reactivos en una escala Likert de 5 intensidades (1= totalmente en desacuerdo *vs.* 5= totalmente de acuerdo). La siguiente figura muestra la distribución de la media de los reactivos, ordenados de menor a mayor puntaje promedio:



Como es posible apreciar en la figura anterior, existen reactivos cuya media se encuentra por debajo del punto intermedio de la escala. En estos casos, los cuatro reactivos de menor puntaje corresponden a reactivos de direccionalidad inversa (ejemplo: «R7. Sólo asigna calificaciones, sin decirme qué tengo que mejorar»). De igual forma se observan reactivos cuya media se encuentra cercana a los niveles superiores de la escala.

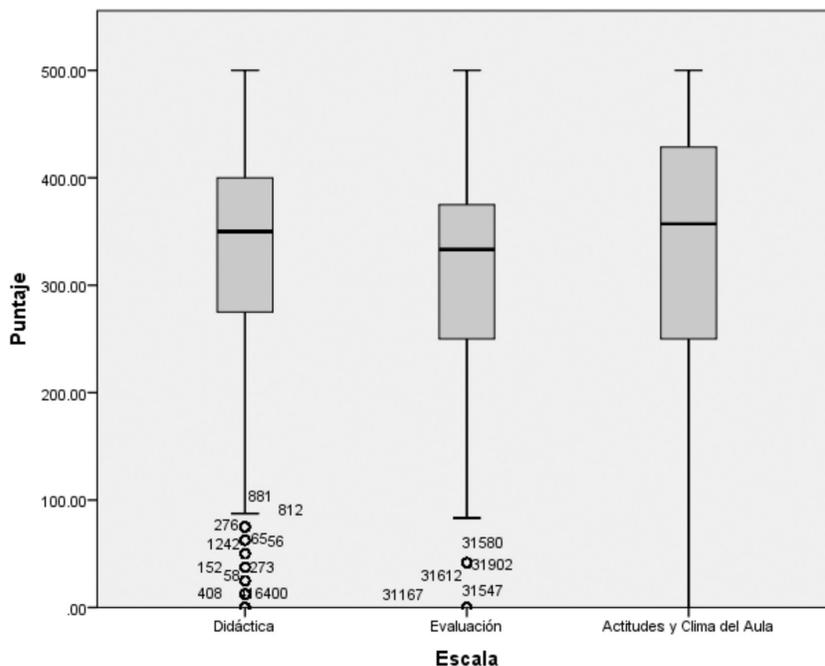
Con el fin de realizar reportes a los colegios involucrados, se optó por recodificar la escala Likert de cinco intensidades. Para ello se agruparon los ítems en tres áreas diferentes: didáctica (diez ítems, Alfa de Cronbach de 0.861), evaluación (tres ítems, Alfa de Cronbach de 0.527), y actitudes y clima del aula (siete ítems, Alfa de Cronbach de 0.848). Una vez agrupados se recodificó la escala desde una lógica criterial de la siguiente manera:

CUADRO 2. ESCALA Y NIVELES DE DESEMPEÑO DEL INSTRUMENTO

Puntaje en escala 1-5	Puntaje en escala 0-500	Nivel de desempeño	Descripción	Porcentaje ideal de distribución de profesores
<2	<125	No satisfactorio	El docente muestra un rendimiento que permite cuestionar la permanencia del docente frente al grupo.	0%
[2-3)	125-249	Seguimiento	El docente requiere seguimiento continuo por parte del coordinador.	20%
[3-4)	250-374	Satisfactorio	El docente muestra un desempeño suficiente para los estándares de la red.	40%
[4-4.5)	375-429	Sobresaliente	El docente muestra un desempeño más allá de los estándares básicos esperados por la red.	40%
>=4.5	>450	Excepcional	El docente muestra un desempeño que amerita seguimiento y reconocimiento.	20%

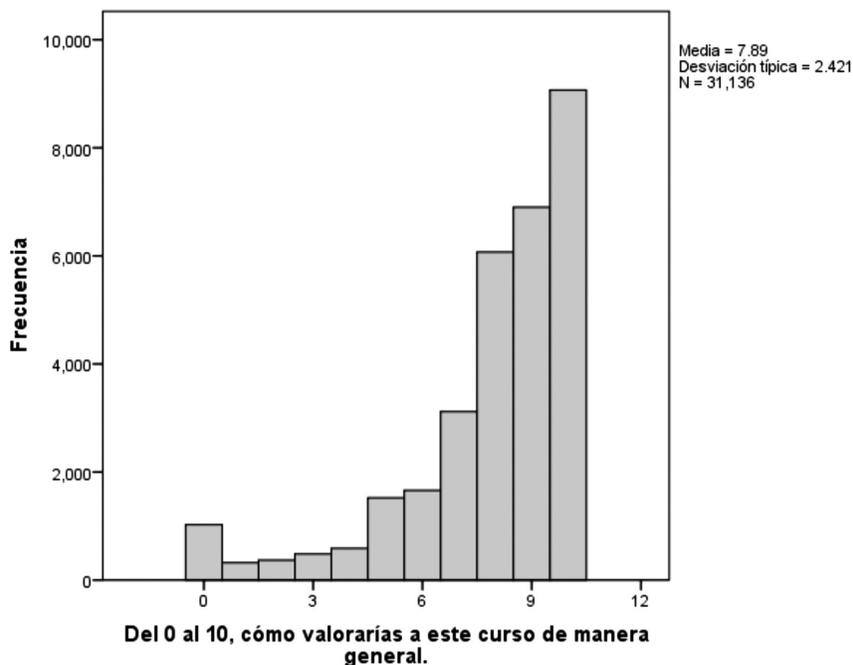
Como es posible observar en la gráfica siguiente, la mediana de los puntajes de las escalas empleadas se encuentra cercana a los 350 puntos. De igual forma, el segundo cuartil se halla por arriba de los 250 puntos, lo que ubicaría a tres cuartas partes de los docentes en puntajes «satisfactorios». También es posible advertir la existencia de mayor heterogeneidad en la escala correspondiente al área de «actitudes y clima del aula», en comparación a las otras dos escalas.

DISTRIBUCIÓN DE LOS PUNTAJES DE LAS ESCALAS (N=31136)



Finalmente, se incluyeron tres ítems de validación que buscaban una valoración general del curso por parte del estudiante. Por los fines de este estudio, únicamente se empleará el ítem que requería que el estudiante valorara el curso, de manera global, asignando una calificación que podría ir del cero al diez (un total de 11 intensidades). La gráfica que se muestra a continuación sugiere que dicho ítem parece encontrarse cargado a la derecha de la distribución con algunos casos extremos del lado izquierdo (alumnos que valoraban con el mínimo puntaje el curso del profesor). Realizando un ejercicio de correlación entre este ítem de validación y las escalas mostradas, se observan coeficientes de correlación Pearson medianos y altos (*vs.* didáctica de 0.68, *vs.* evaluación de 0.5 y *vs.* actitudes y clima del aula de 0.67).

DISTRIBUCIÓN DEL PUNTAJE ASIGNADO POR LOS ESTUDIANTES AL ÍTEM DE VALIDACIÓN (N=31136)



Análisis Diferencial

Con el fin de contestar a la pregunta de investigación inicial, se realizaron distintos tipos de análisis. Unos primeros, consistentes en análisis clásicos de inferencia con el fin de explorar la existencia de diferencias según características en la muestra y del área del conocimiento de los docentes. Sin embargo, es importante advertir que, dado el tamaño de la muestra, se espera observar diferencias estadísticamente significativas con relativa facilidad.

**CUADRO 3. PRUEBAS DE HIPÓTESIS ENTRE VARIABLES
DEMOGRÁFICAS Y LAS ESCALAS DE LA PRUEBA**

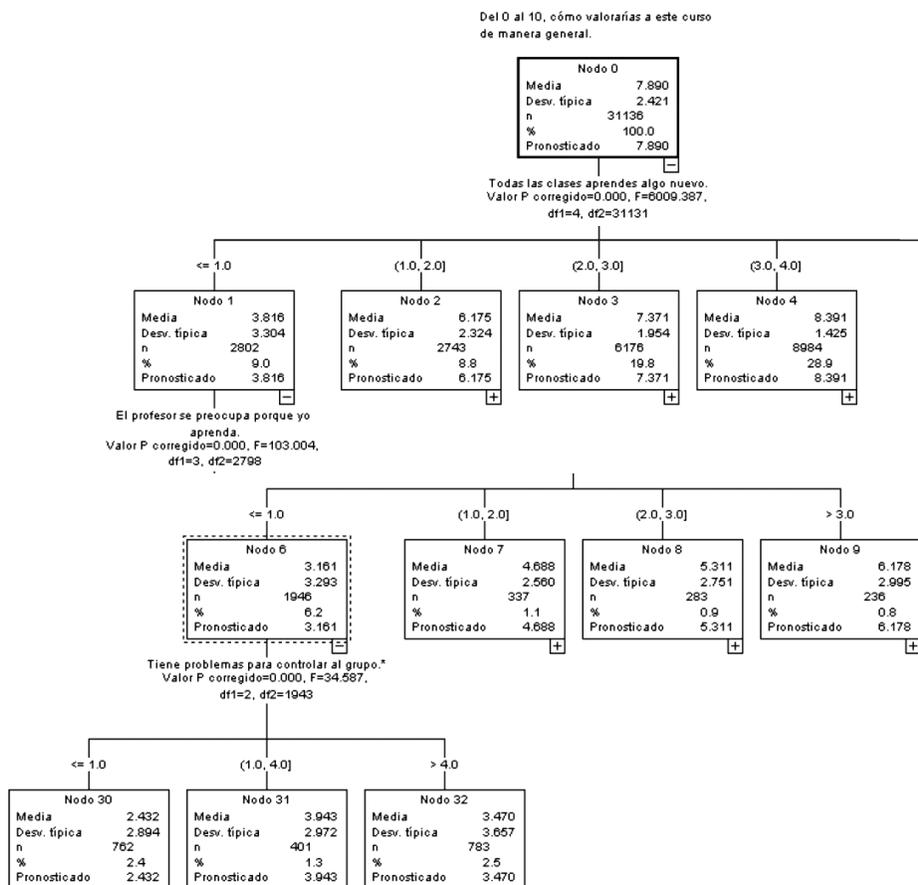
Variable	Prueba	Didáctica			Evaluación			Actitudes y clima del aula		
		vp	gl	p	vp	gl	p	vp	gl	p
Sexo.	t para datos independientes.	3.3	31049.5	0.001	8.6	31022.5	0.000	2.8	31120.9	0.006
Área disciplinar.	ANOVA de un factor.	69.2	31135	0.000	13.2	31135	0.000	46.1	31135	0.000
Grado que cursas.	ANOVA de un factor.	54.9	31135	0.000	36.4	31135	0.000	45.6	31135	0.000
Región del colegio.	ANOVA de un factor.	31.3	31135	0.000	32.4	31135	0.000	33.6	31135	0.000

*vp= valor en prueba, gl= grados de libertad, p=probabilidad asociada.

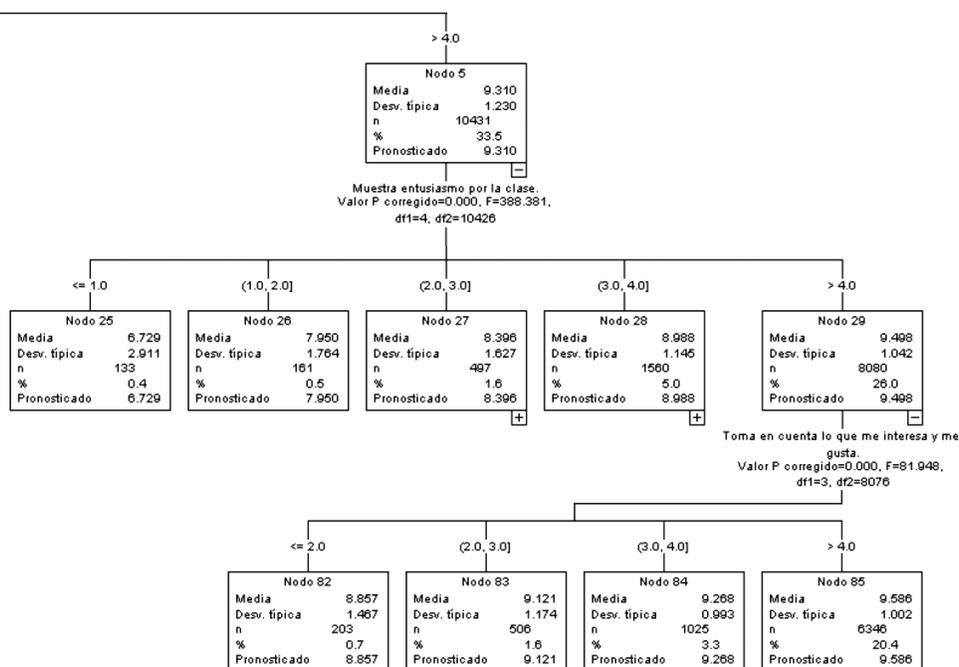
Análisis Supervisado

Debido al tamaño de la muestra, las diferencias encontradas en el apartado sobre estadística inferencial —aunque valiosas para la retroalimentación de los colegios—, requieren complementarse con otro tipo de análisis si se desea responder a la pregunta de investigación inicial. Por este motivo se plantea la posibilidad de recurrir a análisis de árboles, tomando como variable dependiente al ítem de validación descrito en la sección anterior. El análisis de árbol empleó como método de crecimiento CHAID con muestras plegadas de $n=10$ y criterio de poda de probabilidades asociadas de 0.05. Por el tamaño de la muestra, el límite de crecimiento se fijó en 200 para el nodo parental y 100 para el nodo filial.

ANÁLISIS DE ÁRBOL CON VALIDACIÓN CRUZADA DE 10 MUESTRAS, MÉTODO DE CRECIMIENTO CHAID



ANÁLISIS DE ÁRBOL CON VALIDACIÓN CRUZADA DE 10 MUESTRAS, MÉTODO DE CRECIMIENTO CHAID



La figura anterior muestra el diagrama de árbol resultado del procedimiento antes descrito. Aunque la profundidad de todas las ramas es de cuatro nodos, se ha decidido colapsar la mayoría de éstos con fines descriptivos. Únicamente se muestran de manera total los nodos inferior y superior. El inferior corresponde al docente con menor puntaje en el ítem de validación y el superior al docente con mayor puntaje en el ítem de validación. Cada rama genera «reglas» de clasificación que tienen a bien incrementar la probabilidad de acertar el valor que cada estudiante dará a cada profesor en el ítem de validación. Para predecirlo, se utilizan los ítems Likert del cuestionario de evaluación. El procedimiento pone a «competir» los ítems para optimizar la probabilidad de acierto al ítem de validación. Bajo esta lógica, el ítem que tiene mayor importancia de predicción es el que realiza la primera división. Este ítem corresponde a la respuesta a: «Todas las clases aprendes algo nuevo». Los estudiantes que responden «Totalmente de acuerdo» a este ítem, se les predice un valor en el ítem de validación de 9.310; en contraste con los que responden «De acuerdo», con un valor de 8.391; «Indeciso» con 7.371; «En desacuerdo» con 6.175, y «Totalmente en desacuerdo» con 3.816.

En esta tónica y desde este método, se advierte una jerarquía en la importancia de los reactivos. Así, el docente mejor evaluado parece mejorar el puntaje predicho si los alumnos responden «totalmente de acuerdo», en los ítems «Muestra entusiasmo por la clase» y «Toma en cuenta lo que me interesa y me gusta». Por el contrario, las reglas de clasificación sugieren que un docente con baja evaluación posee menor puntaje en el ítem de validación, además del ítem «Todas las clases aprendes algo nuevo», el ítem «El profesor se preocupa porque yo aprenda» y el reactivo «Tiene problemas para controlar el grupo».

Coherente con este método, si se realiza una predicción del ítem de validación a través de una regresión lineal múltiple por pasos sucesivos, se observa que los ítems con mayor varianza explicada son, en primer lugar con el 41.8% de la varianza, el reactivo «Todas las clases aprendes algo nuevo»; en segundo lugar con el 6.6% de la varianza explicada, el reactivo «El profesor se preocupa porque yo aprenda», y en tercer lugar con el 2.5% de la varianza, el reactivo «Domina la

materia que enseña». Así, dos de los tres ítems de mayor importancia a través de este método clásico son coherentes con los resultados del análisis de árbol.

Finalmente, para observar las variables demográficas contempladas en el análisis inferencial, se volvió a realizar un análisis de árbol con la misma variable de validación como dependiente, y como independientes la totalidad de los 20 reactivos del cuestionario y las variables demográficas. A pesar de encontrar diferencias estadísticamente significativas en el apartado de análisis inferencial, se observa que el árbol que se produce es idéntico al árbol descrito con anterioridad en este apartado. Ello corrobora lo argumentado con anterioridad sobre cómo el tamaño de muestra puede facilitar la existencia de probabilidades asociadas bajas.

DISCUSIÓN

El presente trabajo buscar realizar una propuesta «mixta» de aproximación a la *evaluación de la evaluación docente* a través de cuestionarios de opinión. La propuesta «mixta» se refiere a la posibilidad de disponer de un perfil «teórico» e idea de las competencias que desea promover una institución en sus docentes y, a la vez, contar con aportaciones del actor que realizará la evaluación docente: el mismo alumno. Ello se realizó a través del empleo de entrevistas cognitivas (Molina y Smith, 2011).

El tipo de análisis presentado para contestar a la pregunta central del proyecto difirió de la metodología empleada por la literatura presentada en la introducción del escrito. Mientras que los estudios revisados (Aregbeyen, 2010; Catano y Harvey, 2011; Delaney *et al.*, 2010; Jackson, 1999) definen los rasgos que importan a los estudiantes a partir de estudios cualitativos y/o de cuestionarios que expresamente solicitan a los estudiantes el ordenamiento de los rasgos, el presente trabajo partió de un instrumento donde el estudiante aportó de manera activa en su validación y empleó técnicas de modelación supervisadas (en concreto el árbol de decisión) para la exploración de la importancia de los reactivos.

En este sentido, el ítem de mayor importancia («Todas las clases aprendes algo nuevo») para el análisis, pertenece al área identificada como didáctica. En el caso de los profesores evaluados como sobresalientes y excepcionales, los dos ítems que parecen mejorar la probabilidad de ser bien evaluado se refieren a reactivos pertenecientes al área de actitudes y clima del aula («Muestra entusiasmo por la clase») y de didáctica («Toma en cuenta lo que me interesa y me gusta»). En el caso de los profesores con menor evaluación, los ítems que mejoran la probabilidad de ser evaluado con puntajes bajos, corresponden con reactivos del área de actitudes y clima del aula («El profesor se preocupa porque yo aprenda» y «Tiene problemas para controlar el grupo»).

Así, sería posible diferenciar entre lo que al estudiante le importa en los profesores evaluados con puntajes altos y puntajes bajos. En esta tónica, el presente trabajo busca, ante todo, brindar una alternativa metodológica a la construcción de cuestionarios de opinión para la evaluación de la docencia, y también mostrar las posibilidades de análisis sobre los rasgos que interesan a los estudiantes sobre sus profesores (diferenciándolos según si son bien o mal valorados por los estudiantes). Será menester de futuras investigaciones el ahondar en estos rasgos y sus diferentes matices.

ANEXO

CUESTIONARIO DE OPINIÓN PARA LA EVALUACIÓN DE LA PRÁCTICA DOCENTE (ICIF, 2015)

ÁREA	DEFINICIÓN	ÍTEM
Didáctica:	Se refiere a la actuación docente como tal, las estrategias, la estructura de clase, etcétera.	Domina la materia que enseña.
		Indica qué vamos a aprender al inicio de cada clase.
		Busca que participemos en la clase.
		Las estrategias y actividades que emplea son siempre las mismas*.
		Usa de manera adecuada la tecnología para la clase.
		Relaciona la materia con la vida real.
		Me plantea problemas o retos que me hacen pensar.
		Su clase promueve la memorización, no el razonamiento*.
		Las tareas que deja, me ayudan a aprender.
Todas las clases aprendes algo nuevo.		
Evaluación:	Se refiere a los procesos de evaluación y retroalimentación.	Evalúa de manera justa.
		Siempre se sabe cómo va a evaluar el curso.
		Sólo asigna calificaciones, sin decirme qué tengo que mejorar*.
Actitudes y clima del aula:	Se refiere al componente actitudinal y de clima (<i>rapport</i>) que el docente genera con los estudiantes como individuos y como grupo.	El profesor se preocupa porque yo aprenda.
		Muestra entusiasmo por la clase.
		Es coherente con los valores que promueve el colegio.
		Toma en cuenta lo que me interesa y me gusta.
		Tiene problemas para controlar al grupo*.
		Existe apertura y confianza para participar en su clase.
Creo que ama lo que hace.		
De validación:	Ítem que permite la valoración general de la asignatura.	1. Del 0 al 10, qué tanto aprendes en esta clase.
		2. Del 0 al 10, qué tanto te esfuerzas en esta clase.
		3. Del 0 al 10, cómo valorarías este curso de manera general.

REFERENCIAS

Alkin, M. (2004). *Evaluation roots: tracing theorists' views and influences*. USA: Sage Publications.

Aregbeyen, O. (2010). Students perceptions of effective teaching and effective lecturer characteristics at the University of Ibadan, Nigeria. *Pakistan Journal of Social Sciences*, 7(2). Recuperado de <http://www.medwelljournals.com/fulltext/?doi=pjssci.2010.62.69>. DOI: 10.3923/pjssci.2010.62.69

Arregui, I.G; Chaparro, A.L. y Díaz, C. (2015). *Instrumento para evaluar el desempeño docente en educación secundaria desde la percepción de los estudiantes*. Ponencia presentada en el 2do Congreso Latinoamericano de Medición y Evaluación Educativa (COLMEE). Recuperado de <http://www.colmee.mx/public/conferences/1/presentaciones/ponenciasdia3/48Instrumento.pdf>

Catano, V. & Harvey, S. (2011). Student perception of teaching effectiveness: development and validation of evaluation and validation of evaluation of teaching competences. *Assessment and evaluation in higher education*. 36(6). Recuperado de http://www.medsp.umontreal.ca/IRSPUM_DB/pdf/22813.pdf

Delaney, J.; Johnson, A.; Johnson, T. & Treslan, D. (2010). *Students' perceptions of effective teaching in higher education*. Recuperado de http://www.mun.ca/educ/faculty/mwatch/laura_treslan_SPETHE_Paper.pdf

Feldman, K. (1976). The superior college teacher from the student's view. *Research in higher education* (5). Recuperado de <http://link.springer.com/article/10.1007/BF00991967#page-1>

International Center for Integral Formation (2013). *Enfoque curricular de Educación Media Superior*. México: ICIF.

_____ (2015). *Cuestionario de opinión para la evaluación de la práctica docente (tercera edición)*. México: ICIF.

International Society for Technology in Education (2014). *ISTE standards for teachers*. USA: ISTE. Recuperado de http://www.iste.org/docs/pdfs/20-14_ISTE_STANDARDS-T_PDF.pdf

Ministerio de Educación Chileno (2012). *Estándares orientadores para carreras de pedagogía en educación media*. Chile: Centro de Perfeccionamiento, Experimentación e Investigaciones Pedagógicas del Ministerio de Educación de Chile. Recuperado de <http://www.mineduc.cl/usuarios/cpeip/doc/201206011651510.LibroEstAndaresEducaciOnMedia.pdf>

Molina, M. y Smith, V. (2011). La entrevista cognitiva: guía para su aplicación en la evaluación y mejoramiento de instrumentos de papel y lápiz. *Cuaderno Metodológico* (5). Recuperado de <http://iip.ucr.ac.cr/sites/default/files/cuadernosmetodologicos/cuamet5.pdf>

Perrenoud, P. (2007). *Diez nuevas competencias para enseñar*. México: Colofón.

Schmelkes, S. (2015). *Desempeño docente: estado de la cuestión*. Presentación realizada para el Segundo Día del Autor. México: INEE, ICIF, Universidad Anáhuac México Sur.

Secretaría de Educación Pública (2013). *Ley general del servicio profesional docente*. México: SEP. Recuperado de http://servicioprofesionaldocente.sep.gob.mx/content/ba/docs/parametros_indicadores/II_Docentes.pdf

Sutcliff, C. (2011). Secondary students' perceptions of teacher quality. *Electronic theses & dissertations*. USA: Jack N. Averitt College of Graduate Studies. Recuperado de <http://digitalcommons.georgiasouthern.edu/cgi/viewcontent.cgi?article=1391&context=etd>

VV. AA. (1999). *Enciclopedia general de educación*. México: OCEANO.