



FACULTAD DE EDUCACIÓN
PROGRAMA DE MAGÍSTER EN EDUCACIÓN

VALIDEZ DE UN INSTRUMENTO DE EVALUACIÓN DE LA DOCENCIA: EL CASO DE LA UNIVERSIDAD DE SANTIAGO DE CHILE

POR NICOLE EDITH ABRICOT MARCHANT

Proyecto de Magíster presentado a la Facultad de Educación de la Pontificia Universidad Católica de Chile para optar al grado académico de Magíster en educación mención evaluación de aprendizajes

Profesora guía: María Verónica Santelices EtcheGARay

Comisión informante: Carlos González Ugalde

Claudio Núñez Vega

Mayo 2015
Santiago, Chile

©2015, NICOLE EDITH ABRICOT MARCHANT

Se autoriza la reproducción total o parcial para fines académicos, por cualquier medio o procedimiento, incluyendo la cita bibliográfica del documento

DEDICATORIA

*A mi familia y amigos por respetar mis tiempos de
estudio y apoyarme siempre sin pedir nada a cambio.
A mi amado esposo por caminar conmigo en un rumbo
de eterno aprendizaje.*

AGRADECIMIENTOS

La elaboración de este trabajo agradece la colaboración de diversas fuentes y personas. En primer lugar, agradezco el apoyo incondicional y permanente de la profesora Verónica Santelices Etchegaray, quien revisó mi trabajo todas las veces que fue necesario, con gran claridad y dedicación. A los profesores Carlos González Ugalde y Claudio Núñez Vega, quienes impulsaron y apoyaron mi trabajo realizando sugerencias oportunas y enriquecedoras.

Por otra parte, fue muy importante contar con el apoyo de fondos provenientes del Programa Capital Humano Avanzado del Consejo Nacional de Ciencia y Tecnología (CONICYT) y del Proyecto MECESUP USA1307 *“Iniciativas de Innovación en la Docencia Desarrolladas por la Universidad de Santiago de Chile”*.

A través del desarrollo de este proyecto, fue posible contar con el total apoyo de la Universidad de Santiago de Chile, en especial con la Vicerrectora Académica, Dra. Patricia Pallavicini Magnere, con el equipo de la Unidad de Innovación Educativa, de la cual formo parte, y con todos los docentes, jefes de carrera y directores de departamento que participaron del estudio con un amplio sentido de compromiso con la Universidad.

Todos estos apoyos se articularon coherentemente permitiéndome crecer en términos profesionales y académicos, a través de un resultado sólido capaz de guiar la política institucional en la materia. A todos ellos, muchas gracias.

TABLA DE CONTENIDOS

	Página
I. Introducción	1
II. Marco Teórico	2
2.1. Evaluación de la docencia universitaria	2
2.2. Sobre el concepto de validez	3
2.3. Confiabilidad de los instrumentos de evaluación de la docencia	5
2.4. Evidencias sobre la validez de los cuestionarios “Student Evaluation Teaching” (SET)	5
2.4.1. Evidencia basada en el contenido de SET.	5
2.4.2. Evidencia basada en los procesos de respuesta de SET	7
2.4.3. Evidencias sobre la estructura interna de SET.	8
2.4.4. Evidencias sobre la relación de SET con otros criterios de calidad de la docencia.	9
2.4.5. Evidencias sobre las consecuencias generadas por los resultados de SET.	9
III. Contexto del estudio	11
IV. Propósitos del estudio	13
V. Hipótesis que sustentan el estudio.	13
VI. Metodología.	14
6.1. Confiabilidad del IED	14
6.2. Evidencias sobre estructura interna del IED	14
6.3. Evidencias sobre la relación del IED con otros criterios	15
6.3.1. Relación con rendimiento académico de los estudiantes	15
6.3.2. Relación con la experiencia de aprendizaje	15
6.3.3. Relación con la percepción sobre la calidad de la docencia	16
6.4. Evidencias sobre las consecuencias de los resultados del IED	17
6.4.1. Diseño y validación de Encuesta sobre Consecuencias para docentes y directivos	17
6.4.2. Observaciones de docentes y directivos sobre las consecuencias	19
VII. Análisis de resultados	21
7.1. Confiabilidad del IED	21
7.2. Evidencias sobre estructura interna	21
7.3. Evidencias sobre la relación con otros criterios	22
7.3.1. Relación con rendimiento académico	22

7.3.2.	Relación con la experiencia de aprendizaje	23
7.3.3.	Relación con la percepción sobre la calidad de la docencia	24
7.4.	Evidencias sobre las consecuencias de los resultados del IED	26
7.4.1.	Consecuencias sobre uso de resultados del IED para docentes	26
7.4.2.	Consecuencias sobre el uso de resultados para directivos	27
7.4.3.	Observaciones de docentes y directivos sobre el IED	28
VIII.	Discusión de los resultados	31
8.1.	Confiabilidad del instrumento	31
8.2.	Evidencias sobre estructura interna	31
8.3.	Evidencias sobre la relación con otros criterios	32
8.3.1.	Relación con la experiencia de aprendizaje	32
8.3.2.	Relación con la percepción sobre la calidad de la docencia	32
8.3.3.	Relación con rendimiento académico	34
8.4.	Evidencias sobre las consecuencias de los resultados para docentes y directivos	35
IX.	Conclusiones	36
X.	Recomendaciones	38
XI.	Limitaciones del estudio	40
XII.	Referencias bibliográficas	41
XIII.	Anexos	47

Resumen

Este estudio analiza evidencias de validez del Instrumento de Evaluación de la Docencia (IED) aplicado a los estudiantes en la Universidad de Santiago de Chile. Junto con analizar la fiabilidad del instrumento, se aplicó un Análisis Factorial Exploratorio y Confirmatorio a los resultados de la evaluación. Se realizaron correlaciones de Pearson entre las puntuaciones del IED y el rendimiento académico de los estudiantes (calificaciones), así como entre las puntuaciones del Instrumento con las experiencias de aprendizaje de los estudiantes (CEQ). Se analizaron las respuestas abiertas de los estudiantes para relacionar sus percepciones sobre la calidad de la docencia con las puntuaciones del IED. Finalmente, se diseñó y aplicó un cuestionario para evaluar las consecuencias de los resultados de las evaluaciones para docentes y directivos de la Universidad. La evidencia a favor del instrumento se asocia con su alta confiabilidad; relación concurrente y moderada entre el instrumento de evaluación de la docencia y las experiencias de aprendizaje de los estudiantes; y con los usos descriptivos y formativos de los resultados por parte de los docentes. La evidencia en contra tiene se asocia con la unidimensionalidad de la estructura interna del instrumento; las débiles relaciones entre el instrumento de evaluación de la docencia y el rendimiento académico de los estudiantes; la imposibilidad del instrumento para discriminar las percepciones de los estudiantes sobre la calidad de la docencia; y las preocupaciones de la comunidad académica respecto del uso de los resultados. Los análisis aportan evidencias a favor y en contra sobre la validez del uso del IED, lo que sugiere mejorar el instrumento y el proceso que implica. El análisis aporta antecedentes relevantes para revisar los procesos de aseguramiento de la calidad de la docencia en otras universidades.

Palabras clave: evaluación de la docencia universitaria, validez, educación superior.

Abstract

This study analyzes evidence of validity of the Student Evaluation of Teaching's survey (IED) at Universidad de Santiago de Chile. Along with analyzing the reliability of the instrument, Exploratory Factor Analysis and Confirmatory Factor Analyses were applied to results the evaluation. Pearson Correlations were applied between survey results and other criteria such as academic performance (grades) and the learning experiences of student (CEQ). Open responses were coded to relate the perception of students about the teaching quality with the instrument score. Finally, a questionnaire was designed and applied to assess the implications of the evaluation results for teachers and University administrators. The evidence in favor of the instrument is high reliability; concurrent moderate relationship between survey and learning experiences of students; and also, with the formative use teachers make of the results. The negative evidence was the unidimensionality of the instrument, low relationships with academic performance of students, the inability of the instrument to discriminate among student perceptions regarding the quality of teaching, and concerns of the academic community regarding the use of results. The analyses provide evidence in favor and against the validity of survey. Results suggest that the institution should review the conceptual definitions, purposes, application procedures and reporting of the instrument. The study provides relevant information to review the evaluation procedures of teaching quality in others universities.

Key words: Student Evaluation of Teaching, validity, higher education.

I. Introducción

Durante la década del setenta, fuertes ataques políticos contra “instructores incompetentes” de la Educación Superior coincidieron con un mayor interés por mejorar la calidad de la enseñanza universitaria y, particularmente, la calidad del profesorado universitario (Hannan & Silver, 2006). A partir de aquello, las universidades comenzaron a preocuparse por definir estándares de calidad de la docencia y mecanismos de evaluación docente, con el fin de promover la mejora de las experiencias de aprendizaje de los estudiantes (Hannan & Silver, 2006).

Frente a este complejo panorama y en el actual contexto de eficiencia, “la docencia universitaria se constituye como una prioridad estratégica de las instituciones de educación superior en todo el mundo” (Cid-Sabucedo, Pérez-Abellás, & Zabalza, 2009, pp. 2). Tanto a nivel internacional como a nivel nacional, las universidades declaran incorporar a los diversos actores involucrados en el proceso de evaluación de la docencia, con el fin de triangular información desde diversas perspectivas: docentes, estudiantes, directivos y pares (Álvarez, 2009). Sin embargo, la mayoría de las instituciones de educación superior ha considerado como único indicador de calidad de la enseñanza las encuestas aplicadas a los estudiantes a final de cada semestre (Spooren, Brockx, & Mortelmans, 2013).

A pesar de que la mayoría de las investigaciones muestran que los instrumentos entregan información útil, tanto para los docentes como para las autoridades universitarias (Beran, Violato, Kline & Frideres, 2005; Marsh, 1987; Ory, 2001; Penny, 2003), la confiabilidad y validez de tales evaluaciones aún es cuestionada por las comunidades universitarias de diversas casas de estudio (Clayson, 2009)

Este cuestionamiento sobre validez y confiabilidad también está presente en la comunidad académica de la Universidad de Santiago de Chile. A pesar de que la confiabilidad del instrumento resulta alta en todas sus aplicaciones y que se han aplicado algunos análisis de estructura interna al instrumento (Espinoza & González, 2009), los cuestionamientos a su validez continúan presentes. Por ello, previa exposición de la literatura más relevante sobre la validez de estos instrumentos, este estudio analizará diversas evidencias de validez del Instrumento de Evaluación de la Docencia (IED) de la Universidad de Santiago, a través de la estimación de su fiabilidad, estructura interna, relación con otras variables y sus consecuencias para docentes y directivos.

II. Marco Teórico

2.1. Evaluación de la Docencia Universitaria

Durante las últimas décadas, la nueva composición de la matrícula universitaria, los cambios en las prioridades de los estados respecto a la educación superior, la dinámica de rendición de cuentas que se comenzó a instalar en las universidades y los incorporación de la tecnología educativa en los procesos de aprendizaje, impulsaron un nuevo interés por los procesos de enseñanza y aprendizaje en Educación Superior (Silvero, 2006).

En ese contexto, la evaluación de la calidad de la docencia universitaria, como un tópico que pretende aportar a la mejora de la calidad de la enseñanza, se convirtió en un tema de vital relevancia para las instituciones de educación superior, en un escenario que comenzó a instalar progresivamente complejos sistemas de aseguramiento de la calidad en universidades europeas y norteamericanas (Silvero, 2006).

La experiencia en evaluación de la docencia universitaria ofrece propuestas interesantes y novedosas para recoger información respecto del proceso de enseñanza (Berk, 2006; Díaz Barriga y Rigo, 2003; Rueda y Díaz Barriga, 2000). Tanto Berk (2006) como Díaz Barriga y Rigo (2003) proponen procedimientos que van desde una reconceptualización de las evaluaciones de estudiantes hasta las evaluaciones de pares, evaluaciones de expertos, autoevaluaciones, análisis de videos, entrevistas a grupos de estudiantes, opiniones de egresados, encuestas a empleadores, encuestas a directivos (decanos, jefes de carrera, directores), publicación en revistas especializadas en docencia universitaria, premios a la enseñanza efectiva; medición de los resultados de aprendizaje y portafolio docente.

En Estados Unidos, por ejemplo, desde la década del 90' se utiliza el "portafolio docente" para evaluar la docencia universitaria. Para la mayoría, destaca Bain (2007), esto implica "ese horrible proceso de meter todo lo imaginable en una caja y mandárselo al director de departamento o al decano (...). En cambio otros han empezado a tratar el portafolio como un tipo de argumento erudito sobre la calidad de la docencia" (Bain, 2007; pp.185).

En otros países, como en España (Agencia Nacional de Evaluación de la Calidad y Evaluación), en México (Asociación Nacional de Universidad e Instituciones de Educación Superior), y en Chile (Comisión Nacional de Acreditación), las instituciones recomiendan diversificar a los actores y procedimientos implicados en el proceso de evaluación de la docencia universitaria. Se promueve el desarrollo de la autoevaluación docente a través de cuestionarios específicamente diseñados para el contexto europeo (Álvarez, 2009); en México se ha involucrado a autoridades y pares docentes en el proceso de evaluación (Rueda, Elizalde y Torquemada, 2006); y en América Latina existen algunas experiencias documentadas respecto al uso del portafolio y la autoevaluación docente en el último tiempo (Centro Interuniversitario de Desarrollo [CINDA], 2007).

Sin embargo, a pesar de aquellas iniciativas, la mayoría de las universidades del mundo utilizan como el más importante o el único indicador de calidad de la docencia, las puntuaciones que asignan los estudiantes en las encuestas aplicadas cada semestre académico (Zabalza, 2009; Spooren, et al., 2013; Zúñiga & Jopia, 2007).

Para ello las universidades han desarrollado procedimientos e instrumentos relativamente complejos para recoger, analizar e interpretar los resultados de este tipo de instrumentos (Spooren, et al. 2013). En la mayoría de los casos, esos instrumentos definen una serie de dimensiones de calidad docente, se basan en escalas Likert, normalmente están acompañados de espacios para comentarios, son aplicados vía internet a los estudiantes y muchas veces deben ser respondidos de forma obligatoria para acceder a la carga académica semestral de los programas de estudio (Salazar, 2008).

El uso generalizado de instrumentos “Student Evaluation of Teaching” (SET), se debe en gran parte a la aparente facilidad para la recogida de los datos, la presentación e interpretación de los resultados (Penny, 2003; Spooren, et al., 2013) y a la importancia que se atribuye a los estudiantes como evaluadores del proceso de enseñanza (CINDA, 2007; Seldin, 1993). En ese contexto, Seldin (1993) señala que “las opiniones de quienes comen la cena deberían ser consideradas si necesitamos saber cómo sabe” (Seldin, 1993, p. 40).

Sin embargo, a pesar de la coincidencia sobre la importancia de la participación de los estudiantes en el proceso de evaluación docente y que la mayoría de las investigaciones reportan que SET entrega información útil para docentes y autoridades universitarias al momento de tomar decisiones de mejora (Beran, et al., 2005; Marsh, 1987; Ory, 2001; Penny, 2003;), la validez de tales evaluaciones continúa siendo cuestionada en las instituciones de educación superior, sobre todo en términos de sus posibilidades para medir la calidad de la enseñanza debido al escaso consenso respecto de su definición (Clayson, 2009).

2.2. Sobre el concepto de validez

Antes de profundizar en las evidencias sobre la validez de SET, es importante aclarar la perspectiva teórica que se adoptará en este estudio respecto del debatido concepto de validez.

A mediados del siglo XX, es posible identificar que el clásico concepto de validez, entendido como la evidencia empírica que señala que “el test mide lo que debe medir”, es definido en los Standards for Educational and Psychological Testing de 1954, a través de tres tipos: validez referida a criterio, validez de contenido y – por primera vez – la validez de constructo

En 1954, la incorporación de la validez de constructo transformó radicalmente la visión tradicional acerca del concepto. La validez comenzó a ser comprendida como un proceso de acumulación de evidencias en torno al significado de los puntajes de un test. En base a dicha

línea, la aparición del modelo de constructo, prevaleciente desde los años 50 hasta la fecha, considera que un instrumento de medición debe ser analizado empleando todas las herramientas conceptuales y metodológicas que permitan clarificar y respaldar la interpretación y usos de un instrumento (American Educational Research Association[AERA], American Psychological Association [APA] & National Council on Measurement in Education [NCME], 1954).

En dicha línea, desde los años 80 en adelante, la definición de validez comienza a complejizarse. Messick (1995) señala que la validez debe comprenderse como una cuestión permanentemente abierta a nueva evidencia acumulada a través de diversas estrategias, para un propósito determinado. Al respecto el Standards de 1954 (AERA, et al., 1954), indica que la validez es el grado en que el uso del test es capaz de dar cuenta del logro de ciertos objetivos de los instrumentos.

Actualmente se entiende por validez la condición de que un test demuestre evidencias de estar midiendo lo que debe estar midiendo respecto de la finalidad para la cual fue construida (Hogan, 2004). En ese sentido, es posible advertir que la validez no se puede reducir a un índice, sino que se trata de una serie de miradas a los instrumentos que permiten considerar diversas evidencias que avalan (o no) su grado de validez respecto de los fines esperados.

A partir de dichas nociones, los denominados “tipos de validez” entre la década de los 80 y 90, comenzaron a desaparecer, a cambio de la acumulación de evidencia de distinto tipo. En ese sentido, los Standards de 1999 entienden que la “*validez es el grado en el que toda la evidencia acumulada apoya la interpretación deseada de los resultados de las pruebas para la fines previstos*” (AERA, APA & NCME, 1999, pp. 11). En base a lo anterior, el *Standards for Educational & Psychological Testing*, de 1999 reconoce cinco tipos de evidencias acerca de la validez (no tipos de validez como se entendía anteriormente): evidencia basada en el contenido del test; evidencia basada en los procesos de respuesta; evidencia basada en la estructura interna; evidencia basada en relaciones con otras variables; y evidencias basadas en las consecuencias de la medición (AERA, et al., 1999).

Actualmente, Kane (2013) señala que para validar una interpretación o uso de los resultados de los test, es necesario evaluar la plausibilidad de los cuestionamientos basados en los puntajes. La validación es entonces una evaluación de la coherencia y la integridad de la interpretación y uso de los razonamientos y de la plausibilidad de sus inferencias y suposiciones. Además, Lane y Stone (2002) abogan por una evaluación más comprehensiva de las potenciales consecuencias negativas a través de la participación de múltiples agentes implicados (examinados, profesores, responsables académicos, etc.) en la elaboración del argumento interpretativo.

Kane (2001) propuso diferenciar entre la “parte descriptiva” y prescriptiva” del argumento interpretativo, donde la primera implica la estimación de la variable observada en los sujetos,

mientras que en la parte prescriptiva, se destacan las interpretaciones basadas en la decisión o acción derivadas de las puntuaciones obtenidas.

En base a lo anterior, para efectos del presente estudio la validez se entenderá como un concepto único sobre el que se construirán diversos argumentos y evidencias para respaldar las interpretaciones de las puntuaciones de los instrumentos sobre sus consecuencias esperados y no esperadas, descriptivas y prescriptivas, derivadas del uso del instrumento para los actores implicados en la comunidad universitaria. En las secciones que siguen se examinará la evidencia que la literatura entrega sobre la confiabilidad y validez de los instrumentos de evaluación de la docencia.

2.3. Confiabilidad de los instrumentos de evaluación de la docencia

Desde 1927, con la primera investigación de Reamers and Brandenburg, muchos estudios han centrado su quehacer en analizar la validez y confiabilidad de los cuestionarios de evaluación de la docencia, desarrollando la línea de investigación “Student Evaluation of Teaching”.

Respecto de la confiabilidad, a partir del método de bipartición, las investigaciones reportaron altos niveles de consistencia interna de los instrumentos analizados, que van desde 0.77 a 0.94. Los resultados, usando el método de Alpha de Cronbach, también fueron altos (Costin, et al. 1971). Otras investigaciones (Marsh, 1987; Arubayi, 1987; Marsh & Bailey, 1993), utilizando el método Test- Re test, coinciden en reportar altos niveles de confiabilidad en los instrumentos de evaluación de la docencia analizados, aun cuando se presenta un alto período de tiempo entre dos evaluaciones (Richardson, 2005).

Sin embargo, como bien señala Kane (2013), a pesar de que reconocer confiabilidad de las medidas es un aporte relevante a la hora de atender los cuestionamientos a los instrumentos de evaluación, dicho antecedente no es suficiente para dar cuenta sobre la validez de las medidas (Kane, 2013). En ese sentido, la investigación ha generado diversas evidencias para argumentar la validez del uso los instrumentos psicológicos y educativos.

2.4. Evidencias sobre la validez de los instrumentos “Student Evaluation of Teaching” (SET)

Las evidencias sobre la validez de los cuestionarios SET encontrados en la literatura se organizan a continuación en evidencias sobre el contenido, procesos de respuesta, estructura interna, relación con otros criterios y consecuencias del uso de los instrumentos.

2.4.1. Evidencia basada en los contenidos de SET

Al momento de analizar las evidencias de validez respecto del contenido de los instrumentos de evaluación de la docencia, la primera preocupación debe centrarse en conceptualizar la docencia de calidad. En este sentido tanto Coll y Solé (2001) como Shulman (1989), coinciden en que los estudios que han prevalecido en las últimas décadas, son aquellos que destacan que el comportamiento observable del profesor es la variable más significativa en la

determinación del rendimiento académico de los estudiantes. Esta perspectiva, más vinculada al desarrollo de los instrumentos norteamericanos (Richardson, 2005), centra su preocupación en caracterizar el comportamiento observable del profesor y las competencias docentes nutriendo los llamados “modelos de enseñanza eficaz” (Zabalza, 2009).

Estos modelos se relacionan con los tradicionales instrumentos de evaluación docente, que pretenden medir conductas discretas del profesor relacionadas con sus atributos personales o estilos de enseñanza, bajo la premisa de que esos indicadores representan la eficacia docente (Díaz Barriga y Rigo, 2003). Algunos de los ítems que se repiten en este tipo de encuestas son “el profesor es claro”, “el profesor repite y revisa los conceptos e ideas clave”, etc. (Díaz Barriga y Rigo, 2003; pp. 58). Según Richardson (2005) uno de los instrumentos que representa esta perspectiva es el Students’ Evaluations of Educational Quality (CEEQ) de Marsh y Roche (1994).

Sin embargo, criticando esta perspectiva por simple y unidimensional, Shulman (1989) denominó esta línea de investigación como “*paradigma proceso – producto*”, debido a que en ella los estudiantes son “cajas negras”, en que el único factor relevante para resultados exitosos, es la actuación del docente. Otras variables del proceso de enseñanza y aprendizaje – como el contexto de aula, contenidos de la enseñanza y motivaciones de los estudiantes – resultan prácticamente ausentes.

En respuesta a esta crítica, las investigaciones comenzaron a considerar el contexto de aula y las relaciones pedagógicas en el aprendizaje para caracterizar una enseñanza de calidad (Coll & Sánchez, 2008; Coll y Solé, 2001), lo que ha constituido el ámbito académico denominado “*Scholarship of Teaching and Learning*” (Boyer, 1990; Yair, 2008).

En esa línea de investigación, Biggs (2008) planteó el modelo 3P de enseñanza y aprendizaje, como un sistema interactivo en donde se sitúan los factores relacionados con el aprendizaje, desde el Pronóstico, durante el Proceso y desde el Producto. En este Modelo, el resultado de aprendizaje está determinado por muchos factores que interactúan entre sí y que convierten al proceso de enseñanza y aprendizaje en un fenómeno único e irrepetible (Biggs, 2008; Ramsden, 2003).

Ramsden (2003), destaca que los elementos que componen la calidad de la enseñanza se pueden comprender desde dos focos: *docencia centrada en la enseñanza* y *docencia centrada en el aprendizaje*. La primera de ellas (vinculada a los modelos de enseñanza eficaz) considera los contenidos del curso, actividades, materiales, técnicas de enseñanza, características personales del maestro, etc.; mientras que la docencia centrada en el aprendizaje, refiere a la efectividad de la enseñanza para promover aprendizaje profundo en sus estudiantes, lo que quiere decir que no hay una “mejor forma de enseñar” – como se podría entender desde los modelos de docencia efectiva. Según el autor, una buena enseñanza puede seguir distintos

métodos, pero se centra en la preocupación por las necesidades del estudiante, utiliza métodos de evaluación apropiados para promover el aprendizaje profundo y procura definir metas y objetivos claros (Ramsden, 2003).

Esta perspectiva teórica, desarrollada desde Inglaterra y Australia, ha dado paso a una serie de instrumentos que se han centrado en reconocer los distintos factores que componen la interacción en el aula. Entre estos destacan: el Study Process Questionnaire (Richardson, 2004; Biggs, 2001) que recoge la orientación profunda o superficial de los estudiantes hacia el aprendizaje en una asignatura en particular; el Course Experience Questionnaire (Chan, Prosser & Watkins, 2009; Richardson, 2005) que mide la experiencia de aprendizaje de los estudiantes en un curso o programa en particular, a través de las escalas “buena docencia”; “metas y objetivos claros”; “carga de trabajo apropiada” y “evaluación apropiada” (Webster, Chan, Prosser, Watkins, 2009); y el Approaches to Teaching Inventory (Prosser y Trigwell, 2006; González, Montenegro, López, Munita, Collao, 2011) que recoge información sobre cómo abordan la docencia los profesores universitarios.

En cuanto al contenido de los instrumentos de evaluación de la docencia, la dificultad principal está en que muchos administradores universitarios no siguen ninguna de estas perspectivas teóricas de manera clara al momento de desarrollar los instrumentos (Ory y Ryan, 2001; Penny, 2003). Por ello se destaca que, a pesar de que los instrumentos SET normalmente gozan de validez aparente (Onwuegbuzie et al., 2009), existen grandes diferencias entre su conceptualización de la que constituye una buena docencia universitaria (Richardson, 2005).

2.4.2. Evidencia basada en los procesos de respuesta de SET

Desde otro punto de vista, la investigación también ha aportado evidencias sobre los denominados Procesos de Respuesta (AERA, APA, & NCME, 1999), es decir, sobre los procesos psicológicos que se activan en los estudiantes al momento de responder los cuestionarios SET. Al respecto las investigaciones en el área han llegado a dos grandes conclusiones:

En primer lugar, a partir de entrevistas a estudiantes, se ha llegado a determinar aquello que los estudiantes comprenden por docencia efectiva es muy distinto de lo que sus docentes comprenden por ello (Bosshardt y Watts, 2001).

De las investigaciones asociadas a reconocer lo que entienden los estudiantes por una enseñanza de calidad, se destaca que ellos reconocen como docencia efectiva, en términos generales, aquella enseñanza que está centrada en el aprendizaje activo y en la enseñanza no tradicional (Kember and Wong, 2000). En esa línea, el estudio de Onwuegbuzie et al. (2007) señala que los estudiantes reconocen como categorías asociadas a la enseñanza efectiva: la docencia centrada en el estudiante, la experticia, el profesionalismo, el entusiasmo, el

comportamiento ético, la retroalimentación efectiva, la capacidad comunicativa, la accesibilidad, la capacidad de liderazgo, y el ambiente de aprendizaje efectivo. Onwuegbuzie et al. (2007) concluye que las cuatro primeras características son las más relevantes y determinantes para los estudiantes, y que las categorías “docencia centrada en el estudiante”, “experticia”, “entusiasmo” y “comportamiento ético” no están presentes en el cuestionario de evaluación docente que ellos responden en su universidad. Es decir, tres de las cuatro características más relevantes para los estudiantes sobre la docencia efectiva, no están considerados en el instrumento de evaluación de la docencia.

Por otra parte, Barth (2008), Pan et al. (2009) y Pozo et al. (2000) destacan que, a diferencia de lo que el sentido común y las críticas podría señalar, los estudiantes universitarios valoran más la calidad de la enseñanza que las características personales de los docentes, tales como la simpatía o la apariencia física. Sin embargo, otras investigaciones (Ambady & Rosenthal; 1993; Shevlin et al., 2000) concluyen que los estudiantes basan sus respuestas en una sola característica del docente y no atienden las preguntas señaladas en los cuestionarios de respuesta. Es decir, los estudiantes estarían respondiendo bajo un “efecto halo” que los lleva a responder de una determinada forma las evaluaciones. Así, Shevlin et al. (2000), por ejemplo, señala que un 69% de la varianza de SET es explicada por el carisma del profesor; mientras Ambady y Rosenthal (1993) señala que la valoración de los docentes se forma en cuestión de segundos, a partir del atractivo físico y comportamiento no verbal de sus profesores.

2.4.3. Evidencias sobre la estructura interna de SET

Respecto de evidencias sobre la estructura interna de los instrumentos, es posible identificar variadas investigaciones dedicadas a analizar la dimensionalidad de los instrumentos. Los análisis al respecto revelan que, a pesar de que normalmente los instrumentos consideran de 3 a 5 dimensiones en sus diseños, los SET se caracterizan por un solo constructo global (unidimensional) al que se asocian las respuestas de los estudiantes de forma generalizada (Richardson, 2005; García, 2000; Abrami y d’Apollonia, 1991). Ese constructo general, ha sido denominado como “habilidad general para enseñar” (Abrami, et al., 1991).

Spooren et al., (2013) argumentan que dichos antecedentes sobre la unidimensionalidad de los instrumentos SET, deben constituirse en una preocupación para las autoridades universitarias, ya que una estructura unidimensional podría ser apropiada para fines los sumativos de SET – en términos de facilitar el reporte de un puntaje total – pero no para sus fines formativos – por ser más restrictivos en cuanto al tipo de información reportada.

2.4.4. Evidencias sobre relación de SET con otros criterios de calidad de la docencia

Diversas investigaciones (Clayson, 2009; Cohen, 1981; Griffin, Hilton & Barret, 2014) han centrado su quehacer en establecer relaciones entre las puntuaciones de SET y otros criterios asociados a la calidad de la docencia. Algunas de esas investigaciones se han dedicado a analizar la relación entre los resultados de aprendizaje de los estudiantes y las puntuaciones SET que obtienen sus profesores, sin llegar a grandes coincidencias entre ellas, ya que las correlaciones varían entre 0,10 y 0,47 (Cohen, 1981). Sin embargo, una de las últimas investigaciones al respecto destaca que se observan positivas y moderadas correlaciones ($r=0,37$ $p < 0,01$, $n=2073$) entre las calificaciones de los estudiantes y las puntuaciones SET de sus profesores (Griffin, et al., 2014).

Sin embargo, basándose en un meta análisis de la literatura (desarrollado con los estudios generados en la década de los 70), Clayson (2009) encuentra que la percepción que los estudiantes tienen sobre su aprendizaje sí se relaciona con las evaluaciones docentes que asignan a sus profesores. Sin embargo, también destaca que las relaciones entre el aprendizaje de los estudiantes y las evaluaciones SET de sus docentes es pequeña, aun cuando advierte que dicha relación no es aplicable de igual forma a todas las disciplinas y mediciones del aprendizaje (Clayson, 2009). Por ejemplo, las relaciones son más fuertes en el área de la educación y las artes que en las áreas de negocios; y mientras más “objetiva” es la medición del aprendizaje menos fuerte es la relación entre SET y el rendimiento de los estudiantes (Clayson, 2009). Es por ello que Clayson (2009) señala que la pregunta no es *si existe* relación entre el aprendizaje de los estudiantes y SET, sino *cundo* existe dicha relación. Futuros análisis deben considerar dicha perspectiva.

2.4.5. Evidencias sobre las consecuencias generadas por los resultados de SET

Las consecuencias de SET se pueden reconocer en términos descriptivos y prescriptivos. Entre las consecuencias descriptivas la literatura señala que las puntuaciones de las evaluaciones docentes son utilizadas para: diferenciar a los docentes en categorías de calidad (Beran, et al., 2005; Bolivar, 2000; Chen & Hoshower, 2003; Onwuegbuzie et. al., 2007); y diferenciar los niveles de satisfacción de los estudiantes con sus experiencias de aprendizaje (Huybers, 2013).

Respecto de las consecuencias prescriptivas, según la literatura, las puntuaciones de SET deberían ser utilizadas con fines formativos para aportar elementos básicos para la autocrítica docente (Salazar, 2010; Huybers, 2013); ofrecer información para mejorar la calidad de la docencia (Huybers, 2013; Salazar, 2010; Spooren et al. 2013; Onwuegbuzie et. al., 2007; Richardson, 2005); y promover que los docentes busquen instancias de apoyo en formación pedagógica (Moore y Kuol, 2005; Nasser y Fresko, 2010). Respecto de las consecuencias prescriptivas con fines sumativos, es posible encontrar que los resultados se utilizan para

orientar la permanencia, promoción e incentivos de los docentes universitarios (Huybers, 2013; Salazar, 2010; Spooren et al. 2013; Onwuegbuzie et. al., 2007; Richardson, 2005); y reportar la calidad global de la enseñanza de las carreras a instancias de rendición de cuentas como sistemas de acreditación, proyectos, entre otros (Moore y Kuol, 2005).

Cabe señalar que las investigaciones en dicha línea, coinciden en que el uso de los instrumentos con fines sumativos (de promoción, retención e incentivos) trajo consigo que sus propósitos formativos (mejora de la calidad docente) se distorsionaran (García, 2014; Spooren et al, 2013).

A esta variedad de posibles decisiones o acciones derivadas del uso de la encuesta, se suman las diferencias entre los enfoques con los que los agentes, legítimamente implicados en el uso de SET (administradores, decanos, docentes, estudiantes, etc.), valoran el proceso de evaluación juzgando las consecuencias desde diferentes perspectivas, intereses y valores (Padilla, Gómez, Hidalgo y Muñiz, 2007). La mayoría de los estudios dedicados a analizar las consecuencias de los instrumentos, se centran en el análisis cualitativo de entrevistas semi estructuradas o encuestas a actores clave de comunidades universitarias.

Al respecto, el estudio de García (2014) aplicó una encuesta semiestructurada a estudiantes, revelando que ellos declaran que los instrumentos de evaluación de la docencia tienen fines sumativos, de control y supervisión, que son poco válidos y que tienen poca utilidad para mejorar el quehacer docente.

Por su parte, el estudio de Beran, et al., (2005) analizó los usos y consecuencias de los instrumentos SET en estudiantes, docentes y directores, a partir de la aplicación de encuestas a muestras representativas de cada uno de los actores señalados. Cabe señalar que el estudio es realizado en Canadá, en un contexto institucional distinto, donde los resultados de las evaluaciones son públicos para los estudiantes. Entre sus principales conclusiones es posible advertir que un 52% de los estudiantes declaran que nunca han utilizado los resultados de SET, mientras que el resto de ellos indica que utiliza los resultados para contar con información adicional al momento de seleccionar asignaturas.

Por otra parte, el estudio revela opiniones favorables de los profesores, respecto al uso de SET como antecedente que aporta a la mejora de la docencia, sin embargo, paradójicamente ellos señalan que generalmente no los utilizan para incorporar cambios a su enseñanza. Finalmente, los directivos o administradores coinciden en que utilizan SET para tomar decisiones relativas a los incentivos o retención de sus docentes, más no con otros fines (Beran et al. 2005).

En síntesis, la literatura especializada aporta evidencias importantes a la hora de analizar la validez de los instrumentos de evaluación de la docencia que responden sus estudiantes. Sin

embargo, algunas de las evidencias son contradictorias dependiendo del contexto y tipo de investigación – en el caso de la relación entre SET y el aprendizaje.

Por otra parte, la definición de los constructos teóricos sobre los que se desarrollan los instrumentos; la dimensionalidad de los mismos; los procesos cognitivos que ocurren en los estudiantes al momento de responderlos; la relación de los instrumentos con otros criterios de calidad docente; y las consecuencias que tiene el uso de los instrumentos en las universidades, son temas que aún no han sido abordados en las universidades chilenas y sudamericanas. Algunas de esas tensiones serán abordadas en la presente investigación.

III. Contexto de estudio

En la actualidad la Universidad de Santiago de Chile brinda formación profesional a más de 15.000 estudiantes, para los cual cuenta con cuerpo docente de más de 2.000 profesores, compuesto por cerca de 700 académicos jornada y 1300 docentes por hora.

En su afán por mejorar continuamente la calidad de la docencia universitaria impartida a sus estudiantes, la Universidad ha implementado un sistema de evaluación de la docencia que pretende recopilar información respecto del quehacer de sus docentes, para establecer mecanismos de mejora a los procesos pedagógicos al interior de las aulas universitarias.

En ese contexto, se aplica un instrumento de evaluación a los estudiantes y uno autoevaluación del docente, que son administrados vía online cada semestre, por medio de una plataforma especialmente diseñada para ello, durante un período de tres semanas previas a la finalización del semestre académico. Es importante señalar que los estudiantes deben responder de manera obligatoria la encuesta para poder registrar sus asignaturas del semestre próximo, mientras que la autoevaluación del docente es respondida de manera voluntaria en las mismas condiciones y período de tiempo asignado a los estudiantes.

En base a lo anterior, el porcentaje de respuesta de los estudiantes en cada semestre académico es de un 90% aproximadamente, mientras que las autoevaluación de la docencia sólo es respondida por el 10% de los académicos de la institución.

Por ello, a pesar de que existen intenciones de incorporar a nuevos actores en el proceso de recogida de información para la evaluación de la calidad de la docencia, el instrumento más importante del sistema de evaluación – al igual que en la mayoría de las universidades chilenas (Salazar, 2008) – sigue siendo el instrumento aplicado a los estudiantes.

Dicho instrumento está compuesto por cuatro dimensiones de docencia efectiva adaptadas a la Universidad de Santiago de Chile, desde el Modelo de Evaluación del Despeño Docente propuesto por Espinoza y González (2009): planificación de las actividades docentes; ejecución de las actividades docentes; evaluación de los aprendizajes; y relación profesor- alumno (Ver anexo 1). El instrumento contiene 21 ítems organizados en una escala Likert en los siguientes niveles:

1 (*muy en desacuerdo*), 2 (*En desacuerdo*), 3 (*Parcialmente de acuerdo*), 4 (*De acuerdo*) y 5 (*Muy de acuerdo*). La última aplicación del instrumento identificó a un 63% de docentes con promedios superiores a 4, un 32% con promedios superiores a 3 y un 4,2% con promedios inferiores a 3.

Además, se destaca que el instrumento pregunta a los estudiantes, en un formato de respuesta abierta, “*elementos que el docente debería mantener*” y “*elementos que el docente debería mejorar*”. Estas preguntas son voluntarias al momento de responder.

Posterior a la aplicación del instrumento – durante el primer tercio del semestre siguiente – un sistema informático entrega reportes de la información a docentes y directivos, con perfiles asociados a sus niveles de jerarquía. El informe del docente contiene: 1) la autoevaluación docente (si fue contestada); 2) los estadísticos de las respuestas de estudiantes (mínimo, máximo, moda, promedio, promedio departamento, por cada uno de los ítems); 3) comparaciones entre autoevaluación y evaluación de estudiantes (en cada una de las dimensiones); y, finalmente, las respuestas abiertas de los estudiantes. Los informes de los directivos están constituidos por gráficos y tablas que organizan los resultados de los docentes, en los porcentajes de “muy de acuerdo y de acuerdo”, “parcialmente de acuerdo” y “muy en desacuerdo y desacuerdo” que obtiene cada profesor en sus respectivas asignaturas.

Cabe señalar que en los últimos años, los docentes pueden postular a incentivos económicos por su calidad docente. Uno de los insumos considerados en la decisión de dicho incentivo es el resultado de la evaluación de la docencia. Además, el sistema de evaluación del desempeño académico (que considera la extensión y la investigación, además de la docencia) destaca como una de sus variables el resultado obtenido en el sistema de evaluación de la docencia.

La Institución en cuestión ha demostrado particular interés en analizar las propiedades psicométricas del instrumento, para perfeccionarlos constantemente y para dar respuesta a diversos cuestionamientos provenientes del personal académico y de la comunidad universitaria en su conjunto. Si bien los análisis de confiabilidad han arrojado resultados indicando alta consistencia interna en todas las aplicaciones, aun es escasa la evidencia relativa a la validez del instrumento. Dicha evidencia es de importancia considerando los fines para los cuales fue diseñado: la mejora de los procesos docentes en la Universidad de Santiago de Chile.

IV. Propósitos del estudio

El presente estudio pretende analizar diferentes evidencias sobre la validez del instrumento para el uso de los resultados de la evaluación de la docencia, que responden los estudiantes de Universidad de Santiago de Chile. Esto se desarrollará desde los siguientes objetivos específicos.

- Determinar evidencias de confiabilidad y estructura interna del instrumento de evaluación de la docencia de la Universidad de Santiago de Chile.
- Relacionar la encuesta de evaluación docente con el rendimiento académico de los estudiantes, como evidencia convergente o divergente de relación con otras variables.
- Relacionar las puntuaciones de la encuesta de evaluación de la docencia con la percepción de los estudiantes sobre su experiencia de aprendizaje (medida por medio del cuestionario CEQ), como evidencia convergente o divergente de relación con otras variables.
- Relacionar las puntuaciones de la encuesta de evaluación de la docencia con la percepción de los estudiantes sobre la docencia efectiva (expresadas en los comentarios del IED).
- Caracterizar evidencias sobre la validez consecuencial del uso de las puntuaciones de la encuesta de evaluación de la docencia en los directivos y docentes de la Universidad de Santiago de Chile.

V. Hipótesis que sustenta el estudio

La hipótesis del estudio, sustentada en la literatura, plantea que la encuesta de evaluación de la docencia de la Universidad de Santiago de Chile, presentará evidencias en contra y a favor de la validez de sus resultados para los usos actualmente dados. Las evidencias a favor de su validez se relacionarán con la alta estabilidad las puntuaciones (Costin, et al. 1971; Richardson, 2005); su relación concurrente entre la encuesta de evaluación docente y la percepción de los estudiantes sobre su experiencia de aprendizaje (Clayson, 2009); así como la declaración de que los docentes utilizan los resultados de la encuesta con fines descriptivos, más no con fines formativos (Beran, et al., 2005).

Por su parte, se espera que las evidencias en contra de la validez de la encuesta de evaluación docente manifiesten la unidimensionalidad del instrumento y la inexistencia de las cuatro dimensiones predefinidas teóricamente (Richardson, 2005; Spooren, et al., 2013); la relación positiva pero débil de SET y el rendimiento académico de los estudiantes (Clayson, 2009); las escasas consecuencias formativas y sumativas que declaran atribuir los docentes resultados de la encuesta (Beran, et al., 2005); y las escasas consecuencias prescriptivas formativas que realizan los directivos y docentes de los resultados del instrumento, junto con las decisiones sumativas que los directivos declaran atribuir al uso de los resultados del instrumento (Beran, et al., 2005).

V.I. Metodología

El presente estudio no experimental, transversal de alcance descriptivo, cuenta con una metodología de carácter mixta subsecuente, es decir, se abordará de manera cualitativa y cuantitativa, según la naturaleza de los datos de los distintos tipos de evidencias recopiladas.

La investigación pretende alcanzar conclusiones representativas de la comunidad de la Universidad de Santiago de Chile. Esta población está compuesta por cerca de 2000 docentes universitarios y 67 programas de estudio organizados en las siguientes unidades académicas mayores: Facultad de Química y Biología; Facultad de Administración y Economía; Facultad de Ingeniería; Facultad de Humanidades; Facultad de Tecnológica; Facultad de Ciencias Médicas; Facultad de Ciencias; Programa de Bachillerato y Escuela de Arquitectura. La muestra y el tipo de análisis del estudio dependerán de cada una de las evidencias recopiladas, lo cual se explicará en detalle a continuación. La metodología del estudio será descrita para cada uno de los objetivos, con el fin de facilitar la comprensión al lector.

6.1. Confiabilidad del IED

La confiabilidad del instrumento fue revisada en base al análisis de su consistencia interna, a través de Alfa de Cronbach y el método de bipartición.

6.2. Evidencias sobre estructura interna del IED

Se aplicó un Análisis Factorial Exploratorio (AFE) y Análisis Factorial Confirmatorio (AFC) a una muestra de 48.129 respuestas de estudiantes al IED correspondiendo a la totalidad de docentes que realizaron clases el primer semestre del año 2014, en todas las unidades académicas mayores de la Universidad, salvo la Facultad de Ingeniería y la Facultad de Administración e Economía. La omisión de estas dos facultades se debió a que no fue posible conseguir la base de datos desagregada a nivel de las respuestas de los estudiantes.

El AFE fue realizado usando el programa SPSS v20.1 y se empleó como técnica de factorización el “Mínimo Cuadrados No ponderado” –Unweighted Leaste Squares (ULS) y el método de rotación Oblimin (ya que se asume relación entre los factores). En relación al Análisis Factorial Confirmatorio, se utilizó el software STATA/SE v.12, evaluándose en primera instancia la presencia de 4 dimensiones definidas teóricamente y posteriormente la adecuación de los estadísticos al modelo de 1 factor.

6.3. Evidencias sobre la relación del IED con otros criterios.

6.3.1. Relación con rendimiento académico de los estudiantes

A través de un análisis de correlación de Pearson, se determinó la relación entre las puntuaciones IED y las calificaciones de los estudiantes, usando el software SPSS v.20.1. Dicho análisis considerará, las puntuaciones de la evaluación de la docencia y calificaciones del 90% de estudiantes de la Universidad.

Para efectuar el análisis, fue necesario determinar el promedio total de la evaluación docente de cada profesor en cada curso y la media de los promedios de calificaciones de los estudiantes en cada uno de esas asignaturas. Posteriormente, se estandarizaron ambas puntuaciones (en puntaje Z) para que fuesen comparables y se calculó correlación de Pearson entre ambas variables (calificaciones y puntuaciones IED estandarizadas) para el total de la base de datos de la Universidad. Además, este análisis fue aplicado de manera particular para cada una de las facultades de la Universidad, con el fin de determinar posibles diferencias por área de conocimiento.

6.3.2. Relación con la experiencia de aprendizaje

Se analizó la relación entre las puntuaciones de la encuesta de evaluación de la docencia y las puntuaciones de 311 estudiantes (de 19 cursos distintos de 5 áreas de conocimiento) en el instrumento Course Experience Questionnaire (CEQ) (Ver anexo 2).

Dicho instrumento – validado por Webster et al. (2009) – fue traducido y validado en Chile por González et al. (2011), y fue adaptado a las diferencias lingüísticas, psicológicas y culturales, para el contexto de la Universidad de Santiago, por un equipo de especialistas acreditados. La aplicación piloto del instrumento reveló índices adecuados de confiabilidad según alfa de Cronbach, en sus cuatro escalas (buena docencia=.8585; metas y objetivos claros=.7408; carga de trabajo=.7817; y evaluación apropiada=.6977). Por otra parte, el instrumento demostró adecuación al modelo de 4 factores, en base a Análisis Factorial Confirmatorio (KMO = .888; prueba de esfericidad de Bartlett = 4780.88, $p < .001$). Los estadísticos demostraron un ajuste adecuado para la escalas Buena Docencia ($\chi^2 = 10.381$, g.l.= 5.36, $p < .05$; CFI = .997; TLI = .985; SRMR = .011; RMSEA = .060), Metas y Objetivos Claros ($\chi^2=855.841$, g.l.= 483.4, $p < .01$; CFI = 1; TLI = 1; SRMR = .00; RMSEA = .00), Carga de Trabajo ($\chi^2 = 980.307$, g.l.= 586.3, $p < .01$; CFI = 1; TLI = 1; SRMR = .00; RMSEA = .00) y Evaluación Apropiada ($\chi^2 = 522.720$, g.l.= 1994, $p < .01$; CFI = .985; TLI = .956; SRMR = .024; RMSEA = .06

Para efectuar el análisis se estandarizaron las puntuaciones (puntaje Z) de ambos instrumentos y luego se correlacionó el promedio de las puntuaciones que cada estudiante asignó a la encuesta de evaluación de la docencia, con cada una de las medias que logró cada

uno de ellos en las 4 escalas del CEQ. Dichas correlaciones llevaron a cabo a través de correlaciones de Pearson utilizando el software SPSS.

6.3.3. Relación con la percepción sobre la calidad de la docencia

El análisis se hizo en base a la codificación de las respuestas abiertas que los estudiantes responden a las siguientes preguntas de la encuesta de evaluación de la docencia: “señale brevemente algún aspecto que considere positivo en el desarrollo de la asignatura” y “señale brevemente algún aspecto que podría mejorar el desarrollo de la asignatura”.

Los comentarios de 826 docentes en sus respectivas asignaturas correspondientes a siete áreas de conocimiento distintas, sumaron un total de 5991 comentarios codificados. Cabe señalar que cada docente recibió distinta cantidad de comentarios, ya que éstos fueron realizados de manera voluntaria por sus estudiantes. Además, con el fin de focalizar el análisis, se definió que de cada uno de los comentarios se debía extraer hasta un máximo de cuatro códigos distintos.

Tabla 1.

Número total de comentarios y docentes por cada Área de Conocimiento

Área de conocimiento	Docentes	Comentarios
Ciencias Naturales	83	831
Tecnología	68	543
Ciencias de la salud	67	1314
Diseño	76	527
Bachillerato	11	1099
Ciencias	40	401
Humanidades	121	1276
Total	826	5991

La codificación fue deductiva de tipo temática (Braun & Clarke, 2008), a partir de categorías definidas a partir de la literatura (ver anexo 10). Las categorías sobre efectividad docente, fueron definidas principalmente a partir de los resultados del estudio de Onwuegbuzie, et. al. (2007) sobre las percepciones de los estudiantes sobre la efectividad docente de los profesores universitarios; y complementado con otros autores que abordan el constructo (Hajdin et al., 2012; Parcerisa, 2010; Zabalza, 2009; Biggs, 2008; Bain, 2007; Silvero, 2006; Ramsden, 2003; Shevlin et al, 2000; Ambady y Rosenthal, 1993).

El análisis fue realizado por dos investigadores de manera independiente, que codificaron de manera cruzada un 20% de los comentarios de cada carrera. En el proceso de codificación cruzada se obtuvo siempre más de un 70% de índice de acuerdo (con 48 códigos). Posteriormente, se realizó una revisión de los códigos de manera conjunta, con el fin de tomar acuerdo respecto de los códigos que generaban controversia, lo que permitió llegar a acuerdo en un 100% de los casos.

Posteriormente, los códigos fueron contrastados con las puntuaciones de la encuesta en los distintos cursos. Las puntuaciones de la encuesta fueron agrupadas en base a los promedios de las puntuaciones de los estudiantes a sus docentes, en las categorías 0 a 3 (promedios bajos), 3,01 a 4 (promedios medios) y 4,01 a 5 (promedios altos). Esta agrupación está asociada con los rangos definidos en la escala Likert y con la agrupación utilizada en los informes de resultados de docentes y directivos.

Esto permitió caracterizar a cada grupo de docentes (promedios IED bajos, medios y altos) a partir de los códigos más frecuentes. Los códigos con mayor frecuencia en cada categoría de promedios fueron destacados en el análisis, con el fin de advertir las diferencias o similitudes de elementos destacados para los docentes de cada categoría de promedios de IED. El criterio para definir si los códigos destacaban unos de otros, se fijó señalando que los códigos debían contar con frecuencias superiores al 5% en cada categoría definida por IED. Este análisis fue realizado de manera agregada sumando todos los códigos, pero también de manera diferenciada por área de conocimiento. La comparación de los grupos de docentes, se realizó en base al delta entre los promedios de los grupos, para cada uno de los códigos.

6.4. Evidencias sobre las consecuencias de los resultados del IED

El estudio sobre las consecuencias del instrumento se desarrolló en dos partes. En primer lugar, el estudio contempló el diseño, validación y aplicación de dos encuestas: una para docentes y una para directivos de las carreras de la Universidad. Los instrumentos se desarrollaron en base a las consecuencias descriptivas y prescriptivas, esperadas y no esperadas declaradas en la literatura sobre este tipo de instrumentos (Beran, et. al., 2005; Huybers, 2013; Moore y Kuol, 2005; Nasser y Fresko, 2001; Spooren et al. 2013; Onwuegbuzie et. al., 2007; Richardson, 2005; Salazar, 2010).

6.4.1. Diseño y validación de las Encuestas sobre Consecuencias para docentes y directivos

La Encuesta sobre Consecuencias del Instrumento de Evaluación de la Docencia (ECIED) para docentes contó con 19 ítems y la de directivos con 20, ambas organizadas en 3 dimensiones denominadas: “consecuencias descriptivas”, “consecuencias prescriptivas: formativas”, “consecuencias prescriptivas: sumativas”. En ambos casos se trató de ítems de

respuesta cerrada, en base a escala Likert de frecuencia: “siempre”, “casi siempre”, “a veces”, “nunca” (ver anexo 13, 14). Adicionalmente, se solicitaron comentarios y propuesta de mejora, en formato de respuesta abierta.

El contenido de la Encuesta fue analizado por cuatro especialistas en evaluación de programas en Educación Superior de tres universidades nacionales (Universidad de Chile, Universidad Católica de Valparaíso y Universidad de Santiago de Chile), obteniendo un 80% de acuerdo respecto de las aprobación de los ítems del instrumento (validación de jueces) (ver anexo 15); además, fue analizada a través de entrevistas cognitivas aplicadas a una docente y un directivo de la Universidad en cuestión, lo que permitió incorporar algunos cambios de redacción en el instrumento.

La ECIED a docentes fue contestada por 204 docentes (10% del total), mientras que la ECIED a directivos fue respondida por 31 jefes de carrera o directores de departamento de la Universidad (33%). A posteriori los resultados de la Encuesta fueron analizados en términos de su fiabilidad y su estructura interna. La encuesta arrojó una fiabilidad de 0,93 con método de Alpha de Cronbach.

El Análisis Factorial Exploratorio del ECIED se realizó utilizando el software FACTOR v. 9.3, con el fin de resguardar la naturaleza ordinal de las variables y, por tanto, resguardar la condición policórica de la matriz de correlaciones (Freiberg, Stover, de la Iglesia y Fernández, 2013). Este primer análisis empleó como técnica de factorización el “Mínimo Cuadrados No ponderado” –Unweighed Leaste Squares (ULS) y el método de rotación Varimax (ya que no se asume relación entre los factores).

Inicialmente, se encontró que en base a la matriz de correlación es posible realizar un análisis factorial, ya que el índice de adecuación KMO es de 0.92 y el estadístico de Esfericidad de Bartlett es menor a 0,05. Al revisar los Eigenvalue (autovalor) por medio del método de factorización de Mínimos Cuadrados no Ponderados se observan 2 factores mayores a 1, los cuales explican 53% y 17% la varianza respectivamente (Ver anexo 16).

Asimismo, el análisis paralelo basado en el método de Análisis de mínimo factor de rango (MRFA) reportó dos factores, que en total representan el 72% de la varianza. Al revisar la distribución de residuales considerando dos factores se aprecia que la raíz cuadrática residual RMSR (0,0348) permite concluir que el modelo es aceptable.

Cabe señalar que la distribución de los reactivos en los factores coincidió, en términos generales, con la organización definida teóricamente respecto de las dos grandes dimensiones del instrumento: consecuencias formativas y consecuencias sumativas de la encuesta de evaluación de la docencia.

Posteriormente, aunque sin la posibilidad de resguardar la naturaleza ordinal de las variables, el Análisis Factorial Exploratorio fue realizado en el software SPSS v. 20.1, con el

fin de obtener nuevos antecedentes sobre la estructura factorial del instrumento. De igual manera, este análisis encontró que en base a la matriz de correlación es posible realizar un análisis factorial, ya que el índice de adecuación KMO es de 0.93 y el estadístico de Esfericidad de Bartlett es menor a 0,05.

A diferencia del anterior, este análisis reveló 3 autovalores mayores que 1, que explican un 24,7%, 22,1% y 16,2% de la varianza respectivamente, después de la rotación (Varimax). Al evaluar el gráfico de sedimentación, se confirma el análisis apreciándose 3 pendientes claramente identificables (ver Anexo 17).

La relación entre ambas conclusiones, resulta bastante coherente con el modelo teórico. Mientras el primer análisis, sensible a la naturaleza ordinal de las variables, identifica 2 dimensiones, el segundo análisis confirma dichos factores (consecuencias formativas y consecuencias sumativas) pero además plantea la existencia de un tercer factor, que según la distribución de las cargas factoriales, responde a las “consecuencias descriptivas”, que teóricamente debería tener relación con el factor de consecuencias formativas (ver anexo 18).

En conclusión, la ECIED revela la existencia de dos dimensiones claramente identificables y una tercera dimensión que se presenta, aunque de manera más débil. Estas dimensiones podrían denominarse: “consecuencias descriptivas: valoración sobre desempeño”, “consecuencias prescriptivas: cambios y mejoras a las prácticas” y “consecuencias prescriptivas: decisiones sumativas”.

En el caso de la ECIED aplicada a los directivos, se pretendió aplicar el mismo procedimiento de análisis de fiabilidad y estructura interna aplicada a las respuestas de los docentes. La fiabilidad arrojada por los datos fue de 0,923, en base a Alpha de Cronbach. Sin embargo, considerando la escasez de casos (31 de un universo de 94 casos aproximadamente) no es posible realizar un análisis factorial, ya que aunque el estadístico de Esfericidad de Bartlett es menor a 0,05, el índice de adecuación KMO es de 0.624. En base a esa limitación, se utilizó la estructura teórica definida previamente en el instrumento para analizar los resultados.

6.4.2. Observaciones de docentes y directivos sobre las consecuencias del IED.

Por otra parte, con el fin de profundizar en las perspectivas sobre los usos y consecuencias del instrumento, se analizó de manera cualitativa las observaciones de los docentes y directivos a la encuesta, destacando los temas que fueron coincidiendo, pero sin pretender generalización de estos comentarios. El análisis de tipo temático fue organizado en familias de códigos y relacionado con la evidencia proveniente de la literatura, en términos del “aplicación”, “contenido”, “consecuencias” y “reporte de los resultados”. Los comentarios de docentes llegaron a 75, mientras los comentarios de directivos sólo fueron 14.

Tabla 2

Síntesis Aspectos Metodológicos Centrales

Objetivo	Tipo de análisis	Muestra	Recogida de información	Plan de análisis
Estructura interna	Cuantitativo, diseño no experimental transversal, descriptivo.	48.129 respuestas de 6.437 estudiantes (de 7 de 9 facultades)	Encuesta de evaluación de la docencia	Análisis Factorial Exploratorio y Confirmatorio
Relación con otros criterios (rendimiento académico)	Diseño no experimental transversal, correlacional.	48.129 respuestas de 6.437 estudiantes (de 7 de 9 facultades)	Encuesta de evaluación de la docencia Bases de datos institucionales	Análisis de correlaciones (Pearson)
Relación con otros criterios (experiencia de aprendizaje CEQ)	Cuantitativo, diseño no experimental transversal, correlacional.	310 respuestas de estudiantes a 19 docentes	Encuesta de evaluación de la docencia. CEQ	Análisis de correlaciones. (Pearson)
Relación con otros criterios (percepción sobre la calidad de la docencia IED)	Cualitativo.	5.991 comentarios para 826 docentes en sus respectivas asignaturas, de 8 carreras de distintas unidades mayores.	Respuestas abiertas de estudiantes en la evaluación de la docencia.	Análisis de contenido: codificación temática teórica (deductiva
Consecuencias para docentes y directivos (resultados encuesta)	Cuantitativo, diseño no experimental transversal, exploratorio.	32% de jefes de carrera y directivos (31) 10% de los docentes de la Universidad (200 aprox.)	Encuestas a docentes y directivos: respuestas cerradas.	Análisis descriptivos (frecuencias)
Consecuencias por docentes y directivos (comentarios encuesta)	Cualitativo	75 docentes 14 directivos	Encuesta: respuesta abierta	Codificación temática

VII. Análisis de resultados

7.1. Confiabilidad del IED

Respecto a la confiabilidad del IED, Alfa de Cronbach arroja altos niveles de consistencia interna alcanzando .978 (Ver anexo 3). Al mismo tiempo, el método de bipartición arroja una alta confiabilidad de 0,924 (ver anexo 4). Por otra parte, el análisis de Alfa de Cronbach señala que todos los indicadores aportan de forma significativa a la confiabilidad del instrumento. Ninguno de ellos (al eliminarse) aumenta o baja significativamente a la confiabilidad (todos están por sobre 0,977). Las cuatro dimensiones del instrumento arrojan altos niveles de consistencia interna (superiores al .97).

7.2. Evidencias sobre estructura interna

Inicialmente se encontró que en base a la matriz de correlación es posible realizar un análisis factorial, ya que el índice de adecuación KMO es de 0.97 y el estadístico de Esfericidad de Bartlett es menor a 0,05.

El AFE, utilizando Mínimo Cuadrado no Ponderado, arrojó sólo un Eigenvalue (autovalor) mayor que 1, el cual explica el 70% de la varianza, lo que coincide de manera evidente con el gráfico de sedimentación. Al fijar el número de factores en cuatro y aplicar una rotación Oblimin (ya que se asume relación entre los factores) todos los ítems cargan sobre el 0,655 en un solo factor. Sin embargo, cuando se analizan las cargas factoriales para cada uno de los ítems, se observa cierta coincidencia con las dimensiones definidas teóricamente en el instrumento (Ver anexo 5). Los ítems del 4 al 13 están agrupados en un mismo factor; los ítems del 19 al 21 están agrupados en un segundo factor (aunque de manera negativa); los ítems 1,2, 3 y 18 están agrupados en un tercer factor (aunque de manera negativa); y los ítems del 14 al 17 están agrupados en un cuarto factor.

Sin embargo, al probar el modelo de 4 dimensiones en el AFC, se advierte que los índices de ajuste absolutos y relativos no se cumplen, ya que el índice de RMSEA es de 0,092 (lo que supera el rango de 0,08), de modo que el modelo no es aceptable.

Considerando los resultados, se procedió a evaluar el modelo de una dimensión, encontrándose en esta oportunidad los valores esperados para los índices de ajuste, lo que daba cuenta de que el modelo resultaba aceptable. En la Tabla 3 se presentan los estadísticos de ambos análisis.

Tabla 3

Análisis Factorial Confirmatorio Instrumento Evaluación de la Docencia (IED)

			IED (1 escala)	IED (4 escalas)			
				1	2	3	4
Número de ítem			21	3	9	6	3
Coeficiente de confiabilidad			.9786	.9443	.9551	.9386	.9280
Modelo con covarianzas restringidas	RMSEA	< 0.08	.066		0.092		
	CFI	> 0.9	.977		0.952		
	TLI	> 0.9	.972		0.944		
	SRMR	< 0.05	.015		0.026		

En conclusión, no existe evidencia de validez con respecto a las 4 dimensiones que el instrumento de evaluación de la docencia considera teóricamente en su estructura interna, ya que los resultados obtenidos entregan evidencia de que el modelo es unidimensional.

7.3. Evidencias sobre la relación con otros criterios

7.3.1 Relación con rendimiento académico

Respecto de la relación de las puntuaciones del IED con el rendimiento académico de los estudiantes, medido a través de sus calificaciones, es posible advertir correlaciones estadísticamente significativas, pero bajas ($r = 0,214$; $p < 0,01$; $n=2237$). Es decir, parece existir una relación débil entre la valoración de los estudiantes al evaluar a sus docentes y las calificaciones que obtienen en sus respectivas asignaturas.

Tabla 4

Correlación entre evaluación de la docencia y rendimiento académico de los estudiantes

Correlaciones		Rendimiento académico (calificaciones)
Evaluación de la docencia	Correlación de Pearson	,214**
	Sig. (bilateral)	,000
	N	2237

** . La correlación es significativa al nivel 0,01 (bilateral).

El análisis por área disciplinar sobre dicha relación, revela diferencias interesantes en los niveles de correlaciones entre el IED y el rendimiento académico de los estudiantes. Las correlaciones más robustas (aunque siempre moderadas) entre ambas variables se manifiestan en las carreras de humanidades ($r\ 0,32\ p < 0,01; n=549$) y en las ciencias de la salud ($r\ 0,432, p < 0,01; n=48$). Mientras que en las carreras del área de las ingenierías, ciencias matemáticas y ciencias naturales, las correlaciones son muy bajas (ingeniería: $r\ 0,1\ p < 0,01; n=688$; ciencias matemáticas: $r\ 0,1\ p < 0,01; n=247$) e incluso no significativas en términos de significancia estadística (ciencias naturales: $r\ 0,1\ p > 0,01; n=122$). Esto podría indicar que en algunas áreas de conocimiento los docentes bien evaluados alcanzan mejores rendimientos académicos en sus estudiantes; mientras que en otras áreas esa relación es prácticamente inexistente.

7.3.2. Relación con la experiencia de aprendizaje

Las puntuaciones sobre la evaluación de la docencia y la percepción de los estudiantes sobre su experiencia de aprendizaje (medida a través del CEQ) muestran correlaciones significativas en todas las escalas del instrumento, aunque en algunas más fuertes que en otras. La relación del instrumento de evaluación de la docencia con las escalas “buena docencia” y “metas y objetivos claros” es estadísticamente significativa, de manera concurrente y en un nivel moderado ($r=0,50\ p < 0,01$ y $r=0,45\ p < 0,01, p$; $n=311$, respectivamente) mientras que la relación con las escalas “evaluación apropiada” y “carga académica apropiada”, también es estadísticamente significativa y concurrente, aunque en un nivel bajo ($r\ 0,234\ p < 0,01$ y $r\ 0,314\ p < 0,01; n=311$, respectivamente).

Tabla 5

Correlación entre evaluación de la docencia y experiencia de aprendizaje de los estudiantes

		Correlaciones			
		Buena docencia	Metas y Objetivos claros	Evaluación apropiada	Carga de trabajo apropiada
Evaluación de la docencia	Correlación de				
	Pearson	,501**	,451**	,234**	,314**
	Sig. (bilateral)	,000	,000	,000	,000
	N	311	311	311	311

** . La correlación es significativa al nivel 0,01 (bilateral).

Al mirar el análisis diferenciado por área de conocimiento, la tendencia general no cambia significativamente, sin embargo surgen algunas particularidades interesantes. En todas las

áreas de conocimiento analizadas, la relación del IED con la dimensión “buena docencia” del CEQ es superior a $r=0,43$, salvo el área de las ciencias médicas ($r=182$, $p < 0,01$; $n=32$). Esto podría indicar que para los estudiantes de ciencias médicas, la percepción sobre experiencia de aprendizaje desde la “buena docencia”, no está vinculada con lo que responden en IED. Sin embargo, el bajo tamaño de la muestra en las distintas áreas de conocimiento, podría estar sesgando las correlaciones del análisis diferenciado, lo que obliga a mirarlos con cautela (ver anexo 8).

Por otra parte, se destaca que en el caso de las escalas “evaluación apropiada” y “carga de trabajo apropiada”, el área de humanidades y ciencias naturales develan un comportamiento atípico en relación con los resultados del IED. Para los estudiantes de ciencias naturales y humanidades la correlación entre IED y “evaluación apropiada” es alta ($r=0,708$, $p < 0,01$; $n=22$) y moderada ($r=0,483$, $p < 0,01$; $n=26$) respectivamente, a diferencia del resto de las áreas de conocimiento donde las correlaciones son más bien bajas. Esto podría indicar que para los estudiantes de humanidades y ciencias naturales, la calidad de la docencia está más vinculada con su percepción respecto a sus experiencias con la evaluación de sus aprendizajes, que en el caso de las otras áreas de conocimiento. Sin embargo, el bajo tamaño de la muestra en las distintas áreas de conocimiento, podría estar sesgando las correlaciones del análisis diferenciado, lo que obliga a mirarlos con cautela (ver anexo 8).

7.3.3. Relación con la percepción sobre la calidad de la docencia

Tanto el análisis general (agregando las diversas áreas del conocimiento en un análisis agrupado) como el particular (por área de conocimiento), revela que el código “buen profesor” se repite con frecuencia en los docentes con promedios altos y con promedios medios de IED (incluso en los promedios bajos).

En la Tabla 6 es posible advertir que las categorías más frecuentes comentadas a los docentes evaluados con promedios altos en IED – métodos y estrategias apropiadas, claridad para explicar, aprendizaje activo, preocupado por necesidades de estudiantes, características personales (amable, simpático) y responde claramente a las dudas – son también las más frecuentes para el caso de los docentes evaluados con puntuaciones medias, salvo el caso del código *“preocupado por las necesidades del estudiante”*, que se destaca en 3,2 puntos de diferencia respecto del grupo de puntuaciones medias. Este código que recoge comentarios del tipo *“el profesor está muy atento al desempeño personal de cada alumno, haciendo correcciones adecuadas y pertinentes”* ó *“(…) siempre dispuesto a enseñar, bajo un ambiente de respeto y paciencia para con los que menos saben”*, podría estar diferenciando a los profesores de puntuaciones altas y aquellos que obtiene puntuaciones medias.

Tabla 6

Relación evaluación de la docencia y percepción de la calidad de la docencia (respuestas a pregunta: elementos que el docente debería mantener)

	Agrupación promedios IED de docentes				Diferencia entre puntuaciones altas y puntuaciones medias
	Puntuaciones altas		Puntuaciones medias		
		%		%	
	(4 y 5)		(3,01 a 4)		
Buen profesor	425	12,7	81	12,2	0,5
Métodos y estrategias apropiadas	296	8,8	47	7,1	1,7
Claridad para explicar	276	8,2	38	5,7	2,5
Aprendizaje activo	267	8	37	5,6	2,4
Preocupado por necesidades de estudiantes	243	7,3	27	4,1	3,2
Características personales (amable, simpático)	203	6,1	33	5	1,1
Responde claramente a las dudas	167	5	35	5,3	0,3

Ver tabla completa en anexo 11

De todas formas, la inexistencia de diferencias determinantes entre estos tipos de comentarios en los puntajes medios o altos, permitiría esbozar que el instrumento no está identificando a los docentes de acuerdo a las características que sus estudiantes les asignan.

A pesar de que para el análisis general y para algunas áreas de conocimiento, no existen diferencias entre ambos niveles de puntuaciones en el IED, en algunas áreas tales como Bachillerato, Ciencias Naturales, Tecnologías y Ciencias Médicas sí demuestran diferencias entre los docentes con puntajes altos y medios, que deberían ser exploradas en futuras investigaciones (ver anexo 11).

Por otra parte, tal como se plantea en la Tabla 6, sobre las “sugerencias de mejora”, en ambos grupos son frecuentes los comentarios asociados al profesionalismo docente “*métodos y estrategias inapropiadas*” y “*métodos de evaluación inválidos*”. Es decir, elementos de la dimensión “profesionalismo docente”.

Sin embargo, existen algunas diferencias entre docentes con puntuaciones altas y bajas, que tienen que ver con que en los primeros se destacan comentarios que anteceden al docente (“mal diseño del curso horario, nivel, malla” y “problemas de espacio”) y con debilidades en sus métodos de evaluación (métodos de evaluación inválidos). Mientras que en el grupo de docentes con puntuaciones medias, destacan los comentarios sobre la claridad de los docentes para explicar.

Tabla 7

Relación evaluación de la docencia y percepción de la calidad de la docencia (respuestas a pregunta: elementos que el docente debería mejorar)

Categoría	Agrupación promedios IED de docentes				Diferencia entre puntuaciones altas y puntuaciones medias
	Puntuaciones altas (4 y 5)	%	Puntuaciones medias(3,01 a 4)	%	
Mal diseño del curso (horario, malla, nivel)	241	12,8	61	7,1	5,7
Métodos y estrategias inapropiadas	215	11,4	110	12,7	1,3
Métodos de evaluación inválidos	149	7,9	42	4,9	3
No centrado en aprendizaje activo	111	5,9	45	5,2	0,7
Material de apoyo inapropiado	93	5	27	3,1	1,9
Problemas de espacio	93	5	10	1,2	3,8
Irresponsabilidad en el proceso de aprendizaje	67	3,6	52	6,0	2,4
Poca claridad para explicar	47	2,5	49	5,7	3,2
No prepara previamente la enseñanza	87	4,6	47	5,4	0,8

Ver tabla completa en anexo 12

No obstante lo anterior, considerando la totalidad de los antecedentes expuestos, la evidencia reportada por este análisis, se manifiesta en contra de la capacidad del instrumento de evaluación de la docencia al momento de discriminar la calidad de los docentes. El análisis revela que al momento de juzgar la calidad de la docencia, los estudiantes destacan dos elementos: la preocupación por sus necesidades y sus habilidades de profesionalismo docente. Esto al mismo tiempo entrega evidencia interesante respecto de los procesos de respuesta de los estudiantes para este tipo de instrumentos.

7.4. Evidencias sobre consecuencias de los resultados IED

7.4.1. Consecuencias sobre el uso de resultados del IED para docentes

Inicialmente cabe destacar que de los 204 docentes que respondieron la ECIED, un 84,1% declara revisar sus informes de evaluación de la docencia “todos” o “casi todos” los semestres lectivos, lo cual podría evidenciar una preocupación constante por la información emanada de los resultados del IED. Respecto de la dimensión “consecuencias descriptivas: valoración sobre desempeño”, es posible observar que más del 50% de los docentes manifiesta que “siempre” usa los resultados de la encuesta con fines de valorar su desempeño como docente. Es decir, más de la mitad de los docentes señalan que utilizan los resultados de la evaluación de la docencia para, por ejemplo, “reconocer el grado de satisfacción de sus estudiantes con su desempeño como profesor” (71,5% responde siempre) (ver anexo 19). En el caso de la dimensión “consecuencias prescriptivas: cambios y mejoras a las prácticas”, la mayoría de los

docentes señala que “siempre” utiliza los resultados IED con estos propósitos, aun cuando sólo uno de los reactivos supera el 50% en esa categoría (“mejorar mi forma de explicar los contenidos o la materia del curso” con un 56% en la categoría “siempre”). De todas formas, la suma de “siempre” y “casi siempre” sí supera en todos los reactivos el 50%, lo que implica que aun cuando los docentes declaren con mayor convicción utilizar los resultados de las encuestas con fines de reconocer la valoración de sus prácticas, también declaran utilizar los resultados para mejorar sus prácticas docentes. Así, por ejemplo, más del 74% de los docentes declaran que “siempre” o “casi siempre” utilizan los resultados de las encuestas para “mejorar su forma de evaluar los aprendizajes de sus estudiantes” (ver anexo 19).

Finalmente, en la dimensión “consecuencias prescriptivas: decisiones sumativas” la mayoría de los docentes de la Universidad declaran que “nunca” han utilizado la encuesta de evaluación de la docencia con fines de postular a premios, incentivos, y justificar a favor de su continuidad en la institución o promover su promoción a otros cargos. Resulta particularmente interesante que existiendo la figura de premios a la calidad de la docencia en la Universidad, sólo un 20% de los docentes que respondieron la encuesta señala que “siempre” o “casi siempre” ha utilizado los resultados de la evaluación de la docencia para postular a dichos incentivos (ver anexo 19).

En síntesis, los docentes declaran utilizar los resultados de sus evaluaciones docentes con fines descriptivos, para reconocer la percepción sobre su docencia y con fines de mejorar las prácticas que identifican como débiles, es decir, con fines formativos. Por otra parte, parece ser que los fines sumativos no resultan tan habituales en el uso de las evaluaciones docentes.

7.4.2. Consecuencias sobre el uso de resultados del IED para directivos

Se destaca un alto uso de los resultados del IED con fines descriptivos por parte de los directores de departamento y jefes de carrera. La mayoría de los directivos “siempre” o “casi siempre” utilizan los resultados de la evaluación de la docencia para reconocer el grado de satisfacción de los estudiantes con el desempeño de sus docentes (79,31%); reconocer el grado de satisfacción de los estudiantes con la implementación de los cursos (82,74); y para identificar a los docentes con “mejores o peores” habilidades pedagógicas (72,4) (Ver anexo 21).

Cuando se les pregunta por “consecuencias prescriptivas formativas”, es decir, por el uso de los resultados del IED a favor de acciones que promuevan la mejora de las habilidades pedagógicas de los docentes, la tendencia de los directivos es a señalar que sólo “a veces” o “nunca” realizan este tipo de acciones. Particularmente, en aquellos reactivos que declaran acciones asociadas a tareas puntuales de aula, como sugerir cambios en las evaluaciones de aprendizaje, metodologías de enseñanza o intervención en problemas de convivencia, más de un 60% de los directivos declaran “nunca” o “a veces” utilizar los resultados de las

evaluaciones en esa línea. Sin embargo, cuando se les pregunta de manera general respecto de si utilizan los resultados del IED para “orientar mejoras al desempeño de los docentes de su equipo, a partir de la retroalimentación de sus prácticas”, la mayoría de los directivos (68,97%) declara “siempre” o “casi siempre” utilizar los resultados de las encuestas con estos fines (ver anexo 21).

Respecto de las consecuencias prescriptivas de tipo sumativas, es posible observar que en general el uso de los resultados está más bien asociado a la rendición de cuentas y a la gestión del profesorado, más que a la desvinculación, promoción o premiación de los docentes. La mayoría de los directivos declaran que “siempre” o “casi siempre” utilizan los resultados de la evaluación docente para orientar la planeación docente (más del 75,86%), para reportar la calidad de la docencia de su unidad académica a instancias de acreditación o proyectos (62,07%) y para reportar antecedentes a la denominada “Evaluación del Desempeño Académico” (62,07%). Consistentemente, la mayoría de los directivos señalan que sólo “a veces” o “nunca” han utilizado los resultados de las evaluaciones docentes para “sugerir el despido de los docentes de su equipo” (69,96%) o para “entregar premios, reconocimientos o recompensas a la calidad de la docencia” (65,52%) (Ver anexo 21).

Futuros estudios deberían indagar en los usos efectivos de que hacen los docentes y directivos de este tipo de instrumentos, con el fin de tener algunas certezas sobre las declaraciones que ellos realizan.

7.4.3. Observaciones de docentes y directivos sobre las consecuencias de los resultados de la evaluación de la docencia

Respecto de la aplicación del instrumento, los comentarios de algunos docentes coinciden en declarar que éste no es respondido de manera “objetiva” por parte de los estudiantes, debido a que éstos *“no se toman en serio la evaluación de la docencia”* y responden con alta discrepancia a las mismas en elementos “objetivos”, como “puntualidad” o “entrega de programa”. Según ellos, esto está mediado porque los estudiantes responden al final del curso, lo que evidencia sesgos respecto de las calificaciones obtenidas o falta de “conciencia” al responder. Uno de los docentes señala: *“he observado que, en mis cursos, la opinión de los estudiantes está muy correlacionada con el promedio final del curso y el porcentaje de estudiantes aprobados”*.

Otra de las variables que identifican docentes y directivos entre los posibles sesgos, tiene que ver con el exceso de respuestas que deben completar los estudiantes cada semestre y con la presión de “responder rápido” a las encuestas para lograr inscribir asignaturas cada semestre. Adicionalmente, otros docentes atribuyen las respuestas sesgadas de los estudiantes a la “simpatía” de cada profesor y no a sus condiciones para ejercer la docencia (ver anexo 22 y 23).

En relación al contenido del instrumento, a pesar de que algunos docentes declaran que los instrumentos han mejorado en los últimos años, algunos docentes y directivos, también declaran que hay cosas que aparecen en los instrumentos y que “no dependen de mí” o que “sólo se preguntan aspectos formales”. En ese sentido, son habituales los comentarios que solicitan distinciones entre instrumentos para distintos tipos de asignaturas, ya que los cursos con dinámicas distintas a las de aula regular, como el laboratorio o talleres, son evaluados con ítems que no aplican a sus realidades. Por ejemplo, uno de los docentes declara de manera textual que:

“El sistema de evaluación no sirve en el caso de mi Facultad, donde los cursos son impartidos por varios profesores (no hay un único profesor asignado al curso) y donde la mayoría de los cursos tienen actividades de laboratorio asociadas. El sistema de evaluación NO DISTINGUE el grado de participación del profesor (si hace cátedra, laboratorio, o ambas) ni el grado de compenetración del profesor con el curso (no es lo mismo si soy el coordinador del curso o si solo participo dando ciertas clases específicas). Esto produce distorsiones absurdas. Por ejemplo, cuando se pregunta si el profesor entregó el programa del curso, eso vale para el coordinador quien entrega el programa a principio de año, pero el resto de profesores no entrega de nuevo el programa, así que salen mal evaluados en eso. O cuando se pregunta por el uso de TIC...en un laboratorio de ciencia no se usan TIC (se hacen experimentos en el mesón) y por tanto el profesor sale mal evaluado”.

Sobre el reporte de resultados, algunos docentes y directivos declaran no saber cómo acceder a los informes o directamente no tener acceso a ellos. De forma particular algunos docentes reconocen no saber que existían los informes o incluso olvidar revisar los resultados. La mayoría de los docentes así como la mayoría de los directivos, declaran que los reportes son “inoportunos”, ya que son entregados un semestre o un año después cuando las posibilidades de acciones son mucho más limitadas. Al respecto uno de los directivos señala que “[los informes] debieran llegar antes que finalice el semestre para poder tomar decisiones contemplando los resultados”.

Respecto del tipo de informe de resultados, algunos docentes y directivos declaran no acceder a un “informe” propiamente tal, ya que en algunas facultades sólo se están reportando datos de las encuestas sin constituir un informe, lo que dificulta su lectura. Uno de los directivos señala “es complejo porque no se entrega un sintetizado de las encuestas sino que las encuestas de manera individual, en el mejor caso”.

De la misma forma, algunos docentes declaran que reporte de la información resulta complejo de comprender y que los comentarios de respuestas abiertas son los elementos más valorados. Adicionalmente, estos profesores destacan que el reporte de resultados debería

incluir reportes históricos para facilitar las diferencias entre semestres y diferenciación entre los profesores por hora y de los docentes contratados a jornada en la Universidad.

Sobre las consecuencias de los resultados de la evaluación de la docencia, los docentes declaran que en términos descriptivos las encuestas resultan muy importantes para “*saber cómo me ven los estudiantes*”. Sin embargo, tanto docentes como directivos coinciden en que no se puede determinar si un docente es buen o mal profesor por las encuestas, sino que es necesario complementar dichos antecedentes con otras fuentes y métodos de información (ver anexo 22 y 23). Por ejemplo, uno de los docentes señala que:

“Creo que el proceso de evaluación de la docencia, es fundamental. En este caso quizás debería considerarse otros elementos de entrada, si bien es cierto que lo fundamental es el alumno, una manera de hacer que este proceso de evaluación sea íntegro, creo que debe tener la mirada de al menos la dirección de la Facultad. Las entrevistas con directores de carrera, jefes de carrera, etc. (...)”

Respecto de las consecuencias prescriptivas, en términos formativos, es posible identificar que la mayoría de los docentes (al igual que en la encuesta de respuesta cerrada) declaran que los resultados de la encuesta les permite mejorar sus prácticas docentes, incluso algunos de ellos declaran motivarse a tomar cursos de perfeccionamiento para impulsar dicha mejora. Uno de los docentes señala que:

“Leo todos los semestres mis informes de evaluación de la docencia, para corregir algunos aspectos débiles de mi quehacer y potenciar aquellos que representan una fortaleza en mi trabajo frente alumnos”.

Por su parte, algunos directivos de la Universidad declaran utilizar la encuesta de evaluación de la docencia para invitar a los docentes a conversar sobre los resultados y para comparar los progresos de los docentes semestre a semestre.

En términos de consecuencias sumativas del uso de los resultados de la evaluación docente, algunos profesores insisten en la idea de que los resultados se deben utilizar para gestionar la docencia en las unidades académicas, sin embargo los directivos identifican dificultades institucionales para hacer un uso efectivo de las encuestas en esos términos. Por ejemplo, uno de ellos señala que “*en general mis docentes están bien evaluados y a los que no lo están, no hemos podido aplicar medidas por ser de planta*”.

VIII. Discusión de los resultados

La presente investigación recogió una serie de evidencias sobre la validez del instrumento de evaluación de la docencia aplicado en la Universidad de Santiago de Chile, que plantean argumentos a favor y contra de su utilización. Numerosas investigaciones se han dedicado a analizar la validez de este tipo de instrumentos, sin embargo, muy pocas han presentado evidencias diversas para elaborar un argumento de validez del uso de los resultados de un instrumento en particular.

8.1. Confiabilidad del instrumento

La alta consistencia interna del instrumento coincide plenamente con los hallazgos de la literatura respecto de la alta fiabilidad de los resultados del IED (Richardson, 2005). Esto da continuidad a los hallazgos de la literatura, pero además aporta antecedentes a favor de la validez del instrumento.

8.2. Evidencias sobre la estructura interna

Los hallazgos en esta materia están alineados con los antecedentes presentados por la literatura respecto de la estructura interna de los instrumentos de evaluación de la docencia, donde a pesar de que los distintos instrumentos planteen diversas dimensiones asociadas a Modelos de Docencia Efectiva desde un punto de vista teórico, la evidencia empírica reporta la configuración de una estructura unidimensional en los instrumentos (Abrami et al., 1991; García, 2000; Richardson, 2005).

Tal como señala la literatura (Spooren, et al, 2013), este antecedente es negativo para la validez del IED respecto de sus fines formativos, debido a la escasa información que podría reportar el instrumento definido inicialmente para retroalimentar en cuatro dimensiones a los docentes. Sin embargo, podría ser positivo si los fines están asociados a las decisiones sumativas como la promoción o el despido de los académicos. Este antecedente debe ser considerado al momento de redefinir el instrumento de evaluación de la docencia en la Universidad, ya que los fines que guiarán los usos del instrumento deberán ser coherentes con la estructura interna del mismo.

Considerando la literatura sobre los procesos de respuesta de SET, la estructura unidimensional del IED podría estar interferida por dos elementos determinantes para los estudiantes al momento de responder: 1) un sesgo asociado a una única característica del docente que podría tener que ver con su carisma, atractivo físico o comportamiento no verbal de sus profesores (Ambady & Rosenthal; 1993; Shevlin et al., 2000); ó 2) un “efecto halo”, que la literatura ha denominado “habilidad general para enseñar” (Abrami, et al., 1991) o, según la perspectiva de Ramsden (2003), la valoración sobre la “experiencia de aprendizaje de los estudiantes”. Estos antecedentes deben ser profundizados en futuras investigaciones, que

indaguen en los procesos de respuesta de los estudiantes, al momento de responder los cuestionarios.

8.3. Evidencias sobre la relación con otros criterios

8.3.1 Relación con la experiencia de aprendizaje

Siguiendo el argumento sobre la estructura interna, la relación concurrente entre IED y la percepción de los estudiantes sobre su experiencia de aprendizaje en todas las escalas del instrumento (Buena Docencia: $r=.501$; Metas y Objetivos Claros: $r=.451$; Evaluación Apropiaada: $r=.234$; Carga de Trabajo Apropiaada: $r=.314$), podría complementar la idea de que los estudiantes están considerando su experiencia de aprendizaje en la asignatura al momento de evaluar a sus profesores, más que sólo el carisma o atractivo de sus profesores. Además, considerando que la escala “Buena Docencia” de CEQ (Webster, 2009) es la que presenta la correlación más alta con IED ($r=.501$), se podría señalar que el constructo base de las respuestas de los estudiantes es “habilidad general para enseñar” (Abrami, et al., 1991).

Esto puede interpretarse como que el instrumento de evaluación docente estaría efectivamente midiendo variables relacionadas con la experiencia de aprendizaje de los estudiantes, lo que aporta argumentos a favor de la validez del IED.

Además, este análisis permitiría relacionar positivamente un Modelo de Docencia Efectiva (Encuesta de evaluación de la Docencia de la Universidad de Santiago) con un instrumento basado en la percepción de los estudiantes sobre su propio proceso de aprendizaje (CEQ), tal como plantea el Modelo 3P, basado en la teoría de Biggs (2008) y Ramsden (2003). Sin embargo, la correlación entre ambos instrumentos no es alta, lo que reafirma la idea de que ambos instrumentos miden el proceso desde un posicionamiento teórico distinto, que podría ser sensible a las respuestas de los estudiantes.

Este antecedente instala una reflexión importante para la institución, respecto al qué evaluar: la calidad de la docencia o la experiencia de aprendizaje de los estudiantes, como parte constitutiva de la calidad de la docencia. Futuras investigaciones, quizás de orden cualitativo, podrían indagar en la relación de las habilidades de los docentes para enseñar y la experiencia de aprendizaje de los estudiantes, considerando la diversidad de los contextos educativos.

8.3.2. Relación con la percepción de los estudiantes sobre la calidad de la docencia

El análisis sobre la percepción de los estudiantes sobre la calidad de la docencia, aporta nuevos antecedentes sobre la validez del instrumento. Teniendo en cuenta que las dos dimensiones más valoradas por los estudiantes, a la hora de percibir la calidad de la docencia, son “docencia centrada en el estudiante” y “profesionalismo docente” (Ramsden, 2003; Onwuegbuzie et al., 2007), la validez del IED de la Universidad de Santiago resulta

cuestionada, ya que sólo tres de los 21 ítems del instrumento de evaluación de la docencia están relacionados con las dos dimensiones.

Estos tres ítems son los siguientes: 8: “fomenta la participación de los/las estudiantes en clases (mediante preguntas, debates, entre otros); 9: “comunica los contenidos de forma clara”; y “11: utiliza distintas estrategias de enseñanza para facilitar el logro de los aprendizajes”. En lugar de considerar elementos que efectivamente resultan importantes para los estudiantes al momento de evaluar la calidad de la docencia, la encuesta considera otros factores como *“entrega documentos (apuntes, guías, artículos, otros) que complementan el desarrollo de las actividades educativas”* y *“utiliza los recursos bibliográficos definidos en el programa, como referentes para el desarrollo de las clases”*, que no resultan determinantes, ni en el estudio de Onwuegbuzie et al., (2007) ni en los resultados del presente análisis.

Por otra parte, el análisis alerta sobre la escasa capacidad del IED para discriminar la calidad de los docentes. En primer lugar, no es posible diferenciar un número considerable de docentes en las tres categorías de promedios definidas; en otro sentido, el código “buen docente” se repite con frecuencia en docentes con promedios altos, medios y bajos; además, no se observan diferencias respecto de los códigos asociados a promedios altos y medios (salvo en el código “preocupado por las necesidades de los estudiantes”); y los docentes evaluados con promedios altos en el IED, están recibiendo de igual manera comentarios sobre sus debilidades en competencias profesionales docentes.

Por ejemplo, mientras los estudiantes evalúan con altos puntajes a sus docentes en el ítem *“explica a los/las estudiantes con anterioridad, los criterios de evaluación definidos para cada instancia (pruebas, trabajos, disertaciones, etc.)”*; al mismo tiempo realizan comentarios del tipo *“podría haber una pauta de evaluación más clara para desarrollar ficheros, ya que no siempre se sabía por qué las notas obtenidas por el curso”* para ese mismo grupo profesores.

Este tipo de antecedentes entrega recomendaciones importantes a la Institución. Por una parte, se destaca la riqueza de la información cualitativa que entregan las respuestas abiertas de los estudiantes, para retroalimentar la docencia, lo que invita a continuar perfeccionando esta práctica. Además, destaca las dimensiones que podrían constituir las reformulaciones del IED: “la docencia centrada en el estudiante” y el “profesionalismo docente”, desde la experiencia de aprendizaje del estudiante (considerando el análisis previo).

Esta línea de desarrollo debería ser profundizada en futuras investigaciones, con el fin de indagar en las valoraciones de los estudiantes sobre la calidad de la docencia. Este tipo de antecedentes podría iluminar el diseño de mejores instrumentos de evaluación de la calidad de la docencia, desde la opinión de los estudiantes.

8.3.3. Relación con el rendimiento académico

La positiva pero débil relación entre el IED y el rendimiento académico de los estudiantes de la Universidad de Santiago ($r\ 0,214$; $p < 0,01$; $n=2037$) coincide con los hallazgos de Clayson (2009), pero discrepa con uno de los estudios más recientes en la materia (Griffin et al., 2009), que encuentra correlaciones positivas y moderadas entre ambas variables ($r\ 0,37$; $p < 0,01$; $n=2073$).

Por una parte, esto aporta evidencias en contra de la validez del instrumento de evaluación de la docencia, ya que no cumple la lógica de que una buena docencia generaría un positivo rendimiento académico. Sin embargo, esta evidencia al mismo tiempo puede resultar positiva para la validez del instrumento, ya que contradice las afirmaciones de algunos docentes y directivos, que aseguran una relación estrecha entre las calificaciones de los estudiantes y las puntuaciones en el IED.

Por otra parte, al igual que en el estudio de Clayson (2009), se advierten diferencias importantes entre las distintas áreas de conocimiento. Las disciplinas humanistas – quizás con métodos de medición vinculados a situaciones de evaluación de desempeño, como ensayos o demostraciones – se relacionan de manera más fuerte con el IED que las áreas de conocimiento más cercanas a las ciencias exactas – con métodos de evaluación asociados a pruebas de opción múltiple. Esto no sólo aporta evidencias que cuestionan la validez del IED respecto de su relación con el rendimiento académico, sino que además cuestiona su aplicación sin adaptaciones en distintos contextos educativos, debido a las diferencias que podría tener su utilización en las distintas unidades académicas (Clayson, 2009).

Estos antecedentes podrían comprenderse a través de dos hipótesis distintas que pueden ser complementarias: (1) debido a su forma de medición (tipo y calidad de las situaciones evaluativas e instrumentos), las calificaciones no representan calidad en el aprendizaje (Clayson, 2009) por lo cual no tendrían por qué relacionarse de manera fuerte con las puntuaciones IED; y (2) la encuesta de evaluación de la docencia se centra en lo que hacen los docentes, lo que podría o no incidir en el rendimiento académico de los estudiantes, ya que dicho rendimiento está mediado por otros factores que no obedecen sólo al quehacer del docente (Ramsden, 2003). Tal como señala Clayson (2009), es necesario realizar un análisis más detallado para visibilizar algunas diferencias al respecto. Futuras investigaciones deberían focalizar su atención en dichas diferencias, analizando las situaciones evaluativas e instrumentos implicados, por ejemplo.

8.4. Evidencias sobre las consecuencias de los resultados para docentes y directivos

Los antecedentes evidenciados tanto en la encuesta a docentes como en la encuesta a directivos, se muestran parcialmente distintos a los evidenciados en otras investigaciones. Coincidiendo con el estudio de Beran et al. (2005), la mayoría de los docentes encuestados declaran sí revisar y utilizar los resultados de sus evaluaciones de la docencia para reconocer la valoración sobre su enseñanza. Sin embargo, a diferencia del estudio de Beran et al. (2005), donde los profesores universitarios declaran no incorporar cambios a su enseñanza, la mayoría de los docentes encuestados en este estudio declaran sí utilizar la evaluación docente para incorporar cambios en sus labores de enseñanza.

En el caso de los directivos, la evidencia reportada en este estudio no coincide de manera concluyente con el estudio de Beran et al. (2005), ya que a diferencia de esa investigación, la mayoría de los directivos o jefes de carrera de la Universidad de Santiago declaran no utilizar los resultados de las evaluaciones con fines de orientar la permanencia, promoción o incentivos a los docentes de sus unidades académicas. Sin embargo, sí declaran utilizar los resultados con fines de rendición de cuentas en instancias de acreditación o proyectos

Por otra parte, los usos que docentes y directivos declaran asignarle al IED, aportan antecedentes interesantes a la luz de los análisis expuestos anteriormente. Este estudio, a diferencia de lo planteado en la literatura (Beran et al. 2005), señala que los docentes sí declaran utilizar los resultados del IED para la mejora de la enseñanza; y que los directivos no están utilizando estos resultados con fines de retención o incentivos para sus equipos académicos. A diferencia de otras investigaciones (Spooren et al, 2013), los hallazgos de este estudio podrían indicar que los propósitos formativos del uso del IED podrían no estar distorsionados en razón de los fines sumativos. Esto debería ser considerado como un elemento a mantener en posibles mejoras de los procesos, clarificando los propósitos que conducirán los usos del instrumento.

Sin embargo, parece ser que las preocupaciones de los docentes y directivos, que en parte se han evidenciado en este estudio, los han llevado a tener precaución con el uso de los resultados las evaluaciones. Esto se relaciona directamente con los comentarios de docentes y directivos, sobre el reporte inoportuno de los resultados y ciertas críticas a las formas de aplicación del instrumento, que hace a los estudiantes responder a la encuesta por ciertos sesgos o basados en una sola característica o condición, bajo la presión de la inscripción de asignaturas. Este elemento obligaría a replantear las formas de aplicación y entrega de resultados del instrumento, con el fin de reducir las críticas existentes a nivel institucional y reducir los posibles sesgos asociados a la medición.

IX. Conclusión

Las diversas evidencias analizadas permiten señalar que la validez del IED de la Universidad de Santiago de Chile resulta parcial.

Algunas de las evidencias expuestas aportan antecedentes a favor de la validez del instrumento de evaluación de la docencia. La alta confiabilidad de las puntuaciones, que coincide con la literatura al respecto (Richardson, 2005; Costin et al, 1971), advierte sobre la estabilidad los resultados implicados en el proceso. La relación positiva entre el IED y las experiencias de aprendizaje de los estudiantes, es favorable para un instrumento de éstas características. Y por otra parte, los usos descriptivos y formativos que los docentes hacen del IED, permiten determinar potencialidades positivas del instrumento respecto de sus usos para la mejora de la docencia.

Sin embargo, existen diversos antecedentes que cuestionan la validez del instrumento. Las evidencias sobre la unidimensionalidad de su estructura interna del IED, aquejan la validez del mismo, porque a pesar de que facilita la entrega de resultados para fines sumativos (de ranking en la promoción o el despido) reduce la información entregada a los docentes para la mejora focalizada de sus prácticas pedagógicas.

Por otra parte, la débil relación de las puntuaciones del IED con el rendimiento académico de los estudiantes, aunque niega la interferencia de las calificaciones como sesgo en la medición (preocupación de los docentes), también demuestra que no se cumple la idea de que una buena docencia se relaciona positivamente con buenos rendimientos en el aprendizaje de los estudiantes.

Además, la escasa capacidad del IED para discriminar la percepción de la calidad de la docencia de los estudiantes, podría evidenciar que existen otros constructos determinantes al momento de responder, que no se relacionan con los elementos que se preguntan en el IED.

Finalmente, la alta dispersión de los usos de IED (descriptivos, formativos y sumativos) por parte de los directivos, y las preocupaciones de docentes y directivos sobre la aplicación, contenido y consecuencias del uso de los resultados del instrumento, obligan a revisar las definiciones conceptuales, procedimientos de aplicación y deseabilidad de relación con otras variables del instrumento; e invitan a clarificar sus propósitos en relación a sus prácticas, para promover un uso efectivo del instrumento en el proceso de aseguramiento de la calidad de la docencia de la Universidad.

Las evidencias analizadas abren nuevas perspectivas que demandan continuidad en la línea de investigación, promoviendo el análisis y comprensión de los fenómenos que implican la calidad de la docencia en la Universidad. Nuevas investigaciones podrían aventurarse a profundizar en las consecuencias derivadas del uso de SET, más allá de las declaraciones de docentes y directivos especialmente en los procesos de enseñanza y aprendizaje; a diferenciar la calidad de la docencia

en distintos contextos de enseñanza y áreas del conocimiento; a analizar el constructo calidad de la docencia en el contexto regional y nacional; a conocer la opinión de los estudiantes respecto a los procesos de evaluación de la docencia; y a analizar nuevos modelos para evaluar la docencia en la Universidad, entre otros elementos.

En términos específicos, este estudio pretende alertar a otras instituciones de educación superior de Chile, sobre la necesidad abordar de manera crítica y con cautela el uso de los instrumentos de evaluación de la docencia que tan difundidos se encuentra en la mayoría de las universidades del mundo. Las conclusiones de este estudio proveen de un marco teórico y metodológico, que podría promover análisis más profundos y complejos sobre la validez de los instrumentos de evaluación de la docencia en otras universidades chilenas y latinoamericanas. Los hallazgos presentados invitan a analizar los diseños y aplicaciones que circunscriben a estos instrumentos, para orientar con responsabilidad las estructuras internas que los componen, las relaciones convergentes y divergentes que se debería esperar de ellos y los alcances de sus consecuencias y usos en las comunidades universitarias.

Este estudio plantea desafíos a las políticas de mejora de las universidades y abre nuevos cuestionamientos sobre los acelerados procesos de cambio que ocurren en las dinámicas de desarrollo de la educación superior, en pleno proceso de transición desde un modelo de universidades de elite a un modelo de educación superior masificado. La necesidad de repensar la enseñanza para este nuevo escenario, requiere de instrumentos ampliamente fundamentados, que ayuden a avanzar en un camino de mejora continua.

X. Recomendaciones

La evidencia a favor y en contra de la validez del instrumento de evaluación de la docencia de la Universidad de Santiago de Chile, sugiere tomar decisiones institucionales dirigidas a reformular el cuestionario, manteniendo los hallazgos positivos y mejorando las debilidades del proceso.

De manera concreta se sugiere actualizar el IED de la Universidad de Santiago por uno que avance desde un modelo de docencia efectiva, hacia modelos que focalicen su atención en las experiencias de aprendizaje de los estudiantes, considerando los diversos factores contextuales que implican las diversas asignaturas. Es decir, se propone avanzar a definir con claridad el qué evaluar con el IED, si la calidad de la docencia en sí mismo o la experiencia de aprendizaje de los estudiantes, considerando la base teórica del Course Experience Questionnaire (CEQ).

En ese escenario, los estudiantes deberían constituirse en uno de los actores que opinan acerca de sus experiencias de aprendizaje en sus respectivas asignaturas; y no sobre las competencias que sus docentes deberían demostrar al momento de enseñar. En términos concretos, la opinión de los estudiantes debería circunscribirse a aquellos elementos que ellos sean capaces de juzgar en su rol de aprendices, como por ejemplo, si el curso les facilitó el proceso de aprendizaje o las actividades les permitieron participar de la experiencia; y no en elementos del tipo *“realiza actividades que le permiten conocer los aprendizajes previos de sus estudiantes al inicio de la asignatura (ya sea de manera escrita, oral, presencial u online)”* (Anexo 1), entre otros factores.

De acuerdo a los hallazgos del estudio, el nuevo instrumento debería responder a un constructo unidimensional del tipo “buena docencia” o podría constituir una estructura bidimensional, asociada a los constructos “centrado en el estudiante” y “profesionalismo docente” (desde la experiencia de aprendizaje) considerando los elementos más importantes en cuanto a la percepción de los estudiantes sobre la calidad de la docencia en los comentarios de la encuesta IED.

Al mismo tiempo, se sugiere avanzar hacia mejores procedimientos y formas de aplicación del IED, buscando matizar la preocupación del profesorado y directivos en el uso de los resultados. En términos concretos, se sugiere reconsiderar la cantidad de respuestas e ítems que los estudiantes deben responder, además de la obligatoriedad de las respuestas para la inscripción de asignaturas (sin poner en riesgo la alta tasa de respuesta que actualmente mantiene la Universidad). Estas condiciones podrían estar incidiendo negativamente en las respuestas de los estudiantes.

Al mismo tiempo, el estudio sugiere transparentar los fines de la evaluación de la docencia en la comunidad universitaria, con sólidos fundamentos que avalen sus usos en términos descriptivos, formativos y/o sumativos. Esto debería incentivar un diseño democrático de los fines de los instrumentos, la revisión de los informes de directivos y docentes, las normativas

(reglamentos internos, incentivos, por ejemplo), y los medios de difusión y discursos instalados en la cultura institucional, con fin de resguardar las acciones vinculadas al sistema de evaluación de la docencia.

Cabe destacar que actualmente la comunidad universitaria prioriza lo fines formativos por sobre los sumativos IED, por lo cual es necesario resguardar ese equilibrio en nuevos procedimientos. Para profundizar los fines formativos, además de mejorar el instrumento de respuesta cerrada, se debería enriquecer el tipo de información reportada, de manera contextualizada a cada realidad educativa. Para ello se recomienda continuar complementando el instrumento como procedimientos cualitativos, como respuestas abiertas, que profundicen la retroalimentación sobre la calidad de la docencia.

Entendiendo que la calidad de la docencia es un constructo complejo, multidimensional y complejo, no es posible circunscribirlo a la exclusiva opinión de los estudiantes. Por ello, con el fin de enriquecer la información reportada por el IED, una de las estrategias podría estar orientada a diversificar las fuentes de información. Para ello se sugiere vincular a otros agentes evaluativos, como los propios docentes, sus pares y sus directivos, a través de procedimientos cualitativos o cuantitativos – entrevistas, portafolios pedagógicos, encuestas, etc. – que promuevan procesos de reflexión educativa al interior de las unidades académicas, y que se centren en abordar válidamente los fines de la evaluación de la docencia. Si los fines están centrados en promover la mejora de la calidad de la docencia, los instrumentos con sus alcances y limitaciones transparentadas, podrían apoyar acciones formativas y sumativas de manera articulada y coherente, de acuerdo a los fines para los que fueron o serían diseñados.

A modo de proyección este estudio sustentará las innovaciones en el ámbito de la evaluación de la docencia universitaria de la Universidad de Santiago de Chile. De manera inmediata, se socializarán los hallazgos de la investigación en la comunidad universitaria, invitando a actores claves a constituir una mesa de trabajo para definir democráticamente los propósitos del instrumento y las diversas etapas de desarrollo para la mejora del mismo. Este proceso será constitutivo de una nueva investigación en la materia.

En paralelo, se elaborará un artículo de difusión del estudio, que será enviado a una revista científica especializada en el área de la evaluación de procesos en educación superior. Además, al mismo tiempo, se presentará una ponencia con los resultados del estudio en el Encuentro de Centros de Apoyo a la Docencia (ECAD) del año 2015, con el fin de difundir los resultados en la comunidad nacional dedicada a promover la mejora de estos procesos en las universidades chilenas.

XI. Limitaciones del estudio

El presente estudio contó con innumerables ventajas asociadas a la disponibilidad de la mayoría de las bases de datos institucionales, al acceso a muestra del estudio y a la posibilidad de contar con apoyo en la codificación; sin embargo, existieron algunas restricciones específicas que no lograron resolverse y que limitan las conclusiones presentadas.

Un primer elemento tiene que ver con el acceso a las bases de datos de dos facultades de la Universidad. Por encontrarse alojadas en plataforma informática distinta al resto de la información, las bases de datos de estas facultades requerían de permisos especiales que no lograron conseguirse en los tiempos disponibles. Esto obliga a mirar con cautela varias de las conclusiones en éstas unidades académicas.

Otra limitación se asocia a la imposibilidad de lograr participar de la aplicación del instrumento de evaluación de la docencia, lo que no permitió incorporar preguntas específicas que podría haber enriquecido el estudio. Esto obligó a trabajar con las preguntas abiertas que considera el instrumento, que resultan pertinentes pero no completamente apropiadas para el análisis de las percepciones de la calidad de la docencia. En futuros análisis de este tipo, se sugiere reformular las preguntas realizadas, utilizando como referencia el estudio de Onwuegbuzie et al. (2007).

Otro elemento que limitó el estudio, tiene que ver con la cantidad de respuestas obtenidas de parte de los directivos de la Universidad al ECIED. A pesar de que la encuesta fue enviada en al menos tres oportunidades, la cantidad de respuestas fue baja. Considerando el perfil de quienes debían responder, para futuras aplicaciones del instrumento, se sugiere tomar la encuesta de manera presencial.

Finalmente, existió una imposibilidad material en el estudio, asociada a la disponibilidad de otros Software de análisis de datos que permitieran respetar la naturaleza policórica (asociadas a las escalas Likert) de las matrices de correlación, en los Análisis Factoriales de IED y ECIED. Para respetar esta condición sólo se contaba con el software Factor 8.02. El software fue utilizado en el análisis del ECIED, pero no en el de IED porque el tamaño de la muestra no permitió el funcionamiento del programa.

Las limitaciones enunciadas pretenden situarse como antecedentes para futuras aplicaciones de los procedimientos descritos en el estudio, con el fin de mejorar nuevos análisis en la materia en otros contextos de estudio.

XII. Referencias

1. Abrami, P. & d'Apollonia, S. (1991). Multidimensional student's evaluation of teaching effectiveness- generalizability of N=1 research: comments of Marsh (1991). *Journal of Educational Psychology*, 83, 411- 415.
2. American Educational Research Association, American Psychological Association & National Council on Measurement in Education. (1954). *Standards for Educational and Psychological Testing*. Washington: DC: American Psychological Association.
3. American Educational Research Association, American Psychological Association & National Council on Measurement in Education. (1999). *Standards for Educational and Psychological Testing*. Washington: DC: American Psychological Association
4. Álvarez, V. (2009). Diseño de Guías Forcom de Autoevaluación de Competencias Docentes para el EEES. Revista de Docencia Universitaria. En www.redu.um.es/Red_U/4
5. Ambady, N. & Rosenthal, R. (1993). Half a minute: predicting teacher evaluations from thin slices of nonverbal behavior and physical attractiveness. *Journal of Personality and Social Psychology*, 431-441.
6. Arubayi, E. A. (1987) Improvement of instruction and teacher effectiveness: are student ratings reliable and valid? *Higher Education*, 16, 267–278.
7. España, Agencia Nacional de Evaluación de la Calidad y Acreditación (2012). *Informe sobre el estado de la evaluación externa de la calidad de las universidades españolas*. Madrid, España.
8. México, Asociación Nacional de Universidades e Instituciones de Educación Superior (2000). *Evaluación del desempeño del personal académico*. México, D.F.
9. Barron, B. (2006). Interest and Self-Sustained Learning as Catalysts of Development: A Learning Ecology Perspective. *Human Development*, 49:193–224.
10. Bain, K. (2007). *Lo que hacen los mejores profesores universitarios*. Valencia: Editorial Universitat de València.
11. Balam, E., & Shannon, D. (2010). Student ratings of college teaching: A comparison of faculty and their students. *Assessment and Evaluation in Higher Education*, 209-221.
12. Barth, M. (2008). Deciphering student evaluations of teaching: A factor analysis approach. *Journal of Education for Business*, 40-46.
13. Braun, V. & Clarke V. (2008). Using thematic analysis in psychology. *Qualitative Research in Psychology*. 3:2, 77-101
14. Beran, T., Violato, C., Kline, D., & Frideres, J. (2005). The utility of student ratings of instruction for students, faculty and administrators: a "consequential validity" study. *The Canadian Journal of Higher Education*, 2: 49-70.

15. Berk, R. (2006). Thirteen strategies to measure college teaching. Sterling, Virginia: Stylus.
16. Biggs. J. (2001). The revised two-factor Study Process Questionnaire: R-SPQ-2F. British Journal of Educational Psychology, 71, 133–149.
17. Biggs. J. (2008). *Calidad del aprendizaje universitario*. Madrid, España: NARCEA.
18. Boyer, E. (1990). Scholarship reconsidered: Priorities of the professoriate. Princeton, NJ: The Carnegie. Foundation for the Advancement of Teaching.
19. Bosshardt, W. & Watts, M. (2001). Comparing student and instructor evaluations of teaching. Journal of Economic Education, 3-17.
20. Cid-Sabucedo, A., Pérez-Abellás, A. & Zabalza, M. (2009). Las prácticas de enseñanza declaradas de los “mejores profesores”. RELIEVE, 15(2), 1-29.
21. Chile, Centro Interuniversitario de Desarrollo (2007). Evaluación del desempeño docente y la calidad de la docencia universitaria. Chile, Santiago.
22. Chile, Universidad de Santiago de Chile (2013). Modelo Educativo Institucional. Santiago, Chile: Vicerrectoría Académica.
23. Clayson, D. (2009). Student evaluations of teaching: Are they related to what students learn? A meta-analysis and review of the literature. Journal of Marketing Education, 31, 16–30.
24. Cohen, P. (1981). Student ratings of instruction and student achievement: A metaanalysis of multisection validity studies. Review of Educational Research, 281-309.
25. Coll, C. & Solé, I. (2001). “Enseñar y aprender en el contexto de aula”, en C. Coll, J. Palacios y A. Marchesi (Eds.), Desarrollo Psicológico y Educación 2. Psicología de la Educación. España, Madrid: Alianza Editorial.
26. Coll, C & Sánchez, E. (2008) El análisis de la interacción alumno-profesor: líneas de investigación. Revista de Educación, 346, 15-32.
27. Costin, F., Greenough, T., & Menges, R. (1971). Student Ratings of College Teaching: Reliability, Validity, and Usefulness. Review of Educational Research, 41(5), 511-535.
28. Cronbach, Meehl. (1955). Construct Validity in Psychological Test. Psychological Bulletin.
29. Díaz Barriga, F. & Rigo, M. (2003). Realidades y paradigmas de la función docente. Revista de Educación Superior, 127, 53-61.
30. Espinoza, O. & González, L. (2009). Evaluación del desempeño en la educación superior: un modelo de análisis. En N, Fleet, (Eds.), Desafíos y perspectivas de la dirección estratégica de las universidades (pp. 193-316), Chile, Santiago: Comisión Nacional de Acreditación.
31. Freiberg, A., Stover, J., de la Iglesia, G. y Fernández, M. (2013). Correlaciones Policóricas y Tetracóricas en Estudios Factoriales Exploratorios y Confirmatorios. Ciencias Psicológicas, 7(2), 151-164.

32. García, G. (2000). Las dimensiones de la efectividad docente, validez y confiabilidad de los cuestionarios de evaluación de la docencia: síntesis de investigación internacional. En M. Rueda, & D.-B. F (Eds.), *Evaluación de la docencia. Perspectivas Actuales*. México, DF: PAIDOS.
33. García, J. (2014). ¿Para qué sirve la evaluación de la docencia? Un estudio exploratorio sobre las creencias de los estudiantes. *Archivos Analíticos de Política Educativa*, 22(15), 2-24.
34. Ghedin, E. & Aquario, D. (2008). Moving towards multidimensional evaluation of teaching in higher education: A study across four faculties. *Higher Education*, (56), 583–597.
35. Giné, N. (2009). Cómo mejorar la docencia universitaria: el punto de vista del estudiantado. *Revista Complutense de Educación*, 20(1).
36. González, C., Montenegro, H., López, L., Munita, I., Collao, P. (2011). Calidad de las experiencias de aprendizaje y enseñanza en la educación universitaria de pregrado: una mirada desde los profesores y estudiantes. (Proyecto Consejo Nacional de Educación) Santiago, Chile: Pontificia Universidad Católica de Chile, Facultad de Educación.
37. Griffin, T., Hilton, J., P, P. & Barret, D. (2014). Correlation between point averages and student evaluation of teaching scores: taking a closer look. *Assessment & Evaluation in Higher Education*, 39, 339-348.
38. Hajdin, K., & Pažur, G. (2012). Differentiating Between Student Evaluation of Teacher and Teaching Effectiveness. *Journal of Information and Organizational Sciences*, 36(2), 123-134.
39. Hannan, A. & Silver, H. (2006). La innovación en la enseñanza superior: enseñanza, aprendizaje y culturas institucionales. España, Madrid: NARCEA.
40. Hogan, T. (2004). Pruebas psicológicas: una introducción práctica. Colombia, Bogotá: Manual Moderno.
41. Huybers, T. (2013). Student evaluation of teaching: the use of best–worst scaling. *Assessment & Evaluation in Higher Education*, 39:4, 496-513.
42. Jackson, D., Teal, C., Raines, S., Nansel, T., Force, R., & Burdsal, C. (1990). The dimensions of student's perceptions of teaching effectiveness. *Educational and Psychological Measurement* (59), 580-596.
43. Kane, M. (2014). Validating the Interpretations and Uses of Test Scores. *Journal of Educational Measurement*, 50(1), 1-73.
44. Kane, M. (2001). Current concerns in validity theory. *Journal of Educational Measurement*, 38, 319-342.
45. Kember, D. & Wong, A. (2000). Implications for evaluation from a study of students' perceptions of good and poor teaching. *Higher Education*, 40, 67-97.

46. Kember, D., Biggs, J., Leung, D., (2004) Examining the multidimensionality of approaches to learning through the development of a revised version of the Learning Process Questionnaire. *British Journal of Educational Psychology*, 74, 261–280
47. Lane, S. & Stone, C. (2002). Strategies for examining the consequences of assessment and accountability programs. *Educational Measurement: Issues and Practice*, 21, 23-30.
48. Leung, D., & Kember, D. (2005). Comparability of data gathered from evaluation questionnaires on paper and through the internet. *Research in Higher Education*, 46, 571-591.
49. López, L. (2013). La retención y la experiencia de aprendizaje de los estudiantes en el primer año de ingeniería (Tesis de doctorado no publicada). Pontificia Universidad Católica de Chile, Santiago.
50. Marsh, H. W. (1987) Students' evaluations of university teaching: research findings, methodological issues, and directions for future research, *International Journal of Educational Research*, 11, 253–388.
51. Marsh, H. W. & Bailey, M. (1993) Multidimensional students' evaluations of teaching effectiveness, *Journal of Higher Education*, 64, 1–18
52. Marsh H. & Roche, L. (1994) Making Students' Evaluations of Teaching Effectiveness Effective. *American Psychologist*. 52 (11), 1187-119.
53. Messick, S. (1995). Standards of Validity and the Validity of Standards in Performance Assessment. *Educational Testing Service*, 5-8.
54. Mondragón, H. (2005). Los profesores universitarios en escena: un estudio sobre la cultura profesional académica de los buenos profesores. Colombia, Cali: Sello Editorial Javeriano.
55. Moore, S. Kuol, N. (2005) a punitive bureaucratic tool or a valuable resource? using student evaluations to enhance your teaching. En: *Emerging Issues in the Practice of University Learning and Teaching*. O'Neill, G., Moore, S., McMullin, B. (Eds).Dublin: AISHE, 2005
56. Nasser, F. & Fresko, B. (2001). Interpreting student ratings: Consultation, instructional modification, and attitudes towards course evaluation. *Studies in Educational Evaluation*, 27, 291–305.
57. Onwuegbuzie, A., Daniel, L., & Collins, K. (2009). A meta-validation model for assessing the score-validity of student teaching evaluations. *Quality & Quantity*, 43, 197–209.

58. Onwuegbuzie, A., Witcher, A., Collins, K. Filer, J., Moore, C. (2007). Student's 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, 113-160.
59. Ory, J. (2001). Faculty thoughts and concerns about student ratings. *New Directions for Teaching and Learning*, 87, 3-15.
60. Ory, J., & Ryan, K. (2001). How do student ratings measure up to a new validity framework? *New Directions for Institutional Research*, 109, 27-44.
61. Pan, D., Tan, G., Ragupathi, K., Booluck, K., Roop, R., & Ip, Y. (2009). Profiling teacher/teaching using descriptors derived from qualitative feedback: Formative and summative applications. *Research in Higher Education*, 50, 73-100.
62. Padilla, J., Gómez, J., Hidalgo, D & Muñoz, J. (2007) Esquema conceptual y procedimientos para analizar la validez de las consecuencias del uso de los test. *Psicothema*, 19 (1), 173-178.
63. Penny, A. (2003). Changing the agenda for research into students' views about university teaching: Four shortcomings of SRT research. *Teaching in Higher Education*, 8, 399-411.
64. Prosser, M. & Trigwell, K. (2006) Confirmatory factor analysis of the Approaches to Teaching Inventory. *British Journal of Educational Psychology*, 76, 405-419
65. Pozo, C., Reboloso, E., & Fernandez, B. (2000). The "Ideal Teacher". Implications for student evaluations of teaching effectiveness. *Assessment & Evaluation in Higher Education*, 25, 253-263.
66. Ramsden P. (2003). *Learning to teach in higher education*. London: Routledge Falmer.
67. Renaud, R. & Murray, H. (2005). Factorial validity of student ratings of instruction. *Research in Higher Education* (46), 929-953.
68. Richardson, J. (2004). Methodological Issues in Questionnaire-Based Research on Student Learning in Higher Education. *Educational Psychology Review*, 16, 4.
69. Richardson, J. (2005). Instruments for obtaining student feedback: a review of the literature. *Assessment & Evaluation in Higher Education*, 30(4), 387-415.
70. Roche, L., & Marsh, H. (2000). Multiple dimensions of university teacher selfconcept. *Instructional Science*, 28, 439-468.
71. Rueda, M. & Díaz Barriga, F. (2010). *Evaluación de la Docencia, Perspectivas Actuales*. México, DF: Editorial Paidós.

72. Rueda, M., Elizalde, L., Torquemada, A., (2006) La evaluación de la docencia en las universidades mexicanas. *Revista de Educación Superior*, 127, 71-77
73. Salazar, A. (2008). Diagnóstico preliminar sobre evaluación de la docencia universitaria. Una aproximación a la realidad en las universidades públicas y/o estatales en Chile. *Revista Iberoamericana de Evaluación Educativa*, 1(3), 68-84.
74. Salazar, A. (2010). Encuesta de Satisfacción Estudiantil Versus Cultura Evaluativa de la Docencia. *Revista Iberoamericana de Evaluación Educativa*, 3(1), 121-132.
75. Seldin, P. (1993). The use and abuse of student ratings of professors. *Chronicle of Higher Education*, 39 (46). A-40.
76. Shevlin, M., Banyard, P., Davies, M., & Griffiths, M. (2000). The validity of student evaluation in higher education: Love me, love my lectures? *Assessment & Evaluation in Higher Education*, 25, 397-405.
77. Silvero. (2006). Motivación y calidad docente en la universidad. España, Navarra: Universidad de Navarra.
78. Shulman, L., (1989). Paradigmas y programas de investigación en el estudio de la
79. enseñanza: una perspectiva contemporánea” En: M. Wittrock (Eds). *La investigación en la enseñanza*. España, Barcelona: Paidós.
80. Spooren, P., Brockx, B., & Mortelmans, D. (2013). On the Validity of Student Evaluation of Teaching: The State of the Art. *Review of Educational Research*.83, 598-642.
81. Tucker, B. (2014). Student evaluation surveys: anonymous comments that offend or are unprofessional. *Higher Education*, 68, 347-358.
82. Webster, J. Chan W. Prosser, M. & Watkins, D. (2009). Undergraduates’ learning experience and learning process: quantitative evidence from the East. *Higher Education*, 58, 375–386.
83. Yair, G. (2008). Can we administer the scholarship of teaching? Lessons from outstanding professors in higher education. *Higher Education*, 55,447–459.
84. Zabalza, M. (2009). *Competencias Docentes del Profesorado Universitario*. España, Madrid: NARCEA.
85. Zuñiga, M., Jopia, B. (2007). La evaluación del desempeño docente en las universidades chilenas: diagnóstico desde la perspectiva de las autoridades. Centro Interuniversitario de Desarrollo (Eds). *Evaluación del desempeño docente y la calidad de la docencia universitaria*. Chile, Santiago.

XII. Anexos

Anexo 1.

Evaluación de la Docencia Universitaria aplicada a estudiantes

Dimensiones		Indicadores	5	4	3	2	1
D1: PLANIFICACIÓN DE ACTIVIDADES DOCENTES	1.	Da a conocer a los/las estudiantes la planificación de sus actividades docentes (clases, talleres, evaluaciones, salidas a terreno, etc.) al inicio del curso.					
	2.	Organiza las actividades docentes en forma coherente con los objetivos establecidos.					
	3.	Cumple con el programa según lo planificado, justificando aquellos contenidos que no fueron cubiertos (en caso de que así ocurra).					
D2: EJECUCIÓN DE ACTIVIDADES DOCENTES	4.	Entrega documentos (apuntes, guías, artículos, otros) que complementan el desarrollo de las actividades educativas.					
	5.	Utiliza recursos tecnológicos (presentaciones, internet, etc.) y/o materiales (equipamiento, laboratorios, etc.) que facilitan la comprensión de los contenidos.					
	6.	Utiliza los recursos bibliográficos definidos en el programa, como referentes para el desarrollo de las clases.					
	7.	Sugiere recursos bibliográficos complementarios para el desarrollo de los contenidos.					
	8.	Fomenta la participación de los/las estudiantes en clases (mediante preguntas, debates, ejemplos u otros).					
	9.	Comunica los contenidos de forma clara.					
	10.	Contextualiza los contenidos del curso al desempeño profesional futuro de los estudiantes.					
	11.	Utiliza distintas estrategias de enseñanza para facilitar el logro de los aprendizajes.					
	12.	Utiliza el horario de clase eficientemente, optimizando el tiempo disponible.					
	13.	Realiza actividades que le permiten conocer los aprendizajes previos de sus estudiantes al inicio de la asignatura (ya sea de manera escrita, oral, presencial u online).					
D3: EVALUACIÓN DE APRENDIZAJES	14.	Explica a los/las estudiantes con anterioridad, los criterios de evaluación definidos para cada instancia (pruebas, trabajos, disertaciones, etc.).					
	15.	Entrega el resultado de las evaluaciones dentro de los 15 días establecidos por reglamento.					
	16.	Aplica pautas de corrección claras para la revisión de las evaluaciones.					
	17.	Retroalimenta las evaluaciones de los/las alumnos(as), permitiéndoles reconocer sus posibles errores.					
	18.	Utiliza procedimientos evaluativos (pruebas, trabajos, disertaciones, etc.) coherentes con los objetivos del curso.					
	19.	Favorece un clima de respeto en la relación con sus estudiantes.					
D4:	20.	Establece una relación cordial con sus estudiantes.					

RELACIÓN CON LOS ESTUDIANTES	21. Es accesible para atender las consultas de los/las estudiantes (vía correo electrónico, horario de oficina, entre otros).
------------------------------------	---

5: *Muy de acuerdo*; 4: *De acuerdo*; 3: *Parcialmente de acuerdo*; 2: *En desacuerdo*; 1: *Muy en desacuerdo*

Anexo 2

Course Experience Questionnaire

	5	4	3	2	1
(1) La cantidad de trabajo asociado a este curso impide que toda la materia se pueda entender en profundidad.					
(2) El profesor de este curso me motiva a hacer mi mejor trabajo.					
(3) Es fácil conocer las exigencias del trabajo esperado en este curso.					
(4) Sólo se necesita buena memoria para que a uno le vaya bien en este curso.					
(5) El profesor se esfuerza para que la materia sea interesante.					
(6) El profesor dejó en claro desde el inicio del curso lo que se esperaba de los estudiantes.					
(7) El profesor parece más interesado en evaluar lo que he memorizado que lo que he comprendido en el curso.					
(8) El profesor es muy bueno explicando la materia.					
(9) Usualmente he tenido una idea clara de qué se espera de mí en este curso.					
(10) Siento mucha presión como estudiante de este curso.					
(11) El profesor me da retroalimentación sobre mi progreso en el curso.					
(12) Generalmente ha sido difícil descubrir qué se espera de mí en este curso.					
(13) La carga de trabajo en el curso es muy pesada.					
(14) El profesor dedica mucho tiempo para entregarme comentarios sobre mi desempeño en el curso					
(15) En las evaluaciones el profesor sólo hace preguntas sobre información textual de la materia.					
(16) En este curso se otorga suficiente tiempo para comprender los contenidos propuestos.					
(17) El profesor se esfuerza por entender las dificultades que pudiera estar teniendo con mi desempeño en el curso					

5: *Muy de acuerdo*; 4: *De acuerdo*; 3: *Ni de acuerdo ni desacuerdo*; 2: *En desacuerdo*; 1: *Muy en desacuerdo*

Anexo 3

Fiabilidad del instrumento de evaluación de la docencia (Método Alfa de Cronbach)

	N	%
Válidos	48129	100,0
Casos Excluidos ^a	0	,0
Total	48129	100,0

a. Eliminación por lista basada en todas las variables del procedimiento.

Estadísticos de fiabilidad	
Alfa de Cronbach	N de elementos
,978	21

	Estadísticos total-elemento			
	Media de la escala si se elimina el elemento	Varianza de la escala si se elimina el elemento	Correlación elemento-total corregida	Alfa de Cronbach si se elimina el elemento
P1	81,99	375,135	,827	,977
P2	82,03	372,953	,869	,977
P3	82,03	372,951	,852	,977
P4	82,16	373,674	,790	,978
P5	82,19	374,308	,743	,978
P6	82,18	372,828	,812	,977
P7	82,26	372,651	,781	,978
P8	82,08	372,761	,842	,977
P9	82,15	370,216	,859	,977

P10	82,13	371,775	,833	,977
P11	82,28	369,068	,853	,977
P12	82,09	371,870	,834	,977
P13	82,33	371,734	,791	,978
P14	82,15	371,251	,864	,977
P15	82,14	374,682	,741	,978
P16	82,26	370,597	,822	,977
P17	82,26	370,379	,823	,977
P18	82,03	373,201	,862	,977
P19	81,88	376,998	,806	,978
P20	81,91	376,606	,800	,978
P21	82,01	374,094	,801	,978

Anexo 4

Fiabilidad del instrumento de evaluación de la docencia (Método Bipartición)

Estadísticos de fiabilidad			
Alfa de Cronbach	Parte 1	Valor	,963
		N de elementos	11 ^a
	Parte 2	Valor	,957
		N de elementos	10 ^b
	N total de elementos		21
Correlación entre formas			,924
Coeficiente de Spearman-Brown	Longitud igual		,960
	Longitud desigual		,960
Dos mitades de Guttman			,957

a. Los elementos son: P1, P2, P3, P4, P5, P6, P7, P8, P9, P10, P11.

b. Los elementos son: P11, P12, P13, P14, P15, P16, P17, P18, P19, P20, P21.

Anexo 5

Análisis Factorial Exploratorio Evaluación de la Docencia

KMO y prueba de Bartlett		
Medida de adecuación muestral de Kaiser-Meyer-Olkin.		,979
	Chi-cuadrado	1185763,868
Prueba de esfericidad de	aproximado	
Bartlett	Gl	210
	Sig.	,000

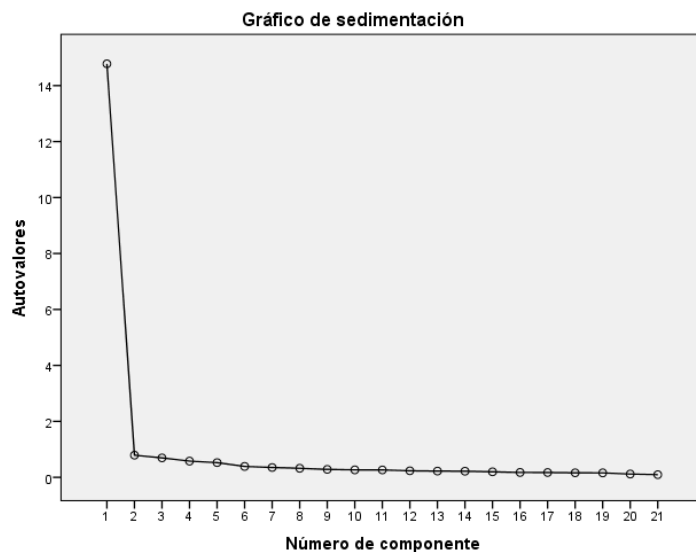
	Comunalidades	
	Inicial	Extracción
P1	1,000	,718
P2	1,000	,785
P3	1,000	,757
P4	1,000	,657
P5	1,000	,586
P6	1,000	,689
P7	1,000	,641
P8	1,000	,738
P9	1,000	,766
P10	1,000	,724
P11	1,000	,754
P12	1,000	,728
P13	1,000	,657
P14	1,000	,773
P15	1,000	,586
P16	1,000	,705
P17	1,000	,706
P18	1,000	,771
P19	1,000	,686

P20	1,000	,677
P21	1,000	,676

Método de extracción: Análisis de Componentes principales.

Componente	Varianza total explicada					
	Autovalores iniciales			Sumas de las saturaciones al cuadrado de la extracción		
	Total	% de la varianza	% acumulado	Total	% de la varianza	% acumulado
1	14,779	70,378	70,378	14,779	70,378	70,378
2	,791	3,765	74,143			
3	,695	3,310	77,453			
4	,580	2,762	80,215			
5	,526	2,502	82,717			
6	,390	1,858	84,576			
7	,352	1,676	86,252			
8	,324	1,542	87,794			
9	,283	1,348	89,142			
10	,266	1,268	90,410			
11	,263	1,250	91,660			
12	,235	1,117	92,777			
13	,223	1,060	93,838			
14	,219	1,043	94,880			
15	,196	,935	95,816			
16	,173	,826	96,641			
17	,172	,817	97,458			
18	,162	,774	98,232			
19	,158	,753	98,985			
20	,118	,561	99,545			
21	,095	,455	100,000			

Método de extracción: Análisis de Componentes principales.



Matriz de estructura

	Factor			
	1	2	3	4
P1	,758	-,714	-,895	,699
P2	,799	-,753	-,940	,731
P3	,775	-,732	-,923	,732
P4	,808	-,647	-,732	,677
P5	,780	-,618	-,675	,615
P6	,874	-,659	-,732	,667
P7	,850	-,629	-,679	,659
P8	,837	-,768	-,746	,721
P9	,831	-,757	-,788	,763
P10	,836	-,729	-,743	,724
P11	,858	-,726	-,742	,779
P12	,777	-,734	-,803	,751
P13	,783	-,645	-,695	,759
P14	,793	-,734	-,823	,832
P15	,655	-,627	-,705	,765
P16	,744	-,673	-,735	,903
P17	,763	-,690	-,714	,876
P18	,786	-,770	-,815	,807
P19	,713	-,939	-,742	,658
P20	,712	-,956	-,723	,650
P21	,730	-,813	-,721	,709

Método de extracción: Mínimos cuadrados no ponderados.

Método de rotación: Normalización Oblimin con Kaiser.

Anexo 6

Correlación entre evaluación de la docencia y rendimiento académico de los estudiantes.

Correlaciones			
		Evaluación de la docencia	Rendimiento académico (calificaciones)
Evaluación de la docencia	Correlación de Pearson	1	,214**
	Sig. (bilateral)		,000
	N	2237	2237

** La correlación es significativa al nivel 0,01 (bilateral).

Anexo 7

*Correlación entre evaluación de la docencia y rendimiento académico de los estudiantes: por
área de conocimiento.*

Humanidades			
		Evaluación de la docencia	Rendimiento académico (calificaciones)
Evaluación de la docencia	Correlación de Pearson	1	,320**
	Sig. (bilateral)		,000
	N	549	549

** La correlación es significativa al nivel 0,01 (bilateral).

Ciencias Naturales

Correlaciones		Evaluación de la docencia	Rendimiento académico (calificaciones)
Evaluación de la docencia	Correlación de	1	,124
	Pearson		
	Sig. (bilateral)		,174
	N	122	122

Ciencias Matemáticas

Correlaciones		Evaluación de la docencia	Rendimiento académico (calificaciones)
Evaluación de la docencia	Correlación de	1	,196**
	Pearson		
	Sig. (bilateral)		,002
	N	247	247

** . La correlación es significativa al nivel 0,01 (bilateral).

Ciencias médicas

Correlaciones		Evaluación de la docencia	Rendimiento académico (calificaciones)
Evaluación de la docencia	Correlación de	1	,432**
	Pearson		
	Sig. (bilateral)		,002
	N	48	48

** . La correlación es significativa al nivel 0,01 (bilateral).

Diseño			
Correlaciones			
		Evaluación de la docencia	Rendimiento académico (calificaciones)
Evaluación de la docencia	Correlación de	1	,211
	Pearson		
	Sig. (bilateral)		,247
	N	32	32
Bachillerato			
Correlaciones			
		Evaluación de la docencia	Rendimiento académico (calificaciones)
Evaluación de la docencia	Correlación de	1	-,252
	Pearson		
	Sig. (bilateral)		,748
	N	4	4
Tecnología			
Correlaciones			
		Evaluación de la docencia	Rendimiento académico (calificaciones)
Evaluación de la docencia	Correlación de	1	,095
	Pearson		
	Sig. (bilateral)		,069
	N	368	368

Ingeniería			
Correlaciones			
		Evaluación de la docencia	Rendimiento académico (calificaciones)
Evaluación de la docencia	Correlación de	1	,101**
	Pearson		
	Sig. (bilateral)		,008
	N	688	688

** . La correlación es significativa al nivel 0,01 (bilateral).

Administración y economía			
Correlaciones			
		Evaluación de la docencia	Rendimiento académico (calificaciones)
Evaluación de la docencia	Correlación de	1	,246*
	Pearson		
	Sig. (bilateral)		,016
	N	96	96

*. La correlación es significativa al nivel 0,05 (bilateral).

Anexo 8

Correlación entre evaluación de la docencia y experiencia de aprendizaje de los estudiantes

		Correlaciones			
		Metas y		Carga de	
		Buena	Objetivos	Evaluación	trabajo
		docencia	claros	apropiada	apropiada
Evaluación de la docencia	Correlación de Pearson	,501**	,451**	,234**	,314**
	Sig. (bilateral)	,000	,000	,000	,000
	N	311	311	311	311

** . La correlación es significativa al nivel 0,01 (bilateral).

Anexo 9

Correlación entre evaluación de la docencia y experiencia de aprendizaje de los estudiantes: por área de conocimiento

		Humanidades			
		Correlaciones		Carga de	
		Metas y		trabajo	
		Buena	Objetivos	Evaluación	
		docencia	claros	apropiada	apropiada
Evaluación de la docencia	Correlación de Pearson	,533**	,057	,483*	-,001
	Sig. (bilateral)	,005	,783	,012	,998
	N	26	26	26	26

** . La correlación es significativa al nivel 0,01 (bilateral).

* . La correlación es significativa al nivel 0,05 (bilateral).

Ciencias Naturales

		Correlaciones Metas y		Carga de	
		Buena docencia	Objetivos claros	Evaluación apropiada	trabajo apropiada
Evaluación de la docencia	Correlación de Pearson	,435*	,268	,708**	-,480*
	Sig. (bilateral)	,049	,241	,000	,028
	N	21	21	21	21

** . La correlación es significativa al nivel 0,01 (bilateral).

* . La correlación es significativa al nivel 0,05 (bilateral).

Ciencias Matemáticas

		Correlaciones Metas y		Carga de	
		Buena docencia	Objetivos claros	Evaluación apropiada	trabajo apropiada
Evaluación de la docencia	Correlación de Pearson	,498**	,493**	,222	,331*
	Sig. (bilateral)	,000	,000	,109	,016
	N	53	53	53	53

** . La correlación es significativa al nivel 0,01 (bilateral).

* . La correlación es significativa al nivel 0,05 (bilateral).

Ciencias Médicas

		Correlaciones Metas y		Carga de	
		Buena docencia	Objetivos claros	Evaluación apropiada	trabajo apropiada
Evaluación de la docencia	Correlación de Pearson	,182	,452**	,220	,029
	Sig. (bilateral)	,318	,009	,227	,876
	N	32	32	32	32

* . La correlación es significativa al nivel 0,05 (bilateral).

** . La correlación es significativa al nivel 0,01 (bilateral).

Tecnología

		Correlaciones			
		Metas y		Carga de	
		Buena	Objetivos	Evaluación	trabajo
		docencia	claros	apropiada	apropiada
Puntuación Z:	Correlación de Pearson	,467**	,415**	,178*	,280**
Evaluación docente	Sig. (bilateral)	,000	,000	,017	,000
	N	178	178	178	178

** . La correlación es significativa al nivel 0,01 (bilateral).

* . La correlación es significante al nivel 0,05 (bilateral).

Anexo 10

Listado de códigos percepción de la calidad de la docencia

Categoría	Subcategoría	Código
Centrado en el estudiante Onwuegbuzie et al. (2007)	Preocupación por necesidades del estudiante (Ramsden, 2003; Onwuegbuzie et al.2007)	CEPRE
	Aprendizaje activo (fomenta participación) (Biggs, 2008; Kember and Wong, 2000)	CEACT
	Aprendizaje profundo	APRO
Experticia disciplinar Onwuegbuzie et al. (2007)	Actualización disciplinar (Bain, 2007)	EXAC
	Manejo del contenido o disciplina	EXMA
Profesional docente Onwuegbuzie et al. (2007)	Preparación y organización previa de la enseñanza (Zabalza, 2009; Bain, 2007; Crumbley, Henry, Kratchman, 2001)	PROPLA
	Cumple con lo planificado	PROCUM
	Metas y objetivos claros (Ramsden, 2003)	PROME
	Cumple con los objetivos o resultados de aprendizaje	PROCUO
	Vínculos con el desarrollo profesional (Bain, 2007; Biggs, 2008; Zabalza, 2009)	PROVIN
	Métodos y estrategias apropiadas (Zabalza, 2009; Bain, 2007)	PROMET
	Uso efectivo del tiempo	PROTIME
	Realiza preguntas claves (Bain, 2007)	PROPRE
	Realiza ejemplos claros (Bain, 2007)	PROEJE

	Responde claramente a las dudas en clases	PRODUD
	Focaliza su atención en los conceptos clave (Ramsden, 2003).	PROCONC
	Fomenta la colaboración como método de enseñanza (Ramsden, 2003)	PROCOL
	Métodos de evaluación válidos (Ramsden, 2003)	PROEVA
	Criterios de evaluación claros (Zabalza, 2009; Bain, 2007)	PROCRI
	Resultados de evaluación a tiempo (Zabalza, 2009)	PRORESEV
	Uso de TIC la docencia (todo lo tecnológico) (Zabalza, 2009)	PROTIC
	Material de apoyo apropiado (Zabalza, 2009; Hajdin y Pažur, 2012).	PROMAT
	Orienta el estudio fuera de la clase (Bain, 2007).	PROOUT
Retroalimentación efectiva Onwuegbuzie et al. (2007)	Retroalimentación efectiva (Ramsden, 2003; Bain, 2007; Zabalza, 2009; Onwuegbuzie et al., 2007;).	RETROEF
Accesibilidad fuera de la clase Onwuegbuzie et al. (2007)	Disponibilidad para tutorizar (Zabalza, 2009)	ACCETUT
	Disponibilidad a responder dudas fuera de la clase (Onwuegbuzie et al. (2007).	ACCEDUD
Entusiasmo Onwuegbuzie et al. (2007)	Entusiasmo por lo que enseña (Onwuegbuzie et al. 2007)	ENTENS
	Motiva a los estudiantes (Ramsden, 2003)	ENTMOT
Claridad comunicativa Onwuegbuzie et al. (2007)	Claridad en las explicaciones (Ramsden, 2003; Onwuegbuzie et al., 2007; Greimel-Fuhrmann y Geyer's, 2003)	CLAEX
	Capacidad para presentar (Crumbly, Henry, Kratchman, 2001)	CLAPRES
	Uso adecuado de la voz (tono y modulación)	CLAVOZ
Comportamiento ético Onwuegbuzie et al. (2007)	Imparcialidad (Zabalza, 2009).	COIMP
	Toma de decisiones fundamentadas (Bain, 2005).	COTOM
	Responsabilidad con el proceso de aprendizaje (Zabalza, 2009) (puntualidad, asistencia)	CORESP
Ambiente de aprendizaje efectivo Onwuegbuzie et al. (2007)	Ambiente de aprendizaje efectivo Clima de confianza en el aula (Bain, 2007)	AMB
	Buena relación profesor alumno	AMBCONF
		AMBRELA

Apariencia	Atractivo físico	APAFÍS
Ambady and Rosenthal (1993)	(Ambady and Rosenthal, 1993)	
	Comportamiento no verbal	APANOV
	(Ambady and Rosenthal, 1993)	
	Características personales (carisma, simpatía, amabilidad, encantador) (Shevlin et al., 2000)	APACAR
Altas expectativas en sus estudiantes	Altas expectativas en que sus estudiantes aprendan. (Bain, 2007; Silvero, 2006).	ALTAPRE
(Bain, 2007)		
Vínculos con la institución	Vínculos con el Modelo Educativo Institucional (Parcerisa, 2010).	ALTAMEI
(Parcerisa, 2010).		
Comentarios ofensivos	Lenguaje ofensivo (malas palabras) (Tucker, 2014)	COOFEN
(Tucker, 2014)		
	Discriminación (género, raza o personales) (Tucker, 2014)	CODISC
	Denuncias de mala conducta o conducta criminal (Tucker, 2014)	CODEN
Condiciones de contexto	Dificultades del espacio	CONTEXT
	Problemas de diseño del curso (malla, carga académica, etc.)	COPREV
Excelencia	Buen docente	BUEN

Anexo 11

Relación evaluación de la docencia y percepción de la calidad de la docencia (suma total de áreas de conocimiento)

“Elementos que debería mantener”

Códigos	Descripción	Promedios IED					
		4,1 a 5	%	3,1 a 4	%	0 a 3	%
BUEN	Buen profesor	425	12,7	81	12,2	3	5,3
PROMET	Métodos y estrategias apropiadas	296	8,8	47	7,1	4	7,0
CLAEX	Claridad para explicar	276	8,2	38	5,7	4	7,0
CEACT	Aprendizaje activo	267	8,0	37	5,6	3	5,3
CEPRE	Preocupado por necesidades de estudiantes	243	7,3	27	4,1	3	5,3
APACAR	Características personales (amable, simpático)	203	6,1	33	5,0	5	8,8
PRODUD	Responde claramente a las dudas	167	5,0	35	5,3	2	3,5
PROVIN	Vínculos con el mundo profesional	149	4,5	48	7,2	2	3,5
CLAPRES	Claridad para presentar	134	4,0	15	2,3	1	1,8
AMBRELA	Buena relación con estudiantes	118	3,5	23	3,5	0	0,0
EXMA	Manejo del contenido	113	3,4	41	6,2	4	7,0
AMB	Ambiente de aprendizaje	91	2,7	13	2,0	2	3,5
PROMAT	Material de apoyo apropiado	79	2,4	22	3,3	1	1,8
CORESP	Responsabilidad en proceso de aprendizaje	66	2,0	15	2,3	2	3,5

ACCEDUD	Disponibilidad para responder dudas	63	1,9	20	3,0	4	7,0
AMBCONF	Ambiente de confianza	61	1,8	9	1,4	0	0,0
RETROEF	Retroalimentación efectiva	56	1,7	11	1,7	0	0,0
ENTMOT	Motiva a los estudiantes	55	1,6	2	0,3	0	0,0
PROEJE	Plantea ejemplos claros	52	1,6	5	0,8	0	0,0
PROEVA	Métodos de evaluación válidos	45	1,3	6	0,9	0	0,0
ENTENS	Entusiasmo por lo que enseña	43	1,3	2	0,3	2	3,5
PROTIC	Uso apropiado de la tecnología	42	1,3	17	2,6	2	3,5
PROCUM	Cumple con lo planificado	41	1,2	12	1,8	4	7,0
PROTIME	Uso efectivo del tiempo	35	1,0	11	1,7	3	5,3
PROPLA	Preparación previa de la enseñanza	30	0,9	4	0,6	0	0,0
PROCOL	Fomenta trabajo colaborativo	29	0,9	2	0,3	5	8,8
PROCUO	Cumple con los objetivos del curso	26	0,8	5	0,8	0	0,0
PROOUT	Orienta el estudio fuera de la clase	26	0,8	10	1,5	0	0,0
PROCONC	Focaliza atención en conceptos clave	24	0,7	14	2,1	0	0,0
PRORESEV	Entrega resultados de evaluación oportunamente	20	0,6	7	1,1	0	0,0
ALTAPRE	Altas expectativas	16	0,5	1	0,2	0	0,0
PROCRI	Criterios de evaluación claros	15	0,4	7	1,1	1	1,8
ACCETUT	Disponibilidad para tutorizar	9	0,3	4	0,6	0	0,0
PROME	Metas y objetivos claros	9	0,3	3	0,5	0	0,0
COPREV	Buen diseño del curso	7	0,2	0	0,0	0	0,0
CLAVOZ	Claridad en la voz, pronunciación	5	0,1	1	0,2	0	0,0
EXACT	Actualización disciplinar	4	0,1	3	0,5	0	0,0
COIMP	Imparcialidad	3	0,1	1	0,2	0	0,0
COPREV	Diseño de curso (malla, carga, nivel, horario)	3	0,1	0	0,0	0	0,0
CONTEXT	Espacio	1	0,0	1	0,2	0	0,0
PRORESP	Preparación previa de la enseñanza	1	0,0	0	0,0	0	0,0
Total		3348	100	663	100	57	4068

“Elementos que se debería mejorar”

		Promedios IED					
Códigos	Descripción	4,1 a 5	%	3,1 a 4	%	0 a 3	%
COPREV-	Mal diseño del curso (horario, malla, nivel)	241	12,8	61	7,1	9	5,7
PROMET-	Métodos y estrategias inapropiadas	215	11,4	110	12,7	10	6,4
PROEVA-	Métodos de evaluación inválidos	149	7,9	42	4,9	7	4,5
CEACT-	No centrado en aprendizaje activo	111	5,9	45	5,2	2	1,3
PROMAT-	Material de apoyo inapropiado	93	5,0	27	3,1	3	1,9
CONTEXT-	Problemas de espacio	93	5,0	10	1,2	0	0,0
PROCRI-	Criterios de evaluación poco claros	87	4,6	28	3,2	24	15,3
PROPLA-	No prepara previamente la enseñanza	87	4,6	47	5,4	11	7,0
PROTIC-	No utiliza o mal utiliza la tecnología	71	3,8	27	3,1	3	1,9
PROTIME-	Uso poco efectivo del tiempo	68	3,6	35	4,0	4	2,5
CORESP-	Irresponsabilidad en el proceso de aprendizaje	67	3,6	52	6,0	14	8,9

RETROEF-	Retroalimentación no efectiva	66	3,5	18	2,1	11	7,0
PRORESEV-	Resultados de evaluación retrasados	62	3,3	27	3,1	2	1,3
CLAEX-	Poca claridad para explicar	47	2,5	49	5,7	4	2,5
CLAPRES-	Poca claridad para presentar	42	2,2	33	3,8	5	3,2
PROCONC-	No focaliza atención en conceptos clave	37	2,0	20	2,3	2	1,3
AMBRELA-	Mala relación profesor alumno	36	1,9	36	4,2	7	4,5
CEPRE-	No se preocupaba por las necesidades del estudiante	35	1,9	39	4,5	4	2,5
APACAR-	Características personales negativas	33	1,8	23	2,7	8	5,1
CLAVOZ-	Uso inadecuado de la voz	32	1,7	21	2,4	1	0,6
PROOUT-	No orienta el estudio fuera de la clase	32	1,7	11	1,3	1	0,6
PRODUD-	No responde dudas con claridad	31	1,7	16	1,8	4	2,5
PROCUO-	No cumple con los objetivos del curso	28	1,5	2	0,2	3	1,9
ACCEDUD-	Sin disponibilidad para responder dudas	24	1,3	17	2,0	6	3,8
PROVIN-	Sin vínculos con el mundo profesional	22	1,2	10	1,2	1	0,6
AMBCONF-	Ambiente de desconfianza	8	0,4	4	0,5	0	0,0
ALTAPRE-	Bajas expectativas	7	0,4	1	0,1	1	0,6
COIMP-	Diferenciación entre estudiantes	7	0,4	4	0,5	2	1,3
EXMA-	No maneja el contenido	7	0,4	10	1,2	0	0,0
AMB-	Ambiente de aprendizaje inadecuado	6	0,3	0	0,0	0	0,0
CODIS-	Discriminación	6	0,3	3	0,3	0	0,0
ENTMOT-	No motiva a los estudiantes	5	0,3	7	0,8	0	0,0
PROCOL-	No fomenta trabajo colaborativo	5	0,3	1	0,1	1	0,6
PROEJE-	No plantea ejemplos claros	5	0,3	2	0,2	1	0,6
PROCUM-	No cumple lo planificado	4	0,2	12	1,4	0	0,0
PROME-	No establece metas y objetivos	3	0,2	0	0,0	0	0,0
ENTENS-	Entusiasmo al enseñar	2	0,1	3	0,3	0	0,0
EXAC-	No está actualizado disciplinar	2	0,1	0	0,0	1	0,6
ACCETUT-	Sin disponibilidad para tutorizar	1	0,1	0	0,0	0	0,0
BUEN-	Mal docente	1	0,1	8	0,9	5	3,2
CODEN-	Denuncias de mala conducta	0	0,0	3	0,3	0	0,0
ENTENS-	No se entusiasma con enseñar	0	0,0	1	0,1	0	0,0
		1878	100	865	100	157	100

Anexo 12

Relación evaluación de la docencia y percepción de la calidad de la docencia (por área de conocimiento)

Humanidades: “Elementos que debería mantener”

Códigos	Descripción del código	Promedios IED					
		4,1 a 5	%	3,1 a 4	%	0 a 3	%
BUEN	Buen profesor	117	15,9	15	11,0	0	0
PROMET	Métodos y estrategias apropiadas	76	10,3	16	11,8	0	0
APACAR	Características personales (amable, simpático)	49	6,7	3	2,2	0	0

CEACT	Aprendizaje activo	48	6,5	7	5,1	0	0
PROVIN	Vínculos con el mundo profesional	48	6,5	11	8,1	0	0
CEPRE	Centrado en necesidades de estudiantes	47	6,4	8	5,9	0	0
EXMA	Manejo del contenido	45	6,1	10	7,4	0	0
AMB	Ambiente de aprendizaje	32	4,4	4	2,9	0	0
AMBRELA	Buena relación con estudiantes	32	4,4	4	2,9	0	0
PROMAT	Material de apoyo apropiado	30	4,1	8	5,9	1	100
CLAEX	Claridad para explicar	23	3,1	2	1,5	0	0
PRODUD	Responde claramente a las dudas	22	3,0	7	5,1	0	0
CLAPRES	Claridad para presentar	21	2,9	4	2,9	0	0
PROEVA	Métodos de evaluación válidos	14	1,9	3	2,2	0	0
RETROEF	Retroalimentación efectiva	14	1,9	3	2,2	0	0
ENTENS	Entusiasmo por lo que enseña	12	1,6	0	0,0	0	0
CORESP	Responsabilidad en proceso de aprendizaje	10	1,4	2	1,5	0	0
ACCEDUD	Disponibilidad para responder dudas	9	1,2	8	5,9	0	0
ENTMOT	Motiva a los estudiantes	9	1,2	0	0,0	0	0
PROCUO	Cumple con los objetivos	9	1,2	2	1,5	0	0
PROPLA	Preparación previa de la enseñanza	9	1,2	0	0,0	0	0
PROTIC	Uso apropiado de las tecnologías	9	1,2	9	6,6	0	0
AMBCONF	Ambiente de confianza	8	1,1	2	1,5	0	0
PROCOL	Fomenta trabajo colaborativo	7	1,0	1	0,7	0	0
PROCUM	Cumple con lo planificado	7	1,0	0	0,0	0	0
PROTIME	Uso efectivo del tiempo	6	0,8	2	1,5	0	0
PROCRI	Criterios de evaluación claros	4	0,5	2	1,5	0	0
PROOUT	Orienta el estudio fuera de la clase	4	0,5	0	0,0	0	0
PROCONC	Focaliza en conceptos clave	3	0,4	1	0,7	0	0
PROME	Metas y objetivos claros	3	0,4	1	0,7	0	0
ALTAPRE	Altas expectativas	2	0,3	0	0,0	0	0
PROEJE	Plantea ejemplos claros	2	0,3	0	0,0	0	0
ACCETUT	Disponibilidad para tutorizar	1	0,1	0	0,0	0	0
PRORESEV	Entrega resultados de evaluación oportunamente	1	0,1	0	0,0	0	0
EXACT	Actualización disciplinar	0	0,0	1	0,7	0	0
		735	100	136	100	1	100

Humanidades: “Elementos que se deberían mejorar”

Promedios IED							
Códigos	Descripción del código	4,1 a 5	%	3,1 a 4	%	0 a 3	%
COPREV-	Mal diseño del curso (horario, malla, nivel)	61	18,7	12	6,6	0	0
PROMET-	Métodos y estrategias inapropiadas	43	13,1	15	8,2	1	8,3
PROEVA-	Métodos de evaluación inválidos	25	7,6	8	4,4	0	0,0
CONTEXT-	Problemas de espacio	22	6,7	7	3,8	0	0
PROPLA-	No prepara previamente la enseñanza	20	6,1	17	9,3	1	8,3

CORESP-	Irresponsabilidad en el proceso de aprendizaje	14	4,3	23	12,6	4	33,3
PROCRI-	Criterios de evaluación poco claros	14	4,3	3	1,6	0	0
RETROEF-	Retroalimentación no efectiva	13	4,0	3	1,6	0	0
PROTIME-	Uso poco efectivo del tiempo	12	3,7	6	3,3	0	0
CEACT-	No centrado en aprendizaje activo	12	3,7	2	1,1	0	0
PROMAT-	Material de apoyo inapropiado	12	3,7	2	1,1	1	8,3
PROTIC-	No utiliza o mal utiliza la tecnología	11	3,4	3	1,6	0	0
AMBRELA-	Mala relación profesor alumno	6	1,8	12	6,6	0	0
ACCEDUD-	Sin disponibilidad para responder dudas	6	1,8	4	2,2	0	0
PROVIN-	Sin vínculos con el mundo profesional	6	1,8	3	1,6	1	8,3
EXMA-	No maneja el contenido	5	1,5	4	2,2	0	0
PRORESEV-	Resultados de evaluación retrasados	5	1,5	4	2,2	1	8,3
CLAVOZ-	Uso inadecuado de la voz	5	1,5	1	0,5	0	0
CLAEX-	Poca claridad para explicar	4	1,2	10	5,5	0	0
AMB-	Ambiente de aprendizaje inadecuado	4	1,2	0	0,0	0	0
CEPRE-	No se preocupaba por las necesidades del estudiante	3	0,9	8	4,4	0	0
CLAPRES-	Poca claridad para presentar	3	0,9	8	4,4	0	0
PROOUT-	No orienta el estudio fuera de la clase	3	0,9	4	2,2	0	0,0
AMBCONF-	Ambiente de desconfianza	3	0,9	2	1,1	0	0
APACAR-	Características personales negativas	2	0,6	2	1,1	2	16,7
CODISC-	Discriminación	2	0,6	1	0,5	0	0
ENTENS-	Entusiasmo al enseñar	2	0,6	1	0,5	0	0
COIMP-	Diferenciación entre estudiantes	2	0,6	0	0,0	0	0
PROCUO-	No cumple objetivos del curso	2	0,6	0	0,0	0	0
PROCUM-	No cumple lo planificado	1	0,3	9	4,9	0	0
ENTMOT-	No motiva a los estudiantes	1	0,3	1	0,5	0	0
ALTAPRE-	Bajas expectativas	1	0,3	0	0,0	0	0
PROEJE-	No realiza ejemplos	1	0,3	0	0,0	1	8,3
PROME-	No establece metas y objetivos	1	0,3	0	0,0	0	0,0
PROCONC-	No focaliza atención en conceptos clave	0	0,0	5	2,7	0	0
BUEN-	Mal docente	0	0,0	2	1,1	0	0
PRODUD-	No responde dudas con claridad	0	0,0	1	0,5	0	0
		327	100	183	100	12	100

Ciencias naturales: “Elementos que debería mantener”

Códigos	Descripción del código	Promedios IED					
		4,1 a 5	%	3,1 a 4	%	1 a 3	%
CLAEX	Claridad para explicar	72	14,9	4	11,1	0	0
BUEN	Buen profesor	65	13,4	6	16,7	0	0
PROMET	Métodos y estrategias apropiadas	45	9,3	2	5,6	0	0
CEPRE	Centrado en necesidades de estudiantes	43	8,9	3	8,3	0	0
PRODUD	Responde claramente a las dudas	33	6,8	3	8,3	2	40
CEACT	Aprendizaje activo	24	5,0	1	2,8	0	0
APACAR	Características personales (amable, simpático)	20	4,1	1	2,8	1	20
ACCEDUD	Disponibilidad para responder dudas	17	3,5	1	2,8	1	20
CLAPRES	Claridad para presentar	14	2,9	0	0,0	0	0
EXMA	Manejo del contenido	14	2,9	0	0,0	1	20
PROVIN	Vínculos con el mundo profesional	14	2,9	1	2,8	0	0
AMB	Ambiente de aprendizaje	13	2,7	1	2,8	0	0
AMBRELA	Buena relación con estudiantes	12	2,5	0	0,0	0	0

PROMAT	Material de apoyo apropiado	10	2,1	1	2,8	0	0
ENTENS	Entusiasmo por lo que enseña	9	1,9	0	0,0	0	0
CORESP	Responsabilidad en proceso de aprendizaje	8	1,7	2	5,6	0	0
ENTMOT	Motiva a los estudiantes	8	1,7	1	2,8	0	0
PROPLA	Preparación previa de la enseñanza	8	1,7	0	0,0	0	0
PRORESEV	Entrega resultados de evaluación oportunamente	8	1,7	1	2,8	0	0
AMBCONF	Ambiente de confianza	7	1,4	1	2,8	0	0
PROTIME	Uso efectivo del tiempo	7	1,4	0	0,0	0	0
PROCOL	Fomenta trabajo colaborativo	6	1,2	0	0,0	0	0
PROCUM	Cumple con lo planificado	5	1,0	3	8,3	0	0
RETROEF	Retroalimentación efectiva	4	0,8	2	5,6	0	0
ALTAPRE	Altas expectativas	3	0,6	0	0,0	0	0
PROEJE	Plantea ejemplos claros	3	0,6	1	2,8	0	0
PROOUT	Orienta el estudio fuera de la clase	3	0,6	0	0,0	0	0
ACCETUT	Disponibilidad para tutorizar	2	0,4	0	0,0	0	0
PROCONC	Focaliza en conceptos clave	2	0,4	0	0,0	0	0
EXAC	Actualización disciplinar	1	0,2	0	0,0	0	0
PROCRI	Criterios de evaluación claros	1	0,2	0	0,0	0	0
PROCUO	Cumple con los objetivos	1	0,2	0	0,0	0	0
PROEVA	Métodos de evaluación válidos	1	0,2	0	0,0	0	0
PROTIC	Uso apropiado de las tecnologías	1	0,2	1	2,8	0	0
		484	100	36	100	5	100

Ciencias naturales: “Elementos que debería mejorar”

		Promedios IED					
Códigos	Descripción del código	4,1 a 5	%	3,1 a 4	%	0 a 3	%
COPREV-	Mal diseño del curso (horario, malla, nivel)	25	8,8	6	9,4	0	0,0
PROMET-	Métodos y estrategias inapropiadas	23	8,1	6	9,4	1	16,7
PROCRI-	Criterios de evaluación poco claros	21	7,4	2	3,1	0	0,0
CEACT-	No centrado en aprendizaje activo	20	7,0	2	3,1	0	0,0
PROMAT-	Material de apoyo inapropiado	18	6,3	0	0,0	0	0,0
PRORESEV-	Resultados de evaluación retrasados	16	5,6	2	3,1	0	0,0
PROEVA-	Métodos de evaluación inválidos	15	5,3	5	7,8	0	0,0
PROPLA-	No prepara previamente la enseñanza	15	5,3	0	0,0	0	0,0
RETROEF-	Retroalimentación no efectiva	15	5,3	0	0,0	0	0,0
CONTEXT-	Problemas de espacio	13	4,6	1	1,6	0	0,0
APACAR-	Características personales negativas	9	3,2	1	1,6	0	0,0
CORESP-	Irresponsabilidad en el proceso de aprendizaje	9	3,2	0	0,0	0	0,0
PROTIME-	Uso poco efectivo del tiempo	9	3,2	0	0,0	0	0,0
CLAVOZ-	Uso inadecuado de la voz	8	2,8	6	9,4	0	0,0
PROCONC-	No focaliza atención en conceptos clave	8	2,8	0	0,0	0	0,0

CLAPRES-	Poca claridad para presentar	7	2,5	11	17,2	2	33,3
AMBRELA-	Mala relación profesor alumno	7	2,5	4	6,3	0	0,0
PRODUD-	No responde dudas con claridad	7	2,5	4	6,3	0	0,0
PROTIC-	No utiliza o mal utiliza la tecnología	7	2,5	2	3,1	0	0,0
PROVIN-	Sin vínculos con el mundo profesional	7	2,5	1	1,6	0	0,0
CLAEX-	Poca claridad para explicar	5	1,8	3	4,7	0	0,0
CEPRE-	No se preocupaba por las necesidades del estudiante	5	1,8	2	3,1	0	0,0
ACCEDUD-	Sin disponibilidad para responder dudas	3	1,1	1	1,6	1	16,7
PROOUT-	No orienta el estudio fuera de la clase	3	1,1	1	1,6	1	16,7
AMBCONF-	Ambiente de desconfianza	2	0,7	0	0,0	0	0,0
PROME-	No establece metas y objetivos	2	0,7	0	0,0	0	0,0
BUEN-	Mal docente	1	0,4	2	3,1	1	16,7
EXMA-	No maneja el contenido	1	0,4	1	1,6	0	0,0
COIMP-	Diferenciación entre estudiantes	1	0,4	0	0,0	0	0,0
PROCUM-	No cumple lo planificado	1	0,4	0	0,0	0	0,0
PROEJE-	No realiza ejemplos	1	0,4	0	0,0	0	0,0
ENTENS-	Entusiasmo al enseñar	0	0,0	1	1,6	0	0,0
		284	100	64	100	6	100

Ciencias matemáticas: “Elementos que debería mantener”

Códigos	Descripción del código	Promedios IED					
		4,1 a 5	%	3,1 a 4	%	0 a 3	%
BUEN	Buen profesor	46	20,6	8	13,6	0	0,0
CEACT	Aprendizaje activo	18	8,1	4	6,8	0	0,0
PROMET	Métodos y estrategias apropiadas	18	8,1	5	8,5	0	0,0
CLAPRES	Claridad para presentar	16	7,2	3	5,1	0	0,0
AMBRELA	Buena relación con estudiantes	15	6,7	4	6,8	0	0,0
CEPRE	Centrado en necesidades de estudiantes	13	5,8	6	10,2	0	0,0
APACAR	Características personales (amable, simpático)	11	4,9	5	8,5	2	22,2
PRODUD	Responde claramente a las dudas	11	4,9	2	3,4	0	0,0
CLAEX	Claridad para explicar	9	4,0	1	1,7	0	0,0
EXMA	Manejo del contenido	8	3,6	2	3,4	0	0,0
PROTIC	Uso apropiado de las tecnologías	7	3,1	1	1,7	0	0,0
AMB	Ambiente de aprendizaje	4	1,8	0	0,0	0	0,0
CORESP	Responsabilidad en proceso de aprendizaje	4	1,8	3	5,1	0	0,0
ENTENS	Entusiasmo por lo que enseña	4	1,8	0	0,0	1	11,1
PROOUT	Orienta el estudio fuera de la clase	4	1,8	1	1,7	0	0,0
PROTIME	Uso efectivo del tiempo	4	1,8	2	3,4	1	11,1
RETROEF	Retroalimentación efectiva	4	1,8	2	3,4	0	0,0
ACCEDUD	Disponibilidad para responder dudas	3	1,3	2	3,4	3	33,3

AMBCONF	Ambiente de confianza	3	1,3	1	1,7	0	0,0
PROEVA	Métodos de evaluación válidos	3	1,3	0	0,0	0	0,0
PROVIN	Vínculos con el mundo profesional	3	1,3	2	3,4	2	22,2
ALTAPRE	Altas expectativas	2	0,9	0	0,0	0	0,0
ACCETUT	Disponibilidad para tutorizar	1	0,4	0	0,0	0	0,0
ENTMOT	Motiva a los estudiantes	1	0,4	0	0,0	0	0,0
PROCOL	Fomenta trabajo colaborativo	1	0,4	0	0,0	0	0,0
PROCRI	Criterios de evaluación claros	1	0,4	2	3,4	0	0,0
PROCUM	Cumple con lo planificado	1	0,4	0	0,0	0	0,0
PROCUO	Cumple con los objetivos	1	0,4	1	1,7	0	0,0
PROEJE	Plantea ejemplos claros	1	0,4	0	0,0	0	0,0
PROMAT	Material de apoyo apropiado	1	0,4	1	1,7	0	0,0
PROME	Metas y objetivos claros	1	0,4	0	0,0	0	0,0
PRORESP	Preparación previa de la enseñanza	1	0,4	0	0,0	0	0,0
PROCONC	Focaliza en conceptos clave	0	0,0	1	1,7	0	0,0
		223	100	59	100	9	100

Ciencias matemáticas: “Elementos que debería mejorar”

		Promedios IED					
Códigos	Descripción del código	4,1 a 5	%	3,1 a 4	%	0 a 3	%
COPREV-	Mal diseño del curso (horario, malla, nivel)	39	14,9	14	10,7	4	6,3
PROEVA-	Métodos de evaluación inválidos	35	13,4	16	12,2	1	1,6
PROMET-	Métodos y estrategias inapropiadas	26	10,0	15	11,5	5	7,8
PROTIC-	No utiliza o mal utiliza la tecnología	19	7,3	9	6,9	2	3,1
PROCRI-	Criterios de evaluación poco claros	17	6,5	10	7,6	9	14,1
PRORESEV-	Resultados de evaluación retrasados	15	5,7	2	1,5	1	1,6
PROMAT-	Material de apoyo inapropiado	14	5,4	9	6,9	1	1,6
CONTEXT-	Problemas de espacio	13	5,0	1	0,8	0	0,0
PROPLA-	No prepara previamente la enseñanza	11	4,2	5	3,8	5	7,8
CEACT-	No centrado en aprendizaje activo	10	3,8	7	5,3	1	1,6
RETROEF-	Retroalimentación no efectiva	9	3,4	3	2,3	6	9,4
PROCUO-	No cumple con los objetivos del curso	8	3,1	2	1,5	3	4,7
PROTIME-	Uso poco efectivo del tiempo	7	2,7	4	3,1	3	4,7
CORESP-	Irresponsabilidad en el proceso de aprendizaje	6	2,3	3	2,3	1	1,6
PRODUD-	No responde dudas con claridad	6	2,3	3	2,3	1	1,6
PROCONC-	No focaliza atención en conceptos clave	5	1,9	8	6,1	2	3,1
CLAEX-	Poca claridad para explicar	5	1,9	3	2,3	1	1,6
PROOUT-	No orienta el estudio fuera de la clase	3	1,1	0	0,0	0	0,0
CLAPRES-	Poca claridad para presentar	2	0,8	2	1,5	2	3,1

CLAVOZ-	Uso inadecuado de la voz	2	0,8	1	0,8	1	1,6
ALTAPRE-	Bajas expectativas	2	0,8	0	0,0	1	1,6
CEPRE-	No se preocupaba por las necesidades del estudiante	2	0,8	0	0,0	3	4,7
ACCEDUD-	Sin disponibilidad para responder dudas	1	0,4	3	2,3	5	7,8
ENTMOT-	No motiva a los estudiantes	1	0,4	2	1,5	0	0,0
CODEN-	Discriminación	1	0,4	1	0,8	0	0,0
PROCOL-	No fomenta trabajo colaborativo	1	0,4	1	0,8	1	1,6
APACAR-	Características personales negativas	1	0,4	0	0,0	0	0,0
AMBRELA-	Mala relación profesor alumno	0	0,0	5	3,8	4	6,3
AMBCONF-	Ambiente de desconfianza	0	0,0	1	0,8	0	0,0
PROVIN-	Sin vínculos con el mundo profesional	0	0,0	1	0,8	0	0,0
COIMP-	Diferenciación entre estudiantes	0	0,0	0	0,0	1	1,6
		261	100	131	100	64	100

Ciencias médicas: “Elementos que se deberían mantener”

		Promedios IED					
Códigos	Descripción del código	4,1 a 5	%	3,1 a 4	%	0 a 3	%
CLAEX	Claridad para explicar	89	11,2	7	4,5	0	0
BUEN	Buen profesor	77	9,7	10	6,4	1	33,3
CEACT	Aprendizaje activo	69	8,7	15	9,6	0	0
PROMET	Métodos y estrategias apropiadas	64	8,1	10	6,4	0	0
CEPRE	Preocupación por necesidades del estudiante	50	6,3	3	1,9	0	0
PROVIN	Vínculos con el mundo profesional	49	6,2	24	15,4	0	0
APACAR	Características personales (amable, simpático)	48	6,1	7	4,5	0	0
PRODUD	Responde claramente a las dudas	39	4,9	7	4,5	0	0
CLAPRES	Claridad para presentar	31	3,9	2	1,3	0	0
ENTMOT	Motiva a los estudiantes	29	3,7	1	0,6	0	0
AMBCONF	Ambiente de confianza	24	3	3	1,9	0	0
PROEJE	Plantea ejemplos claros	23	2,9	2	1,3	0	0
CORESP	Responsabilidad en proceso de aprendizaje	19	2,4	5	3,2	0	0
AMBRELA	Buena relación con estudiantes	18	2,3	1	0,6	0	0
EXMA	Manejo del contenido	16	2	10	6,4	0	0
AMB	Ambiente de aprendizaje	16	2	4	2,6	0	0

PROMAT	Material de apoyo apropiado	15	1,9	3	1,9	0	0
PROCUM	Cumple con lo planificado	12	1,5	7	4,5	2	66,6
RETROEF	Retroalimentación efectiva	11	1,4	3	1,9	0	0
PROCUO	Cumple con los objetivos del curso	10	1,3	2	1,3	0	0
ENTENS	Entusiasmo por lo que enseña	10	1,3	0	0	0	0
ACCEDUD	Disponibilidad para responder dudas	9	1,1	3	1,9	0	0
PROTIC	Uso de la tecnología	9	1,1	0	0	0	0
PRORESEV	Resultados de evaluación a tiempo	8	1	3	1,9	0	0
PROCOL	Fomenta trabajo colaborativo	8	1	0	0	0	0
PROEVA	Métodos de evaluación válidos	7	0,9	1	0,6	0	0
ACCETUT	Disponibilidad para tutorizar	5	0,6	3	1,9	0	0
PROCONC	Focaliza atención en conceptos clave	5	0,6	3	1,9	0	0
PROOUT	Orienta el estudio fuera de la clase	5	0,6	3	1,9	0	0
PROCRI	Criterios de evaluación claros	4	0,5	2	1,3	0	0
PROPLA	Preparación previa de la enseñanza	3	0,4	3	1,9	0	0
ALTAPRE	Altas expectativas	3	0,4	1	0,6	0	0
COPREV	Diseño de curso (malla, carga, nivel, horario)	3	0,4	0	0	0	0
PROTIME	Uso efectivo del tiempo	2	0,3	5	3,2	0	0
EXAC	Actualización disciplinar	2	0,3	2	1,3	0	0
CONTEXT	Espacio	1	0,1	0	0	0	0
PROME	Metas y objetivos claros	0	0	1	0,6	0	0
		793	100	156	100	3	100

Ciencias médicas: “Elementos que se deberían mejorar”

		Promedios IED					
Código	Descripción del código	4,1 a 5	%	3,1 a 4	%	0 a 3	%
PROMET-	Métodos y estrategias inapropiadas	52	12,9	29	12,4	0	0
COPREV-	Diseño del curso (horario, malla, nivel)	41	10,2	7	3	0	0
CEACT-	Promover aprendizaje activo	40	9,9	23	9,8	0	0
PROMAT-	Material de apoyo inapropiado	21	5,2	6	2,6	0	0
CONTEXT-	Problemas de espacio	21	5,2	0	0	0	0
PROEVA-	Métodos de evaluación inválidos	20	5	5	2,1	0	0
PROTIME-	Uso poco efectivo del tiempo	18	4,5	11	4,7	0	0
CLAPRES-	Poca claridad para presentar	18	4,5	5	2,1	0	0
CORESP-	Irresponsabilidad en el proceso de aprendizaje	15	3,7	18	7,7	1	14,2
PROOUT-	No orienta el estudio fuera de la clase	15	3,7	3	1,3	0	0
CLAEX-	Poca claridad para explicar	14	3,5	26	11,1	3	42,8
CLAVOZ-	Uso inadecuado de la voz	14	3,5	9	3,8	0	0
APACAR-	Características personales negativas	14	3,5	7	3	1	14,2
PROPLA-	No prepara previamente la enseñanza	13	3,2	11	4,7	0	0

AMBRELA-	Mala relación profesor alumno	10	2,5	9	3,8	0	0
PRODUD-	No responde claramente a las dudas	10	2,5	5	2,1	0	0
CEPRE-	No se preocupaba por las necesidades del estudiante	9	2,2	15	6,4	1	14,2
PROTIC-	No utiliza o mal utiliza la tecnología	8	2	4	1,7	0	0
PROVIN-	Sin vínculos con el mundo profesional	7	1,7	4	1,7	0	0
RETROEF-	Retroalimentación no efectiva	7	1,7	4	1,7	0	0
PRORESEV-	Resultados de evaluación retrasados	6	1,5	13	5,6	0	0
ACCEDUD-	Sin disponibilidad para responder dudas	6	1,5	2	0,9	0	0
PROCRI-	Criterios de evaluación poco claros	5	1,2	2	0,9	1	14,2
PROCONC-	No focaliza atención en conceptos clave	4	1	4	1,7	0	0
AMBCONF-	Ambiente de desconfianza	2	0,5	1	0,4	0	0
COIMP-	Diferenciación entre estudiantes	2	0,5	1	0,4	0	0
PROEJE-	No plantea ejemplos claros	2	0,5	1	0,4	0	0
EXAC-	No está actualizado disciplinar	2	0,5	0	0	0	0
PROCUO-	No cumple con los objetivos del curso	2	0,5	0	0	0	0
ENTMOT-	No motiva a los estudiantes	1	0,2	2	0,9	0	0
PROCUM-	No cumple con lo planificado	1	0,2	2	0,9	0	0
ACCETUT-	Sin disponibilidad para tutorizar	1	0,2	0	0	0	0
AMB-	Ambiente de aprendizaje inadecuado	1	0,2	0	0	0	0
PROCOL-	No fomenta trabajo colaborativo	1	0,2	0	0	0	0
CODEN-	Denuncias de mala conducta	0	0	3	1,3	0	0
EXMA-	No maneja el contenido o disciplina	0	0	1	0,4	0	0
		403	100	234	100	7	100

Diseño: “Elementos que debería mantener”

		Promedios IED					
Códigos	Descripción del código	4,1 a 5	%	3,1 a 4	%	0 a 3	%
BUEN	Buen profesor	69	19,0	12	17,1	00	0
APACAR	Características personales (amable, simpático)	42	11,6	10	14,3	00	0
CEACT	Aprendizaje activo	34	09,4	04	05,7	00	0
CEPRE	Centrado en necesidades de estudiantes	34	09,4	04	05,7	00	0
PROMET	Métodos y estrategias apropiadas	29	08,0	05	07,1	02	66,6
CLAPRES	Claridad para presentar	17	04,7	03	04,3	00	0
AMBRELA	Buena relación con estudiantes	17	04,7	02	02,9	00	0
EXMA	Manejo del contenido	14	03,9	06	08,6	00	0
CLAEX	Claridad para explicar	14	03,9	03	04,3	00	0
CORESP	Responsabilidad en proceso de aprendizaje	10	02,8	00	00,0	00	0
AMB	Ambiente de aprendizaje	08	02,2	02	02,9	00	0
AMBCONF	Ambiente de confianza	08	02,2	01	01,4	00	0
RETROEF	Retroalimentación efectiva	08	02,2	00	00,0	00	0

PROVIN	Vínculos con el mundo profesional	06	01,7	05	07,1	00	0
PROMAT	Material de apoyo apropiado	06	01,7	02	02,9	00	0
ACCEDUD	Disponibilidad para responder dudas	05	01,4	00	00,0	00	0
PROCUO	Cumple con los objetivos del curso	05	01,4	00	00,0	00	0
PROTIME	Uso efectivo del tiempo	04	01,1	01	01,4	00	0
ENTMOT	Motiva a los estudiantes	04	01,1	00	00,0	00	0
PRODUD	Responde claramente a las dudas	04	01,1	00	00,0	00	0
PROEVA	Métodos de evaluación válidos	04	01,1	00	00,0	00	0
PROOUT	Orienta el estudio fuera de la clase	03	00,8	02	02,9	00	0
ENTENS	Entusiasmo por lo que enseña	03	00,8	01	01,4	01	33,3
PROPLA	Preparación previa de la enseñanza	03	00,8	00	00,0	00	0
PROCONC	Focaliza en conceptos clave	02	00,6	02	02,9	00	0
PROTIC	Uso de la tecnología	02	00,6	02	02,9	00	0
COIMP	imparcialidad	02	00,6	00	00,0	00	0
PROCUM	Cumple con lo planificado	02	00,6	00	00,0	00	0
ALTAPRE	Altas expectativas	01	00,3	00	00,0	00	0
CLAVOZ	Claridad en la voz, pronunciación	01	00,3	00	00,0	00	0
EXACT	Actualización disciplinar	01	00,3	00	00,0	00	0
PROEJE	Plantea ejemplos claros	01	00,3	00	00,0	00	0
CONTEXT	Condiciones de espacio	00	00,0	01	01,4	00	0
PROCOL	Fomenta trabajo colaborativo	00	00,0	01	01,4	00	0
PROCRI	Criterios de evaluación claros	00	00,0	01	01,4	00	0
		363	100	70	100	03	100

Diseño: “Elementos que se deberían mejorar”

Códigos	Descripción del código	Promedios IED					
		4,1 a 5	%	3,1 a 4	%	0 a 3	%
PROMET-	Métodos y estrategias inapropiadas	19	11,7	13	20,3	0	0,0
CORESP-	Irresponsabilidad en el proceso de aprendizaje	14	8,6	3	4,7	2	14,3
PROPLA-	No prepara previamente la enseñanza	13	8,0	2	3,1	3	21,4
COPREV-	Diseño del curso (horario, malla, nivel)	12	7,4	3	4,7	4	28,6
CEACT-	no centrado en aprendizaje activo	11	6,7	3	4,7	0	0,0
CEPRE-	No se preocupaba por las necesidades del estudiante	11	6,7	2	3,1	0	0,0
CONTEXT-	Problemas de espacio	9	5,5	0	0,0	0	0,0
CLAPRES-	Poca claridad para presentar	8	4,9	5	7,8	0	0,0
CLAEX-	Poca claridad para explicar	7	4,3	0	0,0	0	0,0
PROTIME-	Uso poco efectivo del tiempo	7	4,3	0	0,0	0	0,0
APACAR-	Características personales negativas	6	3,7	5	7,8	0	0,0

PROCRI-	Criterios de evaluación poco claros	5	3,1	5	7,8	1	7,1
AMBRELA-	Mala relación profesor alumno	5	3,1	3	4,7	1	7,1
RETROEF-	Retroalimentación no efectiva	5	3,1	1	1,6	0	0,0
PROCONC-	No focaliza atención en conceptos clave	5	3,1	0	0,0	0	0,0
PROMAT-	Material de apoyo inapropiado	4	2,5	3	4,7	1	7,1
PROOUT-	No orienta el estudio fuera de la clase	3	1,8	3	4,7	0	0,0
PROCUO-	No cumple con los objetivos del curso	3	1,8	0	0,0	0	0,0
COIMP-	Diferenciación entre estudiantes	2	1,2	2	3,1	0	0,0
PRORESEV-	Resultados de evaluación retrasados	2	1,2	1	1,6	0	0,0
PROTIC-	No utiliza o mal utiliza la tecnología	2	1,2	1	1,6	0	0,0
ALTAPRE-	Bajas expectativas	2	1,2	0	0,0	0	0,0
PROEVA-	Métodos de evaluación inválidos	1	0,6	3	4,7	1	7,1
ACCEDUD-	Sin disponibilidad para responder dudas	1	0,6	2	3,1	0	0,0
EXMA-	No maneja el contenido	1	0,6	1	1,6	0	0,0
AMB-	Mal ambiente para aprender	1	0,6	0	0,0	0	0,0
CLAVOZ-	Uso inadecuado de la voz	1	0,6	0	0,0	0	0,0
CODISC-	Discriminación	1	0,6	0	0,0	0	0,0
PROCUM-	No cumple con lo planificado	1	0,6	0	0,0	0	0,0
PROEJE-	No plantea ejemplos claros	1	0,6	0	0,0	0	0,0
ENTENS-	No se entusiasma con enseñar	0	0,0	1	1,6	0	0,0
ENTMOT-	No motiva a los estudiantes	0	0,0	1	1,6	0	0,0
PROVIN-	Sin vínculos con el mundo profesional	0	0,0	1	1,6	0	0,0
EXAC-	No está actualizado disciplinar	0	0,0	0	0,0	1	7,1
		163	100,0	64	100,0	14	100

Bachillerato: “Elementos que debería mantener”

		Promedios IED					
Códigos	Descripción del código	4,1 a 5	%	3,1 a 4	%	0 a 3	%
CEACT	Aprendizaje activo	67	10,0	2	4,4	2	13,3
CLAEX	Claridad para explicar	59	8,8	6	13,3	0	0
PROMET	Métodos y estrategias apropiadas	58	8,7	6	13,3	2	13,3
PRODUD	Responde claramente a las dudas	55	8,2	2	4,4	0	0
CEPRE	Centrado en necesidades de estudiantes	53	7,9	2	4,4	0	0
CLAPRES	Claridad para presentar	34	5,1	0	0	1	6,6
APACAR	Características personales (amable, simpático)	31	4,6	1	2,2	0	0
BUEN	Buen profesor	25	3,7	2	4,4	0	0

PROVIN	Vínculos con el mundo profesional	25	3,7	0	0	0	0
AMBRELA	Buena relación con estudiantes	22	3,3	4	8,9	0	0
PROEJE	Plantea ejemplos claros	22	3,3	1	2,2	0	0
AMB	Ambiente de aprendizaje	18	2,7	2	4,4	2	13,3
ACCEDUD	Disponibilidad para responder dudas	17	2,5	0	0	0	0
PROEVA	Métodos de evaluación válidos	16	2,4	1	2,2	0	0
RETROEF	Retroalimentación efectiva	15	2,2	0	0	0	0
PROTIC	Uso de la tecnología	14	2,1	3	6,4	0	0
PROCUM	Cumple con lo planificado	13	1,9	0	0	0	0
PROMAT	Material de apoyo apropiado	13	1,9	1	2,2	0	0
CORESP	Responsabilidad en proceso de aprendizaje	12	1,8	1	2,2	1	6,6
EXMA	Manejo del contenido	12	1,8	5	11,1	0	0
PROTIME	Uso efectivo del tiempo	12	1,8	0	0	1	6,6
PROCONC	Focaliza en conceptos clave	11	1,6	2	4,4	0	0
AMBCONF	Ambiente de confianza	10	1,5	1	2,2	0	0
COPREV	Buen diseño del curso	7	1,0	0	0	0	0
PROCOL	Fomenta trabajo colaborativo	7	1,0	0	0	5	33,3
PROPLA	Preparación previa de la enseñanza	6	0,9	0	0	0	0
ALTAPRE	Altas expectativas	5	0,7	0	0	0	0
ENTENS	Entusiasmo por lo que enseña	5	0,7	0	0	0	0
PROCRI	Criterios de evaluación claros	5	0,7	0	0	1	6,6
PROME	Metas y objetivos claros	5	0,7	0	0	0	0
PROOUT	Orienta el estudio fuera de la clase	5	0,7	0	0	0	0
CLAVOZ	Claridad en la voz, pronunciación	4	0,6	1	2,2	0	0
ENTMOT	Motiva a los estudiantes	3	0,4	0	0	0	0
PRORESEV	Entrega de evaluaciones a tiempo	3	0,4	2	4,4	0	0
COIMP	Imparcialidad	1	0,1	0	0	0	0
		670	100	45	100	15	100

Bachillerato: “Elementos que debería mejorar”

Códigos	Descripción del código	Promedios IED					
		4,1 a 5	%	3,1 a 4	%	0 a 3	%
PROEVA-	Métodos de evaluación inválidos	49	12,9	0	0,0	3	12,0
COPREV-	Mal diseño del curso (horario, malla, nivel)	47	12,3	10	20,0	0	0,0
PROMET-	Métodos y estrategias inapropiadas	40	10,5	5	10,0	1	4,0
PROTIC-	No utiliza o mal utiliza la tecnología	24	6,3	6	12,0	0	0,0
PROCRI-	Criterios de evaluación poco claros	24	6,3	0	0,0	12	48,0
PROMAT-	Material de apoyo inapropiado	20	5,2	4	8,0	0	0,0
PRORESEV-	Resultados de evaluación retrasados	18	4,7	0	0,0	0	0,0
CEACT-	no centrado en aprendizaje activo	15	3,9	3	6,0	0	0,0
RETROEF-	Retroalimentación no efectiva	15	3,9	0	0,0	3	12,0
PROCONC-	No focaliza atención en conceptos clave	14	3,7	1	2,0	0	0,0

CONTEXT-	Problemas de espacio	14	3,7	0	0,0	0	0,0
PROPLA-	No prepara previamente la enseñanza	13	3,4	7	14,0	1	4,0
PROCUO-	No cumple con los objetivos del curso	13	3,4	0	0,0	0	0,0
PROTIME-	Uso poco efectivo del tiempo	12	3,1	2	4,0	0	0,0
CLAEX-	Poca claridad para explicar	9	2,4	0	0,0	0	0,0
CORESP-	Irresponsabilidad en el proceso de aprendizaje	9	2,4	0	0,0	1	4,0
PRODUD-	No responde dudas con claridad	7	1,8	2	4,0	1	4,0
ACCEDUD-	Sin disponibilidad para responder dudas	6	1,6	3	6,0	0	0,0
AMBRELA-	Mala relación profesor alumno	6	1,6	2	4,0	1	4,0
CEPRE-	No se preocupaba por las necesidades del estudiante	5	1,3	0	0,0	0	0,0
CLAPRES-	Poca claridad para presentar	4	1,0	1	2,0	1	4,0
PROCOL-	No fomenta trabajo colaborativo	3	0,8	0	0,0	0	0,0
PROOUT-	No orienta el estudio fuera de la clase	3	0,8	0	0,0	0	0,0
CLAVOZ-	Uso inadecuado de la voz	2	0,5	2	4,0	0	0,0
ALTAPRE-	Bajas expectativas	2	0,5	1	2,0	0	0,0
ENTMOT-	No motiva a los estudiantes	2	0,5	1	2,0	0	0,0
CODEN-	Discriminación	2	0,5	0	0,0	0	0,0
AMBCONF-	Ambiente de desconfianza	1	0,3	0	0,0	0	0,0
APACAR-	Características personales negativas	1	0,3	0	0,0	0	0,0
PROVIN-	Sin vínculos con el mundo profesional	1	0,3	0	0,0	0	0,0
COIMP-	Diferenciación entre estudiantes	0	0,0	0	0,0	1	4,0
		3,81	100	50	100	25	100

Tecnología: “Elementos que debería mantener”

		Promedios IED							
Códigos	Descripción del código	4,1 a 5	%	3,1 a 4	%	0 a 3	%		
BUEN	Buen profesor	26	30,6	28	21,4	2	9,5		
CLAEX	Claridad para explicar	10	11,8	15	11,5	4	19,0		
CEACT	Aprendizaje activo	7	8,2	4	3,1	1	4,8		
PROMET	Métodos y estrategias apropiadas	6	7,1	3	2,3	0	0		
EXMA	Manejo del contenido	4	4,7	8	6,1	3	14,3		
PROMAT	Material de apoyo apropiado	4	4,7	6	4,6	0	0		
PROVIN	Vínculos con el mundo profesional	4	4,7	5	3,8	0	0		
ACCEDUD	Disponibilidad para responder dudas	3	3,5	6	4,6	0	0		
CEPRE	Centrado en necesidades de estudiantes	3	3,5	1	0,8	3	14,3		
CORESP	Responsabilidad en proceso de aprendizaje	3	3,5	2	1,5	1	4,8		
PRODUD	Responde claramente a las dudas	3	3,5	14	10,7	0	0		
AMBRELA	Buena relación con estudiantes	2	2,4	8	6,1	0	0		
APACAR	Características personales (amable, simpático)	2	2,4	6	4,6	2	9,5		
PROOUT	Orienta el estudio fuera de la clase	2	2,4	4	3,1	0	0		
AMBCONF	Ambiente de confianza	1	1,2	0	0,0	0	0		
CLAPRES	Claridad para presentar	1	1,2	3	2,3	0	0		
ENTMOT	Motiva a los estudiantes	1	1,2	0	0,0	0	0		

PROCONC	Focaliza en conceptos clave	1	1,2	5	3,8	0	0
PROCUM	Cumple con lo planificado	1	1,2	2	1,5	2	9,5
PROPLA	Preparación previa de la enseñanza	1	1,2	1	0,8	0	0
ACCETUT	Disponibilidad para tutorizar	0	0,0	1	0,8	0	0
COIMP	Imparcialidad	0	0,0	1	0,8	0	0
ENTENS	Entusiasmo por lo que enseña	0	0,0	1	0,8	0	0
PROEJE	Plantea ejemplos claros	0	0,0	1	0,8	0	0
PROEVA	Métodos de evaluación válidos	0	0,0	1	0,8	0	0
PROME	Metas y objetivos claros	0	0,0	1	0,8	0	0
PRORESEV	Entrega resultados de evaluación oportunamente	0	0,0	1	0,8	0	0
PROTIC	Uso apropiado de las tecnologías	0	0,0	1	0,8	2	9,5
PROTIME	Uso efectivo del tiempo	0	0,0	1	0,8	1	4,8
RETROEF	Retroalimentación efectiva	0	0,0	1	0,8	0	0
		85	100	131	100	21	100

Tecnología: “Elementos que debería mejorar”

Códigos	Descripción del código	Promedios IED					
		4,1 a 5	%	3,1 a 4	%	0 a 3	%
PROMET-	Métodos y estrategias inapropiadas	12	20,3	27	19,3	2	6,9
PROTIME-	Uso poco efectivo del tiempo	3	5,1	12	8,6	1	3,4
CEPRE-	No se preocupaba por las necesidades del estudiante	0	0,0	12	8,6	0	0
COPREV-	Mal diseño del curso (horario, malla, nivel)	17	28,8	9	6,4	1	3,4
APACAR-	Características personales negativas	0	0,0	8	5,7	5	17,2
CLAEX-	Poca claridad para explicar	3	5,1	7	5,0	0	0
RETROEF-	Retroalimentación no efectiva	2	3,4	7	5,0	2	6,9
PROCRI-	Criterios de evaluación poco claros	1	1,7	6	4,3	1	3,4
PROEVA-	Métodos de evaluación inválidos	4	6,8	5	3,6	2	6,9
CEACT-	No centrado en aprendizaje activo	3	5,1	5	3,6	1	3,4
PROPLA-	No prepara previamente la enseñanza	2	3,4	5	3,6	1	3,4
CORESP-	Irresponsabilidad en el proceso de aprendizaje	0	0,0	5	3,6	5	17,2
PRORESEV-	Resultados de evaluación retrasados	0	0,0	5	3,6	0	0
BUEN-	Mal docente	0	0,0	4	2,9	4	13,8
PROMAT-	Material de apoyo inapropiado	4	6,8	3	2,1	0	0
EXMA-	No maneja el contenido	0	0,0	3	2,1	0	0
ACCEDUD-	Sin disponibilidad para responder dudas	1	1,7	2	1,4	0	0
PROCONC-	No focaliza atención en conceptos clave	1	1,7	2	1,4	0	0
CLAVOZ-	Uso inadecuado de la voz	0	0,0	2	1,4	0	0
PROTIC-	No utiliza o mal utiliza la tecnología	0	0,0	2	1,4	1	3,4

AMBRELA-	Mala relación profesor alumno	2	3,4	1	0,7	1	3,4
PRODUD-	No responde dudas con claridad	1	1,7	1	0,7	2	6,9
CLAPRES-	Poca claridad para presentar	0	0,0	1	0,7	0	0
CODISC-	Discriminación	0	0,0	1	0,7	0	0
COIMP-	Diferenciación entre estudiantes	0	0,0	1	0,7	0	0
CONTEXT-	Problemas de espacio	0	0,0	1	0,7	0	0
ENTENS-	Entusiasmo al enseñar	0	0,0	1	0,7	0	0
PROCUM-	No cumple lo planificado	0	0,0	1	0,7	0	0
PROEJE-	No realiza ejemplos	0	0,0	1	0,7	0	0
PROOUT-	No orienta el estudio fuera de la clase	2	3,4	0	0,0	0	0
PROVIN-	Sin vínculos con el mundo profesional	1	1,7	0	0,0	0	0
		59	100	140	100	29	100

Anexo 13

Encuesta consecuencias de la evaluación de la docencia para docentes

Consecuencias del uso de los resultados de evaluación de la docencia universitaria	
Docentes	
1. Datos personales:	
Nombre	
Apellido	
Unidad Académica	
Asignatura (s) impartidas 1er semestre 2014	
Grado académico más alto	
Experiencia como docente en la USACH	
Edad	
Género	
Condición contractual	
2. Acceso a los informes de resultados: responda SI o No según corresponda a las siguientes preguntas.	

	Si	No
1. ¿Ha revisado la retroalimentación reportada de la evaluación de la docencia, AL MENOS UNA VEZ en su experiencia como docente de la Universidad de Santiago de Chile?		
2. ¿Revisa sus informes de evaluación de la docencia TODOS O CASI TODOS los semestres lectivos?		

3. Marque su grado de acuerdo con la afirmación señalada, desde Muy de acuerdo a Muy en desacuerdo.
En mi calidad de docente de la Universidad de Santiago, he utilizado los resultados de la Evaluación Docente para:

Consecuencias descriptivas	Muy acuerdo	De acuerdo	En desacuerdo	Muy desacuerdo
1. (...) Identificar si me reconocen como un buen o mal docente.				
2. (...) Identificar mi posición relativa como docente entre los académicos de mi unidad académica.				
3. (...) Reconocer el grado de satisfacción de mis estudiantes con mi desempeño como profesor				
4. (...) Reconocer el grado de satisfacción de mis estudiantes con el desarrollo del curso				
Consecuencias prescriptivas				
5. (...) Reflexionar en torno a mis fortalezas y debilidades como docente.				
6. (...) Mejorar la calidad global de mi enseñanza				
7. (...) Modificar la planificación de mi asignatura				
8. (...) Mejorar mi forma de explicar				
9. (...) Mejorar mi forma de evaluar el aprendizaje				
10. (...) Mejorar mi relación interpersonal con los alumnos				
11. (...) Modificar la bibliografía del curso				
12. (...) Actualizar mi conocimiento disciplinar en el área que enseño				
13. (...) Modificar la carga de trabajo asignada a mis estudiantes				
14. (...) Motivarme a tomar cursos o diplomados para mejorar mis habilidades pedagógicas.				
15. (...) Motivarme a buscar apoyo para mejorar mi quehacer docente.				
16. (...) Defender mi promoción a otros cargos o funciones en la Universidad.				
17. (...) Defender mi continuidad en la institución o unidad académica				
18. (...) Postular a incentivos que premian la calidad de la docencia.				

*Aplicación online

Anexo 14

Encuesta consecuencias de la evaluación de la docencia para directivos

Consecuencias del uso de los resultados de encuesta de evaluación de la docencia universitaria
--

jefe de carrera/director de departamento				
1. Datos personales:				
Nombre				
Apellido				
Unidad Académica	Desplegar unidades			
Cargo	Desplegar cargos			
Tiempo como jefe de carrera o director de departamento (en total)	Desplegar rangos de tiempo			
Grado académico más alto	Licenciatura, Magíster, Doctorado, Post doctorado			
Experiencia como docente en la USACH	Desplegar rangos			
Edad	Desplegar rangos			
Género	Masculino/femenino			
Condición contractual	Académico/profesor hora			
2. Acceso a los informes de resultados: responda SI o No según corresponda a las siguientes preguntas.				
	Si	No		
3. ¿Ha revisado los informes de resultados AL MENOS UNA VEZ en su experiencia como jefe de carrera de la Universidad de Santiago de Chile?				
3. Marque su grado de acuerdo con la afirmación señalada, desde Muy de acuerdo a Muy en desacuerdo. "Como jefe de carrera, he utilizado los resultados de <u>la Evaluación Docente</u> para:				
Consecuencias descriptivas	Muy acuerd	De acuerdo	En desacuerdo	Muy desacuerdo
1. (...) Identificar a los mejores y peores docentes del equipo"				
2. (...) Reconocer el grado de satisfacción de los estudiantes con <u>el desempeño de sus docentes</u> "				
3. (...) Reconocer el grado de satisfacción de los estudiantes con <u>el desarrollo del curso</u> "				
Consecuencias prescriptivas				
4. (...) Identificar fortalezas y debilidades del quehacer de los profesores de mi equipo.				
5. (...) Retroalimentar el quehacer docente para orientar mejoras a la enseñanza"				
6. (...) Sugerir cambios a la planificación de las asignaturas".				
7. (...) Sugerir cambios en las metodologías de enseñanza".				
8. (...) Sugerir cambios a la evaluación de aprendizajes"				
9. (...) Establecer propuestas para la actualización curricular de la carrera"				
10. (...) Buscar soluciones a problemas de convivencia entre docentes y estudiantes"				

11. (...) Motivar a los docentes a actualizar sus conocimientos disciplinares”.				
12. (...) Motivar a los docentes a inscribirse en cursos de formación pedagógica”.				
13. (...) Sugerir la promoción de algún docente a otros cargos o funciones en la Universidad”.				
14. (...) Sugerir la continuidad de un docente en la institución o unidad académica”.				
15. (...) Orientar la planeación docente (asignación de asignaturas a académicos, organización de horarios, etc.)”.				
16. (...) Sugerir el despido o no renovación de contrato de algún docente”.				
17. (...) Entregar premios a la enseñanza de calidad”.				
18. (...) Reportar la calidad de la docencia de la unidad académica para fines de rendición de cuentas (acreditación, proyectos)”.				
19. (...) Reportar antecedentes a la Evaluación del Desempeño Académico”.				

**aplicación online*

Anexo 15

Encuesta sometida a juicio de experto

Como DOCENTE, he utilizado los resultados de la Evaluación Docente para:		Experto 1				Experto 2				Experto 3				Experto 4				%acuerdo
Criterios de evaluación		1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	
1.	(...) Identificar si me reconocen como un buen o mal docente.	1	1	1	1	2	1	1	1	1	1	1	2	1	1	1	1	87,5
2.	(...) Identificar mi posición relativa como docente entre los académicos de mi unidad académica.	1	1	1	1	1	1	1	1	2	1	1	2	1	1	1	1	87,5
3.	(...) Reconocer el grado de satisfacción de mis estudiantes con <u>mi desempeño como profesor</u> .	1	1	1	1	1	1	1	1	1	1	1	2	1	1	2	1	87,5
4.	(...) Reconocer el grado de satisfacción de mis estudiantes con <u>el desarrollo del curso</u> .	1	1	1	1	1	1	1	1	1	1	1	2	1	1	1	1	93,75
5.	(...) Reflexionar en torno a mis fortalezas y debilidades como docente.	1	1	1	1	1	1	1	1	2	1	1	2	1	1	1	1	87,5
6.	(...) Mejorar la calidad global de mi enseñanza.	1	1	1	1	1	1	1	1	1	1	1	2	1	1	1	1	93,75
7.	(...) Modificar la planificación de mi asignatura.	1	1	1	1	1	1	1	1	1	1	1	2	1	1	1	1	93,75
8.	(...) Mejorar mi forma de explicar.	1	1	1	1	1	1	1	1	2	1	1	2	2	1	1	1	81,25
9.	(...) Mejorar mi forma de evaluar el aprendizaje.	1	1	1	1	1	1	1	1	1	1	1	2	1	1	1	1	93,75
10.	(...) Mejorar mi relación interpersonal con los alumnos.	1	1	1	1	1	1	1	1	2	1	1	2	1	1	1	1	87,5
11.	(...) Modificar la bibliografía del curso.	1	1	1	1	1	1	1	1	1	1	1	2	1	1	1	1	93,75
12.	(...) Actualizar mi conocimiento disciplinar en el área que enseño.	1	1	1	1	1	1	1	1	2	1	1	2	1	1	1	1	87,5
13.	(...) Modificar la carga de trabajo asignada a mis estudiantes.	1	1	1	1	1	1	1	1	2	1	1	2	1	1	1	1	87,5
14.	(...) Motivarme a tomar cursos o diplomados para mejorar mis habilidades pedagógicas.	2	1	1	1	1	1	1	1	1	2	1	2	1	1	1	1	81,25
15.	(...) Motivarme a buscar apoyo para mejorar mi quehacer docente.	2	1	1	1	1	1	1	1	1	1	1	2	1	1	1	1	87,5
16.	(...) Defender mi promoción a otros cargos o funciones en la Universidad.	1	1	1	1	1	1	1	1	2	1	1	2	1	1	1	1	87,5
17.	(...) Defender mi continuidad en la institución o unidad académica.	1	1	1	1	1	1	1	1	2	1	1	2	1	1	1	1	87,5
18.	(...) Postular a incentivos que premian la calidad de la docencia.	1	1	1	1	1	1	1	1	1	1	1	2	1	1	1	1	93,75

*Puntuación 1= sí 2=no *Criterios de evaluación: 1. Claridad en redacción. 2. Libre de memoria. 3. Libre de sesgos, 4. Adecuación a categoría de respuesta

Anexo 16

*Análisis factorial encuesta consecuencias de la evaluación de la docencia para docentes
(programa FACTOR)*

Adequacy of the correlation matrix

<i>Determinant of the matrix</i>	<i>= 0.000000558574069</i>
<i>Bartlett's statistic</i>	<i>= 2459.6 (df = 171; P = 0.000010)</i>
<i>Kaiser-Meyer-Olkin (KMO) test</i>	<i>= 0.92264 (very good)</i>

Explained variance based on eigenvalues

Variable	Eigenvalue	Proportion of variance	Cumulative proportion of variance
<i>1</i>	10.12704	0.53300	0.53300
<i>2</i>	3.32014	0.17474	0.70775
<i>3</i>	0.99858	0.05256	0.76030
<i>4</i>	0.72247	0.03802	
<i>5</i>	0.58245	0.03066	
<i>6</i>	0.49068	0.02583	
<i>7</i>	0.45853	0.02413	
<i>8</i>	0.39442	0.02076	
<i>9</i>	0.30823	0.01622	
<i>10</i>	0.28337	0.01491	
<i>11</i>	0.22849	0.01203	
<i>12</i>	0.22681	0.01194	
<i>13</i>	0.19111	0.01006	
<i>14</i>	0.17305	0.00911	
<i>15</i>	0.14054	0.00740	
<i>16</i>	0.12618	0.00664	
<i>17</i>	0.10920	0.00575	
<i>18</i>	0.07153	0.00376	
<i>19</i>	0.04718	0.00248	

*Parallel analysis (pa) based on minimum rank factor analysis
(Timmerman & Lorenzo-Seva, 2011)*

Implementation details:

Correlation matrices analyzed: Polychoric correlation matrices

Number of random correlation matrices: 500

Method to obtain random correlation matrices: Permutation of the raw data (Buja & Eyuboglu, 1992)

<i>Variable</i>	<i>Real – data % variance</i>	<i>Mean of random % variance</i>	<i>95 percentile of random % variance</i>
<i>1</i>	<i>55.1*</i>	<i>11.6</i>	<i>13.1</i>
<i>2</i>	<i>18.0*</i>	<i>10.2</i>	<i>11.2</i>
<i>3</i>	<i>5.3</i>	<i>9.3</i>	<i>10.1</i>
<i>4</i>	<i>3.9</i>	<i>8.6</i>	<i>9.2</i>
<i>5</i>	<i>3.0</i>	<i>7.9</i>	<i>8.5</i>
<i>6</i>	<i>2.7</i>	<i>7.3</i>	<i>7.8</i>
<i>7</i>	<i>2.5</i>	<i>6.7</i>	<i>7.2</i>
<i>8</i>	<i>1.7</i>	<i>6.1</i>	<i>6.6</i>
<i>9</i>	<i>1.5</i>	<i>5.6</i>	<i>6.1</i>
<i>10</i>	<i>1.3</i>	<i>5.1</i>	<i>5.6</i>
<i>11</i>	<i>1.2</i>	<i>4.6</i>	<i>5.0</i>
<i>12</i>	<i>1.0</i>	<i>4.0</i>	<i>4.5</i>
<i>13</i>	<i>0.8</i>	<i>3.5</i>	<i>4.0</i>
<i>14</i>	<i>0.6</i>	<i>3.0</i>	<i>3.5</i>
<i>15</i>	<i>0.6</i>	<i>2.4</i>	<i>3.0</i>
<i>16</i>	<i>0.5</i>	<i>1.9</i>	<i>2.5</i>
<i>17</i>	<i>0.2</i>	<i>1.3</i>	<i>1.9</i>
<i>18</i>	<i>0.1</i>	<i>0.8</i>	<i>1.3</i>
<i>19</i>	<i>0.0</i>	<i>0.0</i>	<i>0.0</i>

** Advised number of dimensions: 2*

Anexo 17

Análisis factorial encuesta consecuencias de la evaluación de la docencia para docentes (programa SPSS)

KMO y prueba de Bartlett		
Medida de adecuación muestral de Kaiser-Meyer-Olkin.		,923
Prueba de esfericidad de Chi-cuadrado aproximado		2459,638
Bartlett	Gl	171
	Sig.	,000

Factor	Autovalores iniciales			Varianza total explicada					
				Sumas de las saturaciones al cuadrado de la extracción			Suma de las saturaciones al cuadrado de la rotación		
	Total	% de la varianza	% acumulado	Total	% de la varianza	% acumulado	Total	% de la varianza	% acumulado
1	8,988	47,303	47,303	8,635	45,446	45,446	4,706	24,767	24,767
2	2,862	15,065	62,368	2,520	13,264	58,710	4,216	22,188	46,954
3	1,200	6,318	68,686	,857	4,512	63,222	3,091	16,268	63,222
4	,798	4,198	72,884						
5	,700	3,682	76,566						
6	,594	3,125	79,691						
7	,500	2,630	82,321						
8	,473	2,491	84,812						
9	,380	2,001	86,813						
10	,378	1,991	88,804						
11	,334	1,758	90,563						
12	,315	1,656	92,219						
13	,276	1,451	93,669						
14	,254	1,336	95,005						
15	,238	1,255	96,260						
16	,229	1,208	97,468						
17	,189	,993	98,461						
18	,162	,851	99,312						
19	,131	,688	100,000						

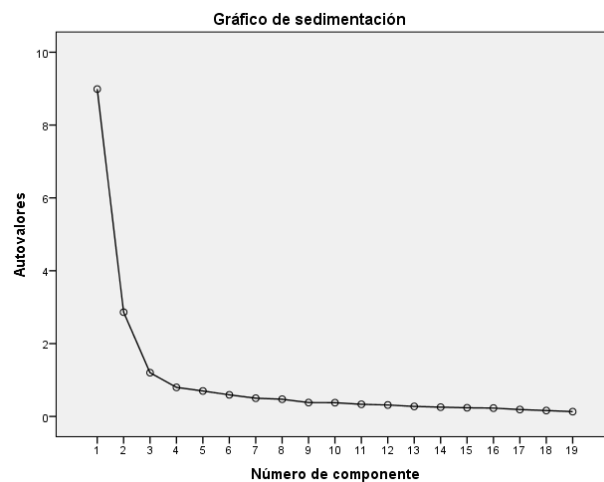
Método de extracción: Mínimos cuadrados no ponderados.

	Matriz de factores rotados ^a		
	Factor		
	1	2	3
p01	0,833	0,135	0,034
p02	0,317	0,323	0,293
p03	0,785	0,356	0,062
p04	0,665	0,448	0,073
p05	0,85	0,2	0,038
p06	0,728	0,146	0,094
p07	0,374	0,663	0,237
p08	0,729	0,404	0,122
p09	0,225	0,619	0,303
p10	0,482	0,675	0,148
p11	0,397	0,764	0,173
p12	0,546	0,427	0,146
p13	0,241	0,696	0,297
p14	0,316	0,675	0,283
p15	0,091	0,494	0,596
p16	0,085	0,481	0,612
p17	0,043	0,138	0,83
p18	0,124	0,225	0,797
p19	0,024	0,105	0,726

Método de extracción: Mínimos cuadrados no ponderados.

Método de rotación: Normalización Varimax con Kaiser.

a. La rotación ha convergido en 6 iteraciones.



Anexo 18

Relación entre dimensiones teóricas y empíricas de encuesta sobre consecuencias sobre evaluación de la docencia

Dimensiones teóricas	Dimensiones empíricas	Nº	ítems
Descriptivas	Valoración sobre mi desempeño	1	Reconocer el grado de satisfacción de mis estudiantes con mi desempeño como profesor.
Formativas	Valoración sobre mi desempeño	3	Reflexionar en torno a mis fortalezas como docente.
Formativas	Valoración sobre mi desempeño	4	Mejorar la calidad global de mi enseñanza.
descriptivas	Valoración sobre mi desempeño	5	Reconocer el grado de satisfacción de mis estudiantes con el desarrollo del curso.
descriptivas	Valoración sobre mi desempeño	6	Identificar si mis estudiantes me reconocen como un buen o mal docente.
Formativas	Valoración sobre mi desempeño	8	Reflexionar en torno a mis debilidades como docente.
formativas	Valoración sobre mi desempeño	12	Revisar mi relación con los alumnos.
descriptivas	Valoración sobre mi desempeño	2	Identificar mi posición relativa en cuanto a la calidad docente de los académicos de mi unidad académica.
Formativas	Cambios a mis prácticas	7	Modificar la planificación de mi asignatura.
formativas	Cambios a mis prácticas	9	Modificar la carga académica (tiempo de dedicación) asignada a mis estudiantes.
formativas	Cambios a mis prácticas	10	Mejorar mi forma de explicar los contenidos o la materia del curso.
formativas	Cambios a mis prácticas	11	Mejorar mi forma de evaluar los aprendizajes de mis estudiantes.
formativas	Cambios a mis prácticas	13	Modificar la bibliografía del curso.
formativas	Cambios a mis prácticas	14	Actualizar mi conocimiento sobre la disciplina que enseño.
formativas	Sumativas	15	Decidirme a tomar cursos o programas de formación en docencia.
formativas	Sumativas	16	Decidirme a buscar apoyo en mis colegas o especialistas en docencia para mejorar mi quehacer docente.
sumativas	Sumativas	17	Fundamentar mi promoción a otros cargos o funciones en la Universidad.

sumativas	Sumativas	18	Fundamentar mi continuidad en la institución o unidad académica.
sumativas	Sumativas	19	Postular a incentivos que premian la calidad de la docencia.

Anexo 19

Resultados consecuencias de evaluación de la docencia para docentes

Dimensiones empíricas	N°	Ítems	Siempre	Casi siempre	A veces	Nunca
Valoración sobre mi desempeño	1	Reconocer el grado de satisfacción de mis estudiantes con mi desempeño como profesor.	71,58% (131)	19,67% (36)	5,46% (10)	3,28% (6)
Valoración sobre mi desempeño	3	Reflexionar en torno a mis fortalezas como docente.	71,58% (131)	19,13% (35)	5,46% (10)	3,83% (7)
Valoración sobre mi desempeño	4	Mejorar la calidad global de mi enseñanza.	71,04% (130)	16,94% (31)	7,1% (13)	4,92% (9)
Valoración sobre mi desempeño	5	Reconocer el grado de satisfacción de mis estudiantes con el desarrollo del curso.	68,85% (126)	24,04% (44)	4,37% (8)	2,73% (5)
Valoración sobre mi desempeño	6	Identificar si mis estudiantes me reconocen como un buen o mal docente.	59,02% (108)	27,32% (50)	7,1% (13)	6,56% (12)
Valoración sobre mi desempeño	8	Reflexionar en torno a mis debilidades como docente.	67,21% (123)	20,22% (37)	7,65% (14)	4,92% (9)
Valoración sobre mi desempeño	12	Revisar mi relación con los alumnos.	52,46% (96)	27,32% (50)	14,75% (27)	5,46% (10)
	2	Identificar mi posición relativa en cuanto a la calidad docente de los académicos de mi unidad académica.	42,62% (78)	21,86% (40)	16,94% (31)	18,58% (34)
Cambios a mis prácticas	7	Modificar la planificación de mi asignatura.	38,8% (71)	32,24% (59)	21,31% (39)	7,65% (14)
Cambios a mis prácticas	9	Modificar la carga académica (tiempo de dedicación) asignada a mis estudiantes.	25,14% (46)	28,96% (53)	26,78% (49)	19,13% (35)
Cambios a mis prácticas	10	Mejorar mi forma de explicar los contenidos o la materia del curso.	56,83% (104)	24,04% (44)	12,57% (23)	6,56% (12)
Cambios a mis prácticas	11	Mejorar mi forma de evaluar los aprendizajes de mis estudiantes.	47,54% (87)	27,32% (50)	16,94% (31)	8,2% (15)
Cambios a mis prácticas	13	Modificar la bibliografía del curso.	26,78% (49)	34,97% (64)	25,14% (46)	13,11% (24)
Cambios a mis prácticas	14	Actualizar mi conocimiento sobre la disciplina que enseño.	46,99% (86)	23,5% (43)	14,21% (26)	15,3% (28)
Sumativas	15	Decidirme a tomar cursos o programas de formación en docencia.	22,53% (41)	20,88% (38)	25,27% (46)	31,32% (57)

Sumativas	16	Decidirme a buscar apoyo en mis colegas o especialistas en docencia para mejorar mi quehacer docente.	12,57% (23)	25,14% (46)	29,51% (54)	32,79% (60)
Sumativas	17	Fundamentar mi promoción a otros cargos o funciones en la Universidad.	10,99% (20)	10,99% (20)	19,78% (36)	58,24% (106)
Sumativas	18	Fundamentar mi continuidad en la institución o unidad académica.	25,27% (46)	13,19% (24)	17,58% (32)	43,96% (80)
Sumativas	19	Postular a incentivos que premian la calidad de la docencia.	8,79% (16)	12,09% (22)	21,98% (40)	57,14% (104)

Anexo 20

Análisis de fiabilidad encuesta a consecuencias para directivos (programa SPSS)

Estadísticos de fiabilidad	
Alfa de Cronbach	N de elementos
,919	20

Estadísticos total-elemento				
	Media de la escala si se elimina el elemento	Varianza de la escala si se elimina el elemento	Correlación elemento-total corregida	Alfa de Cronbach si se elimina el elemento
P1	49,38	147,958	,450	,918
p2	49,69	140,936	,628	,914
p3	49,83	142,148	,734	,912
p4	49,55	142,828	,630	,914
p5	50,07	140,852	,622	,914
p6	50,34	141,091	,615	,914
p7	50,52	139,973	,668	,913
p8	49,28	146,778	,532	,916
p9	50,14	146,766	,471	,917
p10	50,00	143,500	,464	,918
p11	49,90	145,739	,558	,916
p12	50,34	141,305	,629	,914
p13	50,45	145,042	,480	,917
p14	50,83	147,148	,426	,918
p15	50,14	138,195	,683	,912

p16	49,62	137,601	,805	,910
p17	50,62	141,744	,619	,914
p18	50,66	145,091	,412	,919
p19	49,83	139,862	,594	,915
p20	49,93	139,281	,582	,915

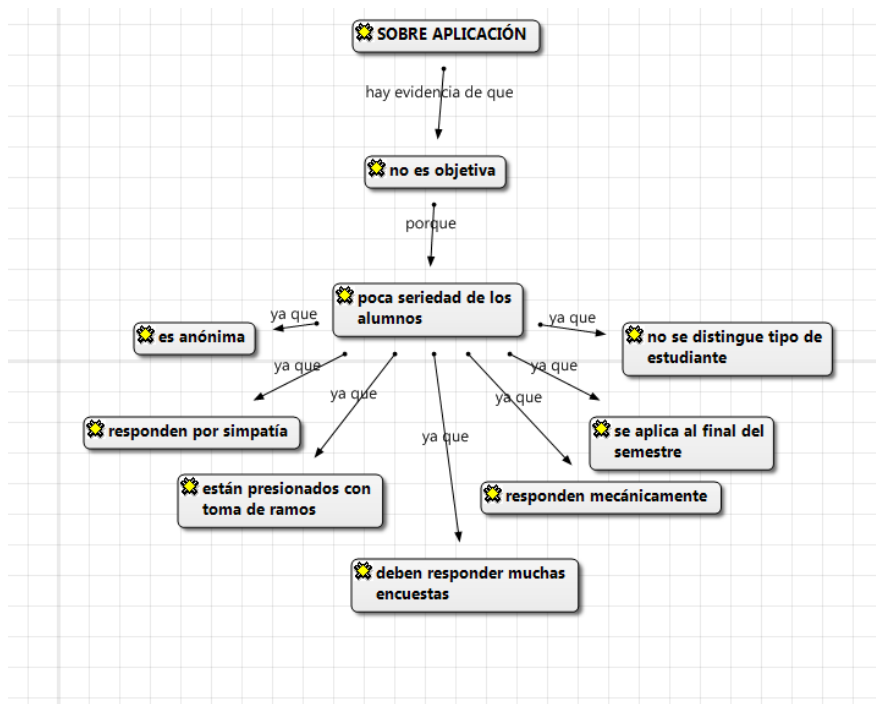
Anexo 21

Resultados consecuencias de evaluación de la docencia para directivos

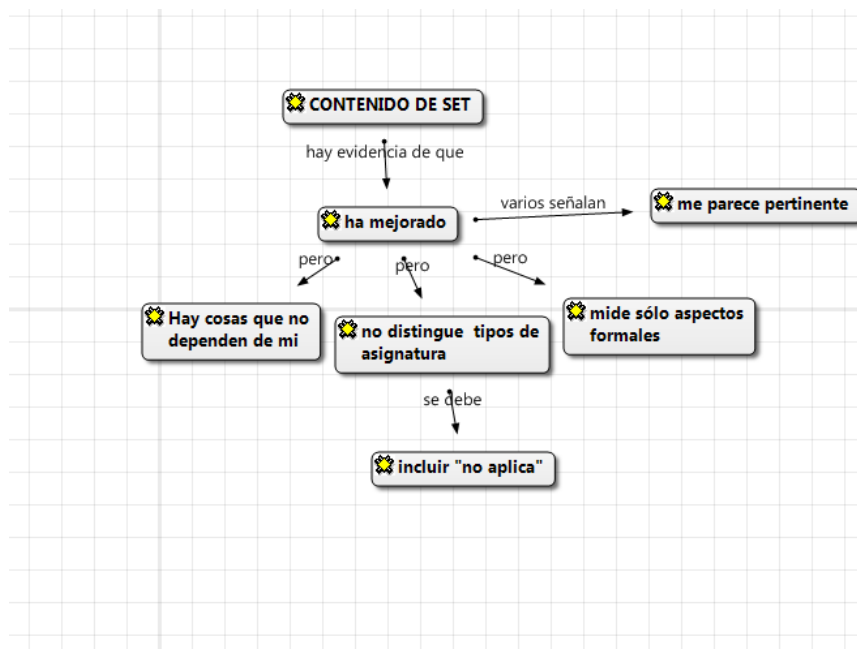
Dimensiones teóricas	N°	Ítems	Siempre	Casi siempre	A veces	Nunca
Descriptivas	1	Reconocer el grado de satisfacción de los estudiantes con el desempeño de sus docentes.	51,72% (15)	27,59% (8)	20,69% (6)	0% (0)
Descriptivas	4	Identificar a los docentes con mejores y peores habilidades pedagógicas de mi unidad académica.	44,83% (13)	27,59% (8)	24,14% (7)	3,45% (1)
Descriptivas	8	Reconocer el grado de satisfacción de los estudiantes con la implementación de los cursos.	58,62% (17)	24,14% (7)	17,24% (5)	0% (0)
Formativas	9	Identificar las debilidades pedagógicas de los docentes de mi equipo.	17,24% (5)	27,59% (8)	48,28% (14)	6,9% (2)
Formativas	3	Identificar las fortalezas pedagógicas de los docentes de mi equipo.	20,69% (6)	51,72% (15)	20,69% (6)	6,9% (2)
Formativas	2	Orientar mejoras al desempeño de los docentes de mi equipo, a partir de la retroalimentación de sus prácticas.	41,38% (12)	27,59% (8)	20,69% (6)	10,34% (3)
formativas	5	Sugerir cambios a la planificación de las asignaturas a los docentes a cargo.	24,14% (7)	31,03% (9)	27,59% (8)	17,24% (5)
formativas	6	Sugerir cambios a los docentes respecto de las evaluaciones de aprendizaje en sus cursos.	17,24% (5)	24,14% (7)	34,48% (10)	24,14% (7)
formativas	7	Sugerir cambios en las metodologías de enseñanza a los docentes a cargo de las asignaturas.	13,79% (4)	20,69% (6)	34,48% (10)	31,03% (9)
formativas	10	Contar con insumos para la actualización curricular de la carrera (organización de asignaturas, carga académica, etc.).	31,03% (9)	27,59% (8)	20,69% (6)	20,69% (6)
formativas	11	Identificar problemas de convivencia entre docentes y estudiantes.	20,69% (6)	41,38% (12)	34,48% (10)	3,45% (1)

formativas	12	Motivar a los docentes a actualizar sus conocimientos disciplinares.	13,79% (4)	31,03% (9)	31,03% (9)	24,14% (7)
formativas	13	Motivar a los docentes a inscribirse en cursos o programas de formación pedagógica.	10,34% (3)	31,03% (9)	31,03% (9)	27,59% (8)
sumativas	14	Sugerir la promoción de algún docente a otros cargos o funciones en la Universidad.	3,45% (1)	24,14% (7)	27,59% (8)	44,83% (13)
sumativas	15	Sugerir la continuidad de un docente en la institución o unidad académica.	24,14% (7)	31,03% (9)	20,69% (6)	24,14% (7)
sumativas	16	Orientar la planeación docente (asignación de asignaturas a académicos, organización de horarios, etc.).	41,38% (12)	34,48% (10)	13,79% (4)	10,34% (3)
sumativas	17	Sugerir el despido o no renovación de contrato de algún docente.	10,34% (3)	20,69% (6)	34,48% (10)	34,48% (10)
sumativas	18	Entregar premios, reconocimientos o recompensas a la enseñanza de calidad.	13,79% (4)	20,69% (6)	20,69% (6)	44,83% (13)
sumativas	19	Reportar la calidad de la docencia de la unidad académica para fines de rendición de cuentas (acreditación, proyectos).	41,38% (12)	20,69% (6)	20,69% (6)	17,24% (5)
sumativas	20	Reportar antecedentes a la Evaluación del Desempeño Académico.	37,93% (11)	24,14% (7)	13,79% (4)	24,14% (7)

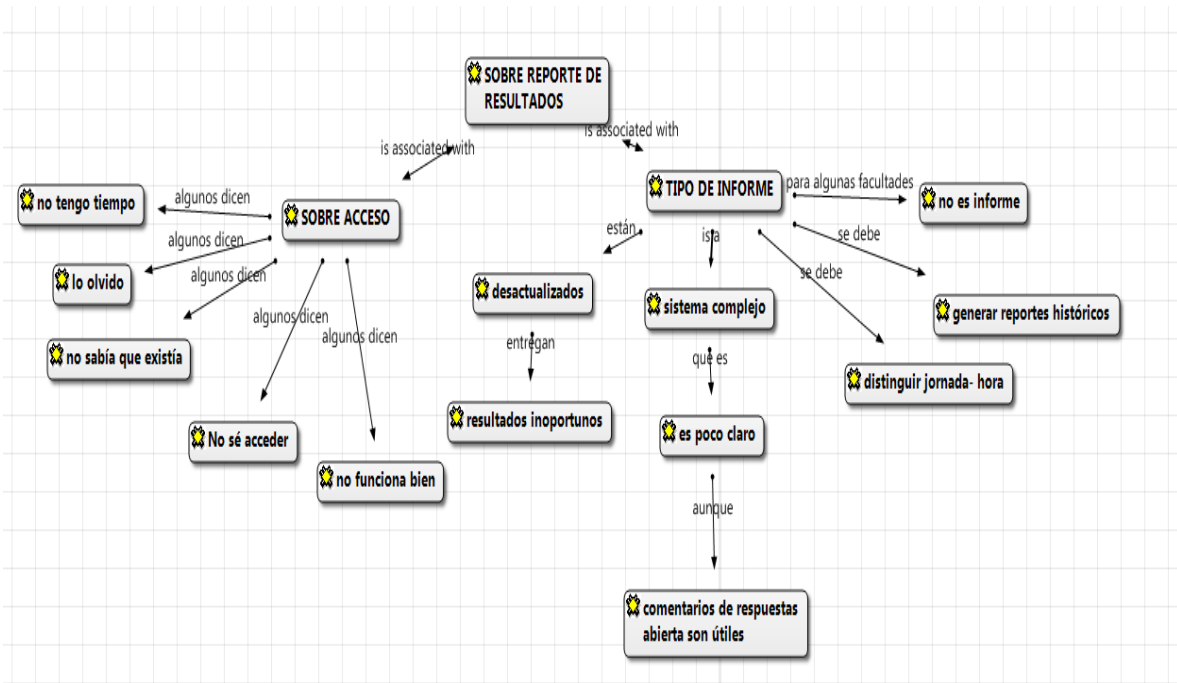
Sobre aplicación del instrumento



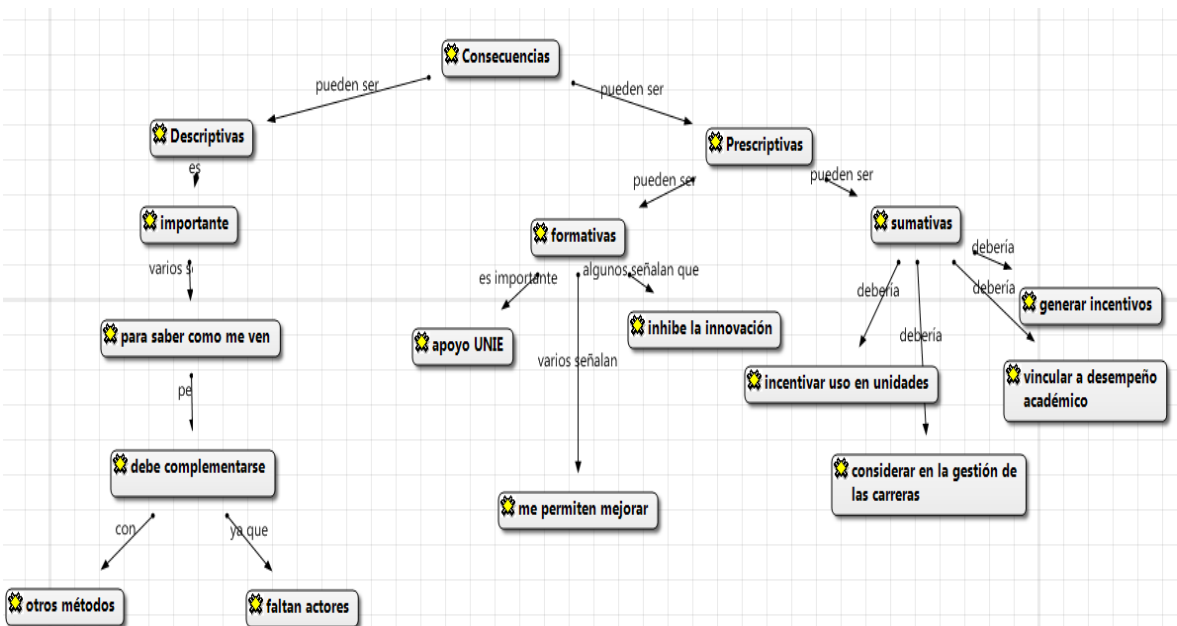
Sobre contenido del instrumento de la evaluación de la docencia



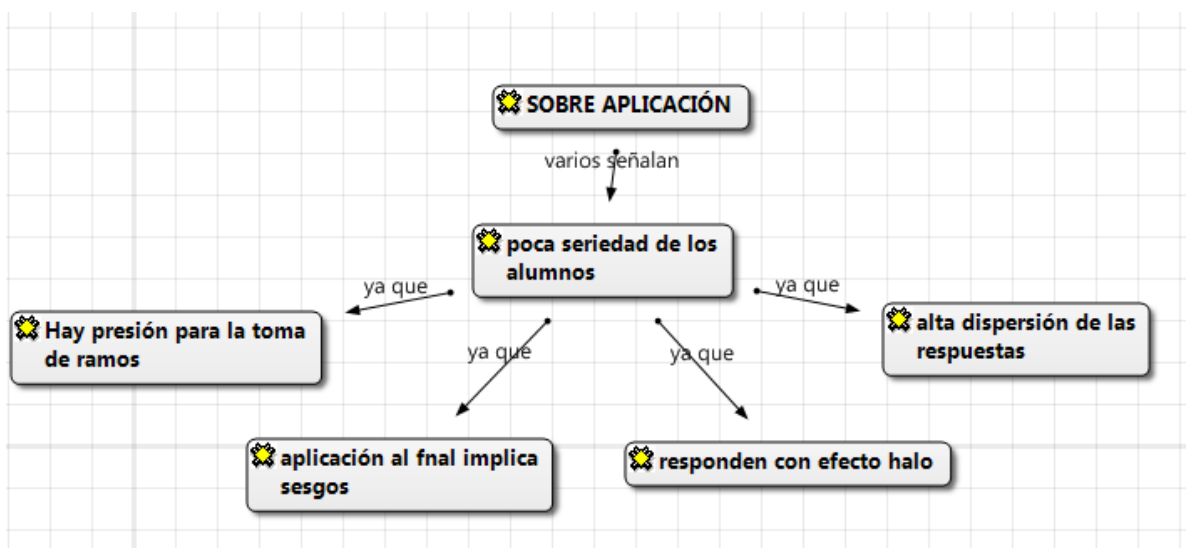
Sobre el reporte de los resultados



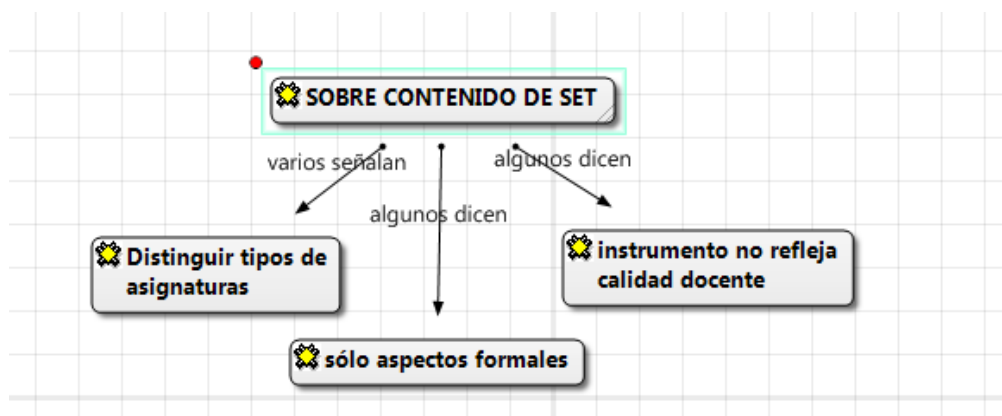
Sobre consecuencias de la evaluación de la docencia



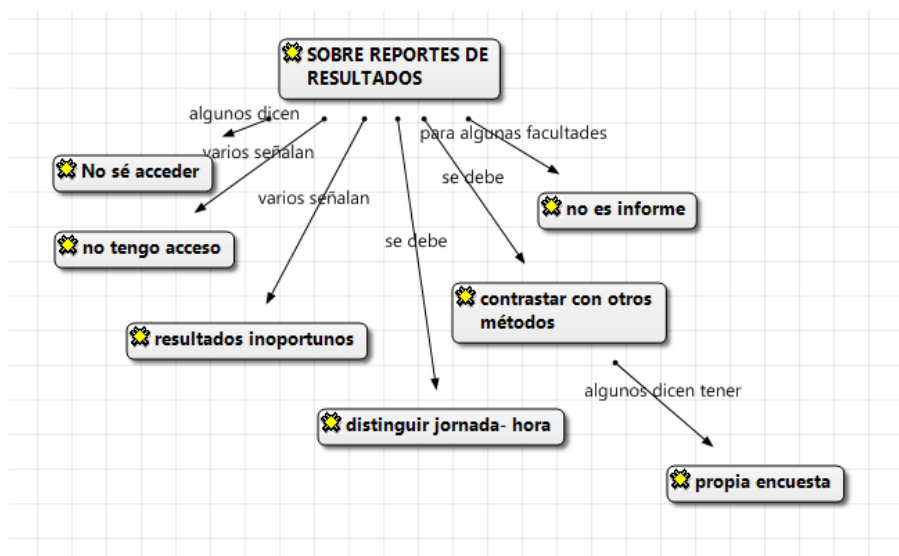
Sobre aplicación del instrumento de evaluación de la docencia



Sobre el contenido del instrumento de evaluación de la docencia



Sobre reporte de resultados de evaluación de la docencia



Sobre consecuencias del instrumento de evaluación de la docencia

