



Revista de estudios y experiencias en educación

ISSN: 0717-6945

ISSN: 0718-5162

Universidad Católica de la Santísima Concepción. Facultad de Educación

Núñez Vives, Mónica; González Bravo, Luis; Zúñiga Peña, José; Melipillán Araneda, Edmundo; Aguayo Torres, Iris; Varas Contreras, Marcela
Propiedades psicométricas de una encuesta genérica de evaluación de la docencia de Magister en una universidad tradicional
Revista de estudios y experiencias en educación, vol. 21, núm. 47, 2022, pp. 254-271
Universidad Católica de la Santísima Concepción. Facultad de Educación

DOI: <https://doi.org/10.21703/0718-5162202202102147014>

Disponible en: <https://www.redalyc.org/articulo.oa?id=243173717015>

- Cómo citar el artículo
- Número completo
- Más información del artículo
- Página de la revista en redalyc.org

UDEM  redalyc.org

Sistema de Información Científica Redalyc
Red de Revistas Científicas de América Latina y el Caribe, España y Portugal
Proyecto académico sin fines de lucro, desarrollado bajo la iniciativa de acceso abierto

Revista de Estudios y Experiencias en Educación

REXE

journal homepage: <http://revistas.ucsc.cl/index.php/rexe>

Propiedades psicométricas de una encuesta genérica de evaluación de la docencia de Magister en una universidad tradicional

Mónica Núñez Vives^a, Luis González Bravo^b, José Zúñiga Peña^c, Edmundo Melipillán Araneda^d, Iris Aguayo Torres^e y Marcela Varas Contreras^f
Universidad de Concepción, Concepción, Chile

Recibido: 09 de junio 2021 - Revisado: 21 de enero 2022 - Aceptado: 07 de marzo 2022

RESUMEN

Este estudio detalla el proceso de validación de una encuesta de evaluación docente para programas de Magister, por a) un equipo multidisciplinario de seis jueces expertos de una unidad de aseguramiento de la calidad de una Universidad, en un trabajo colaborativo con una Dirección de Postgrado, y b) muestras de 272 y 275 estudiantes para dos secciones separadas del cuestionario. Se evaluó la relevancia, contribución e inteligibilidad de los ítems, obteniendo un coeficiente Fleiss Kappa de 0,24 ($z = 6,32$, p -valor $< 0,001$). Los análisis factoriales exploratorios y confirmatorios llevaron a una propuesta integrada de 23 ítems con dos subescalas. Estos resultados se discuten en términos psicométricos, y desde la perspectiva de los equipos de acreditación y las oportunidades que esto implica para las Instituciones.

Palabras clave: Encuesta de evaluación de la docencia; calidad de la docencia; aseguramiento de la calidad.

*Correspondencia: Mónica Núñez Vives (M. Núñez).

^a  <https://orcid.org/0000-0003-1128-3328> (mnunez@udec.cl).

^b  <https://orcid.org/0000-0001-7310-3762> (lgonzalez@udec.cl).

^c  <https://orcid.org/0000-0002-0968-1047> (jozuniga@udec.cl).

^d  <https://orcid.org/0000-0003-3360-433X> (rmelipillan@udd.cl).

^e  <https://orcid.org/0000-0002-9157-539X> (vaguayoster@gmail.com).

^f  <https://orcid.org/0000-0003-2721-9865> (mvaras@udec.cl).

Psychometric properties of a general survey for evaluating Master's degree teaching in a traditional university

ABSTRACT

This study describes the validation process for a Master's teaching evaluation survey conducted by a multidisciplinary team of six expert judges from a university assurance quality unit in collaboration with a postgraduate directorate, as well as samples from 272 and 275 students for two separate questionnaire sections. Relevance, contribution, and intelligibility of the items were assessed, resulting in a Fleiss Kappa coefficient of 0.24 ($z = 6.32$, $p\text{-value} < 0.001$). Exploratory and confirmatory factor analyses led to a 23-item integrated proposal with two subscales. These results are discussed in psychometric terms, and from the perspective of the accreditation teams, as well as the opportunities that this implies for institutions.

Keywords: Teaching evaluation survey; teaching quality; quality assurance.

1. Introducción

Las Universidades utilizan diversos métodos para evaluar el desempeño de los académicos, siendo el predominante las encuestas de evaluación de la docencia que los estudiantes completan al final de cada período académico (Eckhaus y Davidovitch, 2019a). Habitualmente dichos instrumentos –que se espera sean breves, pero bien elaborados (Spooren et al., 2014)– contienen ítems en forma de declaraciones, en torno a las cuales se pide a los estudiantes que indiquen su nivel de acuerdo en formato de respuesta tipo Likert (Huybers, 2014).

En la actualidad, estas encuestas se utilizan para medición del desempeño docente en casi todas las Universidades del mundo, a partir de su aparente facilidad para recopilar datos y presentar e interpretar resultados (Centoni y Maruotti, 2021; Spooren et al., 2014). La evaluación de la enseñanza por parte de los estudiantes (SET) está muy extendida y busca obtener retroalimentación percibida sobre el desempeño en una variedad de aspectos relacionados con el maestro y / o el curso (Huybers, 2014).

1.1 Historia

Las encuestas para la evaluación de la docencia fueron introducidas por primera vez en los Estados Unidos a mediados de la década de 1920 con el objetivo de proporcionar retroalimentación a los académicos sobre su enseñanza. Al día de hoy, las encuestas de evaluación de los estudiantes se realizan de forma rutinaria en universidades de muchas partes del mundo y sus resultados se utilizan tanto para fines formativos (Alderman et al., 2012; Ray et al., 2018), como de evaluación del desempeño docente.

El uso de estas encuestas ha aumentado desde la década de 1970 y con el tiempo han llegado a ser una herramienta con implicaciones de alto valor para la carrera académica, y por ende en los salarios, ascensos y desarrollo profesional. Estas implicancias han transformado las encuestas en objeto de críticas e investigación por parte de mundo académico y de gestión (Eckhaus y Davidovitch, 2019b), que se han traducido en crecientes contribuciones y desarrollos con respecto a la naturaleza y utilidad de estos instrumentos, por ejemplo, multidimensionalidad, validez, y confiabilidad (Alderman et al., 2012; Ray et al., 2018).

1.2 Importancia de la evaluación

Evaluar el desempeño docente es fundamental para mejorar la calidad educativa, en tanto vía para la mejora continua de la enseñanza-aprendizaje. Fundamentalmente focalizadas en las competencias profesionales de los docentes, tradicionalmente se han referido a la preparación de la enseñanza y a la organización del trabajo en el aula y la evaluación de los aprendizajes y logros de los estudiantes (Correa y Rueda Beltrán, 2012).

En general, según indica Marsh (2007), se han utilizado fundamentalmente para cinco propósitos: 1) retroalimentación para los académicos acerca de la efectividad de su enseñanza; 2) medición de la eficacia de la enseñanza para decisiones de carrera académica; 3) información para los estudiantes para elegir secciones o profesores; 4) uso en procesos de aseguramiento de la calidad y 5) investigación.

A pesar del valor atribuido a las encuestas para el desarrollo del currículo y la garantía de calidad, no existe aún suficiente evidencia de que los hallazgos de los estudios se estén utilizando para cambiar o mejorar la experiencia de aprendizaje de los estudiantes (Alderman et al., 2012; Ray et al., 2018).

1.3 Aspectos Psicométricos

Los investigadores recomiendan que las encuestas sobre la calidad de la enseñanza en educación superior se sometan a pruebas de validez (Chamoy, 2018; Tucker et al., 2013); no obstante ello, es relevante señalar que los estudiantes valoran aspectos de la docencia diferentes a los académicos (Tucker et al., 2013).

En particular, los estudiantes valoran las siguientes habilidades y conductas en los profesores: habilidades interpersonales (como la capacidad de motivar, la apertura y la escucha, la accesibilidad, la pasión y el entusiasmo); dominio de la materia presentado de forma clara y lógica; teoría y práctica combinadas; un enfoque justo, ético y comprensivo para los estudiantes; empatía; y, en menor medida, los "aspectos técnicos" de la enseñanza, como la organización. En contraste, los académicos valoran más las siguientes habilidades de enseñanza en otros maestros: habilidades de comunicación, experiencia y habilidades de presentación (Tucker et al., 2013).

Ejemplos de instrumentos para evaluación de la docencia por parte de estudiantes, con características psicométricas apropiadas reportadas en la literatura internacional, son las Evaluaciones de los estudiantes sobre la calidad educativa, SEEQ (Marsh, 1982), el Cuestionario de experiencia del curso o CEQ (Ramsden, 1991), el Cuestionario de experiencia del curso del estudiante o SCEQ (Ginns et al., 2007), el SET37 (Mortelmans y Spooren, 2009), el Cuestionario del curso para maestros ejemplares o ETCQ (Kember y Leung, 2008) y la Lista de verificación de comportamiento (Keeley et al., 2010). Todo ellos, han sido probados en términos de validez y confiabilidad, y usan en la mayoría de los casos escalas tipo Likert (Matosas-López y García Sánchez, 2019).

1.4 Evaluación de la Docencia en Chile

Con la creación de la Comisión Nacional de Acreditación (CNA), la evaluación docente es reforzada y estimulada en las instituciones de educación superior (IES), promovándose como una práctica sistemática y centrada principalmente en el uso de encuestas de opinión a estudiantes. Sin embargo, a menudo hay una inexistencia de políticas efectivas y modelos de evaluación consensuados entre las instituciones (Lamarra y Coppola, 2009; Picco et al., 2013). En Chile, Argentina y México, la evaluación de la docencia universitaria ha sido promovida en tanto requerimiento de la evaluación institucional y acreditación (Loor Ávila et

al., 2018) y en los últimos años se han ido instalando diversos mecanismos de evaluación y control de calidad, dentro de los que destaca la implementación de aquellos vinculados a la calidad de la docencia impartida (Mendoza, 2020).

En Chile, para muchas IES el sistema de evaluación se reduce, en la práctica, a la aplicación de cuestionarios de opinión estudiantil, advirtiéndose, a nivel general, una falta de desarrollo disciplinario de la evaluación de la docencia y de profesionales dedicados al área (Montoya Vargas et al., 2016).

En Chile, México y España, ha predominado el uso de encuestas de opinión cuantitativas para la evaluación del desempeño docente (Lamarra y Coppola, 2009; Montoya Vargas et al., 2016) y en investigaciones realizadas en Argentina, Chile, Colombia, España y Venezuela, se constata que existe una diversidad de mecanismos estatales y locales para llevar a cabo la evaluación de la docencia (Correa y Rueda Beltrán, 2012). En el caso de Chile, en una investigación llevada a cabo por Salazar (2008) con universidades pertenecientes al Consejo de Rectores de las Universidades Chilenas (CRUCH), se encontró que todas ellas coinciden en que la mejora es el fin último de la evaluación docente practicada en la universidad.

En la medida que el proceso de evaluación de la docencia en las IES es reciente, no se observan aún en el caso chileno trabajos que evalúen el devenir de cada institución en materia de evaluación de la docencia, salvo estadísticas generales al interior de cada una.

1.5 El caso del postgrado

Si bien existe una amplia investigación con respecto al valor de la evaluación de la docencia en pregrado, no existe un desarrollo equivalente en postgrado. Además de la diversidad en las metodologías, un elemento importante parece ser el trabajo autónomo que los estudiantes deben realizar en el postgrado, donde a menudo se espera (o se plantea) un rol del académico como supervisor en particular en el período de tesis (Gabalán-Coello et al., 2019).

Más aún, la mayoría de la evidencia en la evaluación de la docencia de postgrado, proviene de programas de doctorado, más que de programas de Magister (Drennan y Clarke, 2009). Esto es significativo, si se considera que los estudiantes que ingresan a un programa de Magister generalmente buscan una experiencia diferente al pregrado: mayor aplicación al mundo laboral, desarrollo de habilidades de pensamiento de orden superior, aprendizaje independiente, erudición, autorreflexión crítica y aplicación del conocimiento, entre otros (Drennan y Clarke, 2009).

A partir de un estudio realizado con estudiantes de postgrado de diferentes niveles (magíster, doctorado, residentes clínicos, académicos, entre otros), Ghadirian et al. (2014) encontraron que las habilidades comunicacionales del académico (conocimientos y habilidades de supervisión, atmósfera, estatutos y reglamentos relacionados con la supervisión, seguimiento y evaluación), juegan un rol muy importante, en particular en lo concerniente a la supervisión de tesis (Ghadirian et al., 2014). Esto es corroborado parcialmente por Comley-White y Potterton (2018), quienes mencionan el apoyo y el ambiente como dos elementos centrales en el nivel de maestría.

Toda esta discusión es muy relevante si se considera la heterogeneidad existente en nuestro país en lo referente al desarrollo de los programas de Magíster, en términos de programas académicos versus profesionalizantes, aseguramiento de la calidad y expectativas de los estudiantes (Rodríguez y Sánchez, 2017).

1.6 Críticas a las evaluaciones

Existen algunas críticas al uso de estos instrumentos, entre las que se pueden mencionar a) advertencias con respecto al uso indiscriminado, aislado e irreflexivo de estos instrumentos, combinando diferentes fines (retroalimentar de la calidad de la enseñanza, monitoreo de la calidad, formulación de políticas, clasificación de los académicos), b) número y naturaleza de las dimensiones que deben ser incluidas en dichos instrumentos, c) ausencia de un marco teórico unificado (Spooren et al., 2014), d) sesgo asociado, por ejemplo a género, calificaciones, tamaño de la clase, tipo de docencia, del aula y número de horas, etc.) nivel de confidencialidad de los resultados (Huybers, 2014).

A pesar de la controversia, hay coincidencia en que se debe mejorar la calidad de la enseñanza, y que estas encuestas deben formar parte de sistemas más integrales dirigidos a una evaluación justa y equitativa, y al desarrollo profesional y académico (Gabalán-Coello et al., 2019).

La presente investigación aplica el conocimiento consolidado por un equipo de asesores en autoevaluación y acreditación, a la validación de un instrumento orientado a evaluar asignaturas de Magíster y el desempeño docente de las mismas, en un trabajo colaborativo entre la Unidad de Calidad y la Dirección de Postgrado de una universidad tradicional.

Objetivo: Obtener las propiedades psicométricas de una escala de evaluación del desempeño docente en programas de Magíster.

2. Método

2.1 Diseño

Se utilizó un diseño de investigación en dos etapas:

I) La primera, de carácter descriptivo, con validación de jueces expertos internos a la Institución, asignando puntuaciones a los ítems de las versiones a considerar.

II) La segunda, correlacional, de corte transversal en un solo momento del tiempo, mediante metodología de encuestas, para identificar la asociación relevante entre las variables y su contribución al instrumento a analizar.

2.2 Muestra

Para la primera parte del estudio (Etapa I), seis jueces expertos profesionales de la Unidad de Aseguramiento de la Calidad Institucional, quienes han asesorado a programas de pre y postgrado en procesos de autoevaluación con fines de acreditación, analizaron tres modelos de encuestas genéricas de evaluación de la docencia de Magíster utilizadas en años anteriores.

Los modelos de encuestas anteriores a validar, se denominaron "Modelo 1" (14 preguntas, orientadas a la asignatura), "Modelo 2" (17 preguntas, estando 5 de ellas orientadas al académico/s) y "Modelo 3" (12 preguntas, 3 de ellas orientadas al académico).

Los jueces expertos evaluaron dichos instrumentos mediante un "Protocolo para Validación de Jueces Expertos", en un archivo Microsoft Excel, en forma individual. El mecanismo para asignación de los puntajes está descrito en la sección "Análisis de Datos" del presente artículo.

Tabla 1*Composición equipo jueces expertos.*

Juez Experto	Niveles de calificación			Vinculación con la organización			Personales		Experiencia en AC* (Años)
	Área Formación Inicial	Título	Grado	Dedica - ción	Contrato en la Unidad	Se desempeña además en alguna unidad académica	Edad	Sexo	
1	Ingeniería	Ingeniero (E) Computación e Informática	Magister	Jornada Completa	Indefini - do	No	57	F	19
2	Estadística	Ingeniero Estadístico	Magister	Parcial (11 horas)	Plazo fijo	Sí	60	F	3
3	Biología	Bióloga	Doctor	Parcial (22 horas)	Indefini - do	Sí	37	F	2
4	Ciencias de la Salud	Enfermera	Magister	Parcial (11 horas)	Plazo fijo	Sí	43	F	3
5	Educación	Profesor Filosofía	Doctor	Parcial (22 horas)	Indefini - do	Sí	39	M	3
6	Ciencias de la Salud	Nutricionista	Magister	Parcial (22 horas)	Plazo fijo	Sí	39	F	2

*Aseguramiento de la Calidad.

Fuente: Elaboración propia.

Una vez que dichos jueces expertos concluyeron la etapa inicial de validación interna (Etapa I), se procedió a la segunda etapa (II). El instrumento resultante de dicha etapa, constaba de 28 preguntas, 16 para evaluación de la calidad de la asignatura, y 12 para evaluación del académico.

Para la etapa II, evaluación cuantitativa del instrumento, en la sección del cuestionario que aborda la calidad de la asignatura se consideró una muestra por accesibilidad constituida por 272 estudiantes/asignaturas de postgrado, que cursaban programas de Magíster. En el caso de la sección del instrumento que considera la evaluación del/la Académico/a, participaron 275 estudiantes/asignatura de Magíster válidos, ya que en algunos casos podía ser evaluado más de un/a académico/a por asignatura. De estos últimos, el 54,55% fue de sexo femenino y el 45,45% de sexo masculino.

Participaron estudiantes de 44 programas de Magister, de los cuales un 70,45% corresponde a programas académicos y un 29,55% a programas profesionales. En el Modelo Educativo de la Universidad de Concepción, un Magíster con orientación académica, incluye una formación orientada a la investigación que puede servir de preparación al Doctorado. Incluye cursos avanzados y un trabajo de investigación. Por otra parte, en el caso de aquellos Magíster con orientación profesional, se considera una formación con orientación a una práctica profesional avanzada. Incluye cursos de formación avanzada y contempla un proyecto final que requiere trabajo personal o una práctica profesional avanzada en una institución ([Universidad de Concepción, 2011](#)).

Como sujeto participante de esta segunda etapa, se consideró estudiante/asignatura de Magíster como el sujeto a ser indagado. Esto significa, que, si un mismo estudiante evaluó dos asignaturas diferentes, fue considerado como dos estudiantes/asignaturas diferentes (sujetos).

Se les envió a los estudiantes, la propuesta integrada de encuesta resultante del análisis de los jueces expertos (Etapa I). La evaluación se realizó mediante encuestas en línea, enviadas en el sistema SurveyMonkey a los estudiantes, a sus correos electrónicos institucionales.

2.3 Análisis de datos

En la primera fase de validación de jueces expertos, los profesionales evaluaron la pertinencia, aporte e inteligibilidad de los ítems (mediante una pauta de respuesta dicotómica para cada ítem), de encuestas utilizadas históricamente en la Universidad para evaluar la docencia de Magíster, teniendo además la opción de agregar comentarios adicionales o propuestas de redacción de ítems.

Una vez concluido dicho trabajo, se asignó puntaje a la evaluación de jueces expertos (entre 0 y 16 puntos), y se eliminaron ítems consensuados por los jueces como irrelevantes, repetidos o redundantes (bajo 10 puntos). Este método, ha sido utilizado en una forma similar con jueces expertos en la educación superior, por [Martínez Iñiguez y Tobón \(2019\)](#).

Esto dio lugar a una propuesta inicial integrada de 29 ítems, que considera dos secciones: Evaluación de la asignatura y Evaluación del/a Académico/a. Posteriormente, se realizó una aplicación piloto en asignaturas de postgrado, y se hizo un análisis estadístico de los resultados.

El análisis del acuerdo inter-jueces, se desarrolló mediante la prueba Kappa de Fleiss utilizando el software R ([R Core Team, 2019](#)), la consistencia interna se analizó mediante el software SPSS (IBM Corp., 2017) y el análisis factorial, mediante el software Mplus ([Muthen y Muthen, 2017](#)).

2.4 Aspectos Éticos

Tanto en el caso de los profesionales que evaluaron el instrumento inicial en el rol de jueces expertos, como en el de los estudiantes, la participación fue absolutamente voluntaria no recibiendo ningún tipo de recompensa o incentivo por ello.

En caso de los estudiantes, se les informó acerca de los objetivos psicométricos centrales de la investigación (mejorar pertinencia y fiabilidad de la medición, para una mejor retroalimentación a las unidades académicas), así como de aspectos éticos relativos al anonimato y confidencialidad de la información. Así mismo, se indicaron las unidades responsables del estudio, a saber, Dirección de Postgrado y Unidad de Aseguramiento de la Calidad Institucional de la Universidad.

Una vez entregada dicha información, los estudiantes libremente elegían si participar o no en el estudio.

3. Resultados

3.1 Validación jueces expertos

Al concluir la validación de jueces expertos se llegó a una propuesta integrada de 29 ítems, que consideró dos secciones: Evaluación de la asignatura y Evaluación del/a Académico/a.

El análisis estadístico de los resultados de la evaluación de dicha etapa, se realizó mediante la prueba Kappa de Fleiss, la cual permite evaluar la confiabilidad del acuerdo entre un número de evaluadores, al asignar clasificaciones categóricas a un número de ítems. El resultado de la prueba Kappa de Fleiss fue de 0.24 ($z = 6.32$, $p\text{-value} < 0.001$), siendo estadísticamente significativa. En el análisis solo se incluyeron las 47 filas con datos completos.

Este valor corresponde a un acuerdo justo según la interpretación propuesta por [Landis y Koch \(1977\)](#).

Una vez realizada esta etapa, tal como se detalló en los métodos, se aplicó esta versión validada por los jueces expertos, a la muestra de asignaturas y estudiantes de postgrado.

3.2 Análisis de Consistencia Interna y estructura factorial del instrumento

Esta sección del análisis, tal como ya se detalló, se realizó con la muestra de 272 estudiantes/asignaturas de postgrado, que cursaban programas de Magister para el caso de la sección de evaluación de la asignatura, y de 275 estudiantes para la sección del instrumento que considera la evaluación del/la Académico/a.

3.3 Consistencia interna

Consistencia interna sección evaluación de la asignatura

Al realizar un análisis de la consistencia interna del instrumento, mediante la prueba Alpha de Cronbach, se encontró que en la sección “Evaluación de la Asignatura”, la consistencia interna del instrumento fue de 0,956, siendo robusta. En el caso de la contribución de los diferentes ítems a dicha consistencia, se encontró que la eliminación de ítems específicos no aumentaba significativamente la consistencia de la prueba.

Consistencia interna sección evaluación del/de la académico/a

En la sección “Evaluación del/de la Académico/a”, la consistencia interna para la versión de 12 ítems, fue de 0,253, siendo no aceptable. Al analizar la matriz de correlaciones, se observó que solo había cuatro ítems que muestran correlación con dos ítems o más de la escala.

Dichos ítems eran los 2, 4, 5 y 8:

i2.- El(la) académico(a) realizó las clases y evaluaciones de acuerdo a lo planificado, cumpliendo la hora de inicio y término de cada una

i4.- El(la) académico(a) escuchó las preguntas e intervenciones de los alumnos y dio espacio para que discrepasen de sus planteamientos

i5.- El(la) académico(a) entregó información completa sobre el programa de la asignatura al inicio de esta

i8.- El(la) académico(a) entregó los resultados de las evaluaciones de forma oportuna dentro del plazo establecido

Al considerar solo estos ítems, la consistencia interna medida con Alpha de Cronbach, ascendió a 0,821, siendo estadísticamente significativa. De esta forma, se consideró solo un factor para esta sub-escala de evaluación del académico.

3.4 Análisis Factorial Exploratorio

A partir de los antecedentes anteriores, que muestran un solo factor para la sección relativa al académico/a, se llevó a cabo un análisis factorial exploratorio (EFA) para la sección Evaluación de la Asignatura.

En dicho análisis, los tres primeros factores arrojan eigenvalues superiores a uno, apuntando a una solución de tres factores. Sin embargo, al analizar los indicadores de ajuste, se encuentra que las soluciones de 3 y 4 factores son estadísticamente significativas con buenos valores de ajuste para ambos modelos, con una leve superioridad de este último, tal como se puede observar en la tabla 2, al observar los valores de Chi-cuadrado, así como los valores RMSEA, CFI, TLI y AIC.

Tabla 2*Valores de ajuste de las soluciones para las escalas de 3 y 4 factores en EFA.*

	Prueba de Chi-cuadrado de ajuste de modelo	RMSEA	CFI	TLI	AIC
3 factores	0.0000	0.091	0.934	0.905	9791.980
4 factores	0.0000	0.079	0.956	0.927	9721.294

Fuente: Elaboración propia.

Como se puede constatar, los valores de RMSEA están sobre lo aceptable (0,05), indicando razonables errores de aproximación a la población (MacCallum et al., 1996). Se ha señalado que este indicador es muy sensible al tamaño muestral, lo cual podría estar ocurriendo en este caso (Hu y Bentler, 1999). Es por esta razón que se sugiere considerar los indicadores en su conjunto para evaluar ambas soluciones.

En el contexto de estos valores de ajuste globales, el ítem 13 “El trabajo autónomo asignado a los estudiantes fue coherente con las exigencias de la asignatura y se ajustó a las cargas de trabajo declaradas y comprometidas”, que estaría asociado al cuarto factor, presenta una innecesaria extensión y redundancia (según juicio de dos jueces expertos), y estaría siendo abordado en forma parcial por la pregunta 12 “El trabajo autónomo asignado a los estudiantes contribuyó al logro de los objetivos o resultados de aprendizaje de la asignatura”.

Debido a ello, se decidió eliminar el ítem 13 de la escala y realizar un análisis factorial confirmatorio con un modelo de 19 ítems, lo que apuntaría en forma preliminar a un modelo tri-factorial.

Esta nueva sub-escala ajustada para la sección evaluación de la asignatura, presenta una consistencia interna, medida a través de Alpha de Cronbach de 0,952.

3.5 Análisis factorial confirmatorio

Al realizar el análisis factorial confirmatorio, se confirman apropiados y estadísticamente significativos indicadores de ajuste para la solución de tres factores, tal como se puede apreciar en la tabla 3, en la prueba de Chi-cuadrado y los valores RMSEA, CFI y TLI.

Tabla 3*Valores de ajuste para el CFA en modelo de tres factores.*

	Prueba de Chi-cuadrado de ajuste de modelo	RMSEA	CFI	TLI
3 factores	0.0000	0.129	0.95	0.949

Fuente: Elaboración propia.

En resumen, la sección “evaluación de la asignatura”, presentó una alta consistencia interna y una estructura interna de tres factores para la versión de 19 ítems. La sección Evaluación del/a Académico/a, por otra parte, solo incorporó cuatro ítems tributando a una sub-escala global. Esto da lugar a una escala global de 23 ítems.

4. Discusión

La idoneidad, conocimientos y experiencia de los jueces expertos permitieron validar y ajustar un instrumento desarrollado históricamente en la institución, ganando eficacia y eficiencia en el proceso de evaluación de la docencia que el postgrado necesitaba. Las características demográficas de dichos jueces expertos, son similares a aquellas mencionadas por [Green y Little \(2016\)](#): mayoritariamente mujeres, combinando docencia y labores de gestión educativa, con alguna carga académica y formación de postgrado.

Este proceso ha demostrado en términos empíricos, el nivel de conocimiento de estos profesionales del tercer espacio, con roles mezclados entre la gestión y lo académico ([Green y Little, 2016](#)), con diferentes grados de experiencia en docencia en investigación ([Berman y Pitman, 2010](#); [Veles y Carter, 2016](#); [Whitchurch, 2009, 2015](#)), capaces de trabajar en proyectos complejos ([Veles y Carter, 2016](#)). Son unidades que ya no solo responden a la planificación y objetivos internos de su unidad, sino que a la organización desde una perspectiva más amplia que puede incluir incluso a las autoridades superiores o a la estrategia global ([Venables y Van Gastel, 2014](#)).

Con respecto al instrumento evaluado, a partir del análisis factorial realizado, es posible obtener dos sub-escalas para evaluar la docencia de postgrado, con adecuadas propiedades tanto en términos de su contenido como de su consistencia interna. El análisis factorial arrojó una solución con apropiados y estadísticamente significativos indicadores de ajuste globales. Aun cuando el tamaño muestral podría haber influido en los resultados de dicho análisis factorial, es posible afirmar que las sub-escalas desarrolladas “Evaluación de la asignatura” y “Evaluación del académico/a”, en sus versiones finales, son confiables y además reflejan la trayectoria de la Institución en la aplicación de instrumentos de evaluación de la docencia.

La sección “evaluación de la asignatura”, presentó una consistencia interna de 0,952 y una estructura interna de tres factores para la versión de 19 ítems: i. Calidad Docencia-Asignatura, ii. Bibliografía y iii. Autoevaluación. En el caso de la sección Evaluación del/a Académico/a, solo cuatro ítems (de los 12 incluidos en la sección) mostraron intercorrelación importante, tributando a una sub-escala global. Al evaluar la consistencia interna de dicho instrumento ajustado, se obtuvo una consistencia de 0,821. Así, una versión de 23 ítems y dos subescalas (evaluación de la asignatura y del académico), tiene las propiedades estadísticas, así como las características necesarias y suficientes para ser utilizada en forma preliminar. Los autores del presente artículo consideran que una aproximación adecuada es aplicar ambas subescalas en conjunto. La sub-escala de cuatro (4) ítems (Evaluación del/la académico) que aparece como breve, tiene buenos antecedentes de consistencia interna y tributa a un cuestionario que en su conjunto presenta adecuados índices, y cuya extensión tiene un antecedente en la investigación de [Marsh et al. \(2002\)](#), quienes para escalas de entre 3 y 6 ítems, reportan consistencias entre 0.79 y 0.91.

Los indicadores de consistencia interna mencionados en el párrafo anterior son similares a los indicados por [Kember y Leung \(2008\)](#), quienes reportan valores que fluctúan entre 0,79 y 0,91, y a aquellos señalados satisfactorios por [Spooren et al. \(2014\)](#). Dada la cantidad de estudiantes considerados en este estudio (272 y 275) para las sesiones de evaluación de la asignatura y evaluación del académico respectivamente), los hallazgos en términos de consistencia interna concuerdan con aquellos reportados por [Marsh \(2007\)](#), quién señala que sobre 50 estudiantes es habitual encontrar índices de consistencia interna sobre 0,95. Por esta razón, aun cuando siempre es posible aumentar los tamaños muestrales, los autores de la presente investigación valoran positivamente el número de estudiantes participantes, considerando que se está trabajando en el nivel de Magister, donde el Universo es menor que en el pregrado.

La existencia de tres factores para la sección “Evaluación de la Asignatura” es coincidente con lo reportado en la literatura. Si bien pudiera parecer llamativo que dentro de la sub-escala evaluación de la asignatura se encuentre una gran dimensión denominada “Calidad Docencia-Asignatura”, separada de por ejemplo, de la “Bibliografía”, no lo es tanto al analizar la literatura. Ya en 1976, Feldman planteaba que “Materiales Suplementarios” pudiera ser una dimensión separada dentro de las categorías de la docencia efectiva (Feldman, 1976), y el SEEQ uno de los instrumentos más utilizados y validados internacionalmente, en la sección de “Tareas” incluye dos preguntas relativas a la pertinencia de la bibliografía (Marsh, 1982), siendo estas aproximaciones que han influido en las escalas propuestas en las décadas posteriores. Con respecto al valor de la autoevaluación en la evaluación de la docencia, éste es rescatado en el análisis realizado por Biaggi (2019) en torno a la evaluación de un modelo de enseñanza-aprendizaje en el contexto universitario

No obstante, los factores encontrados en la presente investigación, no coinciden con aquellos mencionados por Salazar (2008) (desde una perspectiva global de análisis del sistema nacional chileno más que psicométrica), Rivera y Rice (2002) o por Ugarte (2008), señalando este último las dimensiones Evaluación del Profesor, Responsabilidad y Organización, Aspectos del Curso, Aspectos de las Calificaciones y Aspectos del Grupo. Al respecto, cabe señalar que la falta de concordancia entre los resultados de estas investigaciones y los del estudio aquí realizado se pueden atribuir a que: a) la escala que estamos analizando no contempla aspectos didácticos específicos del docente, pues estos están medidos en forma general la escala “Evaluación del/a Académico/a”, y b) el uso de determinadas encuestas para evaluar la docencia en otras organizaciones –por ejemplo, aquella reportada por Ugarte (2008)– no garantiza que en todas ellas se haya realizado un análisis como el aquí planteado. No obstante, la idoneidad de escalas que tengan una detallada diferenciación de factores versus aquellas que ilustren un factor general ha sido una fuente de debate, el cual debe ser abordado dependiendo de la utilización última de los resultados de la evaluación (Gabalán-Coello et al., 2019; Marsh, 2007), lo cual lleva de vuelta a la discusión planteada en el marco conceptual.

Los programas de Magister tienen ciertas especificidades y heterogeneidad en nuestro país, que dificultan el pensar en dimensiones pedagógicas tan específicas como las que se pueden encontrar para el pregrado.

Una heterogeneidad (Rodríguez y Sánchez, 2017) que incluye programas académicos y profesionales, o desde aquellos con un fuerte énfasis en la tesis y por lo tanto en la supervisión, hasta aquellos que se convierten en una especialización de la formación de pregrado en forma de asignaturas más específicas y/o complejas. Desde programas con un fuerte componente práctico y de aplicación, hasta aquellos que son una puerta de entrada a la formación doctoral, van configurando un escenario donde dimensiones más específicas y uniformes de la docencia, tales como las observables en el pregrado, son muy difíciles y quizás sin sentido,

Si bien un 70,45% de los programas participantes corresponden a programas académicos (más que profesionales), la heterogeneidad de la estructura interna de los programas en términos curriculares, didácticos, metodológicos, no alcanza a ser capturada en una investigación como ésta, y se proyecta como futuro desarrollo para la evaluación de la formación de postgraduados.

En cualquier caso, todo este análisis es relevante para el futuro del instrumento y ha sido presentado a la Dirección de Postgrado, y está siendo complementados a partir del Modelo Educativo y requerimientos pedagógicos de la universidad.

Al respecto, se debe tener en cuenta que la evaluación de la docencia debe formar parte de un sistema global de desarrollo de la carrera docente, donde el proyecto educativo de la institución señala características deseables en un docente y, por otra parte, genera expectativas

determinadas en un estudiante. En otras palabras, la evaluación de la docencia es inseparable de la idiosincrasia y la cultura de la IES y los ítems seleccionados sientan una base a ser enriquecida por la Universidad y la Dirección de Postgrado de una forma integral y sistémica (Toró, 2012).

Si bien se espera que los estudiantes de Magíster sean estudiantes con una gran capacidad de autogestión y desarrollen capacidades como pensamiento crítico, resolución de problemas y habilidades de investigación, eso no significa que los programas puedan descansar en dicha perspectiva para eludir su responsabilidad permanente con la mejora de la calidad de la docencia. Los estudiantes siguen esperando de sus asignaturas y sus académicos el cumplimiento de ciertos estándares y expectativas, como los parcialmente identificados en esta investigación.

Al igual como ocurre con los niveles de Doctorado, los programas de Magíster cumplen un rol central y son una contribución fundamental a la generación de conocimiento al menos a nivel nacional. Si se pondera además la globalización y la internacionalización de los programas e instituciones, dicho valor se ve aumentado. Al ir mejorando la calidad de su docencia, van contribuyendo a la calidad de las Instituciones donde se imparten, a la calidad de la formación de postgrado en todo el sistema, y finalmente, a la consolidación de las comunidades científicas y académicas.

5. Conclusiones

Los resultados obtenidos en esta investigación, confirman que el instrumento en su conjunto y las sub-escalas desarrolladas “Evaluación de la asignatura” y “Evaluación del académico/a”, en sus versiones finales, son confiables para su utilización en la docencia de Magíster, sentando una base razonable para desarrollos ulteriores en las Instituciones a partir de los diferentes proyectos educativos, las características de cada programa de Magíster y requerimientos internos en términos de aseguramiento de la calidad.

Como limitación de esta investigación, más allá de los matices ya enunciados en torno al tamaño muestral, se puede mencionar que los ítems considerados para la evaluación de jueces expertos, y que finalmente fueron aplicados, conservan una semántica ajustada al momento histórico en que fueron creados. Este es caso, por ejemplo, de la expresión “alumnos”, para el cual se recomienda en investigaciones futuras preguntar por las/os estudiantes, las/os profesores, compañeros/compañeras.

Finalmente, en otro ámbito, esta investigación es un acotado, pero significativo paso hacia lo descrito por Scharager y Rodríguez Anaiz (2019): la construcción de un espacio común de trabajo colaborativo, un tercer espacio equilibrado entre la participación académica y administrativa (Scharager y Rodríguez Anaiz, 2019).

Agradecimientos

Nuestro más sincero reconocimiento a las profesionales Viviana Ulloa, Denisse Landea y Constanza Greig, sin cuyo trabajo, compromiso y capacidad técnica esta investigación no habría sido posible.

Referencias

- Alderman, L., Towers, S., y Bannah, S. (2012). Student feedback systems in higher education: A focused literature review and environmental scan. *Quality in Higher Education*, 18(3), 261-280. <https://doi.org/10.1080/13538322.2012.730714>.
- Berman, J. E., y Pitman, T. (2010). Occupying a 'third space': research trained professional staff in Australian universities. *Higher Education*, 60(2), 157-169. <https://doi.org/10.1007/s10734-009-9292-z>.
- Biaggi, G. M. (2019). *Efectividad de la enseñanza desde la perspectiva de los alumnos y docentes universitarios: evaluación de un modelo de enseñanza-aprendizaje universitario*. (Tesis Doctoral). Universidad de Morelos, Morelos, México.
- Centoni, M., y Maruotti, A. (2021). Students' evaluation of academic courses: An exploratory analysis to an Italian case study. *Studies in Educational Evaluation*, 70, 101054. <https://doi.org/10.1016/j.stueduc.2021.101054>.
- Chamoy, W. (2018). *Evaluation of the Psychometric Quality and Validity of a Student Survey of Instruction in Bangkok University, Thailand*. (Ph.D.). University of Pittsburgh, Ann Arbor. <https://login.ezpbibliotecas.udel.cl/login?url=https%3A%2F%2Fwww.proquest.com%2Fdissertations-theses%2Fevaluation-psychometric-quality-validity-student%2Fdocview%2F2166831819%2Fse-2> ProQuest One Academic database.
- Comley-White, N., y Potterton, J. (2018). The perceived barriers and facilitators in completing a Master's degree in Physiotherapy. *The South African journal of physiotherapy*, 74(1), 445-445. <https://doi.org/10.4102/sajp.v74i1.445>.
- Correa, M. d. D., y Rueda Beltrán, M. (2012). La evaluación docente en educación superior: uso de instrumentos de autoevaluación, planeación y evaluación por pares. *Voces y silencios. Revista Latinoamericana de Educación*, 3(2), 59-76. <https://doi.org/10.18175/vys3.2.2012.04>.
- Drennan, J., y Clarke, M. (2009). Coursework master's programmes: the student's experience of research and research supervision. *Studies in Higher Education*, 34(5), 483-500. <https://doi.org/10.1080/03075070802597150>.
- Eckhaus, E., y Davidovitch, N. (2019a). How Do Academic Faculty Members Perceive the Effect of Teaching Surveys Completed by Students on Appointment and Promotion Processes at Academic Institutions? A Case Study. *International Journal of Higher Education*, 8(1), 171-180. <https://doi.org/10.5430/ijhe.v8n1p171>.
- Eckhaus, E., y Davidovitch, N. (2019b). Potential for Blocking Advancement: Teaching Surveys for Student Evaluation of Lecturers. *International Journal of Educational Methodology*, 5(3), 401-406. <https://doi.org/10.12973/ijem.5.3.401>.
- Feldman, K. A. (1976). The superior college teacher from the students' view. *Research in Higher Education*, 5(3), 243-288. <https://doi.org/10.1007/BF00991967>.
- Gabalán-Coello, J., Vásquez-Rizo, F. E., y Laurier, M. (2019). *Evaluating teaching practices in graduate programs*. Springer.
- Ghadirian, L., Sayarifard, A., Majdzadeh, R., Rajabi, F., y Yunesian, M. (2014). Challenges for better thesis supervision. *Medical journal of the Islamic Republic of Iran*, 28, 32-32. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4154287/>.
- Ginns, P., Prosser, M., y Barrie, S. (2007). Students' perceptions of teaching quality in higher education: the perspective of currently enrolled students. *Studies in Higher Education*, 32(5), 603-615. <https://doi.org/10.1080/03075070701573773>.

- Green, D. A., y Little, D. (2016). Family portrait: A profile of educational developers around the world. *International Journal for Academic Development*, 21(2), 135-150.
- Hu, L., y Bentler, P. M. (1999). Cutoff Criteria for Fit Indexes in Covariance Structure Analysis: Conventional Criteria Versus New Alternatives. *Structural equation modeling: a multidisciplinary journal*, 6(1), 1-55.
- Huybers, T. (2014). Student evaluation of teaching: the use of best–worst scaling. *Assessment & Evaluation in Higher Education*, 39(4), 496-513. <https://doi.org/10.1080/02602938.2013.851782>.
- IBM Corp. (2017). *IBM SPSS Statistics for Windows, Version 25.0*. Armonk, NY: IBM Corp.
- Keeley, J., Furr, R., M., y Buskist, W. (2010). Differentiating psychology students' perceptions of teachers using the teacher behavior checklist. *Teaching of Psychology*, 37(1), 16-20. <https://doi.org/10.1080/00986280903426282>.
- Kember, D., y Leung, D. Y. (2008). Establishing the validity and reliability of course evaluation questionnaires. *Assessment & Evaluation in Higher Education*, 33(4), 341-353. <https://doi.org/10.1080/02602930701563070>.
- Lamarra, N. F., y Coppola, N. (2009). Aproximaciones a la evaluación de la docencia universitaria en algunos países iberoamericanos. Una perspectiva comparada entre similitudes, diferencias y convergencias. *Perspectivas en Políticas Públicas*, 1(2), 131-163.
- Landis, J. R., y Koch, G. G. (1977). An Application of Hierarchical Kappa-type Statistics in the Assessment of Majority Agreement among Multiple Observers. *Biometrics*, 33(2), 363-374. <https://doi.org/10.2307/2529786>.
- Loor Ávila, K. J., Gallegos Macías, M. R., Rodríguez, I., Monserrate, M., y Guillén Vivas, X. (2018). University faculty evaluation: Ibero-America trends. *Revista Cubana de Educación Médica Superior*, 32(1), 239-252.
- MacCallum, R. C., Browne, M. W., y Sugawara, H. M. (1996). Power analysis and determination of sample size for covariance structure modeling. *Psychological methods*, 1(2), 130-149. <https://doi.org/10.1037/1082-989X.1.2.130>.
- Marsh, H. W. (1982). SEEQ: A reliable, valid, and useful instrument for collecting students' evaluations of university teaching. *British journal of educational psychology*, 52(1), 77-95. <https://doi.org/10.1111/j.2044-8279.1982.tb02505.x>.
- Marsh, H. W. (2007). Students' evaluations of university teaching: Dimensionality, reliability, validity, potential biases and usefulness. In *The scholarship of teaching and learning in higher education: An evidence-based perspective* (pp. 319-383). Springer.
- Marsh, H. W., Rowe, K. J., y Martin, A. (2002). PhD Students' Evaluations of Research Supervision. *The Journal of Higher Education*, 73(3), 313-348. <https://doi.org/10.1080/00221546.2002.11777151>.
- Martínez Iñiguez, J. E., y Tobón, S. (2019). Cuestionario de percepción sobre el impacto de los procesos de acreditación de la calidad educativa en instituciones de Educación Superior. *Conrado*, 15(70), 242-247.
- Matosas-López, L., y García Sánchez, B. (2019). Beneficios de la distribución de cuestionarios web de valoración docente a través de mensajería SMS en el ámbito universitario: tasas de participación, inversión de tiempo al completar el cuestionario y plazos de recogida de datos. *Revista Complutense de Educación*, 30(3), 831-845. <https://doi.org/10.5209/RCED.59224>.

- Mendoza, E. R. (2020). La evaluación del profesorado universitario en tiempos de pandemia: los sistemas online de gestión de encuestas de satisfacción estudiantil. *Campus Virtuales*, 9(2), 61-70.
- Montoya Vargas, J., Arbesú García, I., Contreras Pérez, G., y Conzuelo Serrato, S. (2016). Evaluación de la docencia universitaria en México, Chile y Colombia. *Análisis de experiencias*, 7(2e), 15-42. <https://revistas.uam.es/index.php/riee/article/view/3111>.
- Mortelmans, D., y Spooren, P. (2009). A revalidation of the SET37 questionnaire for student evaluations of teaching. *Educational Studies*, 35(5), 547-552. <https://doi.org/10.1080/03055690902880299>.
- Muthen, B., y Muthen, L. (2017). *Mplus Version 8. US*. Muthen & Muthen.
- Picco, A. M., Ferrandi, M. A., Daruich, N., y Rittatore, M. E. (2013). *Education Official Data University Assessment Within The School Of Civil Engineering*. Universidad Nacional De Rosario. Paper presented at the 2013 Joint IASE / IAOS Satellite Conference. Statistics Education for Progress, Macao, China.
- R Core Team. (2019). *R: A language and environment for statistical computing*. Viena, Austria: R Foundation for Statistical Computing. <https://www.R-project.org/>.
- Ramsden, P. (1991). A performance indicator of teaching quality in higher education: The Course Experience Questionnaire. *Studies in Higher Education*, 16(2), 129-150. <https://doi.org/10.1080/03075079112331382944>.
- Ray, B., Babb, J., y Wooten, C. A. (2018). Rethinking SETs: Retuning Student Evaluations of Teaching for Student Agency. *Composition Studies*, 46(1), 34-56, 192-194.
- Rivera, J. C., y Rice, M. L. (2002). A comparison of student outcomes and satisfaction between traditional and web based course offerings. *Online Journal of Distance Learning Administration*, 5(3), 151-179.
- Rodríguez, M. T. A., y Sánchez, A. M. (2017). Expansión de la oferta y matrícula de Programas de Magíster en universidades chilenas: una mirada desde las políticas de aseguramiento de la calidad. *Pensamiento Educativo, Revista de Investigación Latinoamericana (PEL)*, 54(1), 1-21.
- Salazar, J. (2008). Diagnóstico preliminar sobre evaluación de la docencia universitaria. Una aproximación a la realidad en las universidades públicas y/o estatales de Chile. *Revista Iberoamericana de Evaluación Educativa*, 1(3), 67-84.
- Scharager, J., y Rodríguez Anaiz, P. (2019). Identidad Profesional de los administradores de la calidad en universidades chilenas: entre la invisibilización y la burocratización. (50), 30. <https://doi.org/10.31619/caledu.n50.538>.
- Spooren, P., Mortelmans, D., y Christiaens, W. (2014). Assessing the validity and reliability of a quick scan for student's evaluation of teaching. Results from confirmatory factor analysis and G Theory. *Studies in Educational Evaluation*, 43, 88-94. <https://doi.org/10.1016/j.stueduc.2014.03.001>.
- Toro, J. (2012). *Gestión interna de la calidad en las instituciones de educación superior*. Santiago, Chile: CINDA/ RIL Editores.
- Tucker, B., Oliver, B., y Gupta, R. (2013). Validating a teaching survey which drives increased response rates in a unit survey. *Teaching in Higher Education*, 18(4), 427-439. <https://doi.org/10.1080/13562517.2012.725224>.
- Ugarte, J. J. (2008). *Calidad, Formación y Evaluación de la Docencia en la UC*. Paper presented at the IV Encuentro Nacional de Docentes Universitarios "La Evaluación Docente como Estrategia para el Desarrollo del Profesor Universitario", Bogotá.

- Universidad de Concepción. (2011). *Modelo Educativo*. Universidad de Concepción. Concepción, Chile.
- Veles, N., y Carter, M.-A. (2016). Imagining a future: changing the landscape for third space professionals in Australian higher education institutions. *Journal of Higher Education Policy and Management*, 38(5), 519-533. <https://doi.org/10.1080/1360080X.2016.1196938>.
- Venables, J. P., y Van Gastel, J. (2014). Radiografía de los modelos de acreditación: organización, procesos y prácticas. El caso de las universidades Austral de Chile, De la Frontera y De los Lagos. *Calidad en la Educación*(41), 51-81. <https://doi.org/10.31619/caledu.n41.59>.
- Whitchurch, C. (2009). The rise of the blended professional in higher education: a comparison between the United Kingdom, Australia and the United States. *Higher Education*, 58(3), 407-418. <https://doi.org/10.1007/s10734-009-9202-4>.
- Whitchurch, C. (2015). The Rise of Third Space Professionals: Paradoxes and Dilemmas. In U. Teichler & W. K. Cummings (Eds.), *Forming, Recruiting and Managing the Academic Profession* (pp. 79-99). Cham: Springer International Publishing.

Apéndice 1

Sección Evaluación de la Asignatura, ítems versión final y factor asociado.

Número ítem	Contenido Ítem	Factor
1	Los contenidos fueron pertinentes a los objetivos o resultados de aprendizaje de la asignatura	Factor 1: Calidad Docencia Asignatura
2	Los contenidos de la asignatura estaban actualizados	Factor 1: Calidad Docencia Asignatura
3	La bibliografía de la asignatura estaba actualizada	Factor 2: Bibliografía
4	La bibliografía era pertinente a los objetivos o resultados de aprendizaje de la asignatura	Factor 2: Bibliografía
5	Los materiales y recursos didácticos utilizados contribuyeron al logro de los objetivos de la asignatura o resultados de aprendizaje esperados	Factor 1: Calidad Docencia Asignatura
6	Las actividades realizadas durante la asignatura permitieron lograr los objetivos o resultados de aprendizaje declarados previamente	Factor 1: Calidad Docencia Asignatura
7	Creo que el tiempo asignado a las actividades fue adecuado para el logro de los objetivos o resultados de aprendizaje planteados	Factor 1: Calidad Docencia Asignatura
8	Considero que las actividades de evaluación fueron coherentes con lo abordado en las clases	Factor 1: Calidad Docencia Asignatura
9	Considero que hubo una correcta coordinación de los ¹ docentes al interior de la asignatura	Factor 1: Calidad Docencia Asignatura
10	Fue claro el inicio, desarrollo y cierre de cada sesión	Factor 1: Calidad Docencia Asignatura
11	Existió un tiempo especialmente destinado a resolver dudas o profundizar contenidos	Factor 1: Calidad Docencia Asignatura
12	El trabajo autónomo asignado a los estudiantes contribuyó al logro de los objetivos o resultados de aprendizaje de la asignatura	Factor 1: Calidad Docencia Asignatura
13	Se propició alguna habilidad transversal (por ejemplo, trabajo en equipo, creatividad, investigación, pensamiento crítico) en las distintas actividades de la asignatura	Factor 1: Calidad Docencia Asignatura
14	Usted asistió regularmente a las actividades formativas de la asignatura	Factor 3: Autoevaluación
15	Usted se sintió satisfecho de su actuar como estudiante en esta asignatura	Factor 3: Autoevaluación
16	Usted presentó los trabajos solicitados por sus profesores, en los plazos establecidos	Factor 3: Autoevaluación
17	Usted colaboró con sus compañeros en las actividades del curso	Factor 3: Autoevaluación
18	Usted aportó al desarrollo de las clases con preguntas o comentarios	Factor 3: Autoevaluación
19	Usted trabajó responsablemente en todas las actividades que se le encomendaron en la asignatura	Factor 3: Autoevaluación

1. Se sugiere para versiones posteriores preguntar por las/os estudiantes, las/os profesores, compañeros/compañeras.

Apéndice 2

Sección Evaluación del/a Académico/a.

Número ítem	Contenido Ítem
20	El(la) académico(a) realizó las clases y evaluaciones de acuerdo a lo planificado, cumpliendo la hora de inicio y término de cada una
21	El(la) académico(a) escuchó las preguntas e intervenciones de los alumnos ² y dio espacio para que discrepasen de sus planteamientos
22	El(la) académico(a) entregó información completa sobre el programa de la asignatura al inicio de este
23	El(la) académico(a) entregó los resultados de las evaluaciones de forma oportuna dentro del plazo establecido

2. Se sugiere para versiones posteriores preguntar por estudiantes, y no por alumnos.