

Propiedades psicométricas de instrumento que evalúa perfil de egreso de psicología, basado en evaluación auténtica

Pág. 1 - 26

*Paulina Alejandra
Barrientos Illanes

**Ximena Angélica
Rojas Retamal

***Patricia
Alejandra Mozó
Cabrera

****Shila Scarlett
Peine Grandón

*Universidad San
Sebastián
Facultad de Psicología.
paulina.barrientos@uss.cl

**Universidad San
Sebastián
Facultad de Psicología
Coordinadora Pedagógico
Curricular.
Ximena.rojas@uss.cl

*** Investigadora
Independiente.
patimozo@gmail.com

****Universidad San
Sebastián. Secretaría de
Estudios Carrera
Psicología
shila.peine@uss.cl

Fecha de entrega:
marzo 2020

Fecha de aprobación:
junio 2020

RESUMEN

El alcance de este estudio es instrumental y multimétodo. Tiene por objetivo (1) describir las evidencias psicométricas del instrumento de evaluación, rúbrica analítica, que evalúa el examen de licenciatura de la carrera de Psicología de la Universidad San Sebastián, y (2) evaluar su coherencia con el perfil de egreso y la perspectiva de evaluación auténtica que la sustenta. Los participantes fueron 188 estudiantes en la primera instancia y 237 en la segunda. Se analizaron las evidencias de validez basadas en la estructura interna, por medio de análisis factoriales exploratorios. La validez de contenido se investigó a través del estadístico Kappa y la confiabilidad mediante Alfa de Cronbach. Se concluye que la rúbrica analítica utilizada, evidencia integración del conjunto de conocimientos, destrezas y actitudes que moviliza el estudiante para actuar eficazmente ante las demandas de un determinado contexto, por lo que, al triangular la coherencia de estos con los resultados de aprendizaje y los propósitos de la evaluación, permite evaluar desempeños esperados del perfil de egreso de la carrera. Se demuestra que es posible obtener evidencias de validez y confiabilidad (.91) en instrumentos de respuestas pre construidas, considerando las diferencias esperables entre los juicios de los evaluadores, docentes de la comisión.

Palabras claves:

Rúbrica analítica, validez de contenido (Kappa), validez de constructo, confiabilidad (Alfa de Cronbach), evaluación perfil de egreso.

Psychometric properties of an instrument that evaluates psychology graduation profile, based on authentic evaluation

Abstract

The scope of this study is instrumental and multi-method. The objectives are (1) describe the psychometric evidence of the assessment instrument, analytical rubric, which assesses the undergraduate exam of the Psychology career at the San Sebastian University and (2) evaluate its coherence with the graduation profile and the authentic evaluation perspective that supports it. Evidence of validity based on the internal structure was analyzed by means of exploratory factor analysis. Content validity was investigated using the Kappa statistic and reliability using Cronbach's Alpha. It is concluded that the analytical rubric used shows integration of the set of knowledge, skills and attitudes that the student mobilizes to act effectively in the face of the demands of a certain context, therefore, by triangulating the coherence of these with the learning results and for the purposes of the evaluation, it allows evaluating the expected performance of the graduation profile.

It is shown that it is possible to obtain evidence of validity and reliability (.91) in pre-constructed response instruments, considering the expected differences between the judgments of the evaluators, teachers of the commission.

Keywords:

Analytical rubric, content validity (Kappa), construct validity, reliability (Cronbach's Alpha), graduation profile evaluation.

INTRODUCCIÓN

En un contexto de educación superior donde el énfasis está puesto en la formación por desempeños, tiene cabida un cambio de paradigma en la comprensión de la evaluación, entendiendo ésta como una parte del proceso de formación, donde se busca evidencia respecto a la consolidación de aprendizajes a nivel conceptual, procedimental y actitudinal (Buscá, Suárez, Buset, & Bosch, 2014). El presente artículo tiene por objetivo (1) describir las evidencias psicométricas del instrumento de evaluación, rúbrica analítica, construido para la parte oral del Examen de Licenciatura de la carrera de Psicología de la Universidad San Sebastián (USS), y (2) evaluar la coherencia de la rúbrica con el perfil de egreso y la perspectiva de evaluación auténtica que la sustenta.

Teóricamente, se considera que la evaluación de desempeños en educación superior debe estar asociada a la observación de evidencias en el quehacer mismo, en la praxis profesional, desde donde se valoriza el conocimiento, las habilidades y actitudes de manera integrada, siendo lo necesario para la toma de decisiones en contextos profesionales específicos (Arenas & Gómez, 2013).

A lo largo del desarrollo temático, se define teóricamente el enfoque de evaluación auténtica, precisando y describiendo sus principios teóricos, como la búsqueda de evidencias de un aprendizaje situado, por medio de tareas que simulen ejercicios de la práctica profesional real, ubicando al estudiante en un escenario donde debe tomar decisiones que son propias del ejercicio profesional (López, 2011; Vallejo & Molina, 2014).

Desde la literatura especializada, se conceptualiza y describen las características de la rúbrica analítica como instrumento de evaluación (Föster, 2017). Además, se analizan elementos que permiten establecer coherencia entre la rúbrica analítica y los principios de la evaluación auténtica.

Así mismo, en el contexto de la carrera de Psicología de la Universidad San Sebastián, se describe el Examen de Licenciatura como dispositivo de evaluación de los desempeños esperados del perfil de egreso, con foco en la parte oral del examen.

En lo siguiente, se dan a conocer antecedentes de contexto de la evaluación en educación superior, para pasar a lo específico de la carrera de Psicología de la Universidad San Sebastián. Por otro lado, se presentan elementos de la evaluación auténtica y la coherencia de la rúbrica analítica, ello como fundamentos con el propósito de validar el instrumento propuesto. Finalmente, se da cuenta de sus resultados y conclusiones, las cuales se desprenden de una reflexión más profunda respecto de la evaluación en contextos de perfiles de egreso, construidos en base a desempeños y la complejidad que ello implica para la enseñanza en todas sus dimensiones.

a. Evaluación de Desempeños en Educación Superior

Entenderemos que la formación basada en desempeños está orientada a formar profesionales críticos y reflexivos, con capacidad de toma de decisiones, que les permitan resolver o transformar situaciones y contextos problemáticos, complejos y diversos en el ejercicio profesional (Buscá, et al., 2014).

En este sentido, desde la literatura relativa a la temática, se plantea que el cambio en el proceso de evaluación a nivel universitario implica una modificación en la comprensión de lo que significa la formación. Es decir, respuestas a preguntas relativas al cómo, cuándo y para qué de la evaluación, deben ser redefinidas a la luz de un nuevo concepto de formación universitaria, que considere como foco central, los resultados de aprendizaje de los estudiantes y los instrumentos de evaluación (Buscá et al., 2014). Desde esta mirada, los procesos de enseñanza incorporan la evaluación como parte del mismo proceso de formación y no como un hito separado que tributa únicamente a la calificación.

Por su parte, la Universidad San Sebastián ha tomado definiciones curriculares que se reflejan en el Informe de Autoevaluación de la Carrera de Psicología (2017) en el que se señala:

En la última actualización del Proyecto Educativo (2015), se tomó la decisión que los Perfiles de Egreso se estructuren por *Desempeños Esperados*. Los Desempeños Esperados corresponden a “una declaración de las actuaciones profesionales de distinta complejidad que se espera puedan demostrar los estudiantes en el hito del egreso; que, por encontrarse dentro de un proceso formativo, son de carácter tutelado y contextualizado en ambientes simulados o auténticos.

El concepto de *Desempeños Esperados* surge de una reflexión y aprendizaje institucional que permite designar con mayor precisión aquello que es posible comprometer en la formación, que considera conocimientos, actitudes y habilidades posibles de ser demostrables en el egreso, haciendo la distinción con las competencias –profesionales– reservadas para el ejercicio profesional real.

Lo anterior se fundamenta en definiciones aportadas desde la literatura. En efecto este concepto se complementa con lo que señala Tardif (2004, p.6), quien se refiere a la competencia restringiéndola a lo que es posible lograr en un proceso formativo, siendo ésta más cercana, a la definición USS de “Desempeños Esperados”, ya que la competencia en pleno se observa y manifiesta en el desempeño profesional real.

b. Evaluación Auténtica para la evaluación de desempeños en educación superior

En un contexto social donde las tecnologías cambian la forma en que los estudiantes se relacionan con el conocimiento, lo utilizan y crean, la formación universitaria debe considerar la evaluación como un importante medio que facilita el aprendizaje (Ahumada, 2005). En este marco, la evaluación se podría definir como un proceso a través del cual se obtiene evidencia sobre el aprendizaje del estudiante, pudiendo emitir juicios respecto de éste, de manera válida,

confiable y oportuna (Ahumada, 2005). En un sentido más específico, el enfoque de evaluación alternativo denominado “Evaluación Auténtica” intenta averiguar qué sabe el estudiante o qué es capaz de hacer, utilizando diferentes estrategias y procedimientos evaluativos. En este contexto, es necesario tener dispositivos de evaluación que permitan observar un espectro mucho más amplio de desempeños y donde el estudiante logre desplegarlos de forma integrada y compleja (López, 2011). Este espectro más amplio, debería incluir situaciones de la vida real, en toda su complejidad o cercano a ello, lo anterior implica comprender y asumir que las respuestas cerradas no facilitan este despliegue.

La literatura hace referencia a que las competencias son un conjunto de conocimientos, destrezas y actitudes que movilizan a una persona de forma integrada para actuar eficazmente ante las demandas de un determinado contexto (Perrenoud, 2004; Díaz, 2011; Denyer, Furnémont, Poulain & Vanloubbeeck, 2007), el desempeño por su parte se refiere al manejo de lo anterior para tomar decisiones en contextos específicos. En este sentido, Gallardo, Gil y Govea (2016) apoyándose en varios autores, señalan en su documento, que la evaluación del desempeño debe observar y evidenciar las capacidades de la persona competente en un contexto determinado o específico, integrando los aspectos cognitivos, afectivos y sociales para tomar decisiones frente a demandas específicas de contextos determinados. De esta forma, el estudiante, al enfrentar una situación de evaluación formal, re-construye el conocimiento, crea y propone soluciones o toma decisiones en función de posibles cursos de acción, considerando siempre aquello que da sustento a su decisión. Cada institución educativa debe declarar desde su filosofía el tipo de persona que desea formar, fundada desde su enfoque pedagógico, referentes legales y contexto cultural. Con relación a ello, el perfil de egreso de la Carrera de Psicología de la USS, está construido en base a dominios y desempeños esperados, como referencia para orientar la formación hacia la práctica.

Rúbrica analítica para la evaluación de desempeño en educación superior

Arenas y Gómez (2013) señalan en su documento que la rúbrica es un instrumento que surge del enfoque de formación por desempeños, en tanto evalúa la integración de saberes, habilidades

relacionales y procedimentales que el estudiante evidencia a través del abordaje de situaciones concretas y contextualizadas. Constituye una serie de beneficios tanto para el docente como para el estudiante, así como también es coherente con una formación que busca que el estudiante se prepare para enfrentar situaciones laborales donde deba tomar decisiones dentro de su disciplina

Específicamente, la rúbrica analítica, presenta diversas cualidades que la convierten en una herramienta de evaluación coherente con la evaluación auténtica y la formación basada en desempeños. Esta rúbrica define los criterios, niveles y operacionalizaciones desagregadas de los desempeños puestos en juego en la formación, los que se valoran y tienen como resultado un concepto respecto del resultado.

Lo anterior, implicaría que el uso de la rúbrica resulta coherente con algunos de los fundamentos que la Universidad San Sebastián plantea para orientar los procesos de evaluación del aprendizaje (USS, 2015), tales como proporcionar a los estudiantes orientaciones constructivas sobre cómo mejorar su aprendizaje, dado que con la rúbrica analítica se indica con claridad las fortalezas y debilidades de su desempeño, así como la planificación de las mejoras en etapas previas a la calificación; por otro lado, desarrollar en los estudiantes capacidad de autoevaluarse desde una perspectiva crítica y analítica, dado que la rúbrica promueve la autonomía del estudiante en tanto observe sus progresos y retrocesos con responsabilidad, y por último, tener un énfasis en el cómo aprenden los estudiantes, dado que la rúbrica permitiría tener la visión del progreso de los estudiantes y la integración de conocimientos de forma integrada en la puesta en marcha de un desempeño.

Considerando lo dicho anteriormente, se espera que la evaluación sea parte de la formación y no se constituya por hitos aislados que impliquen sólo una calificación. Lo anterior, supone que los docentes deben trabajar con metodologías evaluativas orientadas a favorecer los aprendizajes profundos, es decir, estamos hablando de evaluación auténtica.

Sistema de evaluación de desempeños para examen de grado en la carrera de Psicología de la Universidad San Sebastián

En el marco de una Educación Superior orientada a la calidad en la formación educativa y con el propósito de asegurar la ecualización de los procesos inter sede, la carrera de Psicología de la USS desarrolló un sistema de evaluación que permite monitorear el logro de los desempeños esperados del perfil de egreso y al mismo tiempo asegurar la ecualización en las cuatro sedes en que se imparte la carrera. Este sistema considera como instrumento de evaluación un examen de licenciatura que consta de dos partes, escrita y oral, composición cuyo propósito fue disminuir los errores de sesgo en la corrección, así como asegurar la posibilidad de otorgar un escenario donde los estudiantes puedan recurrir a distintas habilidades en el contexto evaluativo (Rojas, 2017).

Según los niveles de complejidad profesional definidos por la carrera de Psicología de la USS (Facultad de Psicología, 2017), el examen de licenciatura evalúa los desempeños comprometidos en el perfil de egreso, en un nivel que implica la resolución de problemas y procedimientos a través de conocimientos de diferente naturaleza, teóricos, procedimentales y actitudinales, los que serán descritos en el siguiente apartado (Rojas, 2017). En función de lo anterior, se espera que los estudiantes integren lo aprendido en sus años de formación profesional y lo apliquen en situaciones que representen el ejercicio profesional (Alsina, 2013; Rojas, 2017). En este marco, el estudiante se enfrenta a la descripción de una situación, caso del examen, donde debe tomar decisiones con argumentos y fundamentos coherentes y atingentes para abordarlo desde la disciplina. Se propone de esta forma, dado que es la acción cognitiva que constantemente debemos hacer en nuestro ejercicio profesional.

En definitiva, se pretende la evaluación del desempeño del estudiante en el marco de la toma de decisiones. Para esto, se utiliza la propuesta que realiza Lonergan (Gromi, 2013) para la toma de decisión, donde el supuesto es que, al tomar decisiones, se realiza un ejercicio cognitivo que implica etapas, para llegar a la aplicación o intervención propiamente tal, estas son: Fase 1: Analizamos, distinguimos, Fase 2: Reflexionamos, relacionamos, sintetizamos, Fase 3: Ordenamos las evidencias, Fase 4: Sopesamos las evidencias y seleccionamos, Fase 5: Emitimos juicios, Fase 6: Evaluamos.

La parte oral del examen de licenciatura consiste en la resolución y exposición de un caso correspondiente a una de las áreas de formación profesional (clínica, educacional, organizacional y social comunitaria) del plan curricular, la que es designada de forma aleatoria a cada estudiante. Durante el momento de la exposición del caso, se cuenta con una comisión evaluadora conformada por tres docentes del área, quienes evalúan la presentación y resolución del caso por medio de una rúbrica dispuesta por la carrera. La rúbrica como instrumento de evaluación es de tipo analítica descriptiva y está diseñada como una guía que permite operacionalizar los desempeños esperados del perfil de egreso y descomponer la evaluación de una tarea del ejercicio profesional real (Alsina, 2013; Rojas, 2017). Cabe señalar que los miembros de la comisión deben consignar tanto el puntaje obtenido por el estudiante para cada indicador, como comentarios a estos, contándose así con información cuantitativa y cualitativa del desempeño de los estudiantes, que permite una retroalimentación de éstos de manera integral.

En el caso de este examen, se considera una rúbrica de carácter analítica, por la posibilidad que entrega de evaluar los desempeños esperados del perfil de egreso por separado, así como los criterios que se evaluarán para cada uno y sus respectivos indicadores. Siguiendo la idea planteada por De la Cruz (2011), una medida adicional para facilitar que los estudiantes incorporen estándares de desempeño profesional, consiste en la implementación de dos instancias de retroalimentación a la actuación evidenciada en la parte oral de examen. Por un lado, como etapa final del examen, se realiza una devolución por parte de los tres miembros de la comisión evaluadora, cuyo objetivo es que el estudiante conozca sus debilidades y fortalezas. Asimismo, existe una instancia de carácter individual, en la que un docente de planta, que haya participado en la comisión de examen, reflexiona junto al estudiante acerca de su desempeño, tanto en la parte escrita como en la parte oral. Esta reunión se concreta luego de al menos un mes después de rendido el examen.

Diseño de la rúbrica para evaluar desempeño en examen de licenciatura para la carrera de Psicología Universidad San Sebastián

La rúbrica, como instrumento de evaluación, se construye en función de una serie de criterios establecidos a priori que orientarán la evaluación de los desempeños de los estudiantes en torno a la resolución de un caso/situación profesional. Como lo señala Cubillos-Veja y Ferrán-Aranaz (2019) el proceso de construcción de la rúbrica contiene tres pasos, los que se refieren a la identificación de las dimensiones a valorar, la redacción de los indicadores para medir tales dimensiones y la determinación de los niveles de logro y su puntuación.

Lo primero, fue identificar los desempeños del perfil de egreso de la carrera de Psicología USS que se pretendían abordar con el instrumento. Cabe destacar, que dicha tarea se formula a propósito de los objetivos establecidos para el examen de licenciatura, es decir, se define como parte del proceso global del examen, por lo que las variables a evaluar son los desempeños que evalúa el examen. La parte oral del examen de licenciatura requería de un instrumento que permitiera evaluar desempeños considerando los diferentes elementos que los componen de manera integrada. En este sentido, la rúbrica es un instrumento idóneo ya que permite valorar la evidencia considerando aspectos complejos, imprecisos y subjetivos, siendo de fácil interpretación y transparente tanto para quien evalúa como para quien es evaluado (Cubillos-Veja & Ferrán-Aranaz, 2019). Los criterios a evaluar se constituyen en desempeños del perfil de egreso.

En este marco, se definieron aspectos (indicadores) a evaluar por desempeño, los que son coherentes con lo esperado para el nivel de formación de un licenciado en psicología. La organización de estos aspectos obedece a un consenso de grupos de expertos en el tema, las propuestas de mejora planteadas por los equipos de docentes de la carrera y el análisis estadístico del instrumento utilizado previamente.

Las características de la rúbrica analítica, sugieren la evaluación de una serie de aspectos relevantes en el proceso de examen de grado, aspectos que fueron previamente definidos por una comisión de trabajo y que luego fueron validados por docentes de la Facultad de Psicología de la USS, y docentes externos con experiencia en evaluación de competencias. Los aspectos a

evaluar en relación a los desempeños se relacionan con la capacidad de fundamentar, integrar y organizar la información según criterios lógico - abstractos, de darle coherencia al abordaje que presenta y su pertinencia (Tabla 1). La calidad de las respuestas y la estrategia de puntuación se definieron en 4 niveles de desempeño.

Tabla 1

Desempeños y criterios de evaluación de la rúbrica analítica

Desempeños evaluados Perfil de Egreso 2012	Criterios de evaluación
<p>Dominio diagnóstico</p> <p>a) Identificar la modalidad adecuada de recolección de información según las variables y/o categorías, y área de desarrollo profesional</p> <p>b) Interpretar y analizar la información recolectada en los niveles individual, grupal y organizacional según área de desarrollo profesional.</p> <p>c) Construir una hipótesis diagnóstica que se sustente en los antecedentes recopilados dentro de un marco de referencia conceptual y ético. Se esperaría que el estudiante propusiera hipótesis diagnósticas adecuadamente fundamentadas.</p>	<ul style="list-style-type: none"> - Fundamenta la manera de abordar el caso en cuanto a la modalidad de recolección de información - Integra datos y fuentes considerando la relevancia de los mismos. - Fundamenta la manera de abordar el caso - Pertinencia de la hipótesis diagnóstica formulada
<p>Dominio diseño</p> <p>a) Establecer objetivos para la intervención según los resultados del diagnóstico y las características y condiciones del usuario, según el área de desarrollo profesional. Se esperaría que el estudiante propusiera objetivos adecuadamente fundamentados.</p> <p>b) Diseñar estrategias de intervención (prevención, promoción, tratamiento y rehabilitación, según el caso) que permitan alcanzar de manera coherente los objetivos formulados, incorporando los intereses y perspectivas de los individuos, grupos u organizaciones intervenidos. Se esperaría que el estudiante propusiera estrategias de intervención adecuadamente fundamentados.</p>	<ul style="list-style-type: none"> - Coherencia teórica-práctica con hipótesis - Fundamenta la manera de abordar el caso - Pertinencia entre estrategias/actividades/técnicas con objetivos. - Fundamenta la estrategia de intervención utilizada, asumiendo una posición epistemológica de fondo en su argumento.
<p>Dominio razonamiento y análisis</p> <ul style="list-style-type: none"> - Actuar de acuerdo a estándares éticos de la disciplina. Se espera que el estudiante entregue a partir de su análisis, una evaluación de alguno de los aspectos del caso desde los estándares éticos de la disciplina. 	<ul style="list-style-type: none"> - Utiliza criterios éticos para fundamentar el análisis o razonamiento del caso.
<p>Dominio gestión de relaciones</p> <ul style="list-style-type: none"> - Trabajar de manera efectiva con profesionales de otras disciplinas, valorando el aporte de éstos en su quehacer 	<ul style="list-style-type: none"> - Valoración del trabajo interdisciplinario.

profesional, así como también con otras personas del sistema en el cual participa. Se espera que el estudiante entregue dentro de sus análisis o evaluaciones de las respuestas, una valoración explícita del trabajo interdisciplinario y/o multidisciplinario.

Dominio gestión de relaciones

- Manejar de manera técnica y efectiva, formas de comunicación oral y escrita, adecuadas al contexto profesional y finalidad propuesta.

- Lenguaje claro, formal y técnico.

Fuente: Elaboración propia, documento interno de la Facultad, Perfil de Egreso.

Cada criterio es evaluado con cuatro niveles de desempeño: Deficiente (0 puntos), el estudiante no demuestra la habilidad evaluada; Insuficiente (1 punto), el estudiante demuestra errores/imprecisiones en la habilidad evaluada, no siendo capaz de corregir; Suficiente (2 puntos), el estudiante demuestra la habilidad evaluada parcialmente o con algunos errores, pero es capaz de corregir estos vacíos o errores a partir de las preguntas aclaratorias; y Competente (3 puntos), el estudiante demuestra de manera adecuada la habilidad evaluada.

MÉTODO

Este es un estudio de tipo instrumental, con carácter multimétodo, que se enfoca en el desarrollo de un instrumento, considerando también el diseño y adaptación del mismo.

En cada instancia estudiada del examen de licenciatura, febrero 2017 y febrero 2018, se realizaron análisis factoriales exploratorios para determinar las evidencias de validez basada en la estructura interna de la rúbrica. Lo anterior, dado que, si bien la rúbrica se construyó con base en los desempeños esperados del perfil de egreso de la carrera Psicología USS, se desconocen a priori factores teóricos a la base que expliquen el comportamiento del instrumento (Lloret-Segura, Ferreres-Traver, Hernández-Baeza & Tomás-Marco, 2014).

Para analizar la evidencia de validez basada en el contenido del instrumento, se utilizó el coeficiente Kappa de Cohen, esto para determinar el grado de concordancia inter-observador en los ítems de la rúbrica (Cerdeira & Villarreal, 2008). La fiabilidad de la rúbrica se obtuvo mediante el estadístico alfa de Cronbach.

La muestra consistió en estudiantes de la carrera de Psicología de pregrado de la Universidad San Sebastián, con un avance curricular hasta el octavo nivel en el plan de estudios 2012. En la versión febrero 2017, 188 estudiantes rindieron la parte oral del Examen de Licenciatura y en la versión febrero 2018 el número aumentó a un total de 237 estudiantes.

RESULTADOS Y ANÁLISIS

En este apartado, se dará cuenta del análisis de los resultados obtenidos en la validación de la rúbrica en cuestión, de las evidencias psicométricas del instrumento de evaluación, rúbrica analítica, construido para evaluar el examen de licenciatura de la carrera de Psicología de la Universidad San Sebastián. Para esto, se utilizaron análisis factorial para determinar la validez basada en la estructura interna, el estadístico Kappa de Cohen para analizar la validez de contenido y el coeficiente Alfa de Cronbach para evaluar la consistencia interna de la rúbrica.

Validez basada en la estructura interna de la rúbrica

Los análisis se realizaron con el software estadístico SPSS versión 18. La medida de adecuación muestral Kaiser-Meyer-Olkin (KMO) fue de 0.92 en las dos versiones de la rúbrica analizadas (versión febrero 2017 y febrero 2018) y el test de esfericidad de Bartlett fue altamente significativo en ambas versiones ($p < .001$), por tanto, se pudieron realizar los análisis factoriales previstos (ver tabla 2 y 3).

Tabla 2

Prueba de KMO y Bartlett instancia febrero 2017

Medida Kaiser-Meyer-Olkin de adecuación de muestreo		.92
Prueba de esfericidad de Bartlett	Aprox. Chi-cuadrado	3489
	G1	55
	Sig.	<.001

Tabla 3*Prueba de KMO y Bartlett instancia febrero 2018*

Medida Kaiser-Meyer-Olkin de adecuación de muestreo		.92
Prueba de esfericidad de Bartlett	Aprox. Chi-cuadrado	4029
	Gl	55
	Sig.	<.001

Nota: Fuente salida resultados SPSS.

Del análisis de las comunalidades se desprende que, en la versión 2017 de la rúbrica los ítems que se explican en una menor medida por el modelo son: Fundamenta consideraciones de juicio ético 31%; Lenguaje claro formal y técnico 34%; Fundamenta consideraciones y valoraciones de otros profesionales 37% y Fundamenta manera de recolectar información 38%. Los ítems mejor explicados por el modelo, en la instancia febrero 2017 fueron: Fundamenta manera de abordar el caso 65%; Fundamenta manera de abordar el caso 64%; e Integra datos y fuentes 62% (ver tabla 4).

Tabla 4*Comunalidades febrero 2017 -Método de extracción: máxima probabilidad*

Criterio de evaluación	Inicial	Extracción
Fundamenta manera de recolectar información	.41	.38
Integra datos y fuentes	.64	.62
Fundamenta manera de abordar el caso	.64	.64
Conceptualiza el objeto/foco de análisis	.56	.53
Coherencia teórico-práctica con hipótesis	.55	.55
Fundamenta manera de abordar el caso	.63	.65
Coherencia teórica con hipótesis	.61	.53
Fundamenta estrategia de intervención	.64	.58
Fundamenta consideraciones de juicio ético	.33	.31
Fundamenta consideraciones y valoraciones de otros profesionales	.42	.37
Lenguaje claro, formal y técnico.	.37	.34

Nota 1: Fuente salida resultados SPSS.

Nota 2: cada criterio corresponde a los desempeños evaluados, tal como se muestra en la tabla 1.

En las comunales de la rúbrica oral versión 2018, los ítems que se explican en una menor medida por el modelo son: Fundamenta consideraciones de juicio ético 24%; Fundamenta consideraciones y valoraciones de otros profesionales 36%; Lenguaje claro, formal y técnico 37% y Fundamenta manera de recolectar información 39%. Los ítems mejor explicados por el modelo en la instancia febrero 2018 fueron: Fundamenta manera de abordar el caso 71%; Coherencia teórico-práctica con hipótesis 64% y Fundamenta estrategia de intervención 62% (ver tabla 5).

Tabla 5

Comunalidades febrero 2018 -Método de extracción: máxima probabilidad

Criterio de evaluación	Inicial	Extracción
Fundamenta manera de recolectar información	.45	.39
Integra datos y fuentes	.62	.56
Fundamenta manera de abordar el caso	.61	.61
Conceptualiza el objeto/foco de análisis	.60	.57
Coherencia teórico-práctica con hipótesis	.64	.64
Fundamenta manera de abordar el caso	.68	.71
Coherencia teórica con hipótesis	.61	.55
Fundamenta estrategia de intervención	.66	.62
Fundamenta consideraciones de juicio ético	.29	.24
Fundamenta consideraciones y valoraciones de otros profesionales	.39	.36
Lenguaje claro, formal y técnico.	.38	.37

Nota 1: Fuente salida resultados SPSS.

Nota 2: cada criterio corresponde a los desempeños evaluados, tal como se muestra en la tabla 1.

A través del método de extracción de factores por máxima verosimilitud, se encontró que, los resultados en la versión febrero 2017 de la rúbrica mostraron una escala con un autovalor de 5.61 el cual explica un 51,05% de la varianza (ver tabla 6), y en la versión febrero 2018 se formó una escala con autovalor de 5.62 que explica el 51.05% de la varianza (ver tabla 8). Luego de explorar el gráfico de sedimentación (ver Figura 1 y 2) se observó que en ambas instancias los valores están concentrados alrededor de un solo factor.

Las cargas factoriales de los ítems alcanzaron valores entre 0.55 y 0.80 en la versión febrero 2017 y entre 0.49 y 0.84 en la versión febrero 2018 (ver tabla 10 y 11).

En las instancias febrero 2017 y febrero 2018, se calculó la bondad de ajuste del modelo de un factor, por medio de una prueba de hipótesis con una distribución χ^2 , donde se encontró que la significatividad asociada fue de 0 para ambos momentos, lo que permite verificar el ajuste de los modelos.

Tabla 6*Varianza total explicada versión febrero 2017*

Factor	Autovalores iniciales			Sumas de extracción de cargas al cuadrado		
	Total	% de varianza	% acumulado	Total	% de varianza	% acumulado
1	6.08	55.30	55.30	5.615	51.05	51.05
2	.94	8.57	63.86			
3	.83	7.54	71.40			
4	.60	5.50	76.90			
5	.55	4.99	81.89			
6	.49	4.48	86.38			
7	.40	3.60	89.98			
8	.33	2.95	92.93			
9	.30	2.75	95.69			
10	.28	2.52	98.21			
11	.20	1.79	100.000			

Nota: Método de extracción: máxima probabilidad -Fuente salida resultados SPSS

Tabla 7*Prueba de bondad de ajuste*

Chi-cuadrado	gl	Sig.
450.79	44	<0.001

Nota: Fuente salida resultados SPSS

Tabla 8

Varianza total explicada versión febrero 2018

Factor	Autovalores iniciales			Sumas de extracción de cargas al cuadrado		
	Total	% de varianza	% acumulado	Total	% de varianza	% acumulado
1	6.08	55.30	55.30	5.62	51.05	51.05
2	.94	8.57	63.86			
3	.83	7.54	71.40			
4	.60	5.50	76.90			
5	.55	4.99	81.89			
6	.49	4.48	86.38			
7	.40	3.60	89.98			
8	.33	2.95	92.93			
9	.30	2.75	95.69			
10	.29	2.52	98.21			
11	.20	1.79	100.000			

Método de extracción: máxima probabilidad

Nota: Fuente salida resultados SPSS

Tabla 9

Prueba de bondad de ajuste

Chi-cuadrado	gl	Sig.
482.86	44	<0.001

Nota: Fuente salida resultados SPSS

Tabla 10*Matriz factorial versión febrero 2017*

Criterio de evaluación	Factor 1
Fundamenta manera de abordar el caso	.80
Fundamenta manera de abordar el caso	.80
Integra datos y fuentes	.78
Fundamenta estrategia de intervención	.76
Coherencia teórico-práctica con hipótesis	.74
Conceptualiza el objeto/foco de análisis	.73
Coherencia teórica con hipótesis	.73
Fundamenta manera de recolectar información	.61
Fundamenta consideraciones y valoraciones de otros profesionales	.61
Lenguaje claro, formal y técnico.	.59
Fundamenta consideraciones de juicio ético	.55

Método de extracción: máxima verosimilitud

Nota 1: Fuente salida resultados SPSS

Nota 2: cada criterio corresponde a los desempeños evaluados, tal como se muestra en la tabla 1.

Tabla 11*Matriz factorial versión febrero 2018*

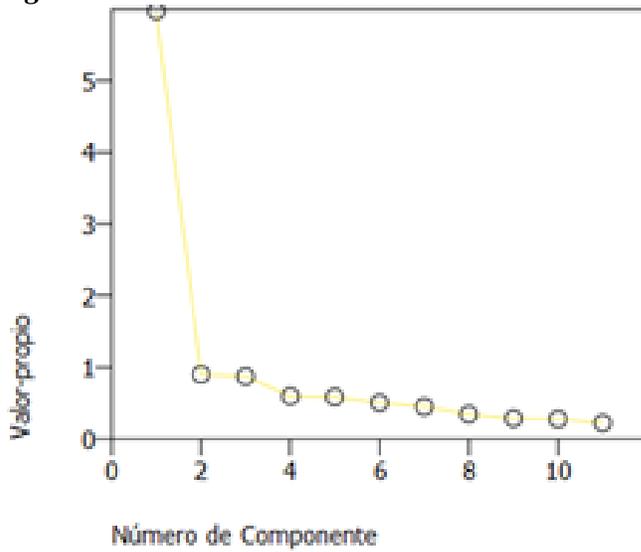
Criterio de evaluación	Factor 1
Fundamenta manera de abordar el caso	.84
Coherencia teórico-práctica con hipótesis	.80
Fundamenta estrategia de intervención	.79
Fundamenta manera de abordar el caso	.78
Conceptualiza el objeto/foco de análisis	.76
Integra datos y fuentes	.75
Coherencia teórica con hipótesis	.74
Fundamenta manera de recolectar información	.63
Lenguaje claro, formal y técnico.	.61
Fundamenta consideraciones y valoraciones de otros profesionales	.60
Fundamenta consideraciones de juicio ético	.49

Método de extracción: máxima verosimilitud.

Nota 1: Fuente salida resultados SPSS

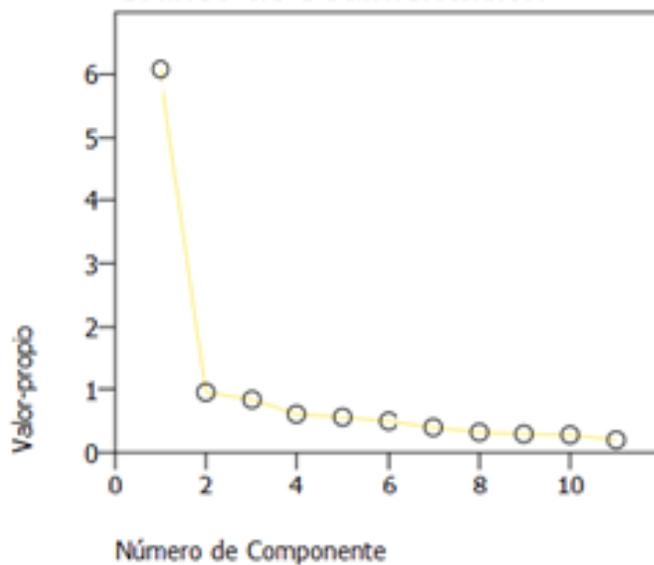
Nota 2: cada criterio corresponde a los desempeños evaluados, tal como se muestra en la tabla 1.

Figura 1. Sedimentación de la Rúbrica versión febrero 2017



Fuente: salida resultados SPSS.

Figura 2. Sedimentación de la Rúbrica versión febrero 2018



Fuente: salida resultados SPSS.

Validez de contenido

Se analizó mediante la prueba estadística Kappa de Cohen, índice Kappa, que corresponde a una medida estadística que ajusta el efecto del azar en la proporción de concordancia interobservador (Cerde & Villarroel, 2008), este índice se calculó con el puntaje asignado por los evaluadores a cada ítem. Considerando que la comisión evaluadora estuvo compuesta por tres docentes, el cálculo del estadístico se realizó entre: evaluador 1 y evaluador 2; evaluador 1 y evaluador 3; evaluador 2 y evaluador 3, y finalmente se promediaron estos índices y se obtuvo el valor kappa por ítem de la rúbrica.

Para la interpretación del índice Kappa se consideraron los siguientes valores de referencia: fuerza de concordancia de 0,01 a 0,20 leve, 0,21 a 0,40 aceptable, 0,41 a 0,60 moderada, 0,61 a 0,80 considerable y 0,81 a 1,00 casi perfecta (Landis & Koch, 1977, citado en Cerda & Villarroel, 2008). De los once ítems que componen la rúbrica, en la versión febrero 2017 se encontró que siete presentan una fuerza de concordancia considerable, un ítem evidenció una fuerza de concordancia moderada y tres presentaron una fuerza de concordancia aceptable. En la versión febrero 2018 nueve ítems presentaron una fuerza de concordancia aceptable y dos evidenciaron una fuerza de concordancia inter evaluador moderada (ver Tabla 12).

Tabla 12

Índice Kappa de Cohen por ítems de la rúbrica en sus dos versiones

Ítems	Índice Kappa en versión febrero 2017	Índice Kappa en versión febrero 2018
Fundamenta manera de recolectar información	.41	.34
Integra datos y fuentes	.48	.39
Fundamenta manera de abordar el caso	.37	.42
Conceptualiza el objeto/foco de análisis	.67	.41
Coherencia teórico-práctica con hipótesis	.35	.34
Fundamenta manera de abordar el caso	.77	.35
Coherencia teórica con hipótesis	.69	.36
Fundamenta estrategia de intervención	.73	.33
Fundamenta consideraciones de juicio ético	.72	.38
Fundamenta consideraciones y valoraciones de otros profesionales	.73	.32
Lenguaje claro, formal y técnico	.75	.32

Nota 1: fuente de elaboración propia

Nota 2: cada ítem (criterio) corresponde a los desempeños evaluados, tal como se muestra en la tabla 1.

Consistencia interna

Se utilizó el coeficiente Alfa para determinar la confiabilidad de la rúbrica por consistencia interna, es decir, para determinar el grado en que todos los ítems de la rúbrica correlacionan entre sí (Oviedo & Campo-Arias, 2005). Se obtuvo un valor de .91 para las dos versiones de la rúbrica (febrero 2017 y febrero 2018) que indica una alta fiabilidad para el instrumento completo de 11 ítems en ambas versiones, considerándose valores por sobre 0.70 como una buena consistencia interna (Oviedo & Campo-Arias, 2005). Se realizó el análisis de correlación de cada ítem respecto al cuestionario encontrándose altos valores de correlación sobre .50, donde los valores menores fluctuaron entre .50 y .60 esto en los últimos tres ítems de la rúbrica, los que corresponden a desempeños genéricos del perfil de egreso. En el análisis de eliminar cada ítem del cuestionario y verificar la modificación en la consistencia interna, se encontró que todos los ítems son estadísticamente válidos (ver Tabla 13).

Tabla 13

Coefficiente Alfa de Cronbach por ítem de la rúbrica en sus dos versiones

Ítems	Correlación ítem/test		Alfa de Cronbach estimado, si se elimina el elemento	
	febrero 2017	febrero 2018	febrero 2017	febrero 2018
Fundamenta manera de recolectar información	.60	.60	.91	.91
Integra datos y fuentes	.74	.72	.90	.91
Fundamenta manera de abordar el caso	.76	.75	.90	.91
Conceptualiza el objeto/foco de análisis	.68	.71	.91	.91
Coherencia teórico-práctica con hipótesis	.69	.75	.91	.91
Fundamenta manera de abordar el caso	.76	.79	.90	.90
Coherencia teórica con hipótesis	.69	.70	.91	.91
Fundamenta estrategia de intervención	.72	.74	.90	.91
Fundamenta consideraciones de juicio ético	.55	.50	.91	.92
Fundamenta consideraciones y valoraciones de otros profesionales	.60	.59	.91	.91
Lenguaje claro, formal y técnico	.57	.59	.91	.91

Nota 1: fuente de elaboración propia

Nota 2: cada ítem (criterios) corresponde a los desempeños evaluados, tal como se muestra en la tabla 1.

CONCLUSIONES

En lo siguiente se concluye respecto de las evidencias psicométricas del instrumento de evaluación, y de la coherencia de la rúbrica con el perfil de egreso y la perspectiva de evaluación auténtica que la sustenta. La expresión oral es una instancia de evaluación, valorada en educación superior en tanto permite evidenciar desempeños en forma integrada. En la parte oral del examen de licenciatura, se busca evaluar desempeños diagnósticos, de interpretación y abordaje de situaciones profesionales en contextos específicos, por tanto, para valorar lo que ocurre en esta instancia se requiere de juicios consistentes y concordantes entre sí, sobre todo considerando que, en este momento la evaluación cuenta con el juicio de una comisión (3 docentes expertos de un área de la disciplina). Para lo anterior se requiere un instrumento que ordene, oriente, estructure y/o matice los juicios, en la línea de lo que se busca evaluar en el examen de licenciatura.

La carrera de Psicología USS optó por una metodología de evaluación del examen de licenciatura que incluyera el análisis de caso abierto y la rúbrica analítica. El primero debido a que este tipo de caso no pre diseña una respuesta única como correcta (no existiendo respuestas correctas o incorrectas), sino que respeta la postura del estudiante en función de las decisiones que toma para abordar la situación disciplinar presentada. La segunda, ya que la resolución del caso permite evidenciar el despliegue de saberes, habilidades relacionales y procedimentales en situaciones concretas y de forma contextualizada por parte de los estudiantes, lo que es calificado a través de una rúbrica analítica construida en base a niveles y criterios. Con base a lo anteriormente expuesto, es posible señalar que la carrera de Psicología de la USS implementa, en la parte oral de su examen de Licenciatura, una metodología de evaluación que responde a los principios de la Evaluación Auténtica para evaluar su perfil de egreso.

Las características de esta evaluación -individual, oral, basada en el análisis de caso abierto, que incluye como dispositivo de calificación una rúbrica analítica y se retroalimenta individualmente en dos instancias, se condicen con el nuevo concepto de Formación Universitaria, es decir, con una mirada del proceso de enseñanza que incorpora la evaluación desde una perspectiva formativa y no solo acreditativa.

La rúbrica construida es válida y confiable como instrumento de evaluación, en el marco de la evaluación auténtica del perfil de egreso basado en desempeños esperados de la carrera Psicología de la Universidad San Sebastián. Lo anterior, debido a que en el estudio se demuestra que es posible obtener evidencias de validez y confiabilidad (.91) en instrumentos de respuestas pre construidas, considerando las diferencias esperables entre los juicios de los evaluadores, docentes de la comisión. En este marco, es importante decir que el proceso de validación del instrumento propuesto cumplió con los elementos y requerimientos usuales para este tipo de instrumentos, de esta manera cumple con ser una guía que orienta a la comisión en la valoración de las respuestas de los estudiantes puestos en situación profesional. Permite evaluar de manera objetiva el abordaje de una situación profesional, que no necesariamente se aborda de manera estandarizada, son respuestas abiertas donde la subjetividad tiene lugar, tanto desde el punto de vista del estudiante que rinde examen, como de los integrantes de la comisión que lo evalúa. El análisis cuantitativo de los juicios subjetivos de los jueces le da a este estudio un carácter multimétodo.

Los resultados de confiabilidad y validez de la rúbrica han sido obtenidos en dos de las versiones anteriores del examen, presentando en ambas versiones, validez de contenido, basada en la estructura interna y demostró ser confiable en la globalidad, así como en sus dimensiones.

La rúbrica analítica permite contar con antecedentes para realizar una retroalimentación reflexiva de la actuación del estudiante en la reunión individual programada para ello, permitiendo que éste visualice la brecha entre su desempeño y el desempeño profesional esperado, lo que facilita la incorporación de estándares de la praxis profesional.

Como reflexiones finales, si bien las características del examen y el tipo de rúbrica utilizada se orientan a la evaluación de un perfil de egreso diseñado en base a desempeños esperados, también es posible que, en estas mismas condiciones, se consiga evaluar el perfil de egreso de una carrera, cuyo diseño se base en competencias.

REFERENCIAS

- Ahumada, P. (2005). *Hacia una evaluación auténtica del aprendizaje*. México: Editorial Paidós.
- Alsina, J. (Coord). (2013). *Rúbricas para la evaluación de competencias*. Cuadernos de Docencia Universitaria, p.26. Barcelona: ICE y Octaedro.
- Arenas, A. & Gómez, K. (2013). *Las rúbricas o matrices de valoración, herramientas de planificación e implementación de una evaluación por desempeños*. UIS Ingenierías, 12(1), pp.81-87.
- Buscá F., Suárez, M., Bursset, S. & Bosch, E. (2014). *La credibilidad de los sistemas de evaluación formativa en docencia universitaria: un estudio de casos múltiple aplicado a la formación del profesorado*. Enseñanza & Teaching, 32, (2), pp.177-193. DOI: [10.14201/et2014321177193/](https://doi.org/10.14201/et2014321177193/).
- Cerda, J. & Villarroel, L. (2008). *Evaluación de la concordancia inter-observador en investigación pediátrica: coeficiente de Kappa*. Revista Chilena de Pediatría, 79(1), pp.54-58. DOI: [10.4067/S0370-41062008000100008/](https://doi.org/10.4067/S0370-41062008000100008/).
- Cubillos-Vega, C. & Ferrán-Aranaz, M. (2019). *Los efectos de una intervención educativa en materia de derechos humanos con alumnado de Trabajo Social: la valoración de las nociones y el compromiso*. Cuadernos De Trabajo Social, 32(2), pp.381-396. DOI: [10.5209/cuts.60617/](https://doi.org/10.5209/cuts.60617/).
- De la Cruz, G. (2011). *La construcción y aplicación de rúbricas: una experiencia en la formación de psicólogos educativos*. Observar, 5, pp.21-41. DOI: [10.21556/edutec.2013.43.3/](https://doi.org/10.21556/edutec.2013.43.3/).
- Denyer, M., Furnémont, J., Poulain, R. & Vanloubbeeck, G. (2007). *Las competencias en la educación*. Un balance. México: Fondo de Cultura Económica.
- Díaz, A. (2011). *Competencias en educación. Corrientes de pensamiento e implicaciones para el currículo y el trabajo en el aula*. Revista Iberoamericana de Educación Superior 2 (5).

Recuperado de: <http://ries.universia.net/index.php/ries/article/view/126/>.

- Föster, C. (2017). *El poder de la evaluación en el aula. Mejores decisiones para promover aprendizajes*. Ediciones UC.
- Gallardo, K., Gil, M. & Govea, A. (2016). *Evaluar el desempeño en el marco del modelo educativo basado en competencias en educación superior con apoyo de redes sociales: un estudio de caso*. *Aula*. (0214-3402), 22, 303-319. DOI: 10.14201/aula201622303319.
- Gromi, A. (2013). *Juzgar. De la opinión no fundamentada al juicio elaborado*. Madrid: Narcea.
- Lloret-Segura, S., Ferreres-Traver, A., Hernández-Baeza, A. & Tomáa-Marco, I. (2014). El Análisis Factorial Exploratorio de los Ítems: una guía práctica, revisada y actualizada. *Anales de Psicología*, 30(3), 1151-1169. Recuperado de: http://scielo.isciii.es/scielo.php?script=sci_abstract&pid=S0212-97282014000300040/.
- López, M. (2011). Evaluación Auténtica de los aprendizajes en Historia y Ciencias Sociales: contexto, problemáticas y alcances metodológicos. *Horizontes Educativos*, 16(1), 73-92. Recuperado de: <http://www.redalyc.org/articulo.oa?id=97923680007>.
- Lòpez, J. (Coord). (2013). *Rúbricas para la evaluación de competencias*. Cuadernos de Docencia Universitaria, p.26. Barcelona: ICE y Octaedro.
- Oviedo, H. & Campo-Arias, A. (2005). *Aproximación al uso del coeficiente alfa de Cronbach*. *Revista Colombiana de Psiquiatría*, 34(4), pp.572-580. Recuperado de: http://www.scielo.org.co/scielo.php?script=sci_arttext&pid=S003474502005000400009&lng=en&tlng=.
- Perrenoud, P. (2004). *Diez nuevas competencias para enseñar*. Invitación al viaje. Barcelona: Graó & México, Secretaría de Educación Pública.
- Rojas, X. (2017). *Orientaciones generales para la construcción de ítems de selección múltiple y caso*. Manuscrito inédito. Facultad de Psicología, Universidad San Sebastián, Chile.

Universidad San Sebastián. (2015). *Orientaciones al proceso de Evaluación para el Aprendizaje, Vicerrectoría Académica, Dirección de Pregrado*. (Documento interno). Chile: Universidad San Sebastián.

Universidad de San Sebastián-Facultad de Psicología USS (2017). *Informe de Autoevaluación. Carrera de Psicología USS*. (Documento interno). Chile: Universidad San Sebastián.

Vallejo, M. & Molina, J. (2014). *La evaluación auténtica de los procesos educativos*. [Monografía]. Revista Iberoamericana de Educación. 64, pp.11-25. Recuperado de: <https://rieoei.org/historico/documentos/rie64a01.pdf>