

Metodología para validar un instrumento de evaluación por competencias en estudiantes de psicología*

Methodology to Validate a Competencies Assessment Tool for Psychology Students

Recibido: 29 de septiembre de 2015 | Aceptado: 05 de noviembre de 2015

ANA ROSENBLUTH**
CLAUDIA CRUZAT-MANDICH***
MARÍA LUISA UGARTE****
Universidad Adolfo Ibáñez, Chile

RESUMEN

Se presenta una metodología para validar una estrategia de evaluación por competencias asociadas al perfil de alumnos de psicología. Mediante un estudio cuantitativo de diseño no experimental-correlacional, con la participación del cuerpo docente de la carrera de psicología, se elaboró un instrumento con base en un grupo de competencias seleccionadas. El proceso de validación se llevó a cabo aplicándolo a estudiantes de segundo año de la carrera. Los resultados muestran su validez y contribuyen a la investigación en el diseño de estrategias para la elaboración y validación de instrumentos que permitan determinar si efectivamente las estrategias de enseñanza-aprendizaje están logrando las competencias propuestas en la formación de los alumnos.

Palabras clave

validación; instrumento; competencias; educación universitaria; formación de psicólogos

ABSTRACT

This paper presents a methodology to validate an instrument for the assessment of competencies associated to a psychology students' profile. Using a quantitative non- experimental- correlational design, with the participation of the school of psychology faculty, a tool was developed based on a group of selected competencies. The validation process was conducted by applying the instrument to second year psychology students. Results demonstrate its validity and contribute to the research on the development and validation of instruments designed to determine whether teaching - learning strategies are achieving the proposed competencies in students' education.

Keywords

validation; instrument; competencies; undergraduate education; Psychology formation

doi: 10.11144/Javeriana.upsy15-1.pppm

Para citar este artículo: Djeriouat, H., & Mullet, E. (2013). Public perception of the motives that lead political leaders to launch interstate-armed conflicts: A structural and cross-cultural study. *Universitas Psychologica*, 15(1), 303-314. <http://dx.doi.org/10.11144/Javeriana.upsy15-1.pppm>

* Artículo de investigación. No tuvo financiación de alguna entidad.

** Directora Docencia y Departamento de Psicología Social, Escuela Psicología. Correo electrónico: ana.rosenbluth@uai.cl

*** Docente y Directora CECA (Centro Estudio Conductas Alimentarias) Escuela Psicología. Correo electrónico: claudia.cruzat@uai.cl

Introducción

Este trabajo pretende contribuir en la construcción de instrumentos que permitan establecer mecanismos de retroalimentación sobre los logros en los procesos de enseñanza-aprendizaje de los alumnos. Específicamente aquí se describe el desarrollo de una metodología para validar un instrumento que evalúa una selección de competencias definidas en el perfil de la carrera de psicología.

Durante los últimos 20 años se registra un número creciente de investigaciones que dan cuenta del cambio en los modelos educativos de la educación superior hacia un enfoque basado en competencias. Este enfoque comenzó en el año 2000 como un proyecto de transformación educativa en Europa con la declaración de Bolonia y el proyecto Tuning Europeo (Tuning, 2006). Posteriormente se transfiere a Latinoamérica con el proyecto Tuning Alfa en el 2004, en el cual participaron varias universidades chilenas (Manríquez, 2012). Junto con esto, en el área de la psicología, la Asociación Americana de Psicología (APA, 2013) establece metas para la educación de pregrado con base en competencias.

El desarrollo de un modelo de educación basado en competencias establece un cambio en el rol que la educación superior tiene en la formación de profesionales que puedan enfrentar los nuevos desafíos y necesidades de la sociedad, junto con nuevas propuestas educativas que asignan un papel activo al estudiante en su proceso de aprendizaje, que apuntan a un tipo de aprendizaje profundo y significativo, y al logro de ciertas habilidades específicas y evaluables.

En general, se utiliza el término competencia en dos amplios sentidos: 1) para definir ámbitos de acción específicos y particulares al rol profesional y 2) para referirse a la formación de un conjunto de habilidades, conocimientos y comportamientos en el ámbito académico. Aquellas competencias asociadas al perfil profesional tuvieron un amplio desarrollo desde los años 60 y hasta la actualidad en diversos campos, especialmente en las áreas de la salud como medicina y enfermería, así como en psicología (Bartram & Roe, 2005; Fouad et al., 2009; Hunsley & Barker, 2011; Roe, 2002). Rodolfa

et al. (2005) definen a una persona competente como un profesional que está calificado, es capaz y puede comprender y hacer ciertas cosas de una manera efectiva y apropiada. Villa y Poblete (2011) enfatizan el rol articulador de la competencia en términos de integrar una serie de aspectos que la persona requiere para analizar y resolver una situación o problema particular. Roe (2002) plantea que la competencia es una capacidad aprendida que se evidencia en la ejecución de tareas específicas en contextos particulares y que se manifiesta en la realización adecuada de una tarea.

Las competencias en el ámbito académico, sin embargo, no son necesariamente homologables a las que se requieren en el perfil profesional (Beneitone, 2007; Blömeke, Gustafsson, & Shavelson, 2015) y, por lo tanto, han tenido un desarrollo diferente. En este ámbito, la competencia tiene que ver con un estilo de aprendizaje que permite resolver situaciones futuras y, sobre todo, con un cambio paradigmático en el rol central que se le otorga al alumno en este proceso (Villa & Poblete, 2011).

Los procesos de cambio hacia un modelo de competencias no están exentos de dificultades. Por una parte, la ya mencionada variedad en las definiciones pero, por otra, las dificultades que conlleva el proceso de implementación en el cual el diseño de los procesos de formación basados en competencias deben ser coherentes con las definiciones de perfil de egreso, con los procesos de enseñanza-aprendizaje y con los procesos evaluativos (Bartram & Roe, 2005; González, Arquero, & Hassall, 2014; Roe, 2002; Villarroel & Bruna, 2014). González et al. (2014) realizan un análisis de caso donde observan los factores que impulsan, facilitan y dificultan los diferentes estadios de un proceso de cambio hacia una formación por competencias. La claridad de los objetivos, recursos y apoyo al sistema de enseñanza son los que impulsan y facilitan el proceso y aquellos que lo dificultan son la falta de capacitación del profesorado en los temas de competencias y la ambigüedad en la formulación y definición de éstas.

Por otra parte, el gran tema pendiente del cambio de los procesos de enseñanza-aprendizaje hacia modelos basados en competencias son los mecanismos de evaluación y medición (Charria, Sarsosa, &

Arenas, 2011). Los métodos empíricos para relacionar competencias con resultados observables son infrecuentes (Blömeke, Zlatkin-Troitschanskaia, Kuhn, & Fege, 2013; Centro Interuniversitario de Desarrollo (CINDA), 2009; Jorquera, 2013; Ryan, Spencer, & Bernhard, 2012; Tremblay, Lallancette, & Roseveare, 2012; Tuning, 2006). Entre aquellos que lo han hecho figuran, por ejemplo, la elaboración y validación de instrumentos para evaluar competencias de docentes universitarios de pregrado (Camargo-Escobar & Pardo-Adames, 2008), otros que analizan el logro de competencias básicas en economía y finanzas en consumidores universitarios (Vargas & Avendaño, 2014) y aquellos que evalúan competencias emocionales, cognitivas y sociales asociadas a beneficios o ganancias empresariales (Ryan et al, 2012). Sin embargo, son escasos los ejemplos de aplicación de instrumentos que permitan evaluar logros de aprendizaje asociados a competencias, sobre todo en el ámbito de las genéricas transversales.

Aun cuando es difícil establecer criterios comunes en la conceptualización de las competencias, el desarrollo de procesos de evaluación y validación constantes permiten retroalimentar al cuerpo docente en la efectividad de los procesos de enseñanza-aprendizaje asociados a los objetivos del programa, la coherencia y comunicación entre las asignaturas y el cumplimiento del perfil asociado a las metas de los diferentes ciclos del progreso académico. La necesidad de contar con estas medidas permite visualizar el cambio curricular como un proceso de gestión del cambio desde una mirada sistémica y con la participación activa del cuerpo docente, orientada hacia el aseguramiento de la calidad en la formación académica de los alumnos.

Tal como sugiere APA (2013), hemos definido dos grandes ciclos en la progresión académica de los alumnos. El primer ciclo representa el cumplimiento de los cursos de primer y segundo año de la carrera. El segundo se refiere a las metas de logro esperadas para quienes terminan el ciclo formativo a nivel de cuarto año. Los lineamientos para el programa de pregrado en psicología de la APA (2013) capturan un conjunto de expectativas óptimas de rendimiento que se deben cumplir al terminar un programa de pregrado en psi-

cología. Estos lineamientos describen un conjunto de metas y resultados de aprendizaje para el programa de pregrado en psicología, los cuales sirven como base para planes de evaluación. Las metas de aprendizaje planteadas por APA hacen referencia a cinco grandes ámbitos: 1) *conocimientos básicos de la psicología* (los estudiantes demostrarán familiaridad con los conceptos principales, perspectivas teóricas, hallazgos empíricos y tendencias históricas de la psicología); 2) *investigación científica y pensamiento crítico* (los estudiantes entenderán y aplicarán métodos de investigación básicos en psicología, así como utilizarán el pensamiento crítico, la indagación escéptica y el abordaje científico para resolver los problemas relacionados el comportamiento y los procesos mentales); 3) *responsabilidad ética y social en un mundo diverso* (los estudiantes serán capaces tolerar la ambigüedad, actuar éticamente y reflejar otros valores que son los fundamentos de la psicología como disciplina en un mundo que involucra crecientemente diversidad); 4) *comunicación* (los estudiantes demostrarán competencia y habilidades comunicacionales tanto en forma oral como escrita), y 5) *desarrollo profesional* (los estudiantes serán capaces de trabajar en equipo, autogestionarse y ser reflexivos, en función de aplicar los contenidos y habilidades psicológicas a objetivos profesionales). Estas metas y resultados de aprendizaje representan la guía de expectativas que debiese tener el cuerpo docente en la formación de los psicólogos para cada uno de los ciclos señalados previamente. En concordancia con los lineamientos de APA, nosotros hemos planteado una serie de competencias muy similares a nivel de primer ciclo (véase Tabla 1). El presente artículo presenta una metodología para validar un instrumento de evaluación por competencias asociadas al perfil del primer ciclo de la carrera de psicología.

Método

Se describen dos grandes etapas. La primera relacionada con el proceso de elaboración del instrumento y la segunda, con su validación. En esta segunda etapa se utilizó un diseño cuantitativo no experimental-correlacional orientado a evaluar la validez y confiabilidad de un instrumento de competencias.

Participantes

La construcción del instrumento contó con la participación de diez docentes de primer y segundo año de la carrera de psicología quienes, a partir de un grupo de competencias definido en el perfil para el estudiante de este nivel, elaboraron un conjunto de preguntas para cada una de éstas.

El instrumento se aplicó a 49 alumnos de dos sedes, 41 mujeres y ocho hombres, cuyas edades fluctuaban entre los 20 y los 24 años, con un promedio de 22 años ($DS = 1.96$ años). Todos los alumnos habían completado todas las asignaturas del primer ciclo (primer y segundo año) al momento de la evaluación.

Procedimiento

Selección de las competencias

En una primera instancia, cada profesor de primer y segundo año revisó sus programas y seleccionó, del conjunto de competencias definidos por la Escuela de Psicología, tres que esperaba desarrollar y evaluar en sus asignaturas. Se seleccionaron las cinco competencias mencionadas con mayor frecuencia en todos los cursos (Tabla 1).

Posteriormente se definieron dos formas para evaluar estas cinco competencias: una modalidad escrita e individual para tres de éstas: 1) *conceptualización del fenómeno psicológico*, 2) *pensamiento crítico* y 3) *análisis comprensivo desde la teoría psicológica*; y una modalidad oral – grupal para otras dos: 4) *trabajo en equipo* y 5) *capacidad de investigar*.

Construcción del instrumento

Para la fase escrita se aplicó la metodología del estudio de caso, la cual ha sido ampliamente utilizada en el aprendizaje de competencias y en la evaluación de habilidades complejas (Sánchez, Santos, Fuentes, & Núñez, 2015). Esta metodología permite el desarrollo de pensamiento y capacidad crítica de los estudiantes. Para evaluar las competencias de *análisis comprensivo*, *pensamiento crítico* y *conceptualización del fenómeno psicológico* se construyó un caso *ad hoc* con base en una revisión bibliográfica que exponía puntos coincidentes y contradictorios establecidos por la literatura en torno al tema del *bullying*.

Para la construcción de las preguntas en relación al caso se les pidió a diez docentes de primer y segundo año que las elaboraran con base en una pauta que establecía las características de éstas, y las definiciones de las competencias de esta fase.

TABLA 1
Competencias seleccionadas y sus definiciones

<p>1. Conceptualización del fenómeno psicológico: a. Capacidad de distinguir y conceptualizar los fenómenos psicológicos b. Capacidad de abstraer y relacionar conceptos para describir, comprender y explicar un fenómeno psicológico</p>
<p>2. Análisis comprensivo desde la teoría psicológica: a. Desagregar un problema de manera de encontrar patrones b. Jerarquizar las aproximaciones y los distintos componentes desde la teoría psicológica</p>
<p>3. Pensamiento crítico: a. Reconocer los modelos de la disciplina b. Argumentar fundadamente</p>
<p>4. Capacidad de investigar en forma guiada: a. Diseñar proyectos de investigación aplicando el método científico concordante con el paradigma desde el que se sitúa</p>
<p>5. Trabajo en equipo: a. Trabaja en pos de metas colectivas b. Administra roles y tiempos c. Desarrolla compromiso y fidelidad hacia el grupo</p>

Fuente: elaboración propia

Producto de esta etapa se generó un banco de 23 preguntas.

Validez de contenido

Se utilizaron los siguientes criterios para evaluar la validez de contenido de las preguntas: (a) nivel de dificultad, (b) calidad, (c) claridad, (d) redacción (e) si contiene información necesaria y suficiente para responder adecuadamente, (f) si puede ser contestada por sentido común, (g) si permite discriminar el desempeño de los alumnos en función de las competencias, (h) si permite que el alumno haga inferencias, entendida como la relación lógica entre la pregunta y el texto y finalmente (i) nivel de adecuación de la pregunta a las competencias.

Los jueces expertos que evaluaron las preguntas fueron diez profesores representantes del área clínica, organizacional y básica que no habían formado parte de la formulación de las preguntas. Este grupo de profesores ha sido parte del desarrollo del currículo por competencias. Dada la gran variedad de interpretaciones y definiciones en el campo de competencias, se pensó que lo más apropiado para evaluar la validez de las preguntas era a través del mismo cuerpo académico que ha sido parte en la definición de éstas. Por otra parte, diez jueces probaron ser suficientes para una estimación confiable de la validez de contenido (Hyrkäs, Appelqvist-Schmidelchner, & Oksa, 2003).

Para evaluar las preguntas se utilizó la metodología de Lynn (1986), que permite evaluar la validez de contenido, es decir, proporción de expertos que evalúan bien un ítem. Para esto se obtuvo el promedio simple del porcentaje de profesores que consideraron que la pregunta cumplía con los criterios explicados más arriba. Así, una pregunta con un índice = 0.9 quería decir que, en promedio, el 90% de los profesores que la evaluaron consideraron que era adecuada en dichos criterios. Se consideraron para la prueba solo preguntas con un índice de calidad igual o superior a 0.85. Basados en ese criterio, se seleccionaron cuatro de las 23 preguntas originales.

La fase oral se diseñó como un trabajo de grupo de cinco alumnos y tres docentes: dos observadores

y un ministro de fe. El estímulo para la interacción de grupo fue generar un proyecto de investigación a partir de un problema relacionado con el tema del *bullying* en un tiempo determinado. Para evaluar las interacciones se construyeron tres pautas: una autoevaluación individual, evaluación de pares y evaluación de la comisión, en función de las competencias definidas para esta fase: *capacidad de investigar y trabajo en equipo*.

Aplicación del instrumento

Para la fase escrita los alumnos respondieron las cuatro preguntas seleccionadas. Estas preguntas entregaban información acerca de cierto concepto psicológico y requerían integrar esa información con una revisión bibliográfica sobre *bullying* entregada a los alumnos y que debían leer antes de responder. Asimismo, el día anterior a la evaluación, se les entregó un listado de publicaciones académicas respecto del tema asociado al caso. Por ejemplo, la pregunta 1 se refería al concepto de *identidad* y solicitaba a los alumnos discutir la relación entre dicho concepto en adolescentes y cierta información entregada en la revisión.

Los alumnos debieron responder en un computador de forma individual las cuatro preguntas seleccionadas en un plazo de cuatro horas. Estas respuestas fueron subidas por los alumnos a la plataforma de la universidad (Moodle-Webcursos). Las respuestas fueron evaluadas por separado (tres docentes por pregunta), es decir, un total 12 docentes de ambas sedes participaron de la corrección (representando las áreas clínica (4), básica (6) y organizacional (2) (6 mujeres y 6 hombres).

Las respuestas fueron evaluadas usando una escala de 1 a 7, considerando las tres competencias eje, más aspectos formales como ortografía, estructura de la respuesta, lenguaje y ajuste a las normas APA.

Cada una de las preguntas fue evaluada en función de doce definiciones operacionales por cada uno de tres docentes en un equipo (cada docente evaluó la misma pregunta para los 49 alumnos). Estas evaluaciones fueron hechas de manera independiente y los docentes eran *ciegos* a la identidad

del alumno que estaban evaluando. Así, cada pregunta y cada definición operacional recibió una nota de cada docente y se calculó además una nota que correspondía al promedio de los tres docentes (Tabla 2).

Para la fase oral los alumnos fueron asignados aleatoriamente para conformar grupos de cinco, quienes se presentaron ante una comisión compuesta por tres docentes de la Escuela. Uno de los docentes tuvo la responsabilidad de ser ministro de fe, con la tarea de dirigir el proceso y velar por la rigurosidad y estandarización de éste. Tuvo una duración de 90 minutos.

Al llegar a los cinco alumnos se les informó que debían generar un proyecto de investigación relacionado al tema de *bullying* (acerca del cual ya habían leído como preparación a la fase escrita y durante la misma). Para esto, dispusieron de papelógrafo y pizarra, pero no tuvieron acceso a otros materiales. Los alumnos recibieron instrucciones de que disponían de 60 minutos para discutir y preparar una respuesta, y que luego dispondrían de 30 minutos para realizar una presentación. En esa instancia, los docentes de la comisión observaron la interacción grupal y la toma de decisiones y evaluaron cuatro competencias eje en una escala de 1 a 5.

TABLA 2.

Pauta de corrección de las competencias y subcompetencias evaluadas en la fase escrita.

Competencia	Subcompetencia	Operacionalización
Aspectos formales	a. Ortografía	Escribe sin faltas de ortografía y puntuación
	b. Estructura de la respuesta	La respuesta tiene una lógica o hilo conductor, con una introducción o contextualización, luego un desarrollo y finalmente un cierre
	c. Lenguaje	Utiliza un lenguaje formal y académico
	d. Citas	Las citas se ajustan al formato APA
Competencias Eje		
Conceptualizar el fenómeno psicológico	a. Capacidad de distinguir y conceptualizar los fenómenos psicológicos	Identifica, define, y explica diferentes conceptos psicológicos tomados de la misma revisión o de otros marcos teóricos en función de la pregunta
	b. Capacidad de abstraer y relacionar conceptos para describir, comprender y explicar un fenómeno psicológico	Establece relaciones entre estos conceptos para describir, comprender y explicar el problema que se pregunta
Análisis comprensivo desde la teoría psicológica	a. Desagregar un problema de manera de encontrar patrones	Usando las relaciones entre los conceptos, desarrolla un argumento reconocible donde organiza los diferentes elementos de su respuesta La respuesta establece relaciones entre conceptos, entre autores, enfoques, y de distinto tipo de evidencia. Además, integra los conceptos implicados en la pregunta (p. ej., identidad, bullying e identidad sexual)
	b. Jerarquizar las aproximaciones y los distintos componentes desde la teoría psicológica	La respuesta enfatiza y jerarquiza ciertos conceptos, en contraste con constituir una discusión de una lista de conceptos
Pensamiento crítico	a. Reconocer los modelos de la disciplina	Reconoce la información entregada para contextualizar la pregunta como parte de una teoría o cuerpo teórico de la Psicología
	b. Argumentar fundamentamente	La respuesta está fundamentada usando conceptos psicológicos, tomados de la misma revisión o de otros marcos teóricos Los argumentos, ideas, evidencia, están adecuadamente apoyados por citas

Fuente: elaboración propia

Source: own work

Al finalizar, los alumnos salieron de la sala e hicieron una evaluación de sus pares y una autoevaluación en las mismas cuatro competencias eje con base en las pautas entregadas (en una escala de 1 a 5). Luego reingresaron a la sala y recibieron retroalimentación sobre su desempeño de parte de la comisión.

Resultados

Resultados fase escrita

Confiabilidad interjueces

En primer lugar se determinó el grado de acuerdo entre los evaluadores respecto de las notas asignadas a la pauta de evaluación de las competencias. Para esto, se calcularon correlaciones bivariadas (Pearson) para cada pregunta por cada grupo de tres docentes respecto de las notas asignadas para el grupo de 49 alumnos, bajo el supuesto de normalidad. Para cada pregunta se calcularon 36 correlaciones (con tres docentes por equipo se obtienen tres pares de docentes x12 definiciones operacionales = 36 correlaciones) (Tabla 3).

Tomando las cuatro preguntas, en general, hubo correlaciones significativas entre las evaluaciones

de los distintos docentes que corrigieron una misma pregunta, a pesar de que este acuerdo interjueces fue modesto desde el punto de vista de un criterio de confiabilidad convencionalmente aceptado ($r \geq 0.7$). Esto permitió realizar los siguientes análisis considerando los promedios a través de los tres profesores que evaluaron cada una de las cuatro preguntas.

Validez de constructo

En relación a la distribución de las notas de la prueba, la nota promedio total de las 4 preguntas muestra una distribución normal (promedio = 4.32; DS = 0.66; min. = 3.16; máx. = 5.79). De hecho, la prueba Shapiro-Wilk no fue significativa ($p = 0.137$), indicando que la distribución no difiere de la normal.

Desde un punto de vista de consistencia interna, se realizaron correlaciones entre las cuatro preguntas de la prueba y entre las doce definiciones operacionales. En el caso ideal, las cuatro preguntas deberían correlacionarse positivamente entre ellas. Efectivamente, se encontró que las cuatro preguntas presentan correlaciones altas y significativas lo que sugiere que todas miden, en cierto sentido, las competencias.

TABLA 3.

Resumen de las correlaciones entre los tres docentes de cada equipo de corrección para cada una de las cuatro preguntas. Cada equipo implica tres pares de jueces a comparar

	Pregunta 1		Pregunta 2		Pregunta 3		Pregunta 4	
	Correlacs.	Significacs.	Correlacs.	Significacs.	Correlacs.	Significacs.	Correlacs.	Significacs.
Media	0.551	0.004	0.452	0.035	0.408	0.043	0.214	0.233
desviación est	0.146	0.009	0.196	0.065	0.150	0.113	0.169	0.281
max	-0.819	0.038	0.872	0.257	0.675	0.597	0.537	0.877
min	0.307	0	0.165	3.3E-16	-0.078	0	-0.183	0
Número de correlacs. no significativas	0		6		6		21	
% de correlacs. no significativas (respecto de 36)	0		17		17		58	

Fuente: elaboración propia

También, entre las subcompetencias idealmente se esperaba encontrar que éstas y sus definiciones operacionales se agruparan de acuerdo a la estructura teórica (donde cada subcompetencia debería estar asociada a su competencia eje). Contrario a esto, todas están fuertemente correlacionadas entre sí, incluyendo los aspectos formales de las respuestas. Adicionalmente, un análisis factorial exploratorio extrajo sólo un factor que explicaba más del 83% de la varianza (extracción de componentes principales), indicando una alta colinealidad entre las subcompetencias.

Para conocer si hubo diferencias entre los promedios de las cuatro preguntas y si hubo diferencias en el comportamiento de las competencias eje, para cada uno de los 49 alumnos, se consideró la nota promedio de cada competencia eje y se sometieron estos datos a un ANOVA factorial de medidas repetidas, con pregunta (las cuatro preguntas) y competencia (las tres competencias eje) como factores. Los resultados muestran un efecto principal de pregunta ($F(3, 144) = 12.9$; $MCe = 1.79$; $p < 0.001$; $Eta^2_p = 0.21$; $(1-\beta) = 0.99$) y un efecto principal de competencia eje ($F(2, 96) = 334.37$; $MCe = 0.394$; $p < 0.001$; $Eta^2_p = 0.87$; $(1-\beta) > 0.99$). Además, la interacción entre los dos factores principales fue también significativa ($F(6, 288) = 7.88$; $MCe = 0.203$; $p < 0.001$; $Eta^2_p = 0.14$; $(1-\beta) = 0.99$).

Para el factor competencia se encontraron diferencias significativas entre la competencia 1 (*conceptualizar el fenómeno psicológico*) y la competencia 2 (*análisis comprensivo*) a favor de la primera ($F(1, 48) = 1470.14$; $MCe = 0.067$; $p < 0.001$; $Eta^2_p = 0.97$; $(1-\beta) > 0.99$) y entre la competencia 2 (*análisis comprensivo*) y la competencia 3 (*pensamiento crítico*), a favor de la segunda, ($F(1, 48) = 98.27$; $MCe = 0.22$; $p < 0.001$; $Eta^2_p = 0.67$; $(1-\beta) > 0.99$). La competencia *análisis comprensivo* fue la que obtuvo notas significativamente más bajas, lo que indica que fue la más compleja de lograr.

El ANOVA indica que aunque las competencias estaban fuertemente correlacionadas entre sí, los evaluadores fueron capaces de discriminar entre ellas en términos de la complejidad de las competencias. Este análisis sugiere que la competencia de *análisis comprensivo* fue más difícil de demostrar para

los alumnos que las competencias *pensamiento crítico* y *conceptualizar el fenómeno* (al menos, tal como fueron definidas operacionalmente en este estudio).

Validez de criterio

Dado que la evaluación por competencias se propone como una medida del desarrollo que han logrado los alumnos luego de los dos años del ciclo básico de su formación, podemos usar las notas de estos primeros años como criterio externo. Con este fin, para los 49 participantes en la evaluación, obtuvimos sus notas de primer y segundo año como promedio general y las desglosamos (cursos fundamentales, de psicología y de talleres). El promedio de la fase escrita se correlacionó positivamente con todas ellas. La más alta de las correlaciones ($r = 0.56$), que corresponde a la correlación con el promedio de notas de segundo año, equivale a una varianza explicada de más de un 30%.

La definición operacional que mejor predice las notas es *análisis comprensivo*, subcompetencia *desagregar*, operacionalizada como *desarrolla un argumento*. Esta subcompetencia se correlaciona con todas las notas calculadas para el ciclo básico, salvo las notas de talleres. Además, en una regresión múltiple (método Stepwise), con las 12 definiciones operacionales como predictores y notas de segundo año como criterio, *desarrolla un argumento* fue la única que logró predecir significativamente notas de segundo año ($Beta = 0.61$; $R^2 = 0.35$; $t(48) = 5.15$; $p < 0.001$). En otras palabras, al descontar el efecto de esta subcompetencia sobre las otras definiciones operacionales, estas últimas ya no logran predecir las notas.

Los análisis anteriores indican que aunque todas las competencias están fuertemente asociadas entre ellas, y aunque los docentes no lograron un acuerdo interjueces más que moderado al evaluarlas, estas sí muestran una asociación con las notas de las asignaturas del primer ciclo de la carrera, tal como se esperaba. Más aún, la competencia de *análisis comprensivo* (y la subcompetencia *desarrolla un argumento*), aunque fue la que recibió notas en promedio más bajas, fue la que mejor se asoció con las notas del primer ciclo.

Resultados fase oral

Dado que los docentes en esta fase podían discutir sus evaluaciones, no tiene sentido calcular una medida de acuerdo interjueces. Para todos los análisis siguientes, se promediaron las evaluaciones de los tres docentes de cada comisión, con el fin de obtener un promedio por subcompetencia. Del mismo modo, las evaluaciones de los pares también se promediaron para obtener una sola nota. Así, cada alumno recibió una nota para cada una de las cuatro subcompetencias, proveniente de la comisión, de los pares y de sí mismo (p. ej., la autoevaluación). Además, para cada alumno se calculó un promedio ponderado de la fase oral, en el que la evaluación de los docentes tuvo una ponderación de 70% y las de los pares y la autoevaluación recibieron una ponderación de 15% cada una.

Sin embargo, sí es interesante conocer la relación entre las notas asignadas por los docentes, por los pares y por los mismos alumnos (p. ej., la autoevaluación). La prueba de Friedman mostró que las notas asignadas por los docentes, por los pares y la autoevaluación, fueron distintas ($p < 0.001$). Las medias indican que la evaluación asignada por los pares y por los mismos alumnos fueron muy similares, mientras que las notas asignadas por los docentes fueron menores (docentes: media = 3.73, DS = 0.85; pares: media = 4.43, DS = 0.44; autoevaluación: media = 4.55, DS = 0.46). Debido a esta diferencia, los análisis que se hacen a continuación consideran solo las notas asignadas por los docentes.

Tal como se hizo para la fase escrita, resulta interesante conocer las correlaciones entre las cuatro subcompetencias de la fase oral. Aunque las correlaciones son todas significativas, este dato debe ser tomado con precaución debido a violaciones de normalidad. Sin embargo, las correlaciones mismas no son afectadas por esto, de modo que son un dato confiable.

Para conocer si hubo diferencias entre las evaluaciones asignadas por los docentes a las cuatro subcompetencias, se consideró la nota promedio de cada uno de los alumnos (a través de docentes) en cada subcompetencia, y estos datos se sometie-

ron a la prueba de Friedman. El resultado indicó que hubo diferencias significativas entre las cuatro subcompetencias ($p < 0.001$). Las medias muestran que *desarrolla compromiso* y *fidelidad hacia el grupo* recibieron una evaluación menor que las restantes subcompetencias.

Dado que la evaluación por competencias se propone como una medida del desarrollo que han logrado los alumnos luego de los dos años del ciclo básico de su formación, tal como se hizo para la fase escrita, se pueden usar las notas de estos primeros años como criterio externo. El promedio de la fase oral se correlacionó positivamente con todas ellas.

Al considerar la capacidad de las competencias por separado para predecir las notas de primer y segundo año, la competencia de *capacidad de investigar* es la que mejor se correlaciona con las notas. Una vez que se descuenta el efecto de esta competencia, las subcompetencias de *trabajo en equipo* no muestran relación con notas. Ingresamos las cuatro subcompetencias en una regresión lineal múltiple (método Stepwise) con notas de las asignaturas de psicología como variable predicha. El análisis mostró que *capacidad de investigar* fue la única subcompetencia que logró ingresar en la ecuación y predecir significativamente las notas de psicología ($Beta = 0.56$; $R^2 = 0.30$; $t(45) = 4.49$; $p < 0.001$). En otras palabras, al descontar el efecto de esta competencia sobre las otras variables, estas últimas ya no logran predecir las notas.

Relación entre fase escrita y fase oral

Aunque existe una correlación entre el promedio final de la fase escrita y las notas puestas por los docentes en la fase oral ($r = 0.346$; $R^2 = 0.12$; $p = 0.019$), una pregunta interesante es si ambas fases entregaron información redundante o no para predecir las notas del ciclo básico. Aparentemente, la fase oral hizo una contribución independiente. Como evidencia de esto, la correlación entre *capacidad de investigar* y las notas de las asignaturas de psicología, al controlar el efecto de la nota promedio de la fase escrita es todavía significativa ($r_{parcial} = 0.454$; $R^2 = 0.21$; $p = 0.002$). Este análisis sugiere que ambas fases de la evaluación por competencias

evalúan competencias distintas. Esto es, la *competencia de análisis* y la de *capacidad de investigar*, aunque relacionadas, tienen características distintas. Finalmente, consideradas en conjunto, el promedio de la fase escrita y las notas puestas por los docentes en la fase oral logran predecir aproximadamente un 37% del total de notas de primer y segundo año ($R = 0.63$; $F(2, 45) = 14.08$; $p < 0.001$).

Discusión y conclusiones

El objetivo central del trabajo reportado fue crear una evaluación que permitiera medir las competencias asociadas al perfil de primer ciclo en estudiantes de psicología. Específicamente se evaluaron competencias generales como *conceptualización del fenómeno psicológico*, *análisis comprensivo desde la teoría psicológica*, *pensamiento crítico*, *capacidad de investigar*, y *trabajo en equipo*. Si bien no era un objetivo de la evaluación propiamente, también se esperaba retroalimentar al cuerpo docente en cuanto a las metodologías y estrategias de evaluación.

Los resultados muestran evidencia en relación al proceso de validación del instrumento, específicamente en relación a la validez de criterio, de contenido y constructo, así como de la confiabilidad interjueces. La incorporación de jueces expertos permitió asegurar la validez de contenido. Por otra parte, la correlación entre las notas de la evaluación por competencias y las notas de primer y segundo año aseguró la validez de criterio externo. La validez de constructo se obtuvo a partir de un análisis factorial el cual arrojó un solo factor. Si bien esto último indica que el instrumento no permite discriminar en forma precisa las diferentes competencias evaluadas, esto es congruente con lo señalado por la literatura dada la complejidad de éstas. En esta línea Blömeke et al. (2015) señalan que los criterios típicos de validez como el de la unidimensionalidad de los ítems, no necesariamente permiten evaluar competencias complejas que son multidimensionales.

Lo anterior permite fundamentar que el instrumento creado mide competencias y no corresponde a una evaluación de contenidos, aunque sí requiere el manejo de herramientas técnicas o instrumenta-

les (por ejemplo, manejo de conceptos, capacidad de análisis de temas de la disciplina, capacidad de investigación). Esto sugiere que a través de este proceso es posible evaluar habilidades adquiridas o desarrolladas durante los dos primeros años de formación.

Los resultados de la construcción y aplicación del instrumento nos permiten conducir una reflexión en torno a los siguientes aspectos: 1) la construcción del instrumento y las definiciones de las competencias y 2) la efectividad de las metodologías docentes para desarrollar las competencias en nuestros estudiantes.

En relación a la construcción del instrumento, la forma en que fueron diseñadas las preguntas parece requerir que se pongan en juego todas las competencias al mismo tiempo, lo que hace difícil discriminarlas por separado. En cuanto a las definiciones de las competencias, los resultados permiten concluir que éstas pueden ser mejoradas. La alta correlación entre las competencias puede deberse a la falta de diferenciación en las definiciones y a que probablemente se consideraron algunas competencias como estrictamente independientes cuando efectivamente pueden estar vinculadas. Por ejemplo las competencias de *análisis comprensivo* y *conceptualización* pueden ser entendidas como dos niveles de la misma competencia. Asimismo, el cuerpo docente consideró que las definiciones para la corrección de las competencias estaban poco diferenciadas entre sí y eran difíciles de usar en el proceso de evaluación.

A partir de estos resultados se pueden delinear ciertas tareas que quedan pendientes. Entre éstas, especificar las definiciones de las competencias para clarificar y diferenciar las habilidades que miden cada una, específicamente las de *análisis comprensivo*, *conceptualización* y *pensamiento crítico*. Por otra parte, es importante medir las competencias del segundo ciclo asociadas al perfil del licenciado, lo que significa un trabajo de seguimiento de los mismos alumnos que participaron en esta primera evaluación. Por último, queda como tarea pendiente reelaborar la evaluación de las competencias de primer ciclo de acuerdo a la retroalimentación obtenida, de manera que el proceso de evaluación sea

continuo y permita obtener resultados que incidan en la planificación del currículo y en las metas de enseñanza–aprendizaje.

En relación a las metodologías, el proceso de construcción del instrumento permitió a los docentes concordar cuáles serían las competencias intencionadas en sus asignaturas, así como explicitar sus estrategias para evaluar las competencias en sus programas. Esto permite que se pueda realizar un esfuerzo transversal por desarrollar e intencionar ciertas competencias a través de metodologías educativas y formas de evaluación, así como la alineación de los programas de cada asignatura del ciclo.

Referencias

- American Psychological Association. (2013). *APA guidelines for the undergraduate psychology major: Version 2.0*. Retrieved from <http://www.apa.org/ed/precollege/undergrad/index.aspx>
- Bartram, D., & Roe, R. A. (2005). Definition and Assessment of Competences in the Context of the European Diploma in Psychology. *European Psychologist*, 10(2), 93–102. <http://dx.doi.org/10.1027/1016-9040.10.2.93>
- Beneitone, P., Esquetini, C., González, J., Marty, M., Siufi, G., & Wagenaar, R. (2007). *Reflexiones y perspectivas de la Educación Superior en América Latina. Informe Final-Proyecto Tuning América Latina 2004-2007*. Bilbao: Universidad de Deusto. Recuperado de <http://tuning.unideusto.org/tuningal/>
- Blömeke, S., Gustafsson, J. E., & Shavelson, R. (2015). Beyond Dichotomies Competence Viewed as a Continuum. *Zeitschrift für Psychologie*, 223(1), 3–13. <http://dx.doi.org/10.1027/2151-2604/a000194>
- Blömeke, S., Zlatkin-Troitschanskaia, O., Kuhn, C., & Fege, J. (Eds.) (2013) *Modeling and Measuring Competencies in Higher Education Tasks and Challenges*. Rotterdam: Sense Publishers. Recuperado de <https://www.sensepublishers.com/media/1531-modeling-and-measuring-competencies-in-higher-education.pdf>
- Camargo-Escobar, I., & Pardo-Adames, C. (2008). Competencias docentes de profesores de pregrado: diseño y validación de un instrumento de evaluación. *Universitas Psychologica* 1(2), 441-456. Recuperado de <http://revistas.javeriana.edu.co/index.php/revPsycho/article/view/441>
- Charria, V. H., Sarsosa, K., & Arenas, F. (2011). Construcción y validación de contenido de un diccionario de las competencias genéricas del psicólogo: Académicas, profesionales y laborales. *Interdisciplinaria*, 28(2), 299-322. Recuperado de <http://www.scielo.org.ar/pdf/interd/v28n2/v28n2a09.pdf>
- CINDA (2009). *Diseño Curricular Basado en Competencias y Aseguramiento de la Calidad de la Educación Superior*. Centro Interuniversitario de Desarrollo, Santiago, Chile. Recuperado de <http://www.cinda.cl>
- Fouad, N., Hatcher, R., Smith, P., Collins, F., Grus, C., Kaslow, N., Madson, M., & Crossman, R. (2009). Competency Benchmarks: A Model for Understanding and Measuring Competence in Professional Psychology Across Training Levels. *Training and Education in Professional Psychology*, 3(4), S5-S26. <http://dx.doi.org/10.1037/a0015832>
- González, J. M., Arquero, J. L. & Hassall, T. (2014). Consolidación de la formación por competencias en la universidad Española: Estudio de un caso. *Educación XXI*, 12 (2), 145-168. <http://dx.doi.org/10.5944/educxx1.17.2.11483>
- Hyrkäs, K., Appelqvist-Schmidlechner, K, & Oksa, L. (2003). Validating an instrument for clinical supervision using an expert panel. *International Journal of nursing studies*, 40(6), 619-625. [http://dx.doi.org/10.1016/S0020-7489\(03\)00036-1](http://dx.doi.org/10.1016/S0020-7489(03)00036-1)
- Hunsley, J., & Barker, K. (2011). Training for Competency in Professional Psychology: A Canadian Perspective. *Australian Psychologist*, 46, 142–145. <http://dx.doi.org/10.1111/j.1742-9544.2011.00027.x>
- Jorquera, M. T. (Ed.) (2013). *Competencias del Psicólogo en Chile: Propuestas desde las Universidades Estatales*. Chile: Ed. Universidad de la Serena.
- Lynn, M. R. (1986). Determination and quantification of content validity. *Nursing Research*, 35, 382-385. <http://dx.doi.org/10.1097/00006199-198611000-00017>
- Manríquez, L. (2012). ¿Evaluación por Competencias? *Estudios Pedagógicos XXXVIII*, 1, 353-366. Recuperado de <http://www.scielo.cl/pdf/estped/v38n1/art22.pdf>

- Rodolfa, E. R., Bent, R. J., Eisman, E., Nelson, P. D., Rehm, L., & Ritchie, P. (2005). A Cube model for competency development: Implications for psychology educators and regulators. *Professional Psychology: Research and Practice*, 36, 347–354.
- Roe, R.A. (2002). What Makes a Competent Psychologist? *European Psychologist*, 7(3), 192–202. <http://dx.doi.org/10.1027//1016-9040.7.3.192>
- Ryan, G., Spencer, L., & Bernhard, U. (2012). Development and validation of a customized competency-based questionnaire. *Cross Cultural Management: An International Journal*, 19(1), 90–103. <http://dx.doi.org/10.1108/13527601211195646>
- Sánchez, S., Santos, L., Fuentes, F., & Núñez, J. (2015). Enseñanza-Aprendizaje por competencias en la educación superior. La construcción de casos de empresa. *Educación XXI*, 18(1), 237-258. <http://dx.doi.org/10.5944/educXXI.18.1.12319>
- Tremblay, K., Lalancette, D., & Roseveare, D. (2012). AHELO Assessment of Higher Education Learning Outcomes Design and Implementation. Feasibility Report, 1. Recuperado de <http://www.oecd.org/edu/skills-beyond-school/AHELOFSReport-Volume1.pdf>
- Tuning (2006). *Una Introducción a Tuning Educational Structures in Europe: La contribución de las Universidades al Proceso de Bolonia*. Recuperado de <http://guayama.inter.edu/peg/proyect-tunning-europa-espanol-2006.pdf>
- Vargas, M., & Avendaño, B. (2014). Diseño y análisis psicométrico de un instrumento que evalúa competencias básicas en Economía y Finanzas: una contribución a la educación para el consumo. *Universitas Psychologica*, 13(4), 1379-1393. DOI: <http://dx.doi.org/10.11144/Javeriana.UPSY13-4.dapi>
- Villa, A., & Poblete, M. (2011). Evaluación de competencias genéricas: principios, oportunidades y limitaciones. *Bordón*, 63(1), 147-170. Recuperado de http://www.innova.deusto.es/images/archivos/evaluaci%C3%B3n_competencias_aurelio_bordon.pdf
- Villarroel, V., & Bruna, D. (2014). Reflexiones en torno a las competencias genéricas en educación superior: Un desafío pendiente. *Psicoperspectivas Individuo y Sociedad*, 13(1), 23-34. <http://dx.doi.org/10.5027/PSICOPERSPECTIVAS-VOL13SSUE1-FULL-TEXT-335>.