

INVESTIGACIONES

Midiendo calidad en educación superior:  
análisis de confiabilidad y validez de una encuesta  
de certificación de calidad del pregrado, versión estudiantes

Measuring quality in higher education: analysis of reliability and validity  
of an undergraduate quality certification scale-student version

*Medindo qualidade na educação auperior: análise de confiabilidade e legitimidade  
de uma pesquisa de certificado de qualidade da licenciatura, versão estudantes*

*Ismael Gallardo, Juan Sánchez, Marcelo Leiva*

Universidad de Talca, e-mail de contacto: igallardo@utalca.cl.  
Teléfono 0056-71-201726

**RESUMEN**

En este trabajo se presentan datos sobre la confiabilidad y validez de constructo (a través de ecuaciones estructurales) de la escala CNA de certificación de calidad educativa en su versión para estudiantes. Los resultados de dos aplicaciones muestran que el instrumento posee una aceptable confiabilidad y validez, indicadores que mejoran al hacer una reducción de los datos con base en criterios estadísticos. Se discuten las implicancias de estos resultados para la medida de la calidad educativa así como algunas reflexiones y recomendaciones para futuros estudios.

*Palabras clave:* educación superior, calidad, medición.

**ABSTRACT**

This research presents data on the reliability and validity of construct (through structural model equations) of the CNA certification scale of educational quality in its student version. The results of two applications show that the instrument has acceptable reliability and that a theoretical model with related dimensions has a better fit of the data in both applications. The values of reliability and validity improve by making a data reduction based on statistical criteria, an improvement that is replicated at both times of measurement. We discuss the implications of these results for measuring the quality of education and provide some reflections and recommendations for future studies.

*Key words:* higher education, quality, measuring.

**RESUMO**

Apresentam-se dados sobre a confiabilidade e legitimidade do instrumento (equações estruturais) da escala CNA de certificação da qualidade educativa, versão para estudantes. Os resultados de duas aplicações revelam que o instrumento possui confiabilidade e legitimidade aceitáveis, indicadores que melhoram ao reduzir os dados com base em critérios estatísticos. Discutem-se as implicações de estes resultados para a verificação da qualidade educativa e algumas reflexões e recomendações para futuros estudos.

*Palavras chave:* educação superior, qualidade, legitimidade.

La calidad de la educación superior es un tema evidente en el proceso de globalización actual. De hecho, la discusión ya no se centra en la necesidad de instalar este tema como hasta hace algunos años, sino que está puesta en los mecanismos por los cuales ésta es más eficiente (CINDA, 2007). En Chile, es el Estado el responsable de la acreditación de logro de la calidad tanto a nivel de las propias instituciones como también de las carreras de pregrado y postgrado. Concretamente, la Comisión Nacional de Acreditación (CNA) tiene la función de verificar y promover la calidad de las universidades, institutos profesionales y centros de formación técnica autónomos, y de las carreras y programas que ellos ofrecen ([www.cnachile.cl](http://www.cnachile.cl)).

Para el caso de las carreras de pregrado, la CNA utiliza nueve criterios de calidad, fruto del trabajo de una comisión especialmente formada para ello, perteneciente a la antigua CNAP (*Comisión Nacional de Acreditación de Pregrado*, Arrau, 2005) y que definen aquellos parámetros que una carrera profesional debe cumplir para acreditar que realiza su trabajo con calidad. En el caso de la acreditación del pregrado, los criterios de evaluación fueron diseñados teniendo como foco de evaluación las carreras en general, los cuales deben complementarse con el perfil profesional que la propia carrera haya tomado ([www.cnachile.cl](http://www.cnachile.cl)). Los criterios propuestos por la CNA corresponden a: 1) propósitos; 2) integridad; 3) estructura organizacional, administrativa y financiera; 4) estructura curricular; 5) recursos humanos; 6) efectividad del proceso enseñanza-aprendizaje; 7) resultados del proceso de formación; 8) infraestructura, y 9) vinculación con el medio.

Si una carrera requiere una certificación de calidad por parte de la CNA, debe evaluar el estado de estos criterios en su propia institución. Una vez iniciado el proceso, estos criterios de calidad son analizados en dos momentos (OCDE, 2009; [www.cnachile.cl](http://www.cnachile.cl)). El primero de ellos es la *autoevaluación*, en donde se evalúa el estado de los criterios de calidad propuestos por la CNA en la carrera que postula a la acreditación de los mismos. Los diferentes actores involucrados en el proceso de enseñanza (estudiantes, docentes, egresados, empleadores) responden a una serie de instrumentos que pretenden evaluar el estado de los criterios de calidad propuestos (e.g., focus group o entrevistas). Con el conjunto de los datos recogidos se redacta un informe que es enviado a una agencia acreditadora para revisión. En un segundo momento se realiza una evaluación externa de los criterios de calidad por pares académicos no pertenecientes a la unidad que solicita acreditación, y que son designados por el organismo acreditador. Los jueces realizan una serie de reuniones para evaluar la fidelidad de los datos de autoevaluación, al final de las cuales se obtiene un primer informe verbal y luego escrito. Del conjunto de informes evaluativos (i.e., autoevaluación e informe de evaluadores externos) surge un acuerdo de acreditación de la agencia acreditadora el cual finalmente define si la institución solicitante es acreditada (o no) y por cuánto tiempo. En esta secuencia de acciones, la autoevaluación constituye uno de los pilares del proceso de aseguramiento de la calidad, ya que es fuente de información fundamental para las instituciones que desean detectar debilidades y fortalezas, así como también monitorear sus propios avances y generar estrategias de mejora (Rodríguez, Fleet y Delgado, 2009).

Para la etapa de autoevaluación de calidad (i.e., las nueve dimensiones ya descritas) en los estudiantes, la CNA propuso hace unos años un instrumento llamado "Encuesta de certificación de calidad educativa". Este instrumento está constituido por 73 afirmaciones que se responden, en su mayoría, de acuerdo con un formato de escala tipo Likert de cuatro opciones, las cuales van desde muy en desacuerdo a muy de acuerdo. Esta encuesta

posee diez áreas de consulta, en línea con los criterios de calidad propuestos. En concreto: 1) propósitos; 2) integridad; 3) estructura organizacional; 4) estructura curricular; 5) competencias generales; 6) recursos humanos; 7) efectividad de la enseñanza; 8) infraestructura y otros recursos; 9) vinculación con el medio, y 10) satisfacción general.

En la actualidad, el uso de este instrumento no es obligatorio y sólo es sugerido a las instituciones para sus procesos de autoevaluación en caso de no poseer otros, lo que de cierta forma les permite ahorrar tiempo y recursos en la elaboración de un instrumento propio, lo que también está autorizado<sup>1</sup>. Esta puede ser una de las posibles razones por las cuales las propiedades métricas (i.e., confiabilidad y validez) de esta encuesta no han sido evaluadas, lo cual es un problema importante si se considera que la autoevaluación es considerada como la etapa más importante en el proceso de aseguramiento de la calidad (www.cnachile.cl; Zúñiga, 2007). Dicha situación permite explicar también el que no existan estudios que evalúen algún aspecto de esta encuesta ni siquiera a un nivel descriptivo, constituyendo en sí mismo un problema para la adecuada medida de los criterios propuestos (véase también Zapata & Tejeda, 2009). Del mismo modo, tampoco está del todo claro que los criterios utilizados sean correctos para evaluar lo que es calidad en la educación (i.e., validez de constructo) y menos aún la mejor forma en que dichos criterios se relacionan para ello.

Dada la importancia del proceso de autoevaluación en el proceso de aseguramiento de la calidad, es necesario contar con instrumentos de evaluación que sean confiables y válidos. Es por esto que el objetivo principal de este trabajo es analizar la confiabilidad y validez de la encuesta de certificación de calidad CNA en su versión estudiantes, en una carrera de pregrado de una universidad pública. Para lograr estos objetivos se plantean dos estudios, los que representan dos momentos diferentes en que se miden los criterios de aseguramiento de la calidad propuestos por la CNA. En el primero de ellos se evalúan las propiedades métricas del instrumento y, en caso de ser atingente, se propondrán alternativas de mejora al mismo y se evaluarán. El segundo estudio replica los análisis del primero con una muestra de estudiantes recogida en un segundo momento, con el objeto de evaluar si las propiedades métricas evaluadas anteriormente se mantienen en el tiempo.

## MÉTODO

### ESTUDIO 1

#### *Participantes*

Los participantes fueron 189 estudiantes de una carrera de ciencias sociales de una universidad pública chilena pertenecientes al ingreso 2003 a 2007 de la misma, siendo 133 mujeres (70.4%) y 56 hombres (29.6%). La muestra fue no probabilística, escogida por conveniencia.

---

<sup>1</sup> Aún cuando en la actualidad estos materiales ya no están disponibles para descarga en la página web institucional de CNA, sí lo están en algunas páginas institucionales universitarias (e.g., departamentos de acreditación), lo que sugiere que sigue siendo una fuente de información de la cual las carreras pueden obtener información para su autoevaluación.

### *Procedimiento*

Durante el mes de mayo de 2007, un académico miembro de la comisión de autoevaluación de la carrera que se sometía a acreditación visitó a los distintos cursos en sus salas de clase para describirles los fundamentos del proceso de acreditación, solicitando posteriormente a los asistentes que respondieran la encuesta de certificación. Dichos cuestionarios fueron auto-administrados y respondidos en el mismo lugar, asegurándose a todos los participantes anonimato y confidencialidad de los datos entregados.

### *Instrumento*

Se utilizó la encuesta de certificación de calidad educativa propuesta por la CNA, en su versión para estudiantes. Está constituida por 73 ítems, los cuales varían entre cuatro a diez ítems por dimensión. Son respondidos en escalas de tipo Likert de 4 puntos, a excepción de la dimensión 5, "Evaluación de Competencias generales", que se responde mediante una escala de 1 a 7, y del ítem 23 de la dimensión 4, "Estructura Curricular", que se responde de manera dicotómica (Sí-No).

### *Plan de análisis*

Para evaluar las propiedades métricas de la encuesta de certificación de calidad educativa se utilizaron tres procedimientos. El primero fue evaluar la confiabilidad de dicha encuesta de acuerdo al coeficiente alfa de Cronbach de consistencia interna. El segundo fue evaluar la validez de constructo siguiendo análisis factorial confirmatorio a través de ecuaciones estructurales. Se compararon una serie de modelos diseñados para reproducir los datos a través de indicadores de ajuste de los mismos. Se entiende que el concepto de ajuste indica la habilidad de un modelo teórico de reproducir los datos (Bollen & Long, 1993). Hu & Bentler (1999) proponen el uso de un conjunto de indicadores para evaluar el ajuste de un modelo más que el uso de un indicador en particular. Siguiendo estas recomendaciones, en este estudio se utilizarán los siguientes:  $X^2$  (Chi Square, donde valores bajos indican un mejor ajuste del modelo), *AIC* (Akaike Information Criterion, donde la medida indica un mejor ajuste cuando los valores son más pequeños)<sup>2</sup>, *RMSEA* (Root Mean Square Error of Approximation, donde valores menores a .08 se consideran como indicadores de ajuste aceptable y valores alrededor de .05 como ajuste óptimo), *CFI* (Comparative Fit Index, valores mayores a 0.95 son considerados como aceptables) y *AGFI* (Adjusted goodness of fit index, valores cercanos a 1 indican un mayor ajuste). Finalmente, se exploró la estructura factorial que proponen los datos a través de un análisis factorial exploratorio por componentes principales y rotación varimax<sup>3</sup>. Se consideró como criterio de corte la obtención de autovalores mayores a 1.

---

<sup>2</sup> Tanto  $X^2$  como *AIC* son indicadores utilizados sobre todo para comparar el mejor ajuste entre dos o más modelos.

<sup>3</sup> Se realizó un análisis siguiendo ejes principales y rotación oblimin, considerando que las dimensiones pudiesen estar relacionadas. Dicho método de extracción y rotación no mostró diferencias importantes con el análisis presentado en los resultados, razón por la que se mantuvo el análisis inicial.

## RESULTADOS

### ANÁLISIS DE CONFIABILIDAD

Al evaluar la consistencia interna general de la versión completa de la escala (ver anexo 2) se obtuvo un alfa de Cronbach de .852, considerado como muy bueno (Nunally, 1987). Al analizar la consistencia interna por dimensiones de la escala los valores fueron aceptables para la mayoría de las escalas, excepto para Estructura Organizacional y Estructura Curricular, considerados como inaceptables.

*Tabla 1*

Valores de confiabilidad por cada dimensión de la encuesta de certificación de la calidad (73 ítems)

Dimensiones	P	I	EO	EC	CG	RH	EE	IF	VM	SG
Valores	.698	.720	.548	.441	.698	.778	.672	.848	.761	.764

\* P = Propósitos; I = Integridad; EO = Estructura Organizacional; EC = Estructura Curricular; CG = Competencias Generales; RH = Recursos Humanos; EE = Efectividad en la Enseñanza; IF = Infraestructura; VM = Vinculación con el Medio; SG = Satisfacción General.

### VALIDEZ DE CONSTRUCTO

Para constatar de forma más precisa la validez del cuestionario y evaluar su comportamiento, se sometió el cuestionario a cuatro Modelos Factoriales Confirmatorios, por medio del programa AMOS 16.0, empleándose el procedimiento de máxima verosimilitud para la estimación de los parámetros. El primer modelo factorial propone que las dimensiones CNA no se relacionan entre sí; el segundo modelo propone que las dimensiones CNA se relacionan entre sí; el tercer modelo señala un modelo de 10 dimensiones (las mismas del instrumento) con un factor de segundo orden llamado Calidad. Finalmente, un último modelo agrupa los diferentes ítems en cuatro nuevas dimensiones (temáticas) con un factor de segundo orden llamado Calidad. El análisis combinado de los indicadores propuestos (Hu & Bentler, 1999) muestra que el modelo teórico que mejor ajuste obtiene a los datos es el de dimensiones correlacionadas.

*Tabla 2*

Indicadores de ajuste para los modelos propuestos

	X <sup>2</sup>	RMSEA	CFI	AIC	GFI	AGFI
Modelo 1	5237.3	.075	.490	5529.3	.529	.502
Modelo 2	4103.4	.058	.697	4485.4	.643	.616
Modelo 3	4753.9	.068	.582	5045.8	.584	.561
Modelo 4	4752.4	.086	.582	5044.446	.580	.541

## ANÁLISIS FACTORIAL EXPLORATORIO

El análisis factorial exploratorio (AFE) se realizó sobre la versión completa de la escala<sup>4</sup>. Dicho análisis identificó un modelo de 22 factores con autovalores mayores a 1, donde la varianza total explicada fue de 76,58% en el conjunto de la escala. La matriz de datos rotados da cuenta que los ítems no se distribuyen de manera homogénea a través de las dimensiones propuestas, existiendo diferencias entre dicha distribución y la propuesta en la escala CNA. Descrito de otro modo, el análisis mostró que los ítems que se definen en la escala como pertenecientes a una dimensión no miden necesariamente aspectos relacionados con ella.

Tras el AFE se eliminó un total de 32 ítems que daban cuenta en menor medida de cada una de las dimensiones (esto es, aquellos con peso factorial menor a 0.40, García, Gil, & Rodríguez, 2000)<sup>5</sup>. El objetivo era utilizar criterios estadísticos de reducción que permitieran evaluar si la cantidad de ítems era un factor por el cual la escala no se comportaba de acuerdo a su diseño original. Con este procedimiento se obtuvo una versión reducida a 41 ítems, la cual fue sometida a los mismos análisis de la escala original. Los resultados se muestran a continuación.

## ANÁLISIS DE CONFIABILIDAD, ESCALA REDUCIDA

Al evaluar la consistencia interna general de la versión reducida de 41 ítems se obtuvo un alfa de Cronbach de .842, considerado como muy bueno (Nunnally, 1987). Los valores alfa por cada dimensión pueden observarse en la tabla 3.

Tabla 3

Valores de confiabilidad por cada dimensión  
de la encuesta de certificación de la calidad (41 ítems)

Dimensiones	P	I	EO	EC	CG	RH	EE	IF	VM	SG
Valores	.741	.730	.650	.725	.792	.718	.772	.787	.761	.764

\* P=Propósitos; I=Integridad; EO=Estructura Organizacional; EC=Estructura Curricular; CG=Competencias Generales; RH=Recursos Humanos; EE=Efectividad en la Enseñanza; IF=Infraestructura; VM=Vinculación con el Medio; SG = Satisfacción General.

<sup>4</sup> Previamente se evaluó la pertinencia del AFE mediante el estadístico de adecuación de la muestra de Kaiser-Meyer-Olkin (*KMO*) y la prueba de esfericidad de Bartlett. Ambos indicadores en conjunto muestran que es pertinente la realización de dicho análisis (*KMO* = .691; Bartlett < .01).

<sup>5</sup> Los ítems seleccionados fueron: 1, 2, 7, 9, 10, 11, 12, 17, 19, 20, 22, 24, 25, 26, 29, 33, 36, 37, 38, 44, 47, 49, 50, 51, 52, 53, 54, 55, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73

#### VALIDEZ DE CONSTRUCTO, ESCALA REDUCIDA

En esta etapa se reprodujo el análisis realizado a la escala de 73 ítems, resultando nuevamente que el modelo de variables inter-relacionadas (el modelo 2) es el que da mejor cuenta de los datos presentados. Los resultados se observan en la Tabla 4.

Tabla 4

Indicadores de ajuste para los modelos propuestos

	X <sup>2</sup>	RMSEA	CFI	AIC	GFI	AGFI
Modelo 1	1877.7	.086	.543	2041.6	.583	.540
Modelo 2	1018.1	.049	.882	1270.0	.795	.760
Modelo 3	1317.8	.061	.776	1481.8	.730	.701
Modelo 4	1389.1	.064	.746	1551.0	.722	.693

#### ANÁLISIS FACTORIAL EXPLORATORIO, ESCALA REDUCIDA

El AFE realizado sobre la versión de 41 ítems<sup>6</sup> identificó un modelo de 10 factores, con una varianza total explicada de 68,37% en el conjunto de la escala. La observación de la matriz de rotación de factores permite dar cuenta que los ítems se distribuyen de manera relativamente homogénea a través de las dimensiones propuestas, existiendo similitud entre dicha distribución y aquella propuesta en la escala CNA.

#### DISCUSIÓN ESTUDIO 1

Los resultados de este primer estudio permiten observar que las dimensiones propuestas por la CNA permiten evaluar correctamente el constructo “calidad” y, de modo más relevante, una versión más reducida de la escala propuesta por dicha comisión entrega mejores indicadores de ajuste a los datos, lo que permite entregar un instrumento menos extenso y por tanto de aplicación más breve.

Para evaluar la persistencia del modelo teórico evaluado, se analizaron los datos de una segunda aplicación realizada durante el segundo semestre del año 2009.

#### ESTUDIO 2

El segundo estudio busca replicar los resultados encontrados en el primer estudio, analizando los datos recogidos en el año 2009 pertenecientes a la misma carrera del estudio 1. Debido a que la solución factorial de dimensiones inter-relacionadas mostró los

<sup>6</sup>  $KMO = .836$ ; Bartlett  $< .01$

valores más adecuados de ajuste, se espera que sea dicha solución la más adecuada en este segundo estudio en comparación con los modelos alternativos. Se espera, asimismo, que la confiabilidad general y por escalas del instrumento sea similar al estudio 1 y que los ajustes propuestos al mismo mejoren las propiedades de modo similar al estudio anterior.

## MÉTODO

### PARTICIPANTES

Los participantes fueron 250 estudiantes de una carrera de ciencias sociales de una universidad pública chilena (la misma del estudio 1), pertenecientes al ingreso 2005 al 2009 de la misma, siendo 190 mujeres (76%) y 60 hombres (24%). La muestra fue no probabilística, escogida por conveniencia. Respecto al procedimiento, tipo y diseño de investigación e instrumento medidos en este estudio, estas fueron realizadas de la misma forma que en el estudio 1.

## RESULTADOS

### Análisis de confiabilidad

Al evaluar la consistencia interna general de la versión completa de la escala se obtuvo un alfa de Cronbach de .872 considerado como muy bueno (Nunally, 1987). Al analizar la consistencia interna por dimensiones de la escala, los valores obtenidos fueron en general buenos, a excepción, nuevamente, de las dimensiones estructura organizacional y estructura curricular, considerados como inaceptables.

*Tabla 5*

Valores de confiabilidad por cada dimensión de la encuesta de certificación de la calidad (73 ítems)

Dimensiones	P	I	EO	EC	CG	RH	EE	IF	VM	SG
Valores	.674	.767	.543	.418	.629	.838	.679	.860	.705	.760

\* P = Propósitos; I = Integridad; EO = Estructura Organizacional; EC = Estructura Curricular; CG = Competencias Generales; RH = Recursos Humanos; EE = Efectividad en la Enseñanza; IF = Infraestructura; VM = Vinculación con el Medio; SG = Satisfacción General.

### VALIDEZ DE CONSTRUCTO

Al igual que en el estudio 1, se propusieron los mismos cuatro modelos teóricos, siendo el modelo con las dimensiones inter-relacionadas (modelo 2) aquel que obtuvo el mejor ajuste a los datos.



Tabla 6

Indicadores de ajuste para los modelos propuestos

	X <sup>2</sup>	RMSEA	CFI	AIC	GFI	AGFI
Modelo 1	6094.5	.074	.523	6376.479	.538	.512
Modelo 2	4559.9	.057	.724	4941.867	.668	.643
Modelo 3	5449.1	.062	.709	5143.233	.655	.639
Modelo 4	5451.3	.067	.609	5743.260	.602	.579

#### ANÁLISIS FACTORIAL EXPLORATORIO

Del mismo modo que en el estudio 1<sup>7</sup>, el análisis de la versión completa de la escala mostró una composición de 18 factores, con una varianza total explicada de 65,25% en el conjunto de la escala. La observación de la matriz de rotación de factores permite dar cuenta que los ítems se distribuyen de manera relativamente heterogénea a través de las dimensiones propuestas, existiendo disparidad entre dicha distribución y aquella propuesta en la escala CNA.

Al igual que en el estudio 1, se buscó reducir la cantidad total de ítems en la escala recurriendo al criterio del peso factorial del ítem sobre la dimensión respectiva. De este modo, se eliminaron los ítems con pesos factoriales menores a 0.40 (García, 2002), identificándose los mismos 32 ítems que en el estudio anterior. De este modo, se obtuvo una escala reducida con los 41 ítems del estudio 1, la que fue sometida nuevamente a análisis de confiabilidad y validez de constructo.

#### ANÁLISIS DE CONFIABILIDAD

Al evaluar la consistencia interna general de la versión reducida de 41 ítems se obtuvo un alfa de Cronbach de .849, considerado como muy bueno. Tal y como se puede observar en la Tabla 7, los valores de confiabilidad se mantuvieron en un nivel aceptable, mientras que los valores de las dimensiones antes consideradas como inaceptables mejoraron ostensiblemente.

Tabla 7

Valores de confiabilidad por cada dimensión de la encuesta de certificación de la calidad (73 ítems)

Dimensiones	P	I	EO	EC	CG	RH	EE	IF	VM	SG
Valores	.728	.703	.693	.734	.739	.792	.779	.769	.705	.760

\* P=Propósitos; I=Integridad; EO=Estructura Organizacional; EC=Estructura Curricular; CG=Competencias Generales; RH=Recursos Humanos; EE=Efectividad en la Enseñanza; IF=Infraestructura; VM=Vinculación con el Medio; SG=Satisfacción General.

<sup>7</sup> *KMO* = .893; Bartlett < .01

## VALIDEZ DE CONSTRUCTO

Nuevamente fueron sometidos a contraste los mismos modelos ya planteados, mejorando los indicadores de todos ellos. Sin embargo, nuevamente es el modelo de dimensiones inter-relacionadas el que mejor da cuenta de los datos recogidos. Los resultados se observan en la Tabla.

Tabla 8

Indicadores de ajuste para los modelos propuestos

	X <sup>2</sup>	RMSEA	CFI	AIC	GFI	AGFI
Modelo 1	5474.5	.093	.542	2638.462	.565	.521
Modelo 2	1250.0	.053	.861	1501.963	.801	.766
Modelo 3	1705.9	.069	.749	1869.948	.740	.712
Modelo 4	1899.1	.076	.697	2061.145	.730	.702

## ANÁLISIS FACTORIAL EXPLORATORIO

El AFE<sup>8</sup> mostró que la versión reducida de 41 ítems identificó un modelo de 10 factores, con una varianza total explicada de 61,60% en el conjunto de la escala. Al igual que en el estudio 1, la distribución de los ítems tendió a ser la propuesta originalmente por la CNA.

## DISCUSIÓN ESTUDIO 2

Al igual que los resultados del estudio 1, los análisis muestran una adecuada confiabilidad global del instrumento (.872), aunque en la consistencia por dimensiones existen valores muy bajos en dos de ellas (Estructura organizacional y Estructura Curricular). Al igual que en el estudio 1, se confirma que el mejor ajuste de los datos se obtiene a partir del modelo de factores interrelacionados y que todos los indicadores de medida mejoran cuando se reduce la cantidad de ítems de la escala.

## DISCUSIÓN GENERAL

Uno de los problemas más importantes a la hora de evaluar o corregir un proceso de calidad es aquel relacionado con la forma de medir los avances o retrocesos del mismo. En el caso del aseguramiento de la calidad en educación superior en Chile esto

<sup>8</sup>  $KMO = .912$ ; Bartlett  $< .01$ .

se hace aún más relevante puesto que si bien ha dado pasos importantes en el proceso de control, aseguramiento y fomento de la calidad (OCDE, 2009), los estudios en que se evalúen los instrumentos con los que se trabaja para medirla son, hasta donde sabemos, inexistentes. El presente trabajo, entonces, se transforma en el primero en dar cuenta de las propiedades de medida de una encuesta de certificación de calidad, en este caso, aquella propuesta por la CNA en su versión para estudiantes.

Los resultados muestran que en general la encuesta posee adecuadas propiedades métricas en su versión original de 73 ítems, a excepción de dos de sus escalas. De manera interesante, estos resultados se replican en una aplicación realizada con posterioridad. Por otro lado, el análisis factorial confirmatorio realizado a partir de ecuaciones estructurales, muestra que el modelo 2 (10 dimensiones interrelacionadas) se ajusta mejor a los datos recogidos que cualquiera de los otros modelos evaluados, pero aún así sus valores de ajuste se alejan de los límites considerados como adecuados (Hu y Bentler, 1999). El análisis factorial exploratorio mostró, además, que la estructura dimensional de 10 factores no se replica del todo en la muestra seleccionada resultando en, al menos, 18 factores para cada medición. Este dato permite inferir que algunos de los ítems de la encuesta evalúan asuntos que no necesariamente corresponden a lo que la CNA define como criterio de calidad.

Sin embargo, al utilizar criterios estadísticos de eliminación de ítems, las propiedades métricas de la encuesta fueron en general mucho mejores que la versión original tanto en confiabilidad como en validez. Aún cuando se podría asumir que los factores interpretables pudieran ser menos de 10, lo cierto es que los ajustes realizados a los ítems permitieron que la cantidad de criterios propuestos por la CNA aparecieran en la encuesta mejorada y, más aún, los ítems tendieron a distribuirse de acuerdo a lo planteado en la encuesta. En conjunto, estos datos sugieren que la encuesta de certificación de calidad propuesta por la CNA versión estudiantes puede reducirse de su versión original de 73 ítems a una más reducida de 41, la cual es confiable y válida para la medición de los criterios de calidad en educación superior<sup>9</sup>.

Todos estos datos apuntan a que es posible tener una escala de medida de los criterios de calidad educativa en población estudiantil universitaria, que es estable a través del tiempo y que parece medir las dimensiones definidas por la CNA. Debido a ello, parece importante que dicha Comisión ponga a disposición de los departamentos de acreditación de instituciones de educación superior escalas de evaluación que permitan levantar información sobre las dimensiones fundamentales a considerar en el proceso de autoevaluación. De hecho, la inexistencia de dichas escalas puede ser perjudicial ya que cada institución puede presentar instrumentos con características de medida que difieren entre sí. Este aspecto es muy relevante dado que la mencionada diferencia podría llevar a que los mejores resultados de la autoevaluación entre una institución respecto a otra se deban a que una de ellas (o ambas) mide las dimensiones de calidad educativa de manera equivocada.

Es por ello que consideramos que los resultados de este trabajo aportan significativamente a la evaluación de la calidad en la educación superior, en lo que atiende a

---

<sup>9</sup> Dado que en el segundo estudio participaron estudiantes del primero, se evaluó si las propiedades psicométricas dependían de la cantidad de veces que una misma persona responde a la encuesta. Para ello, se separó aquellos de la última generación (2009) y se compararon las propiedades métricas con quienes ingresaron en los años anteriores. Los resultados mostraron similitud de ambas submuestras con los resultados generales ya planteados, lo que sugiere que éstos no dependen de la frecuencia de respuesta de una misma persona a través de las aplicaciones.

la autoevaluación de los criterios de calidad de la carrera por parte de los estudiantes. Además, y tal vez más importante, permite evaluar de manera adecuada los criterios propuestos por la CNA y que permiten a los alumnos obtener educación de calidad, lo que se concretiza en la realidad a través del desarrollo de competencias profesionales en los estudiantes que se forman y en el reconocimiento social de la profesión que poseen (Miranda, 2007; Moreno, 2007). Si bien es cierto que la encuesta de certificación de la calidad CNA, versión estudiantes no mide competencias del estudiante, realiza una real contribución en la medición de las dimensiones que las instituciones educativas utilizan para el logro de las mismas. Es por ello que utilizar un adecuado instrumento que mida eficazmente la calidad educativa contribuye a mejorar la calidad de enseñanza que se entrega a los estudiantes por parte de las instituciones educativas y a su vez, garantizar las competencias necesarias para un desempeño óptimo.

Este trabajo posee algunas limitaciones, como el tipo y cantidad de muestra, que deben ser mejorados en futuras investigaciones, las cuales deberán replicar estos resultados en estudiantes de otras carreras profesionales, así como también en otras muestras relevantes para el actuar de cada una de ellas tales como empleadores, académicos o egresados. Por otro lado, se sugiere que próximos trabajos evalúen nuevos criterios conceptuales que puedan entregar un modelo teórico que se ajuste de mejor manera a los datos. Por ejemplo, si bien es cierto la acreditación es un proceso en que se asume la variabilidad de cumplimiento de criterios en función del perfil de cada carrera profesional, este aspecto no está explícitamente inserto en la escala actual. Así, las competencias en medicina se relacionan necesariamente con la ejecución de tareas específicas de esa profesión, lo mismo que un ingeniero forestal, un agrónomo, un psicólogo o un profesor de educación secundaria. Nuevos trabajos, además, podrían evaluar la relación existente entre algunas de las dimensiones de esta escala y otras de salida (o resultados) como, por ejemplo, competencias relacionadas con el perfil de egreso. Esto permitirá tener un nuevo criterio de validez para cada una de las dimensiones: Si las dimensiones CNA correlacionan fuertemente con aquellas medidas (e.g., competencias), entonces será un indicador que dichas escalas realmente miden lo que dicen medir.

Por otro lado, este trabajo ha mostrado que la forma de evaluación de la calidad en educación superior posee debilidades, llevando a preguntarse si ésta no debería asemejarse a otras áreas del conocimiento social en donde los criterios de logro de la calidad o la excelencia están descritos más extensamente. Por ejemplo, el modelo de calidad/excelencia de Chile Calidad (dependiente de CORFO) describe cómo los diferentes criterios que lo componen se relacionan y determinan mutuamente, asignándole además puntuaciones a cada uno de ellos (véase [www.chilecalidad.cl](http://www.chilecalidad.cl)). De un modo similar ocurre con el modelo Malcolm Baldrige del National Institute of Standards and Technology (NIST), de Estados Unidos de América (véase [www.nist.gov](http://www.nist.gov)) o el modelo Asia-Pacific de calidad del Asia-Pacific Quality Network (véase [www.apqn.org](http://www.apqn.org)). Todos ellos proponen formas en que se relacionan los criterios que definen calidad, lo que no ocurre (hasta donde sabemos) para el caso de la educación superior en donde no hay esquema relacional explícito alguno. Esto evitaría uno de los problemas que actualmente posee el sistema de acreditación referente a la vaguedad de definiciones y criterios que permiten a los revisores confiar de manera excesiva en lo que ellos definen como calidad, lo que puede llevar a dudas sobre la justicia de ciertos veredictos de acreditación (OCDE, 2009).

En línea con lo anterior, investigaciones más amplias que la presentada en este trabajo podrían estudiar otros aspectos de medida como, por ejemplo, la relación existente entre la evaluación de los criterios de calidad registrados en la autoevaluación y el juicio que los pares evaluadores realizan de la institución a acreditar. Si bien es cierto la autoevaluación puede enmascarar información que se observa sólo en la visita de pares, no lo es menos el que dicha opinión profesional puede verse afectada por la propia experiencia vivida en cada lugar de trabajo y por las brechas percibidas entre lo observado versus la experiencia cotidiana. Dado que los pares evaluadores pertenecen a instituciones que compiten en un mismo rubro, no es extraño pensar que cuando esa brecha no favorece a la institución de origen del juez la probabilidad de perder objetividad en el proceso pueda ser cada vez mayor. Tal vez, al poseer un modelo de excelencia en la educación superior que describa claramente la relación entre los criterios, sus actores y sus resultados, se pueda incorporar a jueces que incluso sean ajenos a la disciplina profesional que se acredita, reduciendo con ello potenciales sesgos de este tipo en la observación.

## AGRADECIMIENTOS

El presente trabajo fue realizado con el apoyo del programa de investigación en calidad de vida y ambientes saludables de la Facultad de Psicología de la Universidad de Talca y del proyecto FONDECYT 11070074. Agradecemos a Cris Bruel por su valioso aporte a la finalización de este trabajo.

## REFERENCIAS BIBLIOGRÁFICAS

- Arrau, F. (2005). *Indicadores de calidad de la educación superior en Chile*. (Serie Estudios Nº 307, Año XV). Santiago: Biblioteca del Congreso Nacional de Chile.
- Bollen, K. & Long, J. (1993) *Testing structural equation models*. Londres: Sage.
- CINDA (2007). *Educación Superior en Iberoamérica: Informe 2007*. Santiago: RIL Editores.
- García, E., Gil, J. y Rodríguez, G. (2000). *Análisis Factorial*. Cuadernos de Estadística, 7. Madrid: La Muralla.
- Hu, L. y Bentler, P. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives, *Structural Equation Modeling*, 6, 1-55.
- Miranda, C. (2007). Educación superior, mecanismos de aseguramiento de la calidad y formación docente. *Estudios Pedagógicos*, 33, 95-108.
- Moreno, M. (2007). Competencias Profesionales y mejora de la Calidad Educativa en escuelas públicas desde la perspectiva de los derechos de la niñez. *Ciencia y Sociedad*, 32, 309-349.
- Nunnally, J. (1987). *Teoría Psicométrica*. México: Trillas.
- OCDE (2009). *La educación superior en Chile*. Santiago: Ministerio de Educación de Chile.
- Rodríguez, E., Fleet, N. y Delgado, M. (2009). La acreditación en la generación de información sobre la calidad de la educación superior. *Calidad en la Educación*, 31, 212-230.
- Zapata, G. y Tejada, I. (2009). Impactos del aseguramiento de la calidad y acreditación de la Educación Superior. Consideraciones y proposiciones. *Calidad en la Educación*, 31, 192-209.
- Zúñiga, M. (2007). Algunas consideraciones sobre el aseguramiento de calidad. En Ayarza, H. y González, L. (eds.), *Acreditación y dirección estratégica para la calidad en las universidades*. Santiago, Chile: RIL Editores.

