



<Artículo de investigación>

Validación preliminar de la escala de autoeficacia estadística en estudiantes de Grado en Medicina españoles: análisis factorial confirmatorio

Javier Santabárbara¹ , Raúl López-Antón² 

Enviado: 14/04/2019. Aceptado: 26/05/2019. Publicado en prensa: 20/06/2019. Publicado: 01/07/2019

//Resumen

INTRODUCCIÓN: La bioestadística es fundamental en la medicina basada en la evidencia. Para un mayor desempeño en el aprendizaje de esta materia, pensando en los futuros médicos, es crucial el análisis de la autoeficacia en estadística del estudiante de Medicina durante su formación académica. Objetivo: Realizar un estudio preliminar de traducción y validación de una escala de autoeficacia en estadística en una muestra de estudiantes de Grado de Medicina.

MÉTODO: Estudio de validación de la escala *current statistics self-efficacy (CSSE)* para medir la autoeficacia en estadística en una muestra de 40 estudiantes de primer curso del Grado en Medicina en la Universidad de Zaragoza.

RESULTADOS: El análisis factorial confirmatorio permitió verificar que el modelo unidimensional de la CSSE se ajusta adecuadamente a los datos obtenidos en nuestra muestra de estudiantes de Medicina.

CONCLUSIONES: Los resultados encontrados indican que el instrumento posee adecuadas propiedades psicométricas que permiten evaluar la autoeficacia en estadística en estudiantes universitarios españoles que cursan el Grado de Medicina. De esta manera, la escala se constituye como una herramienta importante para los planes de formación en medicina, sugiriéndose su utilización para una mejora en la calidad de la enseñanza y del aprendizaje de este Grado en particular.

//Palabras clave

Autoeficacia en estadística; Validación; Estudiantes de Medicina; Grado.

//Datos de los autores

¹ Departamento Medicina Preventiva y Salud Pública. Facultad de Medicina. Universidad de Zaragoza, España. Autor para la correspondencia: jsantabarbara@unizar.es

² Departamento de Psicología y Sociología. Facultad de Ciencias Sociales y Humanas de Teruel. Universidad de Zaragoza, España.

//Referencia recomendada

Santabárbara, J., y López-Antón, R. (2019). Validación preliminar de la escala de autoeficacia estadística en estudiantes de Grado en Medicina españoles: análisis factorial confirmatorio. *REIRE Revista d'Innovació i Recerca en Educació*, 12(2), 1–13.

<http://doi.org/10.1344/reire2019.12.228505>

© 2019 Javier Santabárbara *et al.* Este artículo es de acceso abierto sujeto a la licencia Reconocimiento 4.0 Internacional de Creative Commons, la cual permite utilizar, distribuir y reproducir por cualquier medio sin restricciones siempre que se cite adecuadamente la obra original. Para ver una copia de esta licencia, visite <http://creativecommons.org/licenses/by/4.0/>



J. Santabàrbara, R. López-Antón. *Validación preliminar de la escala de autoeficacia estadística en estudiantes de Grado en Medicina españoles: análisis factorial confirmatorio*

//Títol

Validació preliminar de l'escala d'autoeficàcia estadística en estudiants del grau de Medicina espanyols: anàlisi factorial confirmatòria

//Resum

INTRODUCCIÓ: La bioestadística és fonamental en la medicina basada en l'evidència. Per a un millor acompliment en l'aprenentatge d'aquesta matèria pensant en els futurs metges, és crucial l'anàlisi de l'autoeficàcia en estadística dels estudiants de Medicina durant la seva formació acadèmica. Objectiu: Dur a terme un estudi preliminar de traducció i validació d'una escala d'autoeficàcia en estadística en una mostra d'estudiants del grau de Medicina.

MÈTODE: Estudi de validació de l'escala *current statistics self-efficacy* (CSSE) per mesurar l'autoeficàcia en estadística en una mostra de quaranta estudiants de primer curs del grau de Medicina a la Universitat de Saragossa.

RESULTATS: L'anàlisi factorial confirmatòria va permetre verificar que el model unidimensional de la CSSE s'ajusta adequadament a les dades obtingudes en la nostra mostra d'estudiants de Medicina.

CONCLUSIONS: Els resultats indiquen que l'instrument té unes propietats psicomètriques adequades que permeten avaluar l'autoeficàcia en estadística a estudiants universitaris espanyols que cursen el grau de Medicina. D'aquesta manera, l'escala es constitueix com una eina important per als plans de formació en Medicina, i se suggereix la seva utilització per a una millora en la qualitat de l'ensenyament i de l'aprenentatge d'aquest grau en particular.

//Paraules clau

Autoeficàcia en estadística; Validació; Estudiants de Medicina; Grau.

//Title

Preliminary validation of the scale of statistics self-efficacy in Spanish medical students: confirmatory factor analysis

//Abstract

INTRODUCTION: Biostatistics is fundamental in evidence-based medicine. To help students learn this subject more effectively, it is crucial to assess their statistics self-efficacy. Objective: To carry out a preliminary study translating and validating a self-efficacy measure in statistics in a sample of students taking a bachelor's degree in medicine.

METHOD: Validation of the *Current Statistics Self-Efficacy Scale* (CSSE) to assess statistics self-efficacy in a sample of 40 first-year students taking the University of Zaragoza's Bachelor's Degree in Medicine.

RESULTS: The confirmatory factor analysis confirmed that the one-dimensional model of the CSSE adequately fits the data obtained in our sample of medical students.

CONCLUSIONS: The results indicate that CSSE's psychometric properties can be used to evaluate statistics self-efficacy in Spanish medical students and should therefore be considered an important tool for improving the quality of teaching and learning in medical degrees.

//Keywords

Statistics self-efficacy; Validation; Medical students; Degree.

1. Introducción

Las creencias de autoeficacia han sido definidas como aquellos juicios de las personas acerca de sus capacidades para alcanzar determinados niveles de rendimiento (Bandura, 1987). Es muy necesaria para hacer que cualquier resultado sea un éxito a través de la conducta y los enfoques que se basan en su creencia de conocimiento en la materia.

En el ámbito académico, se ha demostrado que aquellos estudiantes que confían en sus propias capacidades obtienen un mejor rendimiento académico en diferentes áreas tales como la ciencia (Britner y Pajares, 2001) y las matemáticas (Pajares y Miller, 1994). Así, un metaanálisis calculó un coeficiente de correlación de Pearson de 0,50 para la asociación entre la autoeficacia y el rendimiento académico (Robbins *et al.*, 2004).

En el caso particular de la estadística, además, la autoeficacia aumenta la motivación de los estudiantes para aprender esta materia a través de su conciencia para controlar la motivación. Al tener acceso al desarrollo de conocimientos y habilidades, la autoeficacia tiene un impacto en la participación de los estudiantes en la clase, lo que lleva a su éxito académico (Richardson, Abraham y Bond, 2012).

En ese contexto, Finney y Schraw, (2003) crearon una escala (*current statistics self-efficacy, CSSE*) para medir la autoeficacia en estadística. El estudio psicométrico de la escala arrojó resultados satisfactorios, habiendo sido aplicada internacionalmente principalmente en estudiantes de Psicología, siendo la Medicina un ámbito aún no estudiado.

La bioestadística es fundamental en la medicina basada en la evidencia. Para un mayor desempeño en el aprendizaje de esta materia pensando en los futuros médicos, es crucial el análisis de la autoeficacia en estadística del estudiante de medicina durante su formación. Así, dada la importancia de la evaluación de la autoeficacia en estadística por su potencial relación con el rendimiento académico de la asignatura de Bioestadística, es de gran relevancia adecuar un instrumento que permita medir dicho constructo a la realidad de los estudiantes de medicina de nuestro país. Por ello, se propuso como objetivo realizar un estudio preliminar de traducción y validación de dicha escala en una muestra de estudiantes de medicina.

2. Metodología

Diseño

Estudio observacional, de corte transversal y analítico.

Traducción

Se empleó el método de traducción inversa, traduciendo del inglés al español los ítems de la *CSSE* (Finney y Schraw, 2003) y viceversa. Esta tarea fue realizada por el autor del trabajo y, en forma inversa, por una doctora en psicología bilingüe.

J. Santabàrbara, R. López-Antón. Validación preliminar de la escala de autoeficacia estadística en estudiantes de Grado en Medicina españoles: análisis factorial confirmatorio

Población en estudio y selección de la muestra

Los participantes en el estudio fueron estudiantes del Grado en Medicina que cursaban la asignatura de Bioestadística de primer curso en la Facultad de Ciencias de la Salud y el Deporte (Huesca) de la Universidad de Zaragoza. El reclutamiento se realizó mediante un muestreo no probabilístico de conveniencia. Se solicitó a los alumnos participar voluntariamente en el proyecto de investigación. Todos los que se ofrecieron tuvieron que rellenar la hoja de información demográfica junto con la escala *current statistics self-efficacy (CSSE)* (Finney y Schraw, 2003). El investigador estaba disponible para atender cualquier pregunta de los participantes. Después de completar los cuestionarios, se agradeció a los estudiantes su participación. El cuestionario fue cumplimentado por 40 estudiantes, considerando la muestra representativa de la población de matriculados en primer curso de medicina ($n=45$), ya que representa el 88,9% de la población diana para este estudio de validación preliminar.

Instrumentos

Las características sociodemográficas (edad y sexo) se recogieron mediante un cuestionario realizado "ad hoc" para este estudio.

Para recoger la autoeficacia hacia la estadística, se utilizó la *current statistics self-efficacy (CSSE)* (Finney y Schraw, 2003) (ver Anexo), constructo unidimensional en origen, que evalúa la autoeficacia estadística de forma auto-referida y que consta de 14 ítems. Con este instrumento, se pide a los encuestados que califiquen su creencia actual en su capacidad para completar 14 tareas específicas relacionadas con estadística utilizando una escala de respuesta de 1 (sin confianza) a 6 (confianza completa). Las preguntas incluían declaraciones como "identificar una escala de medición para una variable", "interpretar los resultados de un procedimiento estadístico en términos de la pregunta de investigación" e "identificar los factores que influyen en la potencia". La CSSE se completa generalmente por los estudiantes en menos de cinco minutos. Cuanto mayor sea la puntuación total en la CSSE, mayor autoeficacia en estadística.

Análisis estadístico

Calculamos el coeficiente de correlación de Pearson, para valorar la asociación entre cada ítem y la puntuación total de la CSSE (correlación ítem-total). Así como el coeficiente alfa de Cronbach para estudiar la consistencia interna de la escala.

En la investigación transcultural es frecuente la necesidad de realizar traducciones y adaptaciones de cuestionarios con el fin de comparar poblaciones de sujetos que no comparten el mismo idioma. En estos casos, es conveniente utilizar los métodos y las recomendaciones que sugiere la literatura transcultural para llevar a cabo la traducción y la adaptación de instrumentos de medida (Van de Vijver y Hambleton, 1996). Sin embargo, a pesar de seguir estas recomendaciones, no se puede asumir la equivalencia de las diferentes versiones traducidas. Es necesario confirmarla elaborando análisis que permitan detectar aquellos ítems que presentan un funcionamiento diferencial (Tomás-Marco, González-Romá y Gómez Benito, 2000). Así, con el objetivo de poner a prueba el modelo unifactorial del CSSE reportado por Finney y Schraw (2003), se llevó a cabo un análisis factorial

J. Santabàrbara, R. López-Antón. *Validación preliminar de la escala de autoeficacia estadística en estudiantes de Grado en Medicina españoles: análisis factorial confirmatorio*

confirmatorio (AFC) (Jöreskog, 1969), ya que este representa la aproximación *ad hoc* para abordar el estudio de la equivalencia de versiones de un mismo cuestionario en distinto idioma (Batista-Fogueta, Coenders y Alonso, 2004). Adicionalmente, permite contrastar un modelo construido con antelación, en el que el investigador establece *a priori* el conjunto total de las relaciones entre los elementos que lo configuran.

El AFC se representa mediante diagramas de flujo (*path diagram*), de acuerdo a sus especificaciones particulares. Los rectángulos representan los ítems; las elipses, los factores comunes. Las flechas unidireccionales entre los factores comunes y los ítems expresan saturaciones; las flechas bidireccionales indican la correlación entre factores comunes o únicos (véase la Figura 1) (Jöreskog, 1969). En el AFC es necesario observar las cargas factoriales que permiten establecer la correlación entre las variables y los factores (Garson, 2013). Cuanto más se acerquen a uno, mayor será dicha correlación. Una regla empírica en el AFC establece que las cargas deben ser $\geq 0,7$ (Gefen, Straub y Boudreau, 2000). Pese a que es alto y algunos factores y variables pueden quedar fuera del modelo, estos deben tomarse a criterio del investigador (Raubenheimer, 2004).

Para la evaluación de la bondad del ajuste del modelo se emplearon múltiples indicadores (Hu y Bentler, 1995): el estadístico chi-cuadrado; la razón de chi-cuadrado sobre los grados de libertad (*CMIN/DF*, en inglés); el índice de bondad de ajuste (*GFI*, en inglés); el índice de Tucker-Lewis (*TLI*, en inglés); el índice de ajuste comparativo (*CFI*, en inglés), y el error cuadrático medio de aproximación por grado de libertad (*RMSEA*, en inglés) (Browne y Cudeck, 1989). El estadístico chi-cuadrado indica el ajuste absoluto del modelo pero es muy sensible al tamaño de la muestra. Por consiguiente, usualmente se interpreta también la razón de chi-cuadrado sobre los grados de libertad, con valores inferiores a 3 indicando un buen ajuste. Los índices *GFI*, *TLI* y *CFI* varían entre 0 y 1, con 0 indicando ausencia de ajuste y 1 ajuste óptimo. Valores de 0,95 o superiores son considerados excelentes, y valores superiores a 0,90 sugieren un ajuste aceptable del modelo a los datos (Schermelleh-Engel, Moosbrugger y Müller, 2003). El índice *RMSEA* es considerado óptimo cuando sus valores son de 0,05 o inferiores, y aceptables en el rango 0,05–0,10 (Browne y Cudeck, 1989).

El análisis de los datos se llevó a cabo con el paquete estadístico *R* versión 3.5.1 (<https://www.r-project.org>) para *Mac-OS*.

Aspectos éticos

La coordinadora del Grado en Medicina de la Facultad de Ciencias de la Salud y el Deporte (Huesca) de la Universidad de Zaragoza aprobó el protocolo del estudio. A los estudiantes se les indicó que su participación era voluntaria, que los datos serían empleados en beneficio del cursado, que lo declarado no afectaría la calificación que pudieren obtener en la materia y que el instrumento era anónimo. Una vez obtenido su consentimiento, se les aplicó el cuestionario.

3. Resultados y discusión

Estudio descriptivo

La encuesta fue cumplimentada por 40 de los 45 matriculados en la asignatura de Bioestadística en el Grado en Medicina pertenecientes a la Facultad de Ciencias de la Salud y del Deporte de Huesca (88,9 %). La edad media de los estudiantes fue 18,0 años ($DE = 2,3$), siendo el 67,5 % mujeres.

Fiabilidad y correlación ítem-total

Se calculó el alfa de Cronbach (α) para los datos recogidos en un esfuerzo por analizar la consistencia interna de los ítems de la CSSE (Cronbach, 1951). Se trata de un indicador de la evaluación de fiabilidad adecuada, ya que el constructo se probó de manera unidimensional. Este análisis arrojó niveles de confiabilidad aceptablemente altos con $\alpha = 0,898$ en los 14 ítems del inventario de CSSE (Thompson, 2003). Siendo, por tanto, satisfactoria, ya que fue muy similar a la reportada en su versión original ($\alpha = 0,917$) (Finney y Schraw, 2003). Como se observa en la Tabla 1, este coeficiente no aumenta al eliminar cualquiera de los ítems de la escala.

Es reseñable que la correlación ítem-total es, en todos los casos, elevada ($r \geq 0,45$) y estadísticamente significativa ($p < 0,001$), manteniendo las características de consistencia interna de la versión original (véase la Tabla 1).

Tabla 1. Estudio descriptivo

Ítems (versión española)	media	DE	correlación ítem-total	α de Cronbach si se elimina el ítem
1. Identifico la escala de medida de una variable.	4,75	1,03	0,647	0,743
2. Interpreto el valor-p de un procedimiento estadístico.	3,95	1,11	0,667	0,740
3. Identifico si una distribución es asimétrica cuando me dan los tres valores de tendencia central.	5,10	1,06	0,649	0,742
4. Selecciono el procedimiento estadístico adecuado para responder una pregunta de investigación.	3,55	1,11	0,711	0,739
5. Interpreto los resultados de un procedimiento estadístico en relación con la pregunta de investigación.	4,05	1,01	0,617	0,744
6. Identifico los factores que influyen en la potencia.	2,98	1,05	0,648	0,742
7. Entiendo lo que el valor de la desviación típica significa en términos de la variable que está siendo medida.	4,92	0,89	0,514	0,749
8. Diferencio entre el error de tipo I y de tipo II en un contraste de hipótesis.	4,00	1,18	0,447	0,748

(Continúa en la siguiente página)

J. Santabàrbara, R. López-Antón. Validación preliminar de la escala de autoeficacia estadística en estudiantes de Grado en Medicina españoles: análisis factorial confirmatorio

Ítems (versión española)	media	DE	correlación ítem-total	α de Cronbach si se elimina el ítem
9. Entiendo lo que el valor de error típico significa en términos de la variable que está siendo medida.	3,23	1,21	0,494	0,746
10. Distingo los objetivos de la estadística descriptiva, de los de la inferencia estadística.	4,63	1,06	0,572	0,745
11. Distingo entre la información proporcionada por las tres medidas de tendencia central.	5,30	0,97	0,761	0,740
12. Distingo entre un parámetro poblacional y un estadístico muestral.	5,25	0,90	0,734	0,742
13. Identifico cuando la media, la mediana y la moda deben ser utilizadas como medidas de tendencia central.	4,97	1,14	0,702	0,738
14. Explico la diferencia entre distribución muestral y poblacional.	5,05	0,78	0,655	0,747
Puntuación total CSSE	61,73	9,55	-	0,898

DE: Desviación estándar.

Análisis factorial confirmatorio

Para la obtención de las evidencias de validez sobre la estructura interna se recurre al análisis factorial confirmatorio (AFC), que permite comprobar hasta qué punto los datos ratifican el modelo factorial de la escala generada a partir de la fundamentación teórica.

Este análisis permitirá verificar si el modelo unidimensional propuesto por Finney y Schraw (2003) se ajusta adecuadamente a los datos obtenidos en nuestra muestra de estudiantes de Medicina. En la Figura 1, se representa el resultado de este análisis.

Para estos análisis, el primer ítem de la CSSE ("Identifico la escala de medida para una variable") se fijó a un valor de uno para los fines de convergencia del modelo. Para discutir el ajuste, debemos considerar los criterios de los diversos índices del modelo. Como se reveló en estos análisis, la autoeficacia es un factor unidimensional que converge en los 14 ítems del cuestionario CSSE (véase la Figura 1). El modelo demostró un ajuste razonable según el contraste chi-cuadrado ($\chi^2 = 152$, $gl = 70$, $p < 0,001$) y el $CMIN/DF = 152/70 = 2,17 (< 3)$. El valor obtenido para el error cuadrático medio de aproximación ($RMSEA = 0,107$) se encuentra en el límite razonable (0,10) indicado por Browne y Cudeck (1989). Sin embargo, los valores de $GFI = 0,681$, $CFI = 0,723$ y $TLI = 0,639$ están por debajo de 0,9 para un buen ajuste (Bentler, 1990), pero se conoce que CFI y TLI dependen del tamaño de la muestra (Mulaik *et al.*, 1989) y nuestro estudio únicamente incluye a 40 alumnos. Dado que no existe un acuerdo sobre cuáles son los valores que un buen ajuste debe proporcionar en todos los índices o medidas, en la Tabla 2 se representa la totalidad de estadísticos empleados con los resultados de la estimación de dichos parámetros junto con sus valores recomendados, más extendidos por la literatura para la valoración del ajuste del modelo (Byrne, 2010; Kline, 2016; Ruiz, Pardo y San Martín, 2010). Basado en estos índices, esta muestra tiene un ajuste aceptable al modelo de un factor.

J. Santabàrbara, R. López-Antón. *Validación preliminar de la escala de autoeficacia estadística en estudiantes de Grado en Medicina españoles: análisis factorial confirmatorio*

Figura 1. Modelo de AFC de la escala de autoeficacia estadística (CSSE) en el Grado de Medicina

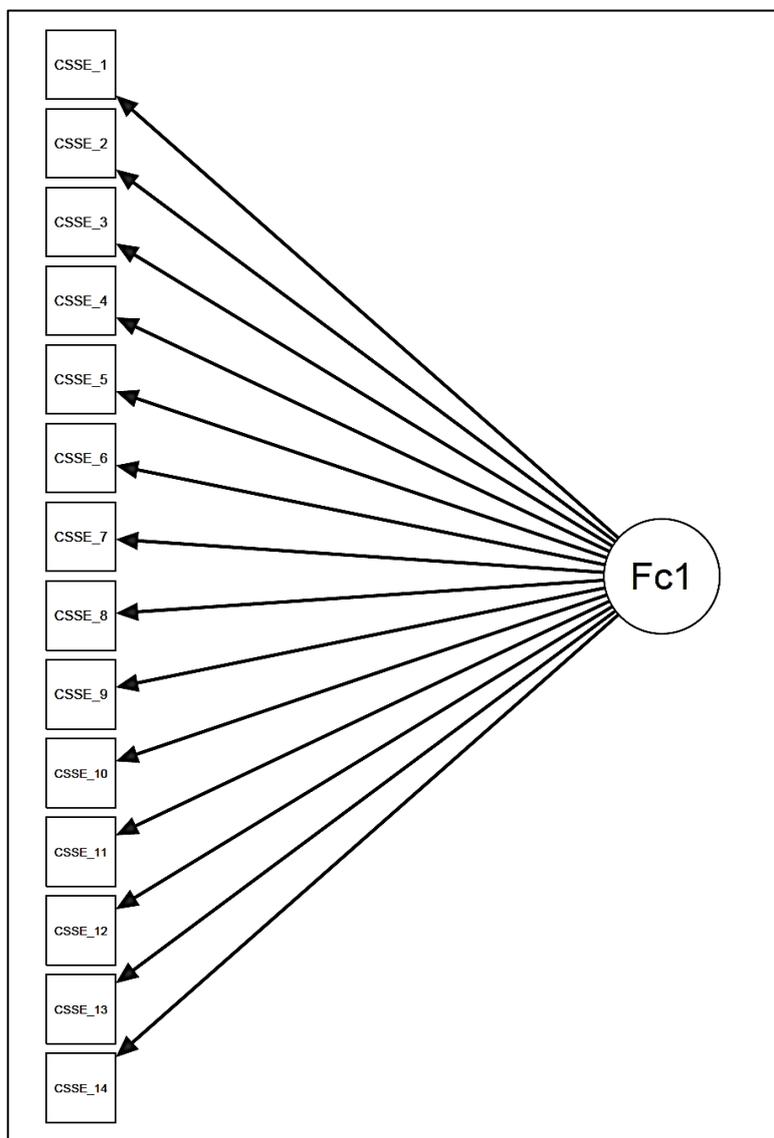


Tabla 2. Resumen de los coeficientes e índices de bondad de ajuste del modelo

Estadísticos		Nivel de ajuste recomendado	Nivel de ajuste obtenido
Chi-cuadrado (χ^2)		$p > 0,05$	152; gl = 70; $p < 0,001$
Modelo <i>fit summary</i> (Ajuste absoluto)	<i>CMIN/DF</i>	Aprox. 2–5	2,17
<i>Baseline Comparisons</i> (ajuste comparativo)	<i>GFI</i>	$> 0,9$	0,681
	<i>TLI</i>		0,639
	<i>CFI</i>		0,723
Otros	<i>RMSEA</i>	$< 0,05$ (buen ajuste) 0,05–0,08 (ajuste aceptable)	0,107

CMIN/DF: χ^2 /gl. *NFI*: *Normed Fix Index* - índice de ajuste no normalizado. *GFI*: *Goodness-of-fit index* - Índice de bondad de ajuste. *TLI*: *Tucker-Lewis Index*. *CFI*: *Comparative Fit Index* - índice de ajuste comparativo. *RMSEA*: *Root Mean Square Error of Approximation* - raíz cuadrada del error cuadrático medio.

J. Santabàrbara, R. López-Antón. Validación preliminar de la escala de autoeficacia estadística en estudiantes de Grado en Medicina españoles: análisis factorial confirmatorio

Los resultados revelan que todas las cargas factoriales entre el factor de primer orden de autoeficacia y sus ítems asociados fueron significativos a un nivel de 0,05. Además, en general, estas deben ser iguales o superiores a 0,7 para una buena validez convergente (Gefen *et al.*, 2000). Del resultado de CFA de este estudio, once cargas son mayores que 0,7 y tres cargas están entre 0,6 y 0,7. Por lo tanto, el modelo teórico propuesto por Finney y Schraw (2003) se ha verificado con la muestra actual de datos (véase la Tabla 3).

Tabla 3. Saturaciones en la muestra española de los ítems en el factor a partir del modelo AFC

Ítems (versión española)	Estimación	EE	IC 95 %		P
			Inferior	Superior	
1. Identifico la escala de medida de una variable.	1,000 ^a				
2. Interpreto el valor-p de un procedimiento estadístico.	0,959	0,261	0,44	1,47	<0,001
3. Identifico si una distribución es asimétrica cuando me dan los tres valores de tendencia central.	1,041	0,252	0,54	1,54	<0,001
4. Selecciono el procedimiento estadístico adecuado para responder una pregunta de investigación.	1,015	0,273	0,47	1,55	<0,001
5. Interpreto los resultados de un procedimiento estadístico en relación con la pregunta de investigación.	0,810	0,243	0,33	1,29	<0,001
6. Identifico los factores que influyen en la potencia.	0,841	0,255	0,34	1,34	<0,001
7. Entiendo lo que el valor de la desviación típica significa en términos de la variable que está siendo medida.	0,640	0,210	0,22	1,05	0,002
8. Diferencio entre el error de tipo I y de tipo II en un contraste de hipótesis.	0,595	0,282	0,04	1,15	0,035
9. Entiendo lo que el valor de error típico significa en términos de la variable que está siendo medida.	0,641	0,288	0,07	1,20	0,026
10. Distingo los objetivos de la estadística descriptiva, de los de la inferencia estadística.	0,945	0,251	0,45	1,44	<0,001
11. Distingo entre la información proporcionada por las tres medidas de tendencia central.	1,092	0,235	0,63	1,55	<0,001
12. Distingo entre un parámetro poblacional y un estadístico muestral.	0,975	0,221	0,54	1,41	<0,001
13. Identifico cuando la media, la mediana y la moda deben ser utilizadas como medidas de tendencia central.	1,207	0,274	0,66	1,74	<0,001
14. Explico la diferencia entre distribución muestral y poblacional.	0,787	0,186	0,42	1,15	<0,001

^a parámetro fijado. EE: error de estimación. IC: Intervalo de confianza. P: valor p.

Limitaciones

La principal limitación reseñable es la utilización de muestreo no probabilístico de conveniencia en un único centro y con escaso tamaño muestral, lo que imposibilita la generalización de los hallazgos. La escala *CSSE* original de 6 puntos definida por Finney y Schraw (2003) es de tipo Likert pero con ausencia de punto medio neutro —como sería el 4 en una escala de 7 puntos—. Dado el diseño transversal del estudio, este se define como validación preliminar, ya que una completa necesitaría de un diseño longitudinal. A pesar de que el CFA ha sido puesto en cuestión en los últimos años a favor del modelo exploratorio de ecuaciones estructurales (*ESEM*, en inglés) (Prudon, 2015), hay autores que defienden que el AFC es el modelo preferido para el ajuste de un factor latente, y el *ESEM* para medidas multidimensionales (Perry, Nicholls, Clough y Crust, 2015).

4. Conclusiones

El objetivo del presente trabajo fue realizar un estudio preliminar de validación de la escala de autoeficacia en estadística *CSSE* (Finney y Schraw, 2003). Los resultados encontrados indican que el instrumento posee propiedades adecuadas psicométricas que permiten evaluar la autoeficacia en estadística en estudiantes universitarios españoles que cursan el Grado de Medicina.

Esto es de suma importancia, dada la necesidad de contar con instrumentos que permitan la evaluación de este constructo en el ámbito universitario, ya que una buena formación en Bioestadística habilita al profesional médico para evaluar críticamente la evidencia médica publicada y mejora la toma de decisiones complejas en la práctica clínica diaria (Dawson, 2009).

De esta manera, la escala se constituye como una herramienta importante para los planes de formación en medicina, sugiriéndose su utilización para una mejora en la calidad de la enseñanza y del aprendizaje de este Grado en particular.

Líneas futuras de investigación

Finalmente, los resultados encontrados, sugieren la futura realización de un análisis factorial confirmatorio en una muestra de estudiantes universitarios de diferentes Grados que permita la utilización general de dicho instrumento en el ámbito universitario, así como evaluar el impacto de la autoeficacia estadística en el rendimiento de la asignatura de Bioestadística en el Grado de Medicina.

<Referencias bibliográficas>

Bandura, A. (1987). *Pensamiento y Acción*. Barcelona: Martínez Roca.

Batista-Fogueta, J. M., Coendersb, G., y Alonso, J. (2004). Análisis factorial confirmatorio. Su utilidad en la validación de cuestionarios relacionados con la salud. *Medicina Clínica (Barcelona)*, 122(1), 1–7.
<https://dx.doi.org/10.1157/13057542>

J. Santabàrbara, R. López-Antón. *Validación preliminar de la escala de autoeficacia estadística en estudiantes de Grado en Medicina españoles: análisis factorial confirmatorio*

- Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin*, 107(2), 238–246. <https://dx.doi.org/10.1037//0033-2909.107.2.238>
- Britner, S. L., y Pajares, F. (2001). Self-efficacy beliefs, motivation, race and gender in middle school science. *Journal of Women and Minorities in Science and Engineering*, 7(4), 271–285. <https://dx.doi.org/10.1615/jwomenminorscieng.v7.i4.10>
- Browne, M. W., y Cudeck, R. (1989). Single sample cross-validation indices for covariance structures. *Multivariate Behavioral Research*, 24(4), 445–455. https://dx.doi.org/10.1207/s15327906mbr2404_4
- Byrne, B. M. (2010). *Structural equation modeling with Amos. Basic concepts, applications and programming* (2ª ed.). Nueva York (NY): Routledge-Taylor and Francis Group.
- Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 16(3), 297–334. <https://dx.doi.org/10.1007/bf02310555>
- Dawson, G. F. (2009). *Interpretación fácil de la bioestadística. La conexión entre la evidencia y las decisiones médicas*. Barcelona: Elsevier.
- Finney, S. J., y Schraw, G. (2003). Self-efficacy beliefs in college statistics courses. *Contemporary Educational Psychology*, 28(2), 161–186. [https://dx.doi.org/10.1016/S0361-476X\(02\)00015-2](https://dx.doi.org/10.1016/S0361-476X(02)00015-2)
- Garson, D. (2013). *Factor Analysis*. Asheboro: North Carolina State University Press.
- Gefen, D., Straub, D., y Boudreau, M. C. (2000). Structural equation modeling and regression: guidelines for research practice. *Communications of the Association for Information Systems*, 4, 1–79. <https://dx.doi.org/10.17705/1CAIS.00407>
- Hu, L., y Bentler, P. (1995). Evaluating model fit. En R. Hoyle (ed.), *Structural equation modelling: Concepts, issues and applications* (pp. 76–99). Thousand Oaks, CA: Sage Publications.
- Jöreskog, K. (1969). A general approach to confirmatory maximum likelihood factor analysis. *Psychometrika*, 34(2), 183–202. <https://dx.doi.org/10.1007/bf02289343>
- Kline, R. B. (2016). *Principles and practice of structural equation modeling* (4.ª ed.). Nueva York: The Guilford Press.
- Mulaik, S. A., James, R. L., Van Alstine, J., Bennett, N., Lind, S., y Stilwell, C. D. (1989). Evaluation of goodness-of-fit indices for structural equation models. *Psychological Bulletin*, 105(3), 430–445. <https://dx.doi.org/10.1037/0033-2909.105.3.430>
- Pajares, F., y Miller, M. D. (1994). The role of self-efficacy and selfconcept beliefs in mathematical problem-solving: A path analysis. *Journal of Educational Psychology*, 86(2), 193–203. <https://dx.doi.org/10.1037/0022-0663.86.2.193>

J. Santabàrbara, R. López-Antón. *Validación preliminar de la escala de autoeficacia estadística en estudiantes de Grado en Medicina españoles: análisis factorial confirmatorio*

- Perry, J. L., Nicholls, A. R., Clough, P. J., y Crust, L. (2015). Assessing Model Fit: Caveats and Recommendations for Confirmatory Factor Analysis and Exploratory Structural Equation Modeling. *Measurement in Physical Education and Exercise Science*, 19(1), 12–21. <https://dx.doi.org/10.1080/1091367x.2014.952370>
- Prudon, P. (2015). Confirmatory factor analysis as a tool in research using questionnaires: a critique. *Comprehensive Psychology*, 4, 1–19. <https://dx.doi.org/10.2466/03.cp.4.10>
- Raubenheimer, J. E. (2004). An item selection procedure to maximize scale reliability and validity. *South African Journal of Psychology*, 30(4), 59–64.
- Richardson, M., Abraham, C., y Bond, R. (2012). Psychological correlates of university students' academic performance: A systematic review and meta-analysis. *Psychological Bulletin*, 138(2), 353–387. <https://dx.doi.org/10.1037/a0026838>
- Robbins, S. B., Lauver, K., Le, H., Davis, D., Langley, R., y Carlstrom, A. (2004). Do psychosocial and study skill factors predict college outcomes? A meta-analysis. *Psychological Bulletin*, 130(2), 261–288. <https://doi.org/10.1037/0033-2909.130.2.261>
- Ruiz, M. A., Pardo, A., y San Martín, R. (2010). Modelos de Ecuaciones Estructurales. *Papeles del Psicólogo*, 31(1), 34–45.
- Schermelleh-Engel, K., Moosbrugger, H., y Müller, H. (2003). Evaluating the fit of structural equation models: Tests of significance and descriptive goodness-of-fit measures. *Methods of Psychological Research Online*, 8(2), 23–74.
- Thompson, B. (2003). *Score Reliability: Contemporary Thinking on Reliability Issues*. Thousand Oaks, CA: Sage.
- Tomás-Marco, T., González-Romá, V., y Gómez Benito, J. (2000). Teoría de respuesta al ítem y análisis factorial confirmatorio: dos métodos para analizar la equivalencia psicométrica en la traducción de cuestionarios. *Psicothema*, 12(2), 540–544
- Van de Vijver, F., y Hambleton, R. K. (1996). Translating tests: Some practical guidelines. *European Psychologist*, 1(2), 89–99. <https://dx.doi.org/10.1027/1016-9040.1.2.89>

<Organismos colaboradores>

Los autores declaran que este trabajo no ha tenido ninguna fuente de financiación.

<Conflicto de intereses>

Los autores declaran no tener ningún conflicto de intereses.

<Agradecimientos>

Los autores expresan su gratitud a los alumnos que participaron en este estudio. Así como a la Dra. Virginia Navascués Howard por su traducción del CSSE.

Anexo: Cuestionario sobre la auto-eficacia en estadística (CSSE)

Por favor califique su confianza en su **capacidad actual** para completar con éxito las siguientes tareas. La escala del ítem tiene 6 respuestas posibles: (1) ninguna confianza en absoluto, (2) poca confianza, (3) buena cantidad de confianza, (4) mucha confianza, (5) muchísima confianza, (6) completa confianza. Para cada tarea, marque la única respuesta que represente su confianza en su **capacidad actual** para completar con éxito la tarea.

Ninguna confianza en absoluto

Completa confianza

	↓			↓		
	1	2	3	4	5	6
1. Identifico la escala de medida de una variable.	1	2	3	4	5	6
2. Interpreto el valor-p de un procedimiento estadístico.	1	2	3	4	5	6
3. Identifico si una distribución es asimétrica cuando me dan los tres valores de tendencia central.	1	2	3	4	5	6
4. Selecciono el procedimiento estadístico adecuado para responder una pregunta de investigación.	1	2	3	4	5	6
5. Interpreto los resultados de un procedimiento estadístico en relación con la pregunta de investigación.	1	2	3	4	5	6
6. Identifico los factores que influyen en la potencia.	1	2	3	4	5	6
7. Entiendo lo que el valor de la desviación típica significa en términos de la variable que está siendo medida.	1	2	3	4	5	6
8. Diferencio entre el error de tipo I y de tipo II en un contraste de hipótesis.	1	2	3	4	5	6
9. Entiendo lo que el valor de error típico significa en términos de la variable que está siendo medida.	1	2	3	4	5	6
10. Distingo los objetivos de la estadística descriptiva, de los de la inferencia estadística.	1	2	3	4	5	6
11. Distingo entre la información proporcionada por las tres medidas de tendencia central.	1	2	3	4	5	6
12. Distingo entre un parámetro poblacional y un estadístico muestral.	1	2	3	4	5	6
13. Identifico cuando la media, la mediana y la moda deben ser utilizadas como medidas de tendencia central.	1	2	3	4	5	6
14. Explico la diferencia entre distribución muestral y poblacional.	1	2	3	4	5	6