



Referencia

Ureta Morales, F. J. (2023). Análisis factorial exploratorio Escala Evaluación del desempeño docente, maestrías en psicología, USAC. *Revista Científica del Sistema de Estudios de Postgrado*. 6(1). 15-29.
DOI: <https://doi.org/10.36958/sep.v6i1.121>

Análisis factorial exploratorio escala evaluación del desempeño docente, maestrías en psicología, USAC

Exploratory factor analysis Scale Teacher performance evaluation, master's degrees in psychology, USAC

Francisco José Ureta Morales

Maestría en medición, evaluación e investigación educativas
Universidad de San Carlos de Guatemala
furetagt@gmail.com
<https://orcid.org/0000-0003-2930-6946>

Recibido: 20/09/2022

Aceptado: 06/06/2023

Publicado: 22/06/2023

Resumen

OBJETIVO: realizar un análisis psicométrico de la confiabilidad y validez de constructo de la Encuesta de Evaluación del Desempeño Docente (EEDD), es parte de los procesos de autoevaluación y mejoramiento de los programas de maestría en la Escuela de Ciencias Psicológicas de la USAC. **MÉTODO:** se obtuvieron 834 respuestas de estudiantes de 5 programas de maestría, con esto se calcularon el Alfa de Cronbach, prueba de normalidad de la distribución de los datos de Kolmogórov-Smirnov, cálculo del análisis factorial exploratorio (AFE) con la medida de Kaiser-Meyer-Olkin de adecuación al muestreo, prueba de esfericidad de Bartlett, rotación Varimax con normalización Kaiser y el método de extracción por análisis de componentes principales. **RESULTADOS:** todas las evidencias indicaron que se tenían las condiciones para la realización del AFE, la EEDD tiene una confiabilidad alta de 0.961, no se confirmó el modelo inicial de 7 factores sino de 3 factores que produjo este análisis. **CONCLUSIÓN:** los tres factores alcanzaron apropiados índices Alfa de Cronbach (0.940, 0.893 y 0.895), que les ofrece consistencia interna y estabilidad. La principal limitación es que la EEDD tiene múltiples aplicaciones a los maestrantes, que podrían asumir patrones de respuesta; se sugieren posteriores análisis factoriales confirmatorios para consolidar el nuevo modelo.

Palabras claves

evaluación del desempeño docente, análisis factorial exploratorio, maestrías en psicología.

Abstract

OBJECTIVE: carrying out a psychometric analysis of the reliability and construct validity of the Teacher Performance Evaluation Survey (EEDD), is part of the self-evaluation and improvement processes of master's programs at the USAC School of Psychological Sciences. **METHOD:** 834 responses from students from 5 master's programs were obtained, with this the Cronbach's Alpha was calculated, the Kolmogorov-Smirnov test of normality of the distribution of the data, calculation of the exploratory factor analysis (EFA) with the Kaiser-Meyer-Olkin measure of adequacy to sampling, Bartlett's test of sphericity, Varimax rotation with Kaiser normalization and the extraction method by principal component analysis. **RESULTS:** all the evidence indicated that the conditions were in place to carry out the exploratory factor analysis, the EEDD has a high reliability of 0.961, the initial 7-factor model was not confirmed, but the 3-factor model produced by this analysis. **CONCLUSION:** the three factors reached appropriate Cronbach's Alpha indices (0.940, 0.893 y 0.895), which offers them internal consistency and stability. The main limitation is that the EEDD has multiple applications to students, who could assume response patterns; further confirmatory factor analyzes are suggested to consolidate the new model.

Keywords

evaluation of teaching performance, exploratory factor analysis, master's in psychology

Introducción

Esta investigación es parte de las autoevaluaciones requeridas en los procesos de mejoramiento académico y educativo de los postgrados en la Universidad de San Carlos de Guatemala (USAC) y en la Escuela de Ciencias Psicológicas (ECPs). Para fundamentar estas autoevaluaciones el departamento de Estudios de Postgrado de la ECPs formuló un modelo de evaluación de sus programas de maestría (Ureta, 2020 y Ureta, 2021) basado en el modelo -CIPP- contexto, insumo, proceso y producto (Stufflebeam, 2001) y el de la Agencia Centroamericana de Acreditación de Postgrado -ACAP- (ACAP, 2015). Uno de los principales indicadores en los procesos de autoevaluación es la evaluación del desempeño de sus docentes por medio de una escala que solicita la opinión de los maestrantes, dicha escala se aplica a todos los estudiantes durante el último mes de cada semestre de estudios. El problema evaluativo que origina este análisis es verificar si la Escala de Evaluación del desempeño docente (EEDD) posee apropiadas características psicométricas, concretamente su confiabilidad y validez de constructo. Problemática que se concretó en las siguientes preguntas: ¿Cuál es el nivel de confiabilidad de la EEDD según el coeficiente Alfa de Cronbach? ¿El modelo teórico sobre el que se construyó la EEDD se confirma empíricamente con el AFE?

En el año de 1999 la USAC aprobó el Reglamento de Evaluación y Promoción del Personal Académico de la Universidad. En ese reglamento el primer artículo se refiere a la definición de los procesos de evaluación docente: Artículo 1. Definición: La evaluación del personal académico es el proceso instituido por la Universidad de San Carlos de Guatemala, con el fin de medir y valorar el desempeño de las funciones del personal académico e impulsar el desarrollo y perfeccionamiento de su función académica (USAC, 2006, p. 75).

La evaluación del desempeño docente pretende que se desempeñen lo mejor posible para favorecer los aprendizajes de sus estudiantes, así como mejorar constantemente su práctica pedagógica al reconocer sus fortalezas y debilidades, por medio de la formación continua en servicio. Estas evaluaciones ofrecen información útil y relevante para la toma de decisiones. En América Latina se han identificado los siguientes aspectos influidos por dichas decisiones: entrega de licencias o acreditaciones, incrementos salariales, promoción del escalafón docente, desarrollo de programas de formación docente por pares, programas de acompañamiento y desarrollo docente y aplicación de sanciones que puede incluir la pérdida del puesto de trabajo (Ureta et al., 2019). Por todas estas situaciones, se justifica analizar las características psicométricas de la EEDD y confirmar que provee información consistente y robusta para la toma de decisiones en las autoevaluaciones de los programas de maestría en la ECPs.

Fundamentación teórica

El desarrollo de la teoría psicométrica se evidencia en sus tres principales teorías, la teoría clásica de los tests (TCT), la teoría de respuesta al ítem (TRI) y la teoría de la generalizabilidad (TG). La EEDD fue elaborada bajo los principios de la TCT, por lo que se explicará brevemente la misma. La TCT fue elaborada a inicios del siglo XX por Spearman casi en su

totalidad, se formalizó con la estadística que predominaba en esos años basada en las correlaciones de Pearson y en el modelo matemático de la curva normal. Este modelo psicométrico de TCT parte de tres ideas básicas, la puntuación observada o real que una persona obtiene en una prueba, la puntuación real que la persona obtendría si se lograran eliminar todas las fuentes que afectan a la confiabilidad del instrumento en un número infinito de aplicaciones en diversos momentos y condiciones, así como la puntuación del error que es la diferencia entre la puntuación real y la verdadera, puede ser negativa y se estima la varianza del error con el cálculo de la confiabilidad de la consistencia interna (Alfa de Cronbach) y la confiabilidad de formas alternativas de los tests o instrumentos de evaluación. Las pruebas o instrumentos de evaluación se definen como el procedimiento o dispositivo evaluativo por medio del cual se capta una muestra conductual u opiniones de evaluados o personas que los responden, en un dominio o concepto específico que luego es evaluado y calificado por medio de un proceso estandarizado. (Tapia, 2018; Villareal et al., 2015; Hogan, 2015; APA, 2010; Martínez et al., 2006).

Como consecuencia de la facilidad de su implementación e interpretación la TCT ha sido la teoría psicométrica dominante, como los principales beneficios se pueden citar su relativamente fácil ejecución y que se puede aplicar a una vasta variedad de disposiciones de medición (APA, 2010).

Confiabilidad de los instrumentos de evaluación

En psicometría se define la confiabilidad de la siguiente forma:

Confiabilidad (reliability) s. capacidad de un instrumento de medición (p. ej., una prueba) para medir un atributo en forma consistente, que arroja los mismos resultados entre múltiples aplicaciones a la misma muestra. El índice básico de confiabilidad es el coeficiente de correlación. (APA, 2010, p. 275).

Afectan la confiabilidad los errores aleatorios de la medición ocasionadas por diversas razones de los evaluados (cansancio, motivación, maduración, etc.) o del ambiente como luz, ruido, instrucciones del evaluador, etc. (Villareal et al., 2015; Martínez et al., 2006). Para realizar los cálculos de un índice de confiabilidad hay varios procedimientos evaluativos, en general, el que se aplica más y que posee facilidad en su comprensión es el índice Alfa de Cronbach, el cual permite realizar el análisis de la consistencia interna de los tests o instrumentos de evaluación. La fórmula que se utiliza para su cálculo es relativamente fácil y puede consultarse en este texto (Celina y Campo, 2005, p. 576).

El coeficiente alfa de Cronbach analiza alternativas de respuestas politómicas, sus resultados se ubican entre 0 y 1, se interpretan de la siguiente forma: 0 representa nula confiabilidad y 1 indica confiabilidad total. Cuando se analizan los ítems individualmente el Alfa de Cronbach se considera como criterio básico con un resultado apropiado entre 0.7 a 1. (Corral, 2009; George & Mallery, 2003). Su interpretación puede hacerse de la siguiente forma: muy alta

confiabilidad de 0,81 a 1,00; alta de 0,61 a 0,80; moderada entre 0,41 y 0,60; baja de 0,21 a 0,40 y muy baja cuando se ubica de 0,01 a 0,20. (Colina et al., 2008).

Validez de los instrumentos de evaluación

A nivel general se dice que la validez es el grado o capacidad de una prueba o medición para reflejar con exactitud y precisión el constructo o concepto que se pretende medir. Específicamente se refiere al grado en que la teoría y evidencia respaldan las interpretaciones hechas, a partir de los puntajes que se obtienen al aplicar instrumentos de evaluación. Inicialmente se definían varios tipos de validez, ahora se habla más de la validez de constructo, como la capacidad que tiene para evaluar un concepto, rasgo o capacidad teórica como la inteligencia, personalidad u opiniones (Villareal et al., 2015; APA, 2010; Martínez et al., 2006).

Los más utilizados para la estimación de la validez de constructo son los análisis factoriales exploratorio (AFE) y confirmatorio (AFC), los cuales analizan y ofrecen evidencias de la estructura interna de los instrumentos, por medio de posibles patrones de asociación entre las respuestas dadas a los ítems o preguntas por las personas que los responden. Este grupo de análisis estadísticos pretenden identificar las dimensiones comunes que subyacen en el desempeño de los evaluados en muchas medidas diferentes, permiten la reducción de un grupo de intercorrelaciones entre las variables manifiestas en el instrumento, a un conjunto más reducido de factores o variables latentes, de allí su nombre. Los resultados se presentan en una matriz factorial, la cual incluye las correlaciones entre las variables originales y los factores recién identificados y establecidos. A estas correlaciones se le denominan pesos o cargas, la correlación de ellos con los nuevos factores, se consideran como viables las cargas que superan el 0.30 como notables, si tienen mayor carga serán más relevantes o notables con el modelo factorial producido (Hogan, 2015; APA, 2010). En este sentido el AFE se puede definir de la siguiente forma:

El análisis factorial exploratorio es un conjunto de técnicas que se aplican cuando:

...se carece de una teoría fuerte y los datos observados se exploran libremente en búsqueda de patrones significativos entre las observaciones. Es decir, los datos son examinados para descubrir las variables subyacentes (latentes) que explican las interrelaciones entre un conjunto mayor de variables observables (manifiestas). (APA, 2010, p. 89).

Modelo pedagógico de base para la construcción de la EEDD

Para la evaluación de los docentes de los programas de maestría de la ECPs se diseñó y aplicó una escala tipo Likert que incluyó 20 reactivos y 4 opciones (nunca, a veces, casi siempre y siempre), se conformaron 7 subescalas: liderazgo docente 4 preguntas, experiencia profesional y/o técnica 4 preguntas, conocimiento del entorno 3 preguntas, creatividad e innovación 2 preguntas, gestión de formación 2 preguntas, relaciones interpersonales 2 preguntas y evaluación personal con 3 preguntas. Se confirieron punteos a las respuestas ofrecidas por los maestrantes, con lo cual el puntaje menor posible es 20 y el mayor posible 80 puntos, las

opciones de respuesta son nunca 1 punto, a veces 2 puntos, 3 puntos casi siempre y siempre con 4 puntos.

Con base en las preguntas formuladas, se planteó el objetivo general de realizar un análisis psicométrico de la confiabilidad y validez de constructo de la EEDD. Los objetivos específicos son analizar el nivel de confiabilidad de la EEDD según el coeficiente Alfa de Cronbach y, confirmar si el modelo teórico de 7 subescalas o factores sobre el que se construyó la EEDD tiene base empírica por medio del AFE.

Materiales y métodos

Se diseñó esta investigación con un enfoque empírico–analítico, con énfasis en el elemento psicométrico y exploratorio, ya que pretende identificar si la EEDD puede medir el constructo o concepto de desempeño docente en la población de maestrantes de los 5 programas de maestrías vigentes en la ECPs. La EEDD se aplicó en dichos programas de maestría desde mayo del 2020 a octubre de 2021, al concluir cada semestre de clases, por lo que la muestra se conformó de los estudiantes inscritos en los programas durante el tiempo arriba indicado. Aunque se considera obligatorio que todos los estudiantes evalúen a sus profesores, no todos lo hacen, por lo que la muestra analizada fue de 834 respuestas por voluntad de los maestrantes, quedó conformada de la siguiente forma: Maestría en Análisis Social de la Discapacidad (MASDIS) 83 respuestas de estudiantes, Maestría en Gestión Humana y Organizacional (MGHO) 534 respuestas de estudiantes, Maestría en Psicología Forense (MAFOR) 157 respuestas de estudiantes, Maestría en Atención Integral de Conductas Adictivas (MACA) 40 respuestas de estudiantes y Maestría en Análisis Psicosocial de las Violencias (MAVIS) con 20 respuestas de estudiantes.

Hay que indicar que los maestrantes llenan en varias ocasiones el instrumento, uno por cada docente que evalúan durante el semestre de estudios. La respuesta se realiza en línea por medio de la aplicación Google Forms, se les motiva para que todos los maestrantes respondan la EEDD para contar con la opinión de todos, pero siempre hay estudiantes que no lo llenan por cada docente o solo contestan una. Es un estudio descriptivo con enfoque cuantitativo de metodología positivista. El procedimiento seguido en este estudio fue realizar la definición del instrumento (conformado por 7 subescalas de liderazgo, experiencia profesional y/o técnica, conocimiento del entorno, creatividad e innovación, gestión de formación, relaciones interpersonales y evaluación personal), pilotaje inicial y ajustes a la EEDD. Con aplicación periódica al concluir cada semestre de estudios durante dos años. Luego la recopilación y limpieza de las bases de datos generadas, traspaso al programa SPSS donde se condujeron los análisis. El siguiente paso fue el cálculo del coeficiente Alfa de Cronbach de cada ítem y del total de la EEDD, cálculo de la prueba de normalidad de la distribución de datos de Kolmogórov-Smirnov. Para concluir con el cálculo del AFE con los siguientes estadísticos, medida de Kaiser-Meyer-Olkin de adecuación al muestreo, prueba de Bartlett de esfericidad, matriz de varianzas rotadas que explican los factores producidos por el modelo, gráfica de sedimentación, rotación Varimax con normalización de Kaiser y el método de extracción por análisis de componentes principales (Gómez et al., 2020).

Resultados, análisis y discusión

El análisis de confiabilidad corrido con el Alfa de Cronbach ofreció el resultado ($\alpha=.961$) que indica una alta consistencia interna de la EEDD con buena confiabilidad, ningún ítem mostró una correlación negativa con la escala completa. Hay que indicar que es posible que la aplicación reiterativa del instrumento a los mismos estudiantes afecte el índice Alfa de Cronbach, facilitando sus resultados altos y la consistencia interna de la EEDD.

La prueba Kolmogórov-Smirnov evidenció que la distribución de los datos de los 20 ítems individualmente, el cálculo del desempeño y los tres factores encontrados responden a los criterios de normalidad, ya que en todos se obtuvo un valor $p < 0.001$ (un valor $p > 0.05$ indicaría que no responden a un comportamiento normal), por lo que se acepta la hipótesis de que los datos se comportan de forma normal. Se procedió a calcular los coeficientes de adecuación del muestreo de Kaiser-Meyer-Olkin (0.961) y la prueba de Bartlett de esfericidad (Aprox. Chi-cuadrado de 13980.08, $gl=190$, $p < 0.001$), para afirmar si el tamaño de la muestra de respuestas es suficiente para realizar el AFE. Los resultados de ambas pruebas indican que la muestra es suficiente para la realización del análisis factorial.

La tabla 1 muestra correlaciones positivas y superiores a 0.553 entre todos los ítems y el total de la EEDD, los Alfas de Cronbach se mantienen estables si se elimina algún ítem o pregunta. Como resultado del AFE corrido con el método de extracción con análisis de componentes principales y rotación Varimax con normalización de Kaiser, se obtuvo evidencia de que la EEDD no conserva la estructura de 7 factores del modelo teórico con que se construyó. El modelo evidenciado incluye 3 factores que pueden explicar el 69.874% del total de la varianza de la EEDD, que al ser mayor del 50% indica que el modelo se adecúa a esos tres factores, la rotación convergió luego de 6 iteraciones. De los 20 ítems 9 se ubican en el primer factor (45%), 5 en el segundo factor (25%) y los restantes 6 (30%) en el tercer factor de los obtenidos en el análisis; por lo que se afirma que los ítems se están relacionando bien con sus factores principales. Los ítems que se ubican en más de un factor, se les confirmó en el factor donde obtuvieron mayor índice luego de la rotación, aparecen en negrita en dicha tabla. Los 20 ítems tienen niveles de extracción de 0.437 para arriba, lo que se considera de que están aportando apropiadamente a la conformación del modelo evidenciado en este análisis.

La tabla 2 muestra las propiedades psicométricas de la EEDD según el modelo producido, el factor 1 explica la mayoría de la varianza común del instrumento (57.980) y tiene también, la mayor cantidad de ítems (9) y el Alfa de Cronbach más alto de los tres factores (0.940). La varianza explicada es la diferencia entre la varianza de error en la predicción del modelo y la varianza del destino (Mavrou, 2015), en este caso este factor puede explicar la mayoría de la varianza con respecto a los dos siguientes factores y conforma el grupo de ítems más consistente del instrumento, referido al apropiado manejo y aplicación de la teoría por los docentes en las aulas. Los otros dos factores tienen un Alfa de 0.893 y 0.895, lo que denota que los tres factores del nuevo modelo son confiables y tienen consistencia interna. La figura 1 muestra la gráfica de sedimentación producida por la rotación, en ella se pueden observar 3 puntos a la izquierda (de color rojo) en caída hasta estabilizarse posteriormente a ellos, estos tres factores tienen autovalores iniciales de 11.596, 1.265 y 1.114, el resto tienen autovalores

menores de 1 por lo que ya no conforman factores estables. Debido a los ítems que se agrupan por factores del nuevo modelo son diferentes al agrupamiento original, se les llamará de la siguiente forma, factor 1 manejo y aplicación de la teoría, factor 2 relaciones con los maestros y factor 3 manejo didáctico del curso.

Tabla 1. Matriz de carga factorial de cada ítem, extracción, correlaciones ítem-prueba y alfa de Cronbach si se suprime el elemento de la EEDD.

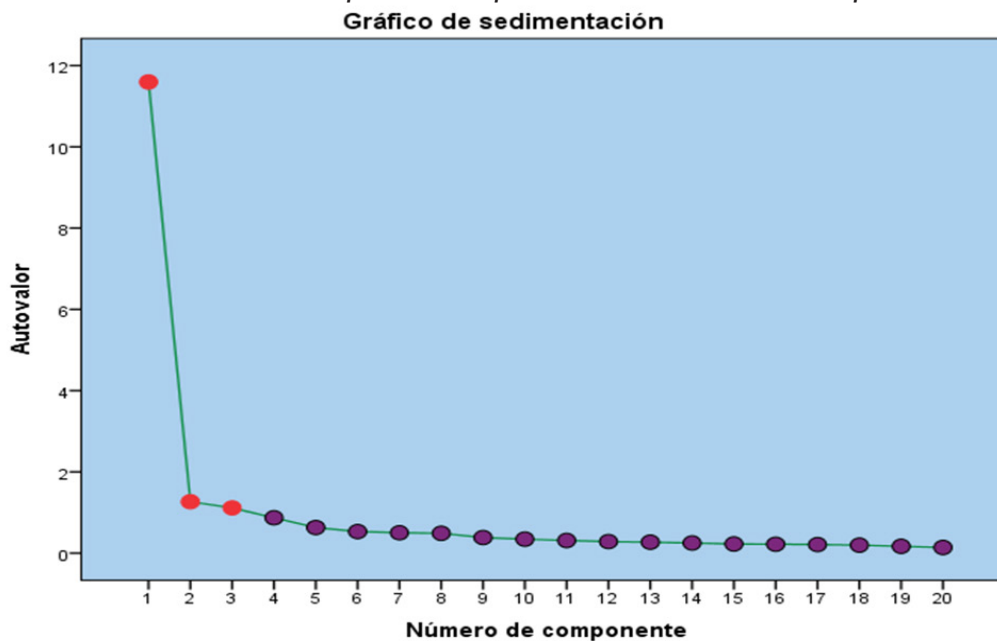
Ítems	Factor 1	Factor 2	Factor 3	% de varianza aportado	Extracción	Correlación total de elementos corregida	Alfa de Cronbach si el elemento se suprime
2.2 Argumenta explicaciones de los temas y su posición frente a ellos.	0.812			2.655	0.766	0.759	0.958
2.1 Demuestra dominio de los contenidos que aborda en el curso	0.762			3.153	0.688	0.724	0.959
2.4 Contextualiza los contenidos de las actividades de aprendizaje con las características y problemas del medio social	0.757			2.438	0.726	0.770	0.958
3.3 Toma situaciones o problemas del medio y los convierte en objeto de aprendizaje. (ej.)	0.719			1.558	0.730	0.798	0.958
2.3 Maneja referencias de material educativo, bibliografía, internet u otras fuentes de información para orientación de estudiantes	0.704			2.507	0.616	0.688	0.959
3.1 Fomenta la crítica y el debate académico como parte de las acciones de formación	0.673	0.403		1.917	0.688	0.780	0.958
4.2 Es creativo, dinámico, genera interés y motivación en el desarrollo de las actividades	0.667			1.353	0.693	0.799	0.958
7.3 Usted se inscribiría nuevamente para recibir un curso con el/la docente evaluado	0.644	0.432	0.403	.694	0.765	0.853	0.957
4.1 Utiliza estrategias didácticas que facilitan la apropiación de conocimientos (Estudio de casos, proyectos, simulaciones y otros)	0.635			1.421	0.626	0.754	0.958
6.1 Trata a estudiantes con equidad y respeto		0.860		1.101	0.806	0.625	0.960
6.2 Se muestra receptivo/a a las críticas y sugerencias planteadas por los estudiantes		0.795		1.053	0.808	0.743	0.958
7.1 Fomenta el vínculo académico e inspira respeto a los estudiantes		0.781		0.985	0.794	0.744	0.958
3.2 Tiene en cuenta y respeta los saberes y opiniones de los aprendices relacionados con el proceso formativo	0.440	0.703		1.723	0.765	0.784	0.958
7.2 Su presentación personal es adecuada		0.501		0.838	0.417	0.561	0.960
1.1 Recibió el programa del curso, con indicaciones precisas del desarrollo de este			0.764	57.980	0.672	0.592	0.960
1.3 Desarrolla las actividades y clases según el plan presentado al comienzo del programa.	0.420		0.734	5.569	0.768	0.755	0.958
1.4 Es puntal para iniciar las actividades? (según la hora programada)			0.732	4.352	0.604	0.553	0.961
5.2 Respeta los acuerdos sobre procesos, fechas y modalidades de evaluación			0.633	1.126	0.623	0.694	0.959
1.2 ¿Dedica el tiempo adecuado a la orientación de las actividades de formación?	0.479		0.608	6.325	0.699	0.779	0.958
5.1 Realiza evaluaciones coherentes con los contenidos, actividades y metodologías desarrolladas en las competencias	0.502		0.557	1.252	0.719	0.812	0.957

Tabla 2. Propiedades psicométricas de la EEDD

Ítems	Ítems	% de varianza explicada	Alfa de Cronbach
Factor 1 Manejo y aplicación de la teoría	9	57.980	0.940
Factor 2 Relaciones con los maestrantes	5	6.325	0.893
Factor 3 Manejo didáctico del curso	6	5.569	0.895
Total EEDD	20	69.874	0.961

Estos resultados coinciden con lo reportado por Colina et al., (2008) en que el modelo original del instrumento de evaluación del desempeño docente aplicado a estudiantes tenía más dimensiones 10 teóricamente, que el producido por su AFE que evidenció 8 dimensiones, refirieron un índice ($\alpha=.9630$). El modelo teórico supuso más dimensiones que las producidas por el modelo producido y un alfa alto, igual que la EED. Martínez-González et al., (2008) validaron un instrumento de desempeño docente en cursos de especialidades médicas, también reportó una reducción de sus 10 factores del modelo teórico inicial a 5 factores evidenciados por el AFE, indicaron que se obtuvo un alto coeficiente Alfa ($\alpha=.9596$), resultados que vuelven a coincidir con lo trabajado en la EEDD; reducción de factores o dimensiones y alta confiabilidad.

Figura 1. Gráfico de sedimentación producido por la rotación de los componentes de la EEDD



Nota. Esta figura muestra la gráfica de sedimentación producida por la rotación, en rojo se aprecian los tres puntos que dan origen a los tres factores producidos, ya que el resto de puntos se alinea de forma horizontal.

Dios et al., (2018) también reportan una reducción de factores o subescalas de 4 del modelo teórico original a 2 del modelo producido por el AFE, indican un coeficiente Alfa ($\alpha=.948$) que ofrece una confiabilidad alta, también coincide con lo encontrado en este estudio. Romero (2014) desarrolló un cuestionario de evaluación del desempeño docente aplicado a estudiantes, nuevamente reportó un ($\alpha=.916$) de confiabilidad alta, sin embargo, en esta ocasión se mantuvo la cantidad de factores del modelo teórico inicial, ocho, luego de correr el AFE, contrastando con lo reportado por anteriores investigadores. Situación que también refirió Prieto (2006) obtuvo un coeficiente Alfa alto con buena confiabilidad ($\alpha=.9541$) y, mantuvo sus factores o dimensiones de la escala de evaluación docente de la Corporación Universitaria Iberoamericana, en el modelo teórico original hay tres factores, que fueron los mismos alcanzados con el AFE.

Finalmente, Ramos y Rueda (2020) indican que los resultados y la opinión de los alumnos solo es un factor por considerar en la evaluación docente, hay otros elementos como la experiencia y su desarrollo profesional, si pretende ser formativa, debiera ser participativa y no jerárquica, positiva ya que busca aumentar la autoestima docente, así como propositiva al generar procesos de reflexión y auto reflexión para mejor el desempeño. También sugieren que estos procesos evaluativos sean supervisados colaborativamente entre sujetos involucrados, realizada de forma abierta y por personas con formación y experiencia, así como con cierto reconocimiento entre el grupo de profesionales, además de ser ejecutada por personal profesional, amable y cortés.

Conclusiones

Este estudio es parte de los procesos de autoevaluación y mejoramiento de los programas de maestría de la ECPs de la USAC en Guatemala, tuvo como objetivo realizar el análisis psicométrico de confiabilidad y validez de constructo de la EEDD aplicada a los maestrantes de los 5 programas de maestrías en curso. La EEDD tuvo un total de 834 aplicaciones con los maestrantes y obtuvo un Alfa de Cronbach de 0.961, el cual indicó una alta consistencia interna con buena confiabilidad, ningún ítem mostró una correlación negativa con la escala completa. La prueba Kolmogórov-Smirnov calculada a las distribuciones de los 20 ítems individualmente, del puntaje del desempeño docente y a los tres factores definidos por el nuevo modelo evidenciaron que responden a los criterios de normalidad. Esta muestra fue apropiada para la realización del AFE, según la evidencia aportada por los coeficientes de adecuación del muestreo de Kaiser-Meyer-Olkin y la prueba de esfericidad de Bartlett.

El modelo factorial original de 7 componentes no pudo ser comprobado, este análisis indicó que se produjo un nuevo modelo factorial con 3 componentes o factores. Este modelo explica el 69.874% del total de la varianza de la EEDD que es considerado apropiado, los tres nuevos factores que asociaron los ítems alcanzaron Alfas de Cronbach de 0.940, 0.893 y 0.895 que los hace confiables y con consistencia interna. La nueva agrupación de los ítems en los tres factores se les designó de la siguiente forma, factor 1 manejo y aplicación de la teoría, factor 2 relaciones con los maestrantes y factor 3 manejo didáctico del curso. En términos generales, se puede afirmar que la EEDD evalúa con propiedad el constructo de desempeño docente,

ya que sus características psicométricas así lo evidencian, buena confiabilidad por el Alfa de Cronbach alto y un modelo factorial de 3 factores comprobado por el AFE. Como se indicó con antelación, es común que los factores del modelo teórico inicial se reduzcan luego del AFE (Colina et al., 2008 y Martínez-González et al., 2008).

Como principal limitación del presente estudio se indica que la EEDD tiene múltiples aplicaciones a los maestrantes de los 5 programas de maestría, ya que es el mismo instrumento que se aplica al concluir los semestres y, deben llenar uno por cada docente evaluado. La repetición y familiaridad con el instrumento y sus respuestas podrían afectar la confiabilidad, ya que se podrían asumir patrones de respuesta por parte de los maestrantes.

Perspectivas de futuro

A futuro se sugiere que se continúen los análisis psicométricos conforme aumente la muestra de maestrantes, así como que se puedan realizar análisis factoriales confirmatorios para consolidar el nuevo modelo de tres factores encontrado. El cual se define como:

...conjunto de procedimientos utilizados en el Análisis Factorial para demostrar que un grupo de variables posee la estructura factorial anticipada por la teoría. En otras palabras, el análisis factorial confirmatorio proporciona pruebas estadísticas formales a favor de una hipótesis previa acerca de las variables específicas subyacentes (latentes) que se cree que explican los datos obtenidos en un conjunto de variables observadas (manifiestas). (APA, 2010, p. 89).

También sería conveniente iniciar los análisis psicométricos de otros instrumentos de evaluación que se aplican en la ECPs como la escala de evaluación del programa de maestría, escala de evaluación de autoridades, escala de evaluación de empleadores, escala de evaluación de la malla curricular y la escala de evaluación de docentes del nivel de licenciatura.

Como indicaron Ramos y Rueda (2020), hay que cuidar detalles en los procesos de evaluación docente, considerar otros factores que solamente la opinión de los estudiantes, realizarla en un ambiente abierto, profesional, de amabilidad y cortesía, por personas con formación y experiencia para que sean resultados apropiados al desempeño realizado por los docentes de maestrías. Que se promuevan las reflexiones y autorreflexiones de los mismos profesores, para que, desde su propia experiencia y opiniones recibidas, se pueda mejorar su desempeño.

Referencias

- Agencia Centroamericana de Acreditación de Postgrado (ACAP). (2015). Guía de Autoevaluación de la ACAP. Consejo de acreditación de la ACAP.
- American Psychological Association [APA]. (2010). Diccionario conciso de psicología. Editorial El Manual Moderno.
- Celina, H. y Campo, A. (2005). Aproximación al uso del coeficiente alfa de Cronbach. *Revista Colombiana de Psiquiatría*, XXXIV (4), 572-580. <https://www.redalyc.org/articulo.oa?id=80634409>
- Colina, Z., Medina, N., Parra, D., Cendrés, J. y Montoya, C. (2008). Modelo para la evaluación del desempeño docente en la función docencia universitaria. *Revista Investigación Educativa*, 12(22), 99-126. <https://revistasinvestigacion.unmsm.edu.pe/index.php/educa/article/view/3867/3098>
- Corral, Y. (2009). Validez y confiabilidad de los instrumentos de investigación para la recolección de datos. *Revista Ciencias de la Educación*, 19(33), 228-247. <http://servicio.bc.uc.edu.ve/educacion/revista/n33/art12.pdf>
- Dios, I., Calmaestra, J. y Rodríguez-Hidalgo, A. (2018). Validación de la escala de competencias docentes organizacionales y didácticas para educadores. *Revista mexicana de investigación educativa*, 23(76), 281-302. http://www.scielo.org.mx/scielo.php?script=sci_arttext&pid=S1405-66662018000100281&lng=es&tlng=es.
- George, D., & Mallery, P. (2003). *SPSS for Windows step by step: A simple guide and reference*. 11.0 update (4th ed.). Allyn & Bacon.
- Gómez, A., Vinaccia, S. y Sierra, W. (2020). Propiedades psicométricas de la Escala de Salud Mental Positiva en jóvenes colombianos: un estudio exploratorio. *Revista CES Psicología*, 13(2), 102-112. <http://dx.doi.org/10.21615/cesp.13.2.7>
- Hogan, T. (2015). *Pruebas psicológicas: una introducción práctica*. 2ª ed. Traducción Jesús Cisneros. Editorial El Manual Moderno.

- Martínez, M.; Hernández, M. y Hernández, M. (2006). *Psicometría*. Alianza Editorial.
- Martínez-González, A., Lifshitz-Guinzberg, A., Ponce-Rosas, R. y Aguilar, V. (2008). Evaluación del desempeño docente en cursos de especialización médica. Validación de un instrumento. *Revista Médica del Instituto Mexicano del Seguro Social*, 46(4), 375-382. <https://www.redalyc.org/articulo.oa?id=457745522006>
- Mavrou, I. (2015). Análisis factorial exploratorio: Cuestiones conceptuales y metodológicas. *Revista Nebrija de Lingüística Aplicada a la Enseñanza de Lenguas*, (19), 71–80. <https://doi.org/10.26378/rnlael019283>
- Prieto, L. (2006). Análisis psicométrico de la escala de evaluación docente de la Corporación Universitaria Iberoamericana. *Suma Psicológica*, 13(2), 193-215. <https://www.redalyc.org/articulo.oa?id=134216870006>
- Ramos, L. y Rueda, M. (2020). Rasgos distintivos de las evaluaciones formativas para el desempeño docente. *Perfiles Educativos*, 42(169), 144-159. <https://doi.org/10.22201/ii-sue.24486167e.2020.169.59287>
- Romero, T. (2014). Cuestionario de opinión para la evaluación del desempeño docente en la UNAN-Managua, Nicaragua, FAREM-Chontales. *Revista Electrónica "Actualidades Investigativas en Educación"*, 14(2), 1-28. <https://www.redalyc.org/articulo.oa?id=44731371010>
- Stufflebeam, D. (2001). The metaevaluation imperative. *American Journal of Evaluation*, 22 (2), 183-209. <https://doi.org/10.1177/109821400102200204>
- Tapia, I. (2018). Análisis psicométrico de exámenes de Matemáticas y Lenguaje y Comunicación en CONALEP Estado de México. *Debates en Evaluación y Currículum /Congreso Internacional de Educación: Evaluación 2018 /Año 4, No. 4. 4 (4)*, 137-147. <https://posgradoeducacionuatx.org/pdf2018/A025.pdf>
- Universidad de San Carlos de Guatemala. (2006). *Leyes y Reglamentos de la Universidad de San Carlos de Guatemala*. USAC, Guatemala.
- Ureta, F.; Espinoza, E. y Mó, R. (2019). Evaluación del desempeño docente: mitos y realidades. Dirección General de Evaluación e Investigación Educativa, Ministerio de Educación, Guatemala. <https://www.mineduc.gob.gt/digeduca/documents/investigaciones/2019/Evaluacion-del-desempeno-docente.pdf>

Ureta, F. (2020). Propuesta de evaluación de maestrías. USAC, Departamento de Estudios de Postgrado, Escuela de Ciencias Psicológicas.

Ureta, F. (2021). Experiencias y lecciones aprendidas de la primera autoevaluación del programa de Maestría en Psicología Forense de la Escuela de Ciencias Psicológicas, USAC. Revista Científica del Sistema de Estudios de Postgrado de la Universidad de San Carlos de Guatemala, 4(1), 21–48. <https://doi.org/10.36958/sep.v4i1.75>

Villareal, M.; Alfaro-Rojas, L. y Brizuela, A. (2015). Construcción de pruebas estandarizadas en el ámbito de la medición educativa y psicológica. Cuadernos metodológicos 8 del Instituto de Investigaciones Psicológicas, Universidad de Costa Rica.

Sobre el autor

Francisco José Ureta Morales

Es estudiante de Doctorado en Psicología Forense Universidad Internacional Iberoamericana - México, Maestría en medición, evaluación e investigación educativas UVG, Licenciado en Psicología USAC, Profesor de licenciatura y postgrado en la Escuela de Ciencias Psicológicas y Facultad de Humanidades, Investigador en DIGI – USAC, con 30 años de experiencia en evaluaciones e investigaciones psicoeducativas con análisis psicométricos.

Financiamiento de la investigación

Con recursos propios.

Declaración de intereses

El autor declara no tener ningún conflicto de intereses, que puedan haber influido en los resultados obtenidos o las interpretaciones propuestas.

Declaración de consentimiento informado

El estudio se realizó respetando el Código de ética y buenas prácticas editoriales de publicación.

Derecho de uso

Copyright (c) (2023) por Francisco José Ureta Morales

Este texto está protegido por la [Licencia Creative Commons Atribución 4.0 Internacional](#).



Este texto está protegido por una licencia
[Creative Commons 4.0](#).

Es libre para compartir, copiar y redistribuir el material en cualquier medio o formato y adaptar el documento, remezclar, transformar y crear a partir del material para cualquier propósito, incluso comercialmente, siempre que cumpla la condición de atribución: debe reconocer el crédito de una obra de manera adecuada, proporcionar un enlace a la licencia, e indicar si se han realizado cambios. Puede hacerlo en cualquier forma razonable, pero no de forma tal que sugiera que tiene el apoyo del licenciante o lo recibe por el uso que hace.