

Propiedades psicométricas del Cuestionario de Evaluación de la Competencia Docente

Psychometric properties of the Evaluation Teacher Competence Questionnaire

Erika Paola Reyes Piñuelas*, Edna Luna Serrano y Joaquín Caso Niebla

Universidad Autónoma de Baja California

El objetivo de este estudio fue reportar los resultados de los análisis de confiabilidad y validez del Cuestionario de Evaluación de la Competencia Docente con base en la opinión de los alumnos, en una población de estudiantes universitarios mexicanos. La estructura teórica del cuestionario corresponde al Modelo de Evaluación de Competencias Docentes y es una escala con 35 reactivos tipo Likert con cinco niveles de respuesta. Se realizó un análisis de unidimensionalidad del test con una muestra de 128,791, un Análisis Factorial Exploratorio (AFE) con 66,148 datos y un Análisis Factorial Confirmatorio (AFC) con 462 datos. Los resultados del análisis de unidimensionalidad con el modelo Rasch mostraron que 32 reactivos evalúan el mismo rasgo latente; en el AFE se obtuvo una estructura bifactorial, planeación y gestión del proceso de enseñanza-aprendizaje, y conducción y valoración del proceso de enseñanza-aprendizaje, ambos factores explican el 75.02% de varianza y un alfa de Cronbach de .98; en el AFC los valores de índices de bondad de ajuste fueron RMSEA=.07, CFI=.93, y SRMR=.02 la estructura del modelo al cual ajustan los datos de manera aceptable quedó conformada por dos factores con 25 reactivos congruentes con el modelo que le subyace.

Palabras clave: Validez de constructo, Evaluación, Competencias, Evaluación docente, Cuestionarios con base en la opinión de los estudiantes.

The objective of this study was to report the results of the analysis of reliability and validity of the Assessment Questionnaire Teaching competence based on the opinion of the students, in a population of Mexican university students. The theoretical structure of the questionnaire corresponds to the evaluation model of teaching competence and is a scale with 35 Likert five response levels. An analysis of unidimensionality of the test with a sample of 128,791, an Exploratory Factor Analysis (EFA) with 66,148 data and Confirmatory Factor Analysis (CFA) with 462 data. The results of the analysis of unidimensionality using Rasch modeling showed 32 items assess the same latent variable; AFE was obtained in the bivariate structure, planning and management of the teaching-learning process, and conduct and evaluation of teaching-learning process, both factors account for 75.02% of variance and a Cronbach's alpha of 0.98, in the AFC, the values of goodness of fit indices were RMSEA=.07, CFI=.93, and SRMR=.02 model structure which fit the data acceptably was formed by two factors with 25 items consistent with the model that underlies it.

Keywords: Construct validity, Evaluation, Competency, Teacher evaluation, Student surveys.

*Contacto: erikareyespinuelas@gmail.com

ISSN: 1989-0397

www.rinace.net/riee/

Recibido: 19 de enero de 2013

1ª Evaluación: 12 de febrero de 2013

Aceptado: 20 de enero de 2014

1. Introducción

El presente estudio reporta los resultados obtenidos del análisis de confiabilidad y validez de un Cuestionario de Evaluación de la Competencia Docente (CECD) con base en la opinión de los alumnos. La argumentación que orienta el estudio, es la necesidad de contar con instrumentos de evaluación docente en educación superior congruentes con el modelo de enseñanza por competencias confiables y con evidencias de validez pertinentes para evaluar la docencia.

Los cuestionarios con base en la opinión de los alumnos son la técnica de recopilación de información del quehacer docente de mayor presencia tanto en el ámbito internacional como nacional (García-Cabrero y Loredó, 2010; Luna, 2000; Paulsen y Feldman, 1995; Rueda, 2006). El creciente interés ha permitido por un lado, dar cuenta de las fortalezas del instrumento y por otro, plantear una serie de recomendaciones sobre su construcción y usos. Con más de ocho décadas de ser utilizados se han ganado un lugar dentro de la evaluación docente en educación superior. Su uso masivo se facilita por el corto tiempo requerido para su aplicación y bajo costo.

Se reconoce que los cuestionarios de apreciación estudiantil tienden a ser estadísticamente confiables y válidos más que cualquier otra fuente de información (Marsh, 2007). Al mismo tiempo, algunas de las críticas persistentes al uso de los cuestionarios señalan que promueven una pedagogía conservadora (Centra, 1993, 1973), al no ser acordes a los modelos educativos actuales, asimismo, que se apegan de manera limitada a la teoría de enseñanza que les subyace (Cisneros-Cohernour y Stake, 2010; Luna, 2000). Además, se cuestiona su uso indiscriminado, sin atender los aspectos básicos de confiabilidad y validez; y la sobrevaloración de su utilidad (Berk, 2006; Luna y Torquemada, 2008).

Un requisito en el desarrollo de instrumentos de medición como los cuestionarios de apreciación estudiantil es que se fundamenten en una teoría de la enseñanza (Gravestock y Gregor-Greenleaf, 2008; Scriven, 1995), congruente con el modelo pedagógico que promueve la institución (Luna, 2000; Red Iberoamericana de Investigadores de la Evaluación de la Docencia [RIIED], 2008). Ory y Ryan (como se citó en Gravestock y Gregor-Greenleaf, 2008) señalaron que este tipo de cuestionarios por lo general se construyen a partir de otros ya elaborados sin reflexionar sobre la teoría que los subyace y en la construcción de dominios. Con frecuencia estos son desarrollados a partir de consenso entre expertos, maestros, miembros de instituciones y/o estudiantes (Marsh, 2007), sobre características específicas relevantes dentro del proceso enseñanza-aprendizaje y culminan en reactivos o factores. Con base en ello, los reactivos comúnmente giran alrededor de factores como la organización del curso, interacción alumno maestro y comunicación verbal sin reflejar una teoría particular (Centra, 1973). Con lo que se promueve la evaluación de un modelo de enseñanza tradicional y en consecuencia de prácticas pedagógicas tradicionales en la que el docente posee el conocimiento y el estudiante desempeña un rol pasivo.

Las teorías en las que se sustentan los cuestionarios se relacionan con el papel asignado a los docentes y alumnos las cuales se reflejan en los reactivos de los cuestionarios, en otras palabras, la concepción de enseñanza que se aborda a través del cuestionario de evaluación docente, en los cuales pueden ser identificadas perspectivas teóricas tales como la cognoscitiva, conductual, humanista y una perspectiva de enseñanza tradicional (Luna, 2000). El hecho de no contar con un cuestionario de evaluación que refleje en sus reactivos las prioridades institucionales y que se encuentre elaborado en congruencia

con la pedagogía que promueve la institución, las evaluaciones pueden tener un impacto negativo en la enseñanza, respecto a su validez y utilidad (Gravestock y Gregor-Greenleaf, 2008).

La base teórica que sustenta al CECD es el Modelo de Evaluación de Competencias Docentes para la Educación Media y Superior (García-Cabrero, Loredó, Luna y Rueda, 2008), los principios de este modelo son: (a) una orientación formativa, privilegia la reflexión y retroalimentación de la práctica docente además del planteamiento de acciones para su mejora; (b) orientación participativa, donde el docente y autoridades se involucran en el diseño e instrumentación; (c) orientación humanista, el docente es visto como un individuo con intereses, motivaciones, emociones, anteponiendo la preservación de su autoestima y dignidad; y (d) enfoque multirreferencial, favorece la diversidad de opiniones y además parte de la complejidad de la evaluación de la docencia con interés en factores que la determinan.

En este modelo se plantean tres momentos considerados secuenciales y paralelos para el desarrollo de la función docente. El primer momento, *Previsión del proceso enseñanza-aprendizaje*, son las actividades previas incluye el pensamiento del docente, el cual hace referencia al conocimiento pedagógico en general y del contenido pedagógico de asignatura, las expectativas del grupo y de sí mismo, y la planeación de la asignatura. Kane, Sandretto y Heath (2004) denominan teorías asumidas a expresiones o justificaciones de los docentes en torno a sus acciones en el proceso de enseñanza, y podrían ser exploradas en este primer momento del modelo. En el segundo momento, *Conducción del proceso enseñanza-aprendizaje*, relacionado con la interacción del docente y alumno en el aula, se observan las teorías en uso, que desde las aportaciones de Kane *et al.* (2004) podrían observarse en esta dimensión ya que se les relaciona con las que sustentan las acciones realizadas en el ejercicio de la docencia dentro del aula. Finalmente, el tercer momento, *Valoración de proceso enseñanza-aprendizaje*, es la reflexión en torno a los resultados alcanzados por el ejercicio de la práctica docente y evaluadas por diferentes actores como alumnos, docentes, los pares y directivos.

El abordaje de la evaluación de la práctica docente a partir de la enseñanza por competencias se persigue desde diferentes acercamientos, por un lado se encuentra la aproximación de competencia global, integradora, que sólo puede evaluarse en contextos auténticos y globalizados; y por otro lado un acercamiento propuesto por Denyer, Furnemont, Poulain y Vanloubbeeck (2007), consideraron que desglosar las competencias permite evaluar aspectos específicos y observables de la actividad docente. Plantearon evaluar las competencias por medio de tareas específicas determinadas a través de indicadores (observables y cuantificables) que miden el nivel de dominio alcanzado de determinada competencia. Estos autores afirmaron que el uso de pruebas resultará objetiva si el desarrollo de las pruebas descansa sobre criterios explicitados por indicadores observables y cuantificables.

En este sentido, el recurso de los indicadores en el uso de las pruebas presenta ventajas alternas, como la facilidad para toma de decisiones sobre el dominio de la competencia docente, favorece el entendimiento por parte del alumno de los elementos invariables de la familia de las tareas, condición importante para la generalización y favorece la interiorización por el alumno sobre cuáles son los criterios que se asocian al éxito y, por tanto, una apropiación del proceso (Denyer *et al.*, 2007). Este planteamiento propuesto por Denyer *et al.* fue retomado por el Modelo de Evaluación de Competencias Docentes para la Educación Media y Superior, modelo que subyace en el CECD.

2. Método

2.1. Participantes

La información objeto de estudio corresponde a los puntajes generados por los estudiantes sobre la evaluación docente y pertenecen al primer periodo semestral de 2011 de la Universidad Autónoma de Baja California (UABC) en México. Los datos analizados involucraron a 37,094 estudiantes adscritos a tres campus. Estos estudiantes representan un 77.3% de participación estudiantil dado que la población de licenciatura en ese periodo fue de 47,943. Como se muestra en la tabla 1, el 53.7% de los participantes lo constituyen mujeres y el 46.3% hombres, la media de la edad fue de 21.35 (DE= 2.318).

Tabla 1. Características de los participantes

VARIABLES	N	PORCENTAJE	
Sexo	Femenino	19,924	53.7
	Masculino	17,170	46.3
Edad	17	16	0.0
	18	2,450	6.6
	19	6,594	17.8
	20	6,882	18.6
	21	6,416	17.3
	22	4,774	12.9
	23	3,090	8.3
	24	1,866	5.0
	25	1,210	3.3
26 en adelante	3,796	10.2	

Fuente: Elaboración propia.

El total de cuestionarios analizados fue de 199,967 número muy por encima de los 37,094 estudiantes, ya que estos contestan un cuestionario por cada asignatura cursada en el semestre. Para el análisis de unidimensionalidad la muestra quedó constituida por 128,791 cuestionarios ya que se prescindió de los cuestionarios con escasa variabilidad es decir, con puntajes totales alrededor de 35 y 175 puntos.

Se utilizó el criterio de replicabilidad por medio del método de validación cruzada, para obtener mayor confianza en los resultados al confirmarlos con diferentes muestras de la misma población (Rodríguez, 1999), por consiguiente se obtuvieron dos muestras aleatorias conformadas una por 66,148 (49.6%) cuestionarios denominada submuestra A y utilizada para el análisis factorial exploratorio (AFE) y 67,111 (50.4%) la submuestra B para el análisis factorial confirmatorio (AFC), de esta última se extrajo una muestra de 462 cuestionarios.

2.2. Instrumento

El CECD con base en la opinión de los alumnos de Luna, Calderón, Caso y Cordero (2012) fue desarrollado a partir de los componentes relevantes del Modelo de Evaluación de Competencias Docentes para la Educación Media y Superior (García, Loredó, Luna y Rueda, 2008). El CECD consta de 35 reactivos tipo Likert con cinco niveles de respuesta de uno (*nunca*) a cinco (*siempre*).

Luna *et al.* (2012) reportan las fases y resultados en el desarrollo y validación del cuestionario considerando la prueba de confiabilidad, análisis validación por un grupo de expertos, análisis de unidimensionalidad con metodología Rasch y AFE. Los resultados relevantes de estos análisis fueron: la confiabilidad con la prueba alfa de Cronbach

resultó de .967 en el total de los reactivos y por factor .753, .876, .920 y .939; en el análisis de unidimensionalidad se concluyó que los reactivos se comportaron en el análisis según las expectativas del modelo Rasch Masters, a excepción de los reactivos 34 y 35 que presentaron valores fuera de los umbrales críticos; los resultados del AFE revelaron que la estructura del CECD se compone de tres dimensiones y un indicador: *Planeación y gestión del proceso enseñanza-aprendizaje*, *Interacción didáctica en el aula*, *Evaluación y comunicación del proceso de enseñanza-aprendizaje*, y como indicador, *Tecnologías de la información y la comunicación* (ver tabla 2). Esta estructura factorial explicaba el 59.3% de varianza. Se concluyó que el cuestionario es válido y confiable para evaluar el constructo de competencia docente.

Tabla 2. Estructura factorial del CECD

REACTIVOS	FACTORES			
	1	2	3	4
5. Enseña procedimientos para usar el conocimiento en la solución de problemas reales.	.724			
1. Menciona la relación de algunos contenidos del curso con temas de otras materias del plan de estudios.	.692			
10. Modifica las actividades o fuentes de información para adecuarlas a las necesidades del grupo.	.682			
2. Explica la utilidad de los conocimientos en situaciones de la vida diaria.	.675			
12. Realiza actividades en clase que me motivan a aplicar lo aprendido fuera del aula.		.468		
15. Pone en operación actividades en el aula que facilitan el aprendizaje de los contenidos.	.641			
11. Plantea situaciones con problemas reales que favorecen mi aprendizaje de los temas.	.626			
14. Me motiva a continuar aprendiendo sobre los temas vistos en el curso.	.618			
4. Puntualiza los conceptos importantes de los temas vistos en el curso.	.613			
9. En las primeras clases identifica los conocimientos de los estudiantes sobre la materia.	.606			
3. Ubica los contenidos del curso en el contexto profesional.	.598			
8. Propone fuentes de información que facilitan el aprendizaje de los temas.	.593			
16. Aborda los temas con una secuencia razonada.		.401		
18. Acepta sugerencias de los estudiantes sobre la manera de organizar las actividades.		.440		
17. El tiempo dedicado para aprender cada tema es suficiente.		.479		
23. Realiza preguntas a los estudiantes cuando expone, para asegurar la comprensión del tema.		.730		
22. Se asegura de que en las actividades en equipo, cada integrante contribuya a su realización.		.717		
21. Estructura tareas de aprendizaje que propician la colaboración entre los estudiantes.		.591		
28. Fomenta la participación del grupo para realizar conclusiones en sus exposiciones.		.560		
19. Da oportunidades iguales a todos los estudiantes para participar en clase.		.558		
25. Cuando habla, expresa claramente sus ideas.		.555		
31. Promueve la elaboración de acuerdos de convivencia en el aula con base en las opiniones de todo el grupo.			.370	
20. Para facilitar la comprensión de conceptos, los define mediante ejemplos		.426		

Tabla 2. Estructura factorial del CECD (continuación)

REACTIVOS	FACTORES			
	1	2	3	4
7. Explica los criterios de evaluación de la materia.			.726	
6. Presenta la lista de temas y su secuencia en el programa de la materia.			.681	
34. Respeta los criterios de evaluación planteados al inicio del curso.			.665	
32. Es respetuoso con todos los estudiantes del grupo.			.620	
35. La evaluación final fue congruente con los contenidos y actividades del curso.			.598	
30. Demuestra ser responsable en su trabajo.			.509	
27. Sus exposiciones se caracterizan por tener un inicio, desarrollo y conclusión.	.434			
26. Manifiesta claramente sus ideas cuando escribe en el pintarrón.	.452			
33. Realiza evaluaciones para retroalimentar sobre las dificultades de aprendizaje.			.458	
29. Incorpora en sus clases el uso de la tecnología de la información y la comunicación (por ej. internet, uso de software, etc.).				.805
24. Utiliza diversos recursos didácticos en clase (por ejemplo presentaciones en computadora, videos, fotos, diagramas, etc.).				.802

Nota: Factor 1: Planeación y gestión del proceso enseñanza-aprendizaje; Factor 2: Interacción didáctica en el aula; Factor 3: Evaluación y comunicación del proceso de enseñanza-aprendizaje; Indicador 4: Tecnologías de la información y la comunicación. Método de extracción de componentes principales, rotación factorial varimax con normalización Keiser. La estructura factorial explica el 59.3% del total de la varianza.

Fuente: Luna *et al.* (2012).

2.3. Procedimiento

Los puntajes de los cuestionarios se obtuvieron de la base de datos institucional a través del Sistema de Evaluación Docente de la UABC y se generaron a partir de respuestas de los estudiantes al completar el CECD por medio de un portal en internet.

La secuencia en cada uno de los análisis realizados se detalla en la figura 1, primero se realizó el análisis de unidimensionalidad con la metodología Rasch Masters en el programa de análisis de datos Winsteps, se requirió procesar la información en dos ocasiones con el objetivo de obtener evidencias que la totalidad de los reactivos del CECD midieran un único constructo.

Posteriormente se realizó el AFE, como se reportó en Luna *et al.* (2012) la estructura factorial propuesta fue a partir de los resultados del AFE, sin embargo aquel estudio presentó condiciones diferentes (tamaño de muestra y aplicación en papel y lápiz) por lo tanto se consideró pertinente en esta investigación identificar la estructura factorial con propósitos exploratorios y obtener suficientes conocimientos para formular hipótesis concretas sobre la relación entre indicadores y dimensiones latentes con una muestra que mantuviera las mismas condiciones de aplicación en los análisis de datos.

El AFE se llevó a cabo con el método de extracción de componentes principales, y se reportaron dos tipos de rotaciones, ortogonal (con varimax) y oblicua (con promax), es recomendable hacer ambas rotaciones, dado que si la agrupación factorial es idéntica en ambas rotaciones entonces la estructura factorial del instrumento es clara y estable (Morales, 2011). Entre las bondades de estos tipos de rotaciones se tiene que los factores ortogonales son más generalizables y es fácil que se repliquen en otras muestras, y en el caso de la rotación oblicua permite una apreciación más real de la estructura subyacente del cuestionario, además, suponen una correlación entre los factores, por lo tanto otorga

congruencia al presente estudio por el análisis de unidimensionalidad donde la finalidad es que los reactivos representen un único constructo.

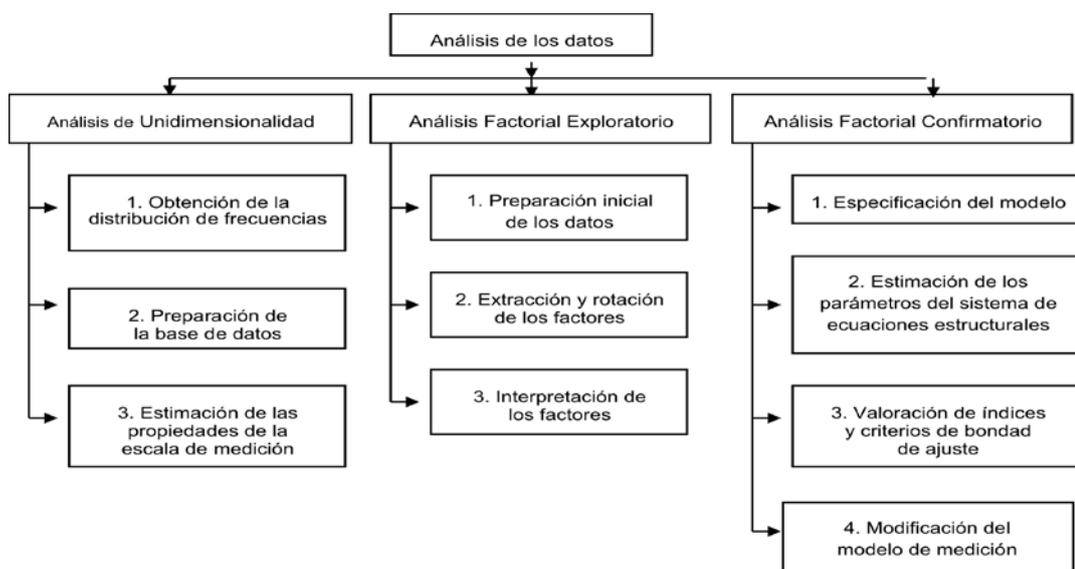


Figura 1. Análisis de los datos
Fuente: Elaboración propia.

Se valoraron los factores susceptibles de ser interpretados a la luz de diferentes criterios de decisión como el criterio de raíz latente y de contraste de caída, la prueba de comunales y la varianza explicada. La identificación de cada reactivo en determinado factor se efectuó con base en la mayor saturación obtenida en las estructuras factoriales rotadas, lo anterior conjuntamente con la revisión y contraste de la teoría que subyace al cuestionario, con el apoyo de la *Reticula de planeación de la segunda versión del CECD* (Luna *et al.*, 2012) para valorar la congruencia teórica.

Finalmente se confirmó la estructura factorial del CECD con el AFC mediante la metodología de modelamiento de ecuaciones estructurales utilizando el programa de cómputo para el análisis de datos Stata.

3. Resultados

3.1. Análisis de unidimensionalidad

El análisis inicial de unidimensionalidad Rasch se realizó a los 35 reactivos con los 199,967 cuestionarios de opinión. Con la finalidad de conocer la variabilidad de los datos se obtuvieron las medias de cada reactivo, estas se ubicaron entre 4.28 la más alta y 4.02 la más baja. La consistencia interna con alfa de Cronbach fue de .97. Como parte de los criterios de unidimensionalidad, las correlaciones punto biserial superaron la expectativa (.20 según Linacre y Rasch, 2008), la menor fue de .75 y la mayor de .90. En cuanto a la discriminación del reactivo, el número 34 se aleja de la expectativa (valores cercanos o superiores a 1.0) con un valor de 0.53.

Otros criterios para la determinación de la unidimensionalidad son: la dificultad del reactivo, se obtuvieron valores entre 0.52 y -0.55 lógitos (es aceptable obtener valores en el rango de +2 a -2 lógitos) esto indica que en términos generales la dificultad de los reactivos es de moderada a baja. En el caso de los criterios de bondad de ajuste próximo

(*infit* en inglés) y ajuste lejano (*outfit* en inglés) los reactivos 34 y 35 presentaron en ambos criterios problemas de ajuste, mientras que el reactivo 1 solo en criterio de ajuste lejano. Los primeros dos reactivos pertenecen al indicador *Tecnologías de la información y la comunicación* y el reactivo 1 al factor *Planeación y gestión del proceso enseñanza-aprendizaje* (según la estructura factorial del estudio que le precede al cuestionario). Dado que los reactivos 1, 34 y 35 muestran valores fuera de la expectativa en uno o varios criterios se decidió omitirlos de subsecuentes análisis. En el mapa de Wright se detectó un número amplio de estudiantes que calificaron con puntuaciones totales máximas que representan una alta competencia docente, así como una proporción menor con puntuaciones que reflejan una baja competencia docente, con el propósito de que los resultados no se encontraran influidos por valores atípicos se decidió excluir los cuestionarios contestados con escasa variabilidad, es decir con puntuaciones totales cercanas a 35 y 175 puntos.

Con el objetivo de obtener un cuestionario donde todos los reactivos midieran un mismo constructo se procedió a un segundo análisis de unidimensionalidad con una muestra de 128,791 cuestionarios. Se retomaron los hallazgos derivados del análisis anterior de unidimensionalidad y se concluyó que los 32 reactivos evalúan el mismo rasgo latente, los valores de los diferentes índices de bondad de ajuste tales como: índice de discriminación, dificultad del reactivo, índices de ajuste lejano y ajuste próximo fueron aceptables y se conservó un alfa de Cronbach de .97.

3.2. Análisis Factorial Exploratorio

La submuestra A constó de 66,148 cuestionarios, con la cual se realizaron las pruebas de KMO y Bartlett, en la primera resulta con un .98, al ser superior a .60 es apropiado factorizar (Zamora, Monroy y Chávez, 2009), y la segunda prueba de esfericidad de Bartlett se admite la existencia de relación entre las variables.

La solución factorial agrupó a los 32 reactivos en dos factores interpretables. Los criterios que permitieron justificar el número de factores a considerarse en la interpretación fueron el criterio de raíz latente y criterio de contraste de caída, en el primero de estos criterios resultó con un valor de 22.85 en el Factor 1 y 1.15 en el Factor 2 (la expectativa es obtener raíces latentes mayores a uno, según Hair, Anderson, Tatham y Black, 1999) y el criterio de contraste de caída indicó dos factores a considerar ya que el punto de quiebre del gráfico de sedimentación dejó de ser pronunciado en el segundo factor. Para la determinación de los factores en la prueba de comunalidades se determinó un mínimo de .50 para ser considerado dentro de un factor (valor deseable según Hair *et al.*, 1999). La mayoría de las cargas factoriales muestran saturaciones altas en uno de los factores y cargas bajas o insignificantes en el otro (comportamiento esperado según Byrne, 2001) y se cumplió la expectativa de al menos tres reactivos por factor.

En la tabla 3 se muestran los pesos y correlaciones obtenidos con la rotación varimax y la promax, ambas rotaciones con un comportamiento idéntico al arrojar dos factores, obtener una varianza explicada de 71.43% en el Factor 1 y de 3.59% en el Factor 2, con un total de 75.02%, la varianza total obtenida superó lo propuesto por Hair *et al.* (1999) para las ciencias sociales al obtener un valor por arriba de 60%.

Posteriormente, se asignó un nombre a la agrupación de las variables observadas que reflejara con mayor precisión el aspecto que representan. En la tabla 1 se muestra la estructura factorial propuesta, el conjunto de reactivos del Factor 1 coincide con las dimensiones de *Conducción y valoración del proceso enseñanza-aprendizaje* por tanto así fue

denominado y los reactivos del Factor 2, en referencia a las dimensiones que coincide se nombró *Planeación y gestión de proceso enseñanza-aprendizaje*.

Tabla 3. Análisis Factorial Exploratorio del CECD

REACTIVOS	VARIMAX				PROMAX	
	Matriz de estructura		Matriz de configuración			
	1	2	1	2	1	2
32. Respeta los criterios de evaluación planteados al inicio del curso.	.803	.894	.720	.925	-.038	
29. Manifiesta claramente sus ideas cuando escribe en el pintarrón.	.794	.884	.712	.915	-.038	
33. La evaluación final fue congruente con los contenidos y actividades del curso.	.781	.886	.731	.873	.016	
6. Propone fuentes de información que facilitan el aprendizaje de los temas.	.769	.860	.697	.879	-.023	
27. Presenta la lista de temas y su secuencia en el programa de la materia.	.758	.842	.676	.877	-.043	
31. Realiza evaluaciones para retroalimentar sobre las dificultades de aprendizaje.	.736	.874	.764	.755	.145	
25. Promueve la elaboración de acuerdos de convivencia en el aula con base en las opiniones de todo el grupo.	.736	.871	.760	.758	.139	
30. Sus exposiciones se caracterizan por tener un inicio, desarrollo y conclusión en sus exposiciones.	.731	.869	.761	.747	.149	
28. Explica los criterios de evaluación de la materia.	.726	.855	.741	.755	.122	
18. Para facilitar la comprensión de conceptos, los define mediante ejemplos.	.713	.850	.748	.722	.157	
19. Estructura tareas de aprendizaje que propician la colaboración entre los estudiantes.	.681	.859	.805	.609	.305	
23. Fomenta la participación del grupo para realizar conclusiones en sus exposiciones.	.681	.854	.794	.618	.287	
26. Es respetuoso con todos los estudiantes del grupo.	.670	.849	.798	.593	.312	
15. El tiempo dedicado para aprender cada tema es suficiente.	.661	.840	.793	.579	.318	
24. Demuestra ser responsable en su trabajo.	.660	.846	.805	.568	.340	
22. Cuando habla, expresa claramente sus ideas.	.654	.829	.780	.577	.307	
20. Se asegura de que en las actividades en equipo, cada integrante contribuya a su realización.	.634	.830	.806	.516	.383	
17. Da oportunidades iguales a todos los estudiantes para participar en clase.	.621	.814	.792	.501	.381	
21. Realiza preguntas a los estudiantes cuando expone, para asegurar la comprensión del tema.	.599	.797	.786	.465	.405	
16. Acepta sugerencias de los estudiantes sobre la manera de organizar las actividades.	.595	.787	.773	.468	.390	
11. Estuvo disponible para dar asesorías individuales cuando lo necesité.	.788	.788	.711	.880	-.030	.904
9. Plantea situaciones con problemas reales que favorecen mi aprendizaje de los temas.	.783	.783	.724	.883	.002	.882
7. En las primeras clases identifica los conocimientos de los estudiantes sobre la materia.	.766	.766	.723	.871	.026	.850
5. Enseña procedimientos para usar el conocimiento en la solución de problemas reales.	.764	.764	.751	.885	.079	.820

Tabla 3. Análisis Factorial Exploratorio del CECD (continuación)

REACTIVOS	VARIMAX		PROMAX		
	Matriz de estructura		Matriz de configuración		
	1	2	1	2	
10. Realiza actividades en clase que me motivan a aplicar lo aprendido fuera del aula.	.762	.755	.886	.089	.813
8. Modifica las actividades o fuentes de información para adecuarlas a las necesidades del grupo.	.758	.728	.869	.049	.829
13. Pone en operación actividades en el aula que facilitan el aprendizaje de los contenidos.	.757	.754	.882	.096	.803
14. Aborda los temas con una secuencia razonada.	.741	.765	.877	.140	.762
2. Explica la utilidad de los conocimientos en situaciones de la vida diaria.	.720	.733	.848	.118	.751
4. Puntualiza los conceptos importantes de los temas vistos en el curso.	.720	.777	.870	.196	.709
3. Ubica los contenidos del curso en el contexto profesional.	.717	.758	.858	.166	.722
12. Me motiva a continuar aprendiendo sobre los temas vistos en el curso.	.703	.728	.833	.136	.722

Nota: Método de extracción: componentes principales. Método de rotación: Varimax con normalización de Kaiser. Método de rotación: Promax con normalización de Kaiser.

Fuente: Luna *et al.* (2012).

La prueba estadística alfa de Cronbach resultó con .972 en el Factor 2, un .981 en el Factor 1 y .987 el total de los reactivos valor alto según diversos sistemas de referencia, además esta solución bifactorial coincidió con la estructura teórica que subyace al cuestionario y concordó de manera general con lo encontrado por Luna *et al.* (2012) por lo tanto, es un instrumento confiable y válido para evaluar la competencia docente producto de su aplicación.

3.3. Análisis Factorial Confirmatorio

Con el propósito de confirmar la estructura factorial del CECD producto del AFE se realizó el AFC con la submuestra B de 67,111 cuestionarios. El análisis se estimó con el método de máxima verosimilitud.

Los índices de bondad de ajuste absolutos fueron: chi cuadrada=203,830.17, con 463 grados de libertad y un valor de $p=.00$, valores fuera de los límites críticos (según Hooper Coughlan y Mullen, 2008; Stapleton, 1997). El valor de la Raíz del residuo cuadrático promedio de aproximación (RMSEA) alcanzó .081, en el límite sugerido (por debajo de .08) según Hair *et al.* (1999), Hooper *et al.* (2008), Ruíz, Pardo y San Martín (2010), y Raíz del residuo cuadrático promedio estandarizado (SRMR), resultó por debajo del .05 recomendado por Hooper *et al.* (2008), con un .026.

En los índice de ajuste incremental, el Índice de Tucker-Lewis (TLI)=.917, se ubicó en el rango deseable (alrededor de .90, según Hair *et al.*, 1999), y en el caso del Índice de bondad de ajuste comparativo (CFI)=.923 (cerca de .95 según Hu y Bentler, 1999). Los valores de índices de bondad de ajuste descritos anteriormente brindaron elementos para la modificación del modelo de medición. La primera estrategia utilizada en esta fase fue disminuir el tamaño de la muestra, dado que algunos índices de ajuste son sensibles al tamaño de muestra, como la chi cuadrada (Hair *et al.*, 1999; Hooper *et al.*, 2008;

Joreskog y Sorbom, 1982; Stapleton, 1997) y CFI (Hooper *et al.*, 2008). Por lo tanto se optó por una muestra de 462 cuestionarios.

El modelo alternativo propuesto del CECD quedó conformado por dos factores con 25 reactivos congruentes con el modelo que le subyace producto del recurso de modificación (ver tabla 4). El procedimiento de modificación del modelo de medición sugirió hacer los siguientes cambios: prescindir de los reactivos 3, 11 y 13 del factor *Planeación y Gestión del proceso enseñanza-aprendizaje*, y los reactivos 6, 18, 21 y 29 del factor *Conducción y Valoración del proceso enseñanza-aprendizaje*. Asimismo, una forma adicional para mejorar el ajuste de los datos fue a través de correlacionar errores de reactivos, no obstante, se recomienda ser conservador en su utilización (Hooper *et al.*, 2008), en este caso se asociaron los errores entre los reactivos 2-4, 30-31 y 32-33.

Tabla 4. Modelo alternativo del CECD

DIMENSIONES	REACTIVOS
Planeación y gestión del proceso de enseñanza-aprendizaje	2. Explica la utilidad de los conocimientos en situaciones de la vida diaria.
	4. Puntualiza los conceptos importantes de los temas vistos en el curso.
	5. Enseña procedimientos para usar el conocimiento en la solución de problemas reales.
	7. En las primeras clases identifica los conocimientos de los estudiantes sobre la materia.
	8. Modifica las actividades o fuentes de información para adecuarlas a las necesidades del grupo.
	9. Plantea situaciones con problemas reales que favorecen mi aprendizaje de los temas.
	10. Realiza actividades en clase que me motivan a aplicar lo aprendido fuera del aula.
	12. Me motiva a continuar aprendiendo sobre los temas vistos en el curso.
	14. Aborda los temas con una secuencia razonada.
	27. Presenta la lista de temas y su secuencia en el programa de la materia.
	28. Explica los criterios de evaluación de la materia.
	15. El tiempo dedicado para aprender cada tema es suficiente.
	16. Acepta sugerencias de los estudiantes sobre la manera de organizar las actividades.
	17. Da oportunidades iguales a todos los estudiantes para participar en clase.
19. Estructura tareas de aprendizaje que propician la colaboración entre los estudiantes.	
Conducción y valoración del proceso de enseñanza-aprendizaje	20. Se asegura de que en las actividades en equipo, cada integrante contribuya a su realización.
	22. Cuando habla, expresa claramente sus ideas.
	30. Sus exposiciones se caracterizan por tener un inicio, desarrollo y conclusión en sus exposiciones.
	23. Fomenta la participación del grupo para realizar conclusiones en sus exposiciones.
	24. Demuestra ser responsable en su trabajo.
	25. Promueve la elaboración de acuerdos de convivencia en el aula con base en las opiniones de todo el grupo.
	26. Es respetuoso con todos los estudiantes del grupo.
	31. Realiza evaluaciones para retroalimentar sobre las dificultades de aprendizaje.
	32. Respeta los criterios de evaluación planteados al inicio del curso.
	33. La evaluación final fue congruente con los contenidos y actividades del curso.

Fuente: Elaboración propia.

Finalmente, el modelo alternativo quedó conformado por dos factores –con lo cual la estructura producto del AFE no fue alterada– congruentes con el modelo teórico que le subyace, Consta de 25 reactivos en el que los índices de bondad de ajuste SRMR=.029, RMSEA=.079, TLI=.931 y CFI=.937, alcanzaron valores que indican buen ajuste del modelo.

4. Conclusiones

Las fortalezas metodológicas del CECD residen en las evidencias de validez y confiabilidad reportadas en el presente artículo que se suman a las obtenidas en el estudio de Luna *et al.* (2012). Los resultados indican valores adecuados en los diversos criterios tanto para la determinación de la unidimensionalidad del test como en el AFE y AFC.

Uno de los propósitos del presente estudio fue obtener a partir de los puntajes del CECD una estructura factorial que revelara las dimensiones del Modelo de Evaluación de Competencias Docentes para la Educación Media y Superior (García *et al.*, 2008) con una muestra que mantuviera las mismas condiciones de aplicación. El proceso de validación del cuestionario ha permitido constatar que dicho modelo que subyace al instrumento se ajusta a los datos.

El CECD quedó conformado por 25 reactivos agrupados en dos factores, Factor 1, *Planeación y gestión del proceso de enseñanza-aprendizaje* y Factor 2, *Conducción y valoración del proceso de enseñanza-aprendizaje*, los cuales se encuentran acordes con el modelo teórico y responden a los tres momentos planteados por este. Dado que el Factor 1 representa en el modelo teórico la dimensión de *Planeación del proceso enseñanza-aprendizaje* y el Factor 2 representa las dos dimensiones restantes (*Conducción del proceso enseñanza-aprendizaje* y la de *Valoración del proceso enseñanza-aprendizaje*).

La estabilidad en la estructura factorial, el porcentaje de varianza explicada, los índices de confiabilidad con alfa de Cronbach y los índices bondad de ajuste práctico resultado del modelamiento de ecuaciones estructurales permiten hacer inferencias válidas y confiables a partir de su aplicación. Así, los resultados proporcionan elementos necesarios para fundamentar su utilización en otras universidades.

A partir de los resultados derivados de procedimientos rigurosos de análisis de validez el CECD se constituye como una herramienta de evaluación de la docencia universitaria con base en la opinión de los alumnos que supera las críticas recurrentes en relación a la validez y confiabilidad, y al desarrollo de instrumentos bajo una pedagogía conservadora.

Finalmente, los cuestionarios de opinión por estudiantes son un instrumento útil y valorado en la evaluación docente, sin embargo no están exentos de sesgo. Por tanto, es recomendable que la información generada a partir de los puntajes del CECD forme parte de una estrategia integral que incorpore otras fuentes de información como autoevaluación, evaluación por pares y directivos.

Referencias

- Berk, R. (2006). *Thirteen Strategies to Measure College Teaching: a consumer's guide to rating scale construction, assessment, and decision making for faculty, administrators, and clinicians*. Sterling, VA: Stylus.

- Byrne, B. (2001). *Structural Equation Modeling With AMOS: Basic Concepts, Applications, and Programming*. Mahwah, NJ: Lawrence Erlbaum Associates, Inc.
- Centra, J. (1973). The Student as Godfather? The Impact of Student Ratings on Academia. En A.L. Sockolff (Ed.), *Proceedings of the First Invitational Conference on Faculty Effectiveness as Measured by Students*. Philadelphia, NJ: Temple University.
- Centra, J. (1993). *Reflective faculty evaluation: enhancing teaching and determining faculty effectiveness*. Nueva York: Jossey-Bass.
- Cisneros-Cohernour, E. y Stake, R. (2010). La evaluación de la docencia en educación superior: de evaluaciones basadas en opiniones de estudiantes a modelos por competencias. *Revista Iberoamericana de Evaluación Educativa*, 3(1), 218-231.
- Denyer, M., Furnemont, J., Poulain, R. y Vanloubbeeck, G. (2007). *Las competencias en la educación, un balance*. México: Fondo de Cultura Económica.
- Ferrando, P. y Anguiano-Carrasco, C. (2010). El análisis factorial como técnica de investigación en psicología. *Papeles del Psicólogo*, 31(1), 18-33.
- García, B., Loredó, J., Luna, E., y Rueda, M. (2008). Modelo de Evaluación de Competencias Docentes para la Educación Media y Superior. *Revista Iberoamericana de Evaluación Educativa*, 1(3e), 96-108.
- García-Cabrero, B., Loredó, J., y Carranza, G. (2008). Análisis de la práctica educativa de los docentes: pensamiento, interacción y reflexión. *Revista Electrónica de Investigación Educativa, Especial*. Recuperado de <http://redie.uabc.mx/NumEsp1/contenido-garcialoredocarranza.html>
- Gravestock, P. y Gregor-Greenleaf, E. (2008). *Student Course Evaluations: Research, Models and Trends*. Toronto: Higher Education Quality Council of Ontario. Recuperado de <http://www.mcgill.ca/files/tls/StudentCourseEvaluations.pdf>
- Hair, J., Anderson, R., Tatham, R. y Black, W. (1999). *Análisis Multivariante*. Madrid: Pearson/Prentice Hall.
- Hooper, D.; Coughlan, J. y Mullen, M. (2008). Structural Equation Modelling: Guidelines for Determining Model Fit. *Electronic Journal of Business Research Methods*, 6(1), 53-60.
- Hu, L. y Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, 6(1), 1-55.
- Joreskog, K. y Sorbom, D. (1982). Recent Developments in Structural Equation Modeling. *Journal of Marketing Research*, 19, 404-16.
- Kane, R., Sandretto, S. y Heath, C. (2004). An investigation into excellent tertiary teaching: Emphasising reflective practice. *Higher Education*, 47(3), 283-310.
- Kieffer, K. (noviembre, 1998). Orthogonal versus Oblique Factor Rotation: A Review of the Literature regarding the Pros and Cons. Trabajo presentado en la *Annual Meeting of the Mid-South Educational Research Association*, New Orleans, EE.UU.
- Linacre, J. y Rasch, G. (2008). *The Expected Value of a Point-Biserial (or Similar) Correlation*. Recuperado de <http://www.rasch.org/rmt/rmt221e.htm>
- Luna, E. (2000). Aspectos implícitos sobre la enseñanza reflejados en los cuestionarios de evaluación de la docencia. En M. Rueda y F. Díaz-Barriga (eds.), *Evaluación de la docencia. Perspectivas actuales* (pp. 63-83). México: Paidós
- Luna, E. y Torquemada, A. (2008). Los cuestionarios de evaluación de la docencia por los alumnos: balance y perspectivas de su agenda. *Revista Electrónica de Investigación Educativa, Especial*. Recuperado de <http://redie.uabc.mx/NumEsp1/contenido-lunatorquemada.html>

- Luna, E., Calderón, N., Caso, J. y Cordero, G. (2012). Desarrollo y validación de un cuestionario de evaluación de la competencia docente con base en la opinión de los estudiantes. En E. J. Cisneros-Cohernour, B. García-Cabrero, E. Luna y R. Marín (coords.), *Evaluación de Competencias Docentes en la Educación Superior* (pp.119-158). México: Juan Pablos Editor.
- Marsh, H. (2007). Students' evaluations of university teaching: Dimensionality, reliability, validity, potential Biases and usefulness. En R.P. Perry y J.C. Smart (eds.), *The Scholarship of Teaching and Learning in Higher Education: An Evidence-Based Perspective* (pp. 319-383). Dordrecht: Springer.
- Morales, P. (2011). *El Análisis Factorial en la construcción e interpretación de tests, escalas y cuestionarios*. Recuperado de <http://www.upcomillas.es/personal/peter/investigacion/AnalisisFactorial.pdf>
- Paulsen, M. y Feldman, K. (1995). *Taking teaching seriously: Meeting the challenge of instructional improvement (ASHE-ERIC Higher Education Informe No. 2)*, Washington, DC. Recuperado de <http://www.eric.ed.gov/PDFS/ED396616.pdf>
- Red Iberoamericana de Investigadores de la Evaluación de la Docencia (2008). Reflexiones sobre el diseño y puesta en marcha de programas de evaluación de la docencia. *Revista Iberoamericana de Evaluación Educativa*, 1(3), 163-168.
- Rodríguez, N. (1999). Validación cruzada de pruebas psicométricas. Trabajo presentado en el *Congreso Interamericano de Medición y Evaluación Psicológica*, Caracas, Venezuela. Recuperado de http://www.psicococonsult.com/download/cdt_36.pdf
- Rueda, M. (2006). *Evaluación de la labor docente en el aula universitaria*. México: UNAM.
- Rueda, M., Luna, E., García, B. y Loredó, J. (2010). Universidades Públicas Mexicanas: Un Diagnóstico Para Su Comprensión y Mejora. *Revista Iberoamericana de Evaluación Educativa* 3(1e), 77-92
- Ruíz, M., Pardo, A., y San Martín, R. (2010). Modelos de ecuaciones estructurales. *Papeles del psicólogo*, 31(1), 34-45.
- Scriven, M. (1995). Student ratings offer useful input to teacher evaluations. *Practical Assessment, Research and Evaluation*, 4(7). Recuperado de <http://PAREonline.net/getvn.asp?v=4&n=7>
- Stapleton, C. (enero, 1997). Basic Concepts and Procedures of Confirmatory Factor Analysis. Trabajo presentado en la conferencia de la *Annual Meeting of the Southwest Educational Research Association*, Texas, EE.UU. Recuperado de <http://www.eric.ed.gov/PDFS/ED407416.pdf>
- Zamora, S., Monroy, L. y Chávez, C. (2009). *Análisis factorial: una técnica para evaluar la dimensionalidad de las pruebas*. Cuaderno técnico 6. México: Centro Nacional de Evaluación para la Educación Superior, A.C.