

**EVALUACIÓN DE LAS PROPIEDADES PSICOMÉTRICAS DEL
CUESTIONARIO DE EVALUACIÓN DOCENTE (DOCENTIA)
PARA SU USO EN UNA MUESTRA MEXICANA**

**EVALUATION OF THE PSYCHOMETRIC PROPERTIES OF THE TEACHER
EVALUATION QUESTIONNAIRE (DOCENTIA) FOR ITS USE IN A MEXICAN
SAMPLE**

**AVALIAÇÃO DAS PROPRIEDADES PSICOMÉTRICAS DO QUESTIONÁRIO DE
AVALIAÇÃO DOCENTE (DOCENTIA) PARA SEU USO NUMA MOSTRA
MEXICANA**

Beatriz Acosta

Revista Iberoamericana de Evaluación Educativa 2012 - Volumen 5, Número 2

<http://www.rinace.net/riee/numeros/vol5-num2/art13.pdf>

Fecha de recepción: 10 de diciembre de 2011

Fecha de dictaminación: 24 de febrero de 2012

Fecha de aceptación: 29 de febrero de 2012

1. CONTEXTO Y PRÁCTICA DE LA EVALUACIÓN DE LA ACTIVIDAD DOCENTE

Los procesos de evaluación educativa en nuestro país y en otras naciones del mundo han acrecentado su importancia en las últimas décadas. Éstos han adquirido dimensiones de relevancia especial, pues surgen como un instrumento para mejorar la calidad del sistema escolar. Del mismo modo, se han convertido en un medio para tener acceso a distintos programas de compensación salarial para el personal académico y como condición para obtener recursos económicos adicionales a nivel institucional (ANECA, 2006; Rueda, 2004). Así, la evaluación educativa se considera un pilar de las políticas actuales en materia de educación superior, puesto que forma parte de los procesos de planeación impulsados por los programas de los organismos del gobierno federal vinculados al sector educativo. Estos mecanismos de evaluación se insertan a su vez en las políticas y recomendaciones que emiten instituciones internacionales como el FMI y Banco Mundial, las cuales han asociado la mejora de los procesos de enseñanza-aprendizaje en la educación superior con estos esquemas de evaluación.

En el caso de las universidades públicas, la evaluación se empezó a implementar de forma generalizada desde la década de los noventa, sobre todo a partir de las políticas promovidas por organismos oficiales como la Secretaría de Educación Pública y la Asociación Nacional de Universidades e Instituciones de Educación Superior (Rueda, Elizalde y Torquemada, 2003). En los últimos años, las políticas del sector educativo han fomentado los procesos de acreditación de los programas de formación profesional; lo anterior derivado del escrutinio minucioso de pares académicos que cuentan con el reconocimiento del organismo oficial encargado de cumplir con esta función, el Consejo para la Acreditación de la Educación Superior (COPAES). De esta forma, las acciones de evaluación institucional de planes y programas, así como las dirigidas a los académicos de educación superior constituyen parte de este nuevo esfuerzo por lograr elevar la calidad del sector (Rueda, 2004).

El contexto descrito forma parte de lo que ha venido ocurriendo con respecto a la evaluación de la docencia en México y de manera particular en el nivel superior. La interacción de los profesores con sus estudiantes en los salones de clase es un aspecto que se ha privilegiado en el marco del programa de formación profesional. Esta actividad es objeto de evaluación con un tratamiento diferencial respecto a otras funciones desempeñadas por los académicos, tales como la investigación, la extensión o la difusión de la cultura (Rueda, 2004).

2. ENFOQUE PARA LA EVALUACIÓN DE LA ACTIVIDAD DOCENTE

La actividad docente se define como el conjunto de actuaciones que se realizan dentro y fuera del aula, destinadas a favorecer el aprendizaje de los estudiantes, con relación a los objetivos y competencias definidas en un plan de estudios y en un contexto institucional determinado. En consecuencia, la actividad docente implica la coordinación y gestión de la enseñanza a clase vacía, el despliegue de métodos de enseñanza, actividades de aprendizaje y de evaluación a clase llena, así como la posterior revisión y mejora de las actividades realizadas (ANECA, 2006).

La evaluación de la actividad docente ha sido abordada y definida desde diversas perspectivas y a partir del contexto institucional en que se desarrolla la misma; en este sentido, la ANECA (2006) define la evaluación de la actividad docente como *“la valoración sistemática de la actuación del profesorado considerando su rol profesional y su contribución para conseguir los objetivos de la titulación en la que está implicado”* (p.9).

Para Llanera (1991, p.22) la evaluación *“es una estrategia que permite identificar y atender en forma pertinente los problemas asociados al desempeño académico y profesional de los docentes, con la finalidad de verificar, retroalimentar y mejorar la calidad de los procesos de enseñanza-aprendizaje”*.

Luviano (2002) la define como *“un proceso temporal que permite identificar eficiencia y eficacia del desempeño docente”*. Además, a través de la evaluación se obtiene un diagnóstico de necesidades, se infieren juicios, se establecen valoraciones y se toman decisiones para mejorar la práctica docente.

Valdés (2000) considera a la evaluación del desempeño docente como *“un proceso sistemático de obtención de datos válidos y confiables, con el objetivo de comprobar y valorar el efecto educativo que produce en los alumnos el despliegue de sus capacidades pedagógicas, su emocionalidad, responsabilidad laboral y la naturaleza de sus relaciones interpersonales con los alumnos”*.

En este sentido, la evaluación de la actividad docente ha permitido a las instituciones educativas reflexionar sobre la enseñanza, su misión, su quehacer y, en general sobre la calidad de la enseñanza que imparten. Asimismo, ésta permite analizar las fortalezas y debilidades de la práctica docente convirtiéndose en un valioso referente de implementación de mecanismos que ayuden a mejorarla (Universidad de Colima, 2006).

3. PROCEDIMIENTOS DE RECOLECCIÓN DE INFORMACIÓN SOBRE LA ACTIVIDAD DOCENTE

Las universidades deben seleccionar aquellas fuentes y métodos de recolección de información, que estimen más adecuadas con base en los objetivos e implicaciones de la evaluación. Éstas se condicionan al cumplimiento de determinados requisitos que aseguren que la calidad de la información sirva de base para la evaluación (ANECA, 2006).

Para dar rigor y consistencia a la evaluación de la calidad de la actividad docente, las universidades hacen uso de diversos procedimientos de recolección de información, tales como cuestionarios de autoevaluación para profesores o autoinforme, evaluaciones a través de pares académicos, directivos o informes de responsables académicos, así como el cuestionario de opinión de los estudiantes (ANECA, 2007; Rueda *et al.*, 2003; Rueda, 2003).

Los cuestionarios de opinión dirigidos a los estudiantes son la única manera para las universidades de dar respuesta a las políticas de evaluación de la docencia. El uso de los cuestionarios, al igual que se hace en la mayor parte de las universidades de otros países, ha tenido un fuerte impulso en México (Rueda, 2003). Estos están dirigidos a aquellos estudiantes matriculados en cualquiera de las enseñanzas que se inscriben en el ámbito definido por las universidades en sus procedimientos de evaluación (ANECA, 2007).

La utilización del cuestionario se debe a que *“el rendimiento o aprendizaje que logra el alumno en el curso es la variable que más afecta la efectividad del maestro(...) lo que más valoran los alumnos al evaluar el desempeño del profesor es lo aprendido en el curso(...), los factores o variables extraclase y su influencia en la evaluación no es tan grande como se presume y por lo tanto, los cuestionarios de evaluación se comportan como instrumentos válidos y confiables para valorar la docencia”* (García, 2000, p. 319).

En esta dirección, el cuestionario recupera las opiniones de los alumnos, pues se consideran los principales agentes del proceso educativo y aquellos que se encuentran inmersos directamente en la

dinámica que se vive en el aula. Los alumnos son los informadores del desempeño docente, ya que son los receptores de la educación que imparten los profesores y quienes, básicamente, reflejan la calidad del trabajo de los docentes (Universidad de Colima, 2006; Rico, Montalvo y Ayala, 2001; Abrami, Cohen y D'Apollonia, 1990; Epstein, 1974).

4. MODELO DE EVALUACIÓN DE LA ACTIVIDAD DOCENTE. PROGRAMA DOCENTIA

Según Rueda (2003) en los cuestionarios prevalecen las dimensiones relativas a la planeación, el desarrollo y la cobertura del programa, así como las formas de evaluación de los aprendizajes. En menor medida se consideran dimensiones sobre el dominio de la asignatura, el manejo de técnicas didácticas, el estilo de interacción profesor-alumno, la asistencia y la puntualidad. En la misma línea, Barrón (2004) considera como base para el análisis de las diversas formas de la evaluación de la docencia: el pensamiento del profesor, la planificación de la enseñanza, la práctica educativa del aula y la evaluación de los resultados de aprendizaje de los alumnos.

El programa DOCENTIA define a la evaluación de la actividad docente como la valoración sistemática de la actuación del profesorado considerando su rol profesional y su contribución para conseguir los objetivos de la titulación en la que está implicado, en función del contexto institucional en que ésta se desarrolla.

El término actividad docente del profesorado contenido en la definición implica diferentes acciones encaminadas a organizar, coordinar, planificar y enseñar a los estudiantes, además de evaluar el aprendizaje. Dichas acciones se despliegan en respuesta a los objetivos formativos y capacidades que se pretenden favorecer en los estudiantes.

Los resultados de la evaluación docente son un fundamento de revisión y mejora en los planes de estudio de las universidades, con el fin de que dichos resultados conlleven a iniciar una nueva manera de formación (ANECA, 2012).

El modelo de evaluación definido en DOCENTIA contempla tres dimensiones en el análisis y valoración de la actividad docente: Planificación de la docencia, Desarrollo de la Enseñanza y Resultados (ver figura 1).

FIGURA 1. DIMENSIONES DE EVALUACIÓN DE LA ACTIVIDAD DOCENTE



Cada una de las dimensiones presentadas en la figura 1 se articula a su vez en un conjunto de elementos, los cuales se presentan en el cuadro 1.

CUADRO 1. DIMENSIONES Y ELEMENTOS DE LA EVALUACIÓN DOCENTE

Dimensiones	Elementos	
Planificación de la docencia	1. Organización y coordinación docentes	Modalidades de organización
		Coordinación con otras actuaciones docentes
	2. Planificación de la enseñanza y del aprendizaje con relación a las materias impartidas	Resultados del aprendizaje previstos
		Actividades de aprendizaje previstas
		Criterios y métodos de evaluación
Desarrollo de la enseñanza	3. Desarrollo de la enseñanza y evaluación del aprendizaje	Materiales y recursos de la docencia
		Actividades de enseñanza y aprendizaje realizadas
Resultados	4. Resultados en términos de objetivos formativos	Procedimientos de evaluación aplicados
		5. Revisión y mejora de la actividad docente: formación e innovación

Fuente: ANECA (2006). DOCENTIA (Programa de apoyo para la evaluación de la actividad docente del profesorado universitario). Modelo de evaluación. España: ANECA

La dimensión **Planificación de la enseñanza** responde cuestiones relacionadas con:

El programa de la asignatura. Los estudiantes opinan sobre la claridad con la que están planteados los objetivos y el consiguiente despliegue de contenidos, actividades a realizar por los estudiantes, criterios y formas de evaluación, bibliografía y otros recursos. De igual modo, los estudiantes opinan sobre la viabilidad del desarrollo efectivo del programa desde la perspectiva del tiempo que debe invertir el estudiante.

La coordinación docente. Los estudiantes opinan sobre la existencia de repeticiones o solapamientos de contenidos en el programa de la asignatura, así como las diferencias manifiestas en los criterios de evaluación con relación a programas de otras asignaturas.

En la dimensión **Desarrollo de la enseñanza**, los estudiantes responden a cuestiones relacionadas con:

Actividades de enseñanza y aprendizaje. Los estudiantes manifiestan sus opiniones sobre el cumplimiento del programa de la asignatura y la realización de las actividades prácticas, y de tutoría. Asimismo, opinan sobre la metodología y los recursos didácticos utilizados por el profesor, y el modo en que ésta ha contribuido a favorecer el aprendizaje.

Procedimientos de evaluación. Los estudiantes opinan sobre el modo en que los procedimientos de evaluación utilizados por el profesor, se ajustan a los recogidos en el programa de la asignatura y han permitido estimar de un modo apropiado, el aprendizaje realizado por el alumno. De igual modo, los estudiantes opinan sobre la transparencia en el desarrollo del proceso de evaluación (aplicación de *criterios* explícitos de corrección, objetividad, publicidad, etc.).

En la dimensión **Resultados** los estudiantes responden a cuestiones relacionadas con los resultados obtenidos en la asignatura. Así:

Logro por los estudiantes de los objetivos formativos. Los estudiantes opinan sobre el progreso alcanzado en la asignatura, con relación a las competencias previstas en el programa y a su situación de partida de ellas. Asimismo, opinan sobre las carencias en su formación, con relación a la prevista en la asignatura (ANECA, 2007; ANECA, 2006).

El resultado de la evaluación de la actividad docente que obtiene un profesor, mediante el sistema de indicadores propuesto, queda determinada por el conjunto de actores inmediatos que intervienen directa o indirectamente en las actividades docentes del profesor, es decir, alumnos, profesor y responsables académicos. En consecuencia, una forma adecuada de presentar el sistema de indicadores para la evaluación docente es mediante tres instrumentos complementarios e interrelacionados que incluyan los indicadores de la actividad docente del profesor desde las tres ópticas determinadas o sugeridas en función de los actores implicados: a) Cuestionario de estudiantes, b) Autoinforme del profesor y, c) Informe de los responsables académicos (MEC, 2010; ver cuadro 2).

CUADRO 2. DIMENSIONES, FUENTES Y PROCEDIMIENTOS DE EVALUACIÓN

Responsable	Fuentes y procedimientos de evaluación		
	Profesor	Responsables académicos	Estudiantes
Dimensiones			
Planificación	Auto-informe	Informe	Encuestas
Desarrollo			
Resultados			

Fuente: ANECA, (2012). Docentia: Programa de apoyo para la evaluación de la actividad docente del profesorado universitario. Modelo de evaluación. España: ANECA

Un antecedente normativo de la evaluación docente para todas las universidades españolas lo establece el Real Decreto 1313/2007 de 5 de octubre, el cual menciona la Acreditación Nacional para el Acceso a los Cuerpos Docentes Universitarios; éste, en el apartado 28 de su Anexo, concreta como uno de los méritos evaluar la calidad de la actividad docente. De forma más específica, en la Guía de Ayuda del Programa ACADEMIA, elaborada por ANECA (2008), se indica que las universidades deberán emitir un certificado de calidad docente en el que conste "... la valoración global de la actividad docente del solicitante con indicación de la metodología utilizada para obtenerla, por ejemplo las establecidas mediante el Programa de Apoyo a la Evaluación de la Actividad Docente del Profesorado Universitario (DOCENTIA)" (Aneca, 2008, p.17).

El programa DOCENTIA ha representado un esfuerzo español para la adaptación de los títulos al nuevo marco del Espacio Europeo de Educación Superior (EEES). La evaluación de la actividad docente se convierte así en un pilar fundamental en el ámbito de los nuevos Sistemas de Garantía de la Calidad (SGC). Por ello, y en este contexto de cambios, la necesidad de emprender la evaluación de la labor docente de forma integral constituye una de las demandas más urgentes para las universidades españolas. El programa DOCENTIA promovido por la ANECA (2007) proporciona un marco de referencia, un modelo y unos procedimientos para abordar la evaluación de la actividad docente que se desarrolla en las universidades de forma que éstas diseñen sus propios procedimientos de evaluación ajustados a su contexto específico y a sus necesidades (Pozo, Giménez y Bretoñes, 2009).

Pozo *et al.*, (2009) realizó la valoración de las propiedades psicométricas del cuestionario de evaluación docente DOCENTIA-ANDALUCÍA. El resultado de la consistencia interna utilizando la prueba alfa de Cronbach reportó una puntuación de $\alpha=0.97$, lo que se traduce en una elevada confiabilidad. Por su parte el análisis factorial exploratorio con rotación varimax permitió identificar dos factores independientes que explican el 68.93% de la varianza total. El primer factor (45.55% de la varianza total) está conformado por un total de 16 reactivos, mientras que el segundo, explica el 23.27% agrupa un total de 7 elementos.

De acuerdo con Acosta Uribe (2011) para evaluar las propiedades psicométricas de un instrumento, suele recurrirse a procedimientos tradicionales: análisis "clásico" de ítems, estudio de la fiabilidad, de la estructura factorial. Entre las razones principales para que estos análisis se hayan convertido en una práctica rutinaria se encuentran: a) que están disponibles en la mayoría de los paquetes estadísticos; y b) que no requieren un tamaño de muestra elevado.

Uno de los primeros análisis para evaluar la equivalencia de las versiones entre las diferentes versiones del test o cuestionario consiste en determinar la igualdad de los coeficientes de fiabilidad. El estadístico más utilizado para este tipo de análisis es el estadístico W de Feldt (Alsawalmeh y Feldt, 1994; Feldt, 1980). En este sentido, si no existen diferencias significativas se considera la igualdad de los coeficientes alfa de Cronbach, en cambio, si se presentan diferencias significativas, dichas diferencias pueden deberse a que el instrumento tiene un conjunto de ítems defectuosos o, bien, que no se mide el mismo constructo en todos los grupos.

La comparación de la correlación ítem-total o índice de discriminación (ID) permite identificar la adecuación de los ítems del test o cuestionario en los diversos grupos culturales para los que se ha adaptado. Si existen marcadas diferencias entre los índices de discriminación pueden deberse a problemas del instrumento, problemas en la administración, como por ejemplo: efectos substanciales del entrevistador, a las propias características de los sujetos, como diferencias transculturales, o a estilos diferenciales de respuestas como la aquiescencia o la deseabilidad social.

El análisis factorial exploratorio es quizás, la técnica más aplicada para evaluar la equivalencia de constructo o estructural, seguida de la rotación y del cómputo de un índice de acuerdo factorial entre las diferentes estructuras factoriales halladas. Índices como el de identidad, el de aditividad, el de proporcionalidad y el de linealidad evalúan el acuerdo factorial. No obstante, el más utilizado es el coeficiente de congruencia de Tucker también conocido como Phi de Tucker (Tucker, 1951 en García Cueto, 1994; van de Vijver y Leung, 1997).

El presente trabajo tiene como propósito evaluar las propiedades psicométricas del cuestionario de evaluación docente DOCENTIA (ANECA, 2006) para su uso en una muestra mexicana; a partir las administraciones que se realizaron durante los ciclos lectivos 2007, 2008 y 2010 a los alumno(a)s de la Universidad del Papaloapan, Campus Tuxtepec y Campus Loma Bonita.

La evaluación de las propiedades psicométricas implica desglosar el objetivo general en los siguientes objetivos específicos:

1. Analizar las propiedades psicométricas del cuestionario de evaluación docente DOCENTIA.
2. Analizar las semejanzas entre la consistencia interna y la estructura factorial de las administraciones realizadas en los ciclos lectivos 2007, 2008 y 2010 del cuestionario de evaluación docente DOCENTIA.
3. Obtener evidencias sobre el grado de equivalencia alcanzado entre las administraciones realizadas en los ciclos lectivos 2007, 2008 y 2010 del cuestionario de evaluación docente DOCENTIA.
4. Evaluar la dimensionalidad del cuestionario de evaluación docente DOCENTIA.

5. METODOLOGÍA

5.1. Participantes

En 2007, la primera evaluación docente se realizó en la Universidad del Papaloapan, Campus Tuxtepec. En este periodo, la comunidad de estudiantes estaba conformada por 81 alumno(a)s. En 2008, la administración del cuestionario se llevó a cabo en los dos campus (Tuxtepec y Loma Bonita). En este momento, la Universidad contaba con 297 estudiantes (130 y 167, respectivamente). En 2010, la Universidad tenía una matrícula de 341 alumno(a)s en los dos campus. De ellos, 167 eran del campus Tuxtepec y 174 del Campus Loma Bonita.

De acuerdo con los informes técnicos 2008a, 2008b y 2010 (Acosta y Ruiz, 2008a; Acosta, Ruiz, Vigil y de la Rosa, 2008b y Acosta *et al.*, 2010). En 2007, la evaluación docente se realizó en el Campus Tuxtepec. En ella, participó el 74.07% del total de la matrícula, es decir 60 alumnos (29 de segundo semestre y 31 de cuarto semestre) de las diversas carreras de la Universidad del Papaloapan, Campus Tuxtepec.

En este periodo, los alumno(a)s proporcionaron su opinión respecto al quehacer docente dentro del aula de 18 profesores-investigadores, todos ellos de tiempo completo. Dichos profesores impartieron 26 materias en el periodo antes mencionado.

En 2008, la evaluación se llevó a cabo en los dos campus, en este periodo, se logró la participación del 89.56% del total de la población estudiantil. En el campus Loma Bonita participaron 152 alumno(a)s, de ellos, 63 estaban inscritos en el primer semestre, 24 en el tercero, 35 en el quinto, 18 en el séptimo y 9 en el noveno semestre de su carrera. En cuanto al Campus Tuxtepec, la evaluación fue contestada por 114 alumno(a)s, de los cuales 51 cursaban el primer semestre, 32 el tercer semestre y 31 el quinto semestre.

En dicha evaluación docente, se calificó a 71 profesores (47 del Campus Loma Bonita y 24 del Campus Tuxtepec), los cuales dieron 102 materias (66 en Loma Bonita y 34 en Tuxtepec).

TABLA 1. DISTRIBUCIÓN DE ALUMNOS PARTICIPANTES Y CUESTIONARIOS ADMINISTRADOS EN LA EVALUACIÓN DOCENTE 2007, 2008 Y 2010

Campus Semestre	Tuxtepec				Loma Bonita					
	2007		2008		2010		2008		2010	
	No. alumno(a)s	No. cuestionarios	No. alumno(a)s	No. cuestionarios	No. alumno(a)s	No. cuestionarios	No. alumno(a)s	No. cuestionarios	No. alumno(a)s	No. cuestionarios
1	0	0	51	273	0	2	63	310	0	427
2	29	186	0	0	85	498	0	0	74	0
3	0	0	32	190	0	0	24	133	0	113
4	31	179	0	0	24	128	0	0	19	0
5	0	0	31	196	0	0	35	115	0	61
6	0	0	0	0	11	66	0	0	10	0
7	0	0	0	0	0	0	18	94	0	46
8	0	0	0	0	19	113	0	0	8	0
9	0	0	0	0	0	0	12	33	0	32
10	0	0	0	0	23	114	0	0	6	0
Total general	60	365	114	659	162	921	152	685	117	679

Durante el ciclo escolar 2010, participó en el estudio el 81.81% (279) de la comunidad estudiantil. De ellos, 117 alumno(a)s del Campus Loma Bonita y 162 alumno(a)s del Campus Tuxtepec. Del primer campus, 74 alumno(a)s cursaban el segundo semestre, 19 el cuarto, 10 el sexto, 8 el octavo y 6 el décimo semestre. Con respecto al segundo campus, 85 alumnos eran de segundo semestre, 24 de cuarto, 11 de sexto, 19 de octavo y 23 de décimo semestre. En este periodo, los alumno(a)s evaluaron a 89 profesores (54 de Loma Bonita y 35 de Tuxtepec), los cuales impartieron un total de 140 materias (89 en Loma Bonita y 51 en Tuxtepec).

Cabe mencionar que, en la primera administración se respondieron 365 cuestionarios, todos del Campus Tuxtepec. En la segunda 1344 (659 Campus Tuxtepec y 685 Campus Loma Bonita) y, en la tercera 1600 (921 Campus Tuxtepec y 679 Campus Loma Bonita, ver tabla 1).

6. RESULTADOS

Los resultados se distribuirán en tres apartados. El primero presenta los procedimientos tradicionales para el estudio de la confiabilidad y de la estructura factorial. En este sentido, en principio se muestran los valores de los estadísticos descriptivos calculados para analizar las distribuciones de respuestas a los ítems, los valores de la correlación ítem-total (Índice de Discriminación, de aquí en adelante por sus siglas "ID"), los valores de la contribución del ítem a los coeficientes alfa de Cronbach para cada una de las escalas y los resultados de la consistencia interna del instrumento a partir del coeficiente alfa de Cronbach.

El segundo apartado contiene las evidencias para probar la equivalencia del cuestionario, en este sentido, se muestra el estadístico W de Feldt para analizar la similitud entre los coeficientes antes mencionados. Enseguida, se incluyen los índices de identidad, de aditividad, de proporcionalidad y de linealidad con el fin de obtener la similitud factorial. El tercer apartado evalúa la dimensionalidad del instrumento a través del criterio de información de Akaike.

6.1. Procedimientos tradicionales (confiabilidad y análisis de la estructura factorial)

TABLA 2. ANÁLISIS DESCRIPTIVOS (MEDIA, DESVIACIÓN ESTÁNDAR, ÍNDICE DE DISCRIMINACIÓN Y ALFA DE CRONBACH DE LOS REACTIVOS) DISTRIBUIDOS POR CICLOS LECTIVOS Y POR ESCALAS

Escala	No.	2007				2008				2010			
		Media	d.s.	ID	Alfa si el ítem se elimina	Media	d.s.	ID	Alfa si el ítem se elimina	Media	d.s.	ID	Alfa si el ítem se elimina
Planeación	pla01	4.82	1.11	0.78	0.91	4.65	1.26	0.78	0.91	4.63	1.19	0.71	0.90
	pla02	5.01	1.09	0.79	0.90	4.87	1.14	0.81	0.91	4.83	1.10	0.79	0.89
	pla03	4.88	1.11	0.68	0.92	4.70	1.18	0.73	0.92	4.60	1.17	0.69	0.90
	pla04	4.85	1.16	0.81	0.90	4.75	1.21	0.82	0.91	4.68	1.18	0.79	0.89
	pla05	4.61	1.16	0.79	0.90	4.69	1.20	0.78	0.91	4.55	1.18	0.74	0.90
	pla06	4.70	1.26	0.78	0.91	4.77	1.20	0.78	0.91	4.61	1.18	0.78	0.89
Desarrollo	des07	4.55	1.35	0.50	0.94	4.57	1.24	0.72	0.95	4.40	1.24	0.60	0.93
	des08	4.91	1.28	0.76	0.92	4.79	1.24	0.80	0.95	4.63	1.26	0.75	0.93
	des09	4.87	1.27	0.83	0.92	4.71	1.32	0.82	0.95	4.67	1.23	0.80	0.92
	des10	4.97	1.24	0.83	0.92	4.82	1.27	0.84	0.94	4.72	1.22	0.80	0.92
	des11	4.86	1.25	0.70	0.93	4.78	1.23	0.77	0.95	4.56	1.26	0.69	0.93
	des12	5.00	1.14	0.77	0.92	4.78	1.22	0.81	0.95	4.65	1.18	0.75	0.93
	des13	4.93	1.15	0.66	0.93	4.73	1.26	0.78	0.95	4.63	1.25	0.74	0.93
	des14	4.97	1.12	0.77	0.92	4.82	1.16	0.78	0.95	4.70	1.17	0.74	0.93
	des15	5.08	1.08	0.76	0.92	4.87	1.22	0.79	0.95	4.78	1.16	0.75	0.93
des16	4.87	1.25	0.81	0.92	4.67	1.31	0.82	0.95	4.58	1.25	0.76	0.93	
Resultados	res17	5.11	1.21	0.81	0.92	4.85	1.24	0.79	0.93	4.74	1.18	0.74	0.91
	res18	4.88	1.23	0.81	0.92	4.85	1.18	0.81	0.92	4.67	1.20	0.77	0.90
	res19	4.90	1.24	0.84	0.91	4.72	1.27	0.87	0.91	4.62	1.20	0.83	0.89
	res20	4.76	1.24	0.82	0.92	4.56	1.22	0.84	0.92	4.57	1.15	0.80	0.90
	res21	4.87	1.38	0.82	0.92	4.72	1.36	0.83	0.92	4.63	1.28	0.80	0.90

La tabla 2 contiene la información relacionada con los análisis descriptivos, los resultados del índice de discriminación y del alfa de Cronbach si se elimina el reactivo. En este sentido, puede apreciarse que en el periodo 2007, los resultados de los índices de discriminación de los ítems en las escalas oscilan entre 0.50 y 0.84; para el periodo 2008 entre 0.60 y 0.83 y para el periodo 2010 entre 0.64 y 0.85. Las diferencias encontradas entre 2007/2008 fluctúan entre 0.02 y 0.22; entre 2007/2010 entre 0.07 y 0.10 y, entre 2008/2010 entre 0.00 y 0.12.

En todos los periodos lectivos, los resultados reportados del alfa de Cronbach si el ítem se elimina permiten corroborar, que todos los enunciados contribuyen de manera eficiente con la consistencia interna de las escalas.

Los resultados de la consistencia interna en los tres ciclos escolares son muy altos, éstos oscilan entre 0.91 y 0.95. Como puede apreciarse, las escalas en el periodo lectivo 2008 son ligeramente superiores con respecto a los otros dos ciclos escolares (entre 0.01 y 0.02, ver tabla 3).

TABLA 3. CONFIABILIDAD DE LAS ESCALAS DE LOS CICLOS LECTIVOS 2007, 2008 Y 2010

Periodo / Escala	2007 (n=365)	2008 (n=1344)	2010 (n=1600)
Planeación	0.92	0.93	0.91
Desarrollo	0.93	0.95	0.93
Resultados	0.93	0.94	0.92
General	0.92	0.94	0.93

Antes de llevar a cabo el Análisis Factorial correspondiente, se calculó el Test de Esfericidad de Bartlett y el Índice KMO a fin de determinar la pertinencia de un análisis factorial.

El Test de Esfericidad de Bartlett contrasta la H_0 de que la matriz de correlaciones es una matriz identidad (incorrelación lineal entre las variables). Si, como resultado del contraste, no pudiésemos rechazar ésta H_0 y el tamaño de la muestra fuese razonablemente grande, deberíamos reconsiderar la realización de un Análisis Factorial.

Por su parte, el Índice KMO (Kaiser-Meyer-Olkin) de adecuación de la muestra miden la correlación existente entre dos variables una vez eliminada la influencia que las restantes variables ejercen sobre ellas. Es comúnmente aceptado que:

Si $KMO < 0.5$ no resultaría aceptable hacer un Análisis Factorial

Si $0.5 < KMO < 0.6$ significa un grado de correlación medio, y habría aceptación media.

Si $KMO > 0.7$ indica alta correlación y, por tanto, conveniencia de Análisis Factorial (Gondar Nores, 2004).

En el ciclo escolar 2007, los resultados del índice $KMO = 0.97$ y del Test de esfericidad de Bartlett $\chi^2 = (7085.68; g.l. 210); p < 0.001$; en el periodo escolar 2008, el índice $KMO = 0.98$ y $\chi^2 = (27\ 639.72; g.l. 210); p < 0.001$ y, en el ciclo lectivo 2010, el índice $KMO = 0.97$ y $\chi^2 = (27304.03; g.l. 210); p < 0.001$ permiten rechazar la Hipótesis nula sobre incorrelación lineal entre las variables. Dichos resultados nos conducen a realizar un Análisis Factorial (ver tabla 4).

En los tres ciclos escolares, el resultado del primer autovalor fue de 13.35, 14.16 y 12.72, respectivamente, el cual explica el 63.58%, 67.46% y 60.58% del total de la varianza, respectivamente.

Los resultados del Análisis de Componentes Principales sin rotar de los ciclos escolares 2007, 2008 y 2010, reportaron cargas factoriales que oscilan entre 0.56 y 0.87; 0.74 y 0.85 y; 0.66 y 0.83,

respectivamente. Como puede apreciarse el primer factor da cuenta totalmente de la matriz de correlación reducida, por lo que podría aceptarse el supuesto de unidimensionalidad de la escala (ver tabla 5).

TABLA 4. TEST DE ESFERICIDAD E BARTLETT, ÍNDICE KMO DE LOS CICLOS LECTIVOS 2007, 2008 Y 2010

Ciclos lectivos	Test de Esfericidad de Bartlett		Índice KMO
	χ^2	g.l.	
2007.00	7085.68	210	0.96
2008.00	27639.72	210	0.98
2010.00	27304.03	210	0.97

TABLA 5. ESTRUCTURA FACTORIAL DE LA EVALUACIÓN DOCENTE ANTES DE ROTAR DE LOS CICLOS ESCOLARES 2007, 2008 Y 2010

Reactivo	2007	2008	2010	
	1	1	1	2
pla01	0.80	0.81	0.74	0.26
pla02	0.83	0.82	0.80	0.26
pla03	0.73	0.74	0.69	0.41
pla04	0.82	0.84	0.78	0.34
pla05	0.80	0.80	0.76	0.32
pla06	0.80	0.82	0.79	0.32
des07	0.56	0.77	0.66	
des08	0.83	0.83	0.80	
des09	0.87	0.85	0.83	
des10	0.86	0.85	0.82	
des11	0.72	0.79	0.70	-0.31
des12	0.79	0.83	0.76	-0.28
des13	0.68	0.81	0.78	
des14	0.78	0.81	0.77	
des15	0.79	0.82	0.78	
des16	0.85	0.85	0.80	
res17	0.82	0.83	0.80	
res18	0.83	0.84	0.79	
res19	0.83	0.86	0.84	
res20	0.84	0.83	0.81	
res21	0.86	0.85	0.83	

Al llevar a cabo el análisis de componentes principales con rotación varimax, los resultados reportaron lo siguiente:

En el periodo 2007, los autovalores de los tres primeros componentes fueron 13.35, 0.94 y 0.80. El primero explica el 63.58% de la varianza, el segundo el 4.47% y el tercero 3.82%, explicando el 71.88% del total de la varianza.

En el periodo 2008, los primeros tres autovalores fueron 14.16, 0.80 y 0.59. El primero da cuenta del 67.47% de la varianza, el segundo del 3.80% y el tercero del 2.81%. Los tres componentes explican el 74.08 del total de la varianza.

En el periodo 2010, los autovalores de los tres primeros componentes fueron 12.72, 1.03 y 0.78. El primero explica el 60.58% de la varianza, el segundo el 4.88 y el tercero el 3.71%. Estos explican el 69.17% del total de la varianza.

Las cargas factoriales del periodo escolar 2007 oscilaron entre 0.55 y 0.75, las del ciclo escolar 2008 entre 0.52 y 0.77 y las del ciclo lectivo 2010 entre 0.53 y 0.77.

Como puede apreciarse, en los tres ciclos escolares, las cargas factoriales de la escala de planeación (plan01-plan06) se agrupan en un componente (III, II y II, respectivamente).

Caso similar ocurre con las cargas factoriales de la escala de resultados. En los periodos 2007 y 2010, éstas se agrupan en el componente I. No obstante, en 2008, las cargas factoriales de dicha escala se agrupan en dos componentes: en el componente III se agrupan res17 y res18 y en el componente II res 19, res 20 y res 21. Al analizar los reactivos puede constatarse que los ítems: res17 "El modo en que evalúa (exámenes, trabajos individuales o de grupo, etc.) guarda relación con el tipo de tareas (teóricas, prácticas, individuales, grupales, etc.) desarrolladas" y, res18 "El profesor aplica de un modo adecuado los criterios de evaluación recogidos en el programa", tienen que ver con "procedimientos de evaluación aplicados" y los otros tres: res19 "El profesor ha facilitado mi aprendizaje, gracias a su ayuda he logrado mejorar mis conocimientos, habilidades o modo de afrontar determinados temas", res20 "He mejorado mi nivel de partida, con relación a las competencias previstas en el programa" y, res21 "En general, estoy satisfecho con la labor docente de este profesor", se relacionan con "resultados en términos de objetivos formativos".

Los reactivos de la escala de desarrollo fueron los que mostraron más dispersión en los tres ciclos lectivos. Sin embargo al analizar el enunciado de cada reactivo, estos pueden agruparse de la siguiente manera: des07, des08, des09 y des10 tienen que ver con dominio del tema; des11, des12, des13 y des14 tienen que ver con habilidades pedagógicas y, des15 y des16 tienen que ver con manejo del grupo (ver tabla 6).

TABLA 6. ESTRUCTURA FACTORIAL DE LA EVALUACIÓN DOCENTE DESPUÉS DE LA ROTACIÓN

Reactivo	2007			2008			2010		
	Componente			Componente			Componente		
	1.00	2.00	3.00	1.00	2.00	3.00	1.00	2.00	3.00
pla01	0.29	0.39	0.70	0.40	0.66	0.35	0.27	0.67	0.36
pla02	0.39	0.41	0.64	0.30	0.71	0.41	0.30	0.70	0.40
pla03		0.45	0.60		0.74	0.34		0.74	
pla04	0.36	0.37	0.69	0.37	0.67	0.42	0.31	0.74	0.31
pla05	0.32	0.34	0.72	0.37	0.68	0.33	0.40	0.68	
pla06	0.41		0.74	0.42	0.63	0.37	0.44	0.70	
des07	0.69		0.25	0.52	0.61		0.63	0.39	
des08	0.51	0.33	0.60	0.51	0.59	0.33	0.56	0.49	0.32
des09	0.55	0.48	0.48	0.63	0.47	0.35	0.63	0.41	0.39
des10	0.56	0.55	0.37	0.59	0.40	0.48	0.49	0.38	0.56
des11		0.75	0.33	0.31	0.35	0.74			0.79
des12	0.31	0.73	0.33	0.46	0.37	0.62	0.30	0.28	0.77
des13		0.72	0.27	0.36	0.41	0.65	0.41	0.38	0.57
des14	0.31	0.64	0.39	0.34	0.41	0.66	0.38	0.35	0.63
des15	0.44	0.56	0.37	0.52	0.40	0.51	0.57	0.27	0.50
des16	0.67	0.40	0.41	0.71	0.39	0.36	0.67	0.30	0.40
res17	0.61	0.53	0.29	0.50	0.34	0.60	0.53	0.39	0.45
res18	0.59	0.58	0.27	0.51	0.38	0.57	0.62	0.33	0.40
res19	0.73	0.42	0.28	0.76	0.30	0.41	0.72	0.32	0.39
res20	0.65	0.44	0.37	0.75	0.32	0.36	0.75	0.29	0.32
res21	0.69	0.31	0.48	0.77	0.34	0.33	0.72	0.33	0.35

6.2. Análisis de la equivalencia

El estadístico W de Feldt indica que no existen diferencias significativas, para el conjunto de los coeficientes alfas obtenido por las escalas en los periodos lectivos 2007-2008 y, en las escalas de Planeación y Desarrollo en los periodos escolares 2007-2010, por lo que no se puede rechazar la hipótesis sobre la igualdad de los coeficientes. Sin embargo, se reportaron diferencias significativas en la escala Resultados de los ciclos escolares 2007-2010 y en todas las escalas en los ciclos 2008-2010, por lo que se rechaza la hipótesis nula sobre la igualdad de los coeficientes (ver tabla 7).

TABLA 7. CONFIABILIDAD DE LAS ESCALAS Y W DE FELDT DE LOS CICLOS LECTIVOS 2007, 2008 Y 2010

Escala/Periodo	Alfa de Cronbach			W de Feldt		
	2007 (n=365)	2008 (n=1344)	2010 (n=1600)	2007-2008	2007-2010	2008-2010
Planeación	0.92	0.93	0.91	1.14	0.88	0.77**
Desarrollo	0.93	0.95	0.93	1.40	1.00	0.71**
Resultados	0.93	0.94	0.92	1.16	0.87*	0.75**
General	0.92	0.94	0.93	1.33	1.14	0.85**

*p<0.05; **p<0.001

Una vez realizado el análisis de componentes principales con rotación varimax, es preciso estimar el acuerdo factorial entre las diversas matrices. van de Vijver y Poortinga (1994) proponen reportar los índices de identidad, de aditividad, de proporcionalidad y de linealidad y posteriormente, comparar sus valores. Lo anterior permitirá obtener un retrato de la similitud factorial. En este sentido, cuando todos los índices son altos, existe una clara evidencia de que el factor en cuestión mide lo mismo en las diversas matrices factoriales. Cuando los valores son bajos o, cuando se encuentran grandes diferencias entre los índices observados, la universalidad de los factores puede ser cuestionada. Valores por encima de 0.95 permiten afirmar que ambas estructuras factoriales son semejantes y valores por debajo de 0.90 (van de Vijver y Poortinga, 1994) o 0.85 (Ten Berge, 1998 en van de Vijver y Leung, 1997) indican incongruencia en las estructuras factoriales.

La tabla 8 presenta los resultados de los coeficientes para evaluar el acuerdo entre las estructuras factoriales entre los diversos periodos escolares. Como puede apreciarse, los índices de identidad y de proporcionalidad obtuvieron valores por encima del valor propuesto. No obstante, los coeficientes de aditividad y de linealidad plantean dudas sobre la semejanza de las estructuras factoriales. Lo anterior, conduce a cuestionarse la universalidad de los factores.

TABLA 8. COEFICIENTES PARA LA EVALUACIÓN DE ACUERDO FACTORIAL DESPUÉS DE LA ROTACIÓN

	2007/2008	2007/2010	2008/2010
Identidad	0.99	1.00	1.00
Aditividad	0.46	0.57	0.66
Proporcionalidad	0.99	1.00	1.00
Linealidad	0.46	0.57	0.66

Debido a que el análisis de componentes principales sin rotación dio indicios a considerar que, el instrumento está compuesto por un componente bien definido y, los resultados del análisis de componentes principales con rotación varimax reportaron una estructura de tres componentes, de acuerdo a la construcción del cuestionario. En este sentido, se realizó el cálculo del criterio de información de Akaike para decidir el mejor modelo. La lógica del modelo es la siguiente: al añadir parámetros libres al modelo (planteando más factores), se consigue un mejor ajuste (menor valor mínimo de la función de discrepancia). El modelo más parsimonioso (el que dará más información), será aquél

que consiga simultáneamente un valor más bajo de la función de discrepancia, utilizando el menor número posible de parámetros libres. (Ferrando, 1996).

6.3. Dimensionalidad del cuestionario

En concreto, se examinó la parsimonia de un modelo unifactorial, de un modelo con dos factores comunes y, por último, de un modelo con tres factores comunes.

La Tabla 9 presenta los resultados del criterio de información de Akaike (recogido en Ferrando, 1996), para los tres modelos en los tres ciclos lectivos.

Como puede apreciarse en los tres periodos, el modelo que más se ajusta es el de un factor común AIC=7169.68, 27723.72 y 27388.03, respectivamente.

TABLA 8. DETERMINACIÓN DEL MEJOR MODELO A PARTIR DEL CRITERIO DE INFORMACIÓN DE AKAIKE

Ciclo escolar	χ^2	g.l.	AIC		
			I	II	III
2007	7085.68	210	7169.68	7210.68	7250.68
2008	27639.72	210	27723.72	27764.72	27804.72
2010	27304.03	210	27388.03	27429.03	27469.03

7. CONCLUSIONES

El objetivo principal de este estudio fue analizar las propiedades psicométricas del cuestionario de evaluación docente.

Los análisis preliminares (estructura factorial, análisis de confiabilidad, correlación ítem-total (ID), distribuciones de respuestas de los reactivos) examinan las propiedades métricas del reactivo y del instrumento en su conjunto, condicionando el posterior nivel de equivalencia (van de Vijver y Leung, 1997). Los resultados obtenidos sobre las propiedades psicométricas en dicho estudio han aportado evidencias relevantes sobre las propiedades métricas del cuestionario en cuestión.

Los niveles de discriminación y la contribución a la consistencia interna de cada una de las escalas han sido excelentes en todos los reactivos. Por un lado, se reportaron coeficientes alfa de Cronbach muy altos, estos se encuentran entre 0.91 y 0.95. Del mismo modo, las puntuaciones de los índices de discriminación para las escalas fueron elevadas de entre 0.50 a 0.85 y el alfa si el ítem se elimina permitió determinar que cada reactivo contribuye a la confiabilidad del instrumento.

Los resultados obtenidos en la estructura factorial, no permiten concluir tajantemente, que las escalas se ajusten exactamente a la planificación prevista en el cuestionario de evaluación docente, sobre todo en la escala de desarrollo de la enseñanza, en la cual, los reactivos se distribuyeron en dos componentes. Sin embargo, dos aspectos se pueden considerar:

1. la estructura factorial sin rotar lleva a pensar que el cuestionario se trata de una estructura unidimensional.
2. las escalas pueden redefinirse en tres escalas y una de ellas, componerse de tres subescalas: planificación de la docencia; desarrollo de la docencia y ésta a su vez redistribuirse en dominio del tema, habilidades pedagógicas y manejo del grupo, tal como lo plantea (Rueda, 2004), y resultados de la enseñanza.

Por consiguiente, los resultados de los estadísticos tradicionales permiten utilizar el cuestionario de evaluación docente con unas garantías suficientes de calidad.

En cuanto a los análisis para obtener evidencias con las que analizar la equivalencia lograda entre las diversas administraciones del cuestionario de evaluación docente. Los resultados para este conjunto de datos fueron muy alentadores, los valores del estadístico W de Feldt señalaron la igualdad de las puntuaciones de los coeficientes alfa, excepto en la comparación de los ciclos lectivos 2008-2010, lo anterior pudo deberse al tamaño de la población encuestada.

Los coeficientes para evaluar el acuerdo entre las estructuras factoriales, plantean dudas respecto a la semejanza de las mismas. Lo anterior, debido a que los valores del índice de aditividad y del índice de linealidad obtuvieron valores bajos en los tres periodos lectivos. Sin embargo, van de Vijver y Leung (1997) mencionan que el estadístico más empleado y más riguroso es el coeficiente phi de Tucker o coeficiente de proporcionalidad, seguido de los coeficientes de proporcionalidad y aditividad y en esta línea, el menos estricto es el coeficiente de linealidad. En este sentido, las puntuaciones del coeficiente de congruencia indicaron que las estructuras factoriales de los tres ciclos lectivos son similares. Por lo tanto, los resultados permiten afirmar que se ha avanzado en la consecución de la equivalencia configural del cuestionario de evaluación docente.

En cuanto a la dimensionalidad del cuestionario de evaluación docente para determinar si resulta más adecuado un modelo unidimensional o tridimensional, tal cual se plantea teóricamente. Se sugiere realizar otra investigación que aporte más información al respecto, ya que los resultados del índice de parsimonia indican que el instrumento se ajusta mejor al modelo de un factor.

En resumen, los resultados de los análisis preliminares, de los coeficientes para obtener evidencias sobre la equivalencia dan garantía suficiente en términos de confiabilidad y validez del cuestionario de evaluación docente. No obstante, se deben realizar futuros estudios para seguir aportando evidencias sobre la validez de dicho instrumento.

REFERENCIAS BIBLIOGRÁFICAS

- Abrami, P.C., Cohen, P.A., and d'Apollonia, S. (1988). Implementation problems in metaanalysis. *Review of Educational Research*, 58, pp.151-179.
- Acosta Uribe, B. (2011). *Adaptación al español del inventario de autogobierno. Del alemán al español hablado en México y España*. USA: Editorial Académica Española.
- Acosta Uribe, B., Ruiz Sánchez, J. (2010). *Evaluación docente de la Universidad del Papaloapan, Campus Tuxtepec. Febrero-Julio 2010*. (Informe técnico No. FO-UNPA/008/07). Oaxaca, México: Universidad del Papaloapan.
- Acosta Uribe, B., Ruiz Sánchez, J., Vigil Rangel, M. y de la Rosa Ramírez, J.C. (2008b). *Evaluación docente de la Universidad del Papaloapan. Octubre 2007-Febrero 2008*. (Informe técnico No. FO-UNPA/008/07). Oaxaca, México: Universidad del Papaloapan.
- Acosta Uribe, B. y Ruiz Sánchez, J. (2008a). *Evaluación docente de la Universidad del Papaloapan, Campus Tuxtepec. Febrero-Julio 2007*. (Informe técnico No. FO-UNPA/008/07). Oaxaca, México: Universidad del Papaloapan.

- Agencia Nacional de Evaluación de la Calidad y Acreditación (ANECA). (2006). *DOCENTIA (Programa de apoyo para la evaluación de la actividad docente del profesorado universitario). Modelo de evaluación*. España: ANECA.
- Agencia Nacional de Evaluación de la Calidad y Acreditación (ANECA). (2007). *DOCENTIA (Programa de apoyo para la evaluación de la actividad docente del profesorado universitario). Orientaciones para la elaboración del procedimiento de evaluación*. España: ANECA.
- Agencia Nacional de Evaluación de la Calidad y Acreditación (ANECA). (2008). *Programa Academia: Acreditación Nacional para el Acceso a los Cuerpos Docentes Universitarios* (versión 17/03/08). España: ANECA.
- Agencia Nacional de Evaluación de la Calidad y Acreditación (ANECA). (2012). *DOCENTIA. Programa de apoyo para la evaluación de la actividad docente del profesorado universitario. Modelo de evaluación*. España: ANECA
- Alsawalmeh, Y.M. y Feldt, L.S. (1994). A modification of Feldt's test of the equality of two dependent alpha coefficients. *Psychometrika*, 59(1), pp. 49-57.
- Barrón Tirado, C. (2004). Reseña de La evaluación de la docencia en la Universidad. Perspectivas desde la investigación y la intervención profesional de Mario Rueda Beltrán y Frida Díaz-Barriga Arceo. *Revista Mexicana de Investigación Educativa*, 9, pp. 777-786.
- Epstein, S. (1974). The Self Concept Revisited. *American Psychologist*, 28, pp. 403-416.
- Feldt, L. S. (1980). A test of the hypothesis that Cronbach's alpha reliability coefficient is the same for two tests administered to the same sample. *Psychometrika*, 45, pp. 99-105.
- Ferrando, P.J. (1996). Evaluación de la unidimensionalidad de los ítems mediante análisis factorial. *Psicothema*, 8(2), pp.397-410.
- García Cueto, E. (1994). Coeficiente de congruencia. *Psicothema*, 6(3), pp. 465-468.
- García Garduño, J.M. (2000). Las dimensiones de le efectividad docente, validez y confiabilidad de los cuestionarios de evaluación de la docencia: síntesis de investigación internacional. En Rueda Beltrán, M. y Barriga, F. (Eds.). *Evaluación de la docencia*, pp. 41-62. México: Paidós.
- Gondar Nores J.E. (2004). *Técnicas estadísticas aplicadas a las ciencias de la salud*. Madrid. Data Mining Institute.
- Llanera de Thierry, R., (1991). Evaluación del personal académico. *Perfiles educativos*, 53-54, pp.18-29.
- Luviano Jiménez, D. (2002). Marco conceptual y metodológico para evaluar el desempeño docente en las maestrías del CENIDET 2002. Consultado el 1 de julio de 2007 en 148.208.209.4/subaca/web-dda/docs/trayectoria2002.pdf.
- Ministerio de Educación y Ciencia (MEC). (2010). Informe final. España: MEC.
- Pozo Muñoz, C., Giménez Torres, M.L. y Bretones Nieto, B. (2009). La evaluación de la calidad docente en el nuevo marco del EEES. Un estudio sobre la encuesta de opinión del Programa DOCENTIA-ANDALUCIA. *Revista de educación*, 11, pp. 43-64.
- Rico García, M.I., Montalvo Córdova, E.Y y Ayala Atrián, S.S. (2001). Utilidad de la evaluación del desempeño docente. *Revista Enferm IMSS*, 9, pp. 237-141.

- Rueda, M. (2004). La evaluación de la relación educativa en la Universidad. *Revista Electrónica de Investigación Educativa*, 6, Consultado el 1 de julio de 2007 en <http://redie.uabc.mx/vol6no2/contenido-rueda.html>.
- Rueda, M., Elizalde, L. y Torquemada, A.D. (2003). La evaluación de la docencia en las universidades mexicanas. *Revista de la Educación Superior*, 321, pp.71-77.
- Universidad de Colima. (2006). *Evaluación Docente en el Nivel Medio Superior. Febrero-Julio 2006*. Universidad de Colima.
- Valdés Velóz, H. (2000). Encuentro Iberoamericano sobre evaluación del desempeño docente. *Revista Electrónica de Organización de Estados Iberoamericanos*. Consultado el 10 de julio de 2007 en <http://www.campus-oei.org/de/rifad02.htm>.
- Van de Vijver F. y Poortinga, Y.H. (1994). Testing across cultures. In H. Ronald K. y J.Jac N. (Eds.) *Advances in educational and psychological testing*. USA: Kluwer Academic Publishers.
- Van de Vijver, F. y Leung, K. (1997). *Methods and data analysis for cross-cultural research*. USA: Sage Publications.

