

ANÁLISIS MULTIRRASGO-MULTIMÉTODO EN LA VALIDACIÓN DE INSTRUMENTOS PARA LA EVALUACIÓN DE LA CALIDAD DOCENTE EN INSTITUCIONES UNIVERSITARIAS

por José Manuel GARCÍA RAMOS
Universidad Complutense de Madrid

1. Introducción

En el contexto de la Educación Superior, la necesidad de establecer Teorías y Constructos que guíen la acción educativa no es menor que la que se precisa en otros niveles educativos, máxime cuando en nuestros días la educación superior se ha universalizado y los recursos para su mejora no son ilimitados. Conectamos así con el tema de la Calidad de las Instituciones de Educación Superior y con su correlato natural, la Evaluación de la Calidad de las Instituciones Universitarias.

No nos vamos a extender ahora en desarrollar estos temas, sobre los cuales han corrido y siguen corriendo ríos de tinta. Baste como justificación y como punto de partida la aceptación común de la necesidad de realizar estudios e investigaciones que faciliten la consolidación de un cuerpo teórico que nos permita guiar, de la mejor manera posible, nuestra intervención educativa en contextos universitarios. Por otra parte, a ello habrá que añadir la necesidad de Evaluación de los sistemas o modelos de calidad universitaria que se definan, en orden a su mejora permanente (evaluación formativa) o bien, a la clasificación de las Instituciones de Educación Superior, a la toma de decisiones Institucional Interna, a la asignación de Recursos Públicos o Privados a las Instituciones o Centros Educativos de nivel superior, ... Ambas preocupaciones, Investigación sobre... y Evaluación de... son importantes y reflejan necesidades distintas pero complementarias (García Ramos, 1997a, 1997b y 1998).

Ciertamente, habrá de definir con la mayor precisión posible estos sistemas de Calidad Educativa Universitaria. Pero habrá que hacerlo dinámicamente, probando modelos teóricos, formas o modelos de acción docente... y observando resultados, entendidos aquí no solamente como aprendizajes individuales e inmediatos, sino como Productos educativos valiosos (individuales y sociales, mediatos e inmediatos) para la sociedad en la que dicha acción educativa se inscribe (satisfacción de la demanda empresarial, cambio de actitud o de valores, calidad de los conocimientos adquiridos, capacidad para aplicar esos conocimientos adquiridos en la realidad socio-profesional-laboral, mejora de la calidad de vida general de la sociedad, etc...).

El presente trabajo de investigación, tiene como objetivo contribuir, aunque sea modestamente, a establecer modelos de acción docente eficaz en el nivel universitario. Utilizamos como instrumentos de recogida de datos y medición el CEDA y el CAD (ampliamente estudiados en otros trabajos previos ya publicados: García Ramos, 1997 a y b y García Ramos, 1998), y otros instrumentos contruidos ad hoc para el presente trabajo.

Concretamente, vimos necesario construir y aplicar dos escalas más (de diez ítems cada una, valorando de 1 a 6 puntos), en las cuales la eficacia docente del profesor era valorada también (además de por el alumno, CEDA y por el propio profesor, CAD) por los responsables académicos del Centro: Decano-Coordinador de Carrera y Director General Académico del Centro Universitario. A cada uno de ellos se le pidió que valorase analíticamente al profesor, en función de los 4 núcleos básicos del CEDA y del CAD (2 ítems por cada núcleo, paralelos a los ítems del CEDA y del CAD) y globalmente (2 ítems más).

Ciertamente, Decanos y Director, mostraban cierta resistencia a realizar valoraciones en algunos ítems, dado que les resultaba difícil apreciar aspectos internos de la clase, en cada docente y las relaciones entre profesores y alumnos. No obstante, accedieron en la medida en que se trataba de una investigación. En este sentido, Decanos y Director valoraron dichos aspectos según su percepción subjetiva indirecta, lo que necesariamente añade márgenes de error a la estimación de la eficacia docente del profesor, en sus diferentes núcleos teóricos. Cada Decano-Coordinador de carrera valoró a todos los profesores que impartían materia en su carrera, el Director valoró a todos los profesores del Centro.

Construimos así un nuevo Banco de Datos que constaba de las siguientes variables siendo la unidad de análisis el profesor (N=63).

Tabla 1: Variables utilizadas en los estudios complementarios

1. Experiencia Docente:	1. 1 a 2 años	Media obtenida por cada profesor en los ítems de cada subescala del CEDA en función de valoraciones de todos sus alumnos.	
	2. 3 a 4 años		
	2. 5 a 10 años		
	3. Más de 10 años		
2. Edad	1. Menos de 30 años		
	2. De 30 a 40 años		
	3. De 40 a 50 años		
	4. Más de 50 años		
3. Sexo	1. Varón		
	2. Mujer		
4. Carrera	1. Economía y Administración y Dirección de Empresas.		
	2. Derecho		
	3. Periodismo		
	4. Otras materias (Idiomas, Formación Humanística, Informática, etc...).		
5. CEDA I	Planificación	Medias obtenidas por cada profesor en estos ítems del CEDA, considerados criterios analíticos de las subescalas (las cuatro primeras) y globales de eficacia docente (los dos últimos).	
6. CEDA II	Aspectos Didácticos		
7. CEDA III	Aspectos Personales y Motivadores.		
8. CEDA IV	Evaluación-Exámenes.		
9. CEDA 23	dominio de Contenidos (Item 23)	Medias obtenidas por cada profesor en los ítems del CEDA, considerados criterios analíticos de las subescalas (las cuatro primeras) y globales de eficacia docente (los dos últimos).	
10. CEDA 26	Motivación de grupo (Item 26)		
11. CEDA 34	Incremento del Interés por la Materia. (Item 34)		
12. CEDA 42	Es justo al Evaluar. (Item 42)		
13. CEDA 43	Eficacia Docente Global. (Item 43)		
14. CEDA 44	Eficacia Docente Global Comparativa. (Item 44)		
15. CAD I	Planificación.	Medias obtenidas por cada profesor en los ítems de cada Subescala del CAD, según la propia valoración de sí mismo de cada profesor.	
16. CAD II	Aspectos Didácticos.		
17. CAD III	Aspectos personales motivacionales		
18. CAD IV	Evaluación-Exámenes.		
19. VALCOOR I	Planificación.	Medias obtenidas por cada profesor en los ítems de cada Subescala de los nuevos instrumentos, según la valoración del coordinador de carrera y del Director Académico del Centro.	
20. VALCOOR II	Aspectos Didácticos.		
21. VALCOOR III	Aspectos Personales y Motivacionales.		
22. VALCOOR IV	Evaluación-Exámenes.		
23. VALDIR I	Planificación.		
24. VALDIR II	Aspectos Didácticos.		
25. VALDIR III	Aspectos personales y motivacionales.		
26. VALDIR IV	Evaluación-Exámenes.		
27. Índice de No Aprobados en junio del 95.	Se calculó el porcentaje medio general de no aprobados (suspensos + no presentados) en las distintas asignaturas y grupos en que cada profesor impartía docencia.		

La fiabilidad de las distintas Sub-escalas, correspondientes a los 4 instrumentos utilizados en la medida del constructo Competencia Docente del Profesor Universitario, fueron las siguientes:

Tabla 2: Fiabilidad de las Sub-escalas de los distintos instrumentos

Instrumento (Fiabilidad Global)	Sub-escalas	Rxx	N	Observaciones
CEDA (valoración del Alumno) Rxx = 0.96	1. CEDA1 - Planificación-Organización.	.84	4109	Muy satisfactorios. indicadores de Fiabilidad de las diferentes subescalas (ver tabla 19)
	2. CEDA2 - Aspectos Didácticos	.87		
	3. CEDA3 - Aspectos Personales y Motivacionales.	.91		
	4. CEDA4 - Evaluación-Exámenes.	.86		
CAD (Auto evaluación del profesor) Rxx = 0.85	5. CAD1 - Planificación-Organización.	.76	63	. Satisfactorios indicadores de fiabilidad de las diferentes sub-escalas.
	6. CAD2 - Aspectos Didácticos.	.79		
	7. CAD3 - Aspectos Personales y Motivacionales.	.73		
	8. CAD4 - Evaluación-Exámenes.	.78		
COORD. (Valoración del Coordinador) Rxx = 0.70	9. COORD1 - Planificación-Organización.	.81	63	. Desiguales niveles de fiabilidad de las diferentes escalas. . Baja en 2-3 subescalas.
	10. COORD2 - Aspectos Didácticos.	.60		
	11. COORD3 - Aspectos Personales y Motivacionales.	.56		
	12. COORD4 - Evaluación-Exámenes.	.42		
DIR. (Valoración del Director) Rxx = 0.80	13. DIR1 - Planificación-Organización.	.68	63	Desiguales niveles de fiabilidad de las diferentes escalas.
	14. DIR2 - Aspectos Didácticos.	.58		
	13. DIR3 - Aspectos Personales y Motivacionales.	.80		
	14. DIR - Evaluación-Exámenes.	.73		

Como puede observarse, la fiabilidad de las subescalas correspondientes a COORD. y DIR. son desigualmente satisfactorias, siendo en general los índices correspondientes a la Escala COORD., relativamente más bajos.

La definición de las variables anteriores se realizó para poder responder a los siguientes objetivos:

- * Analizar las correlaciones entre indicadores de eficacia docente, correspondientes a distintas fuentes de valoración (alumno, profesor, coordinador y director).
- * Estudiar el grado de validez convergente y discriminante de los distintos instrumentos de medida y en especial del CEDA, fuente que consideramos más fiable y objetiva para aproximaciones al constructo analizado.

Se trata de estudiar, mediante *técnicas multirrasgo-multimétodo*, estas características técnicas que los instrumentos aportan. Aunque es verdad que en la validez de constructo el interés del investigador se centra

más en la propiedad o constructo que se mide, las características técnicas de los instrumentos tienen también su importancia. No debemos olvidar que la validación de constructo es un proceso complejo que requiere numerosos estudios —de la teoría y de las medidas— y que en cierto modo está siempre inacabado (Nunnally, 1978). Por lo general, no basta probar un tipo de validez para dar por válidas las inferencias realizadas a partir de un instrumento. Considerando los estudios ya realizados —en especial los realizados mediante análisis de estructuras de covarianza— presentamos en este apartado una aproximación más (complementaria de las anteriores).

La técnica Multirrasgo-Multimétodo fue ideada por Campbell y Fiske (1959) como medio para obtener índices acerca de la validez convergente, de la validez discriminante y de la fiabilidad en un diseño con múltiples rasgos y múltiples métodos. El modelo se basa en la diferencia entre rasgo como atributo, característica o propiedad mensurable y método, como forma de aproximación a la medida del rasgo (López Feal, 1986). Cada ítem de la prueba puede ser un rasgo, aunque lo más frecuente es que *los rasgos vengan definidos* —como en nuestro caso— *por un conjunto de ítems o elementos*. En este trabajo, *los métodos se definen como distintas fuentes de valoración* de una misma variable o conjunto de variables (alumnos, profesores, coordinadores, director).

Aunque Campbell y Fiske (1959) describan esta aproximación como más interesada en la adecuación de los tests como medidas de un constructo, que en la adecuación del constructo mismo, la lógica subyacente implica que si métodos distintos verifican los mismos rasgos, tendremos seguridad de que el constructo posee realidad independiente del método con el que se mida (Thorndike, 1982). El criterio general de Convergencia es fundamental y se define operativamente así (Campbell y Fiske, 1959): los coeficientes de validez convergente, esto es, las correlaciones entre medidas de un mismo rasgo, a través de métodos diversos, *deben ser significativamente distintas de cero y altas*.

Si tomamos como métodos los ítems o sub-escalas de una misma fuente (por ejemplo CEDA), *fiabilidad* —como consistencia interna— y *validez convergente* pueden llegar a ser la misma cosa (Morales, 1987; López Feal, 1986). Esto ya fue estudiado anteriormente. En este momento interesa tomar como métodos las distintas fuentes (más o menos fiables y válidas) utilizadas para apreciar las distintas dimensiones (subescalas o secciones de los tests) del constructo Competencia Docente del Profesor Universitario.

Respecto a la validez divergente, Campbell y Fiske (1967) nos explican que tal concepto se basa en el supuesto de que los tests no deben tener alta correlación entre ellos, cuando se pretende que difieran. Cada sub-

test o subescala es una unidad rasgo-método. Por ello, para examinar la validez discriminante hemos de usar varios métodos (fuentes) y más de un rasgo (cada dimensión del constructo) en nuestro caso las 4 sub-escalas (rasgos) medidas en los 4 instrumentos (fuentes): Planificación, Aspectos Didácticos, Aspectos Personales y Motivacionales y Evaluación-Exámenes, medidos por Alumnos (CEDA), Profesores (CAD), Coordinador y Director. *La validez divergente o discriminante* vendrá apoyada entonces por las *relativamente bajas correlaciones entre rasgos diferentes, medidos con el mismo método* (fuente o instrumento), lo que permite probar el carácter multidimensional del constructo —si ello se pretende— o diferenciar un constructo de otro empíricamente. También tiene utilidad para determinar el grado de validez de rasgos distintos.

2. Descripción de la muestra.

A título indicativo señalamos los siguientes resultados del análisis descriptivo.

CARRERA	fi	%
1.	24	38.1
2. Derecho	12	19.0
3. Periodismo	22	34.9
4. Otras	5	7.9
TOTAL	63	100.0

CATEGORÍA	fi	%	Doble número de profesores varones
1. Varón	42	66.7	
2. Mujer	21	33.3	

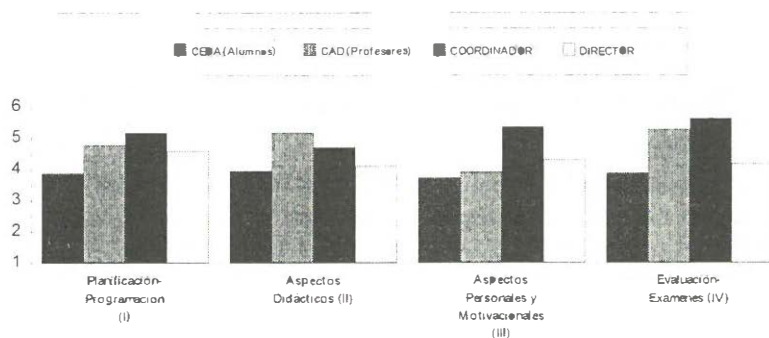
Tabla 5: Distribución de la muestra por nivel de Experiencia Docente y

Experiencia Docente (Años)	Edad		Edad	Edad		Profesorado joven mayoritariamente. Poca experiencia docente, en general. Aunque este dato no aparece aquí, un 55% son Doctores.
	fi	%		fi	%	
1. (1 a 2)	32	50.8	1. Menor de 30 años	22	34.9	
2. (3 a 4)	8	12.7	2. 30/40 años	23	36.5	
3. (5 a 10)	13	20.6	3. 40/50 años	8	12.7	
4. (más de 10)	10	15.9	4. más de 50 años	10	15.9	

La tabla siguiente, nos muestra las *medias y desviaciones típicas* obtenidas por los profesores en las distintas subescalas (rasgos) correspondientes a los distintas fuentes de valoración (métodos) de la eficacia docente.

Tabla 6: Medias y Desviaciones Típicas de los distintos rasgos, valorados por los distintas fuentes (métodos)

Rasgo ↓	Método →	FUENTES DE VALORACIÓN DE LA EFICACIA DOCENTE							
		ALUMNO (CEDA)		PROFESOR (CAD)		COORD. CARRERA		DIRECTOR	
		Media	S	Media	S	Media	S	Media	S
1. Planificación-programación.		3.895	.731	4.790	1.12	5.192	.911	4.598	1.248
2. Aspectos Didácticos.		3.957	.767	5.19	.968	4.709	1.12	4.109	.989
3. Aspectos personales motivacionales.		3.733	.811	3.92	1.39	5.37	.749	4.322	1.13
4. Evaluación-Exámenes.		3.886	.765	5.29	.96	5.639	.65	4.18	.623



El gráfico anterior nos muestra algunos datos interesantes:

- El juicio del alumno es el más duro (puntuaciones medias más bajas) en todos los núcleos teóricos o rasgos. No obstante, son satisfactorias todas las valoraciones medias, próximas al valor 4 las más bajas.
- Los juicios del profesor (autoevaluación) son más blandos (puntuaciones medias más altas), siendo todavía mejores los juicios del Coordinador. En todo caso, el profesor se autovalora siempre mucho mejor de lo que los valora el alumno.
- El juicio del Coordinador tiende a ser muy elevado, lo que indica alto grado de satisfacción de los Coordinadores-Decanos, con su profesorado. Lo mismo puede decirse del Director, aunque con valoraciones algo menos elevadas.
- Comparando rasgos (núcleos), los menos valorados son: planifi-

cación y evaluación-exámenes, aunque para el alumno y el profesor los menores niveles de satisfacción se dan en los aspectos de relación profesor-alumno. Esto es, aceptándose globalmente unos a otros, parecen no estar totalmente satisfechos de la relación que se establece entre ellos, relación que consideran mejorable.

3. Validez Convergente y Discriminante.

Aún cuando la literatura sobre dimensiones del constructo Competencia Docente del profesor muestran que las subdimensiones están, por lo general, muy relacionadas, es útil —como afirman la mayoría de los autores— estudiar la validez convergente y discriminante en la consideración del constructo estudiado.

La validez convergente puede analizarse estudiando las correlaciones mono-rasgo/hetero-método. Correlaciones significativas entre métodos —especialmente sin son altas— apoyan la validez convergente del constructo y de los instrumentos con que se mide.

La validez discriminante se estudia en las correlaciones mono-método/hetero-rasgo y especialmente en las correlaciones heterorrasgo-hetero-método. Sólo si las correlaciones son muy extremas (próximas a la unidad) debe cuestionarse la validez discriminante, especialmente si, como hemos dicho, las subdimensiones del constructo se hallan muy correlacionadas.

3.1 Estudios de Validez Convergente. (monorrasgo-heterométodo)

En función de las correlaciones obtenidas entre las distintas variables podemos concluir lo siguiente:

a Núcleo-rasgo de Planificación-Programación:

Tabla 7: Intercorrelaciones entre fuentes de Valoración del rasgo Planificación-Programación

Núcleo: Planificación.	CEDA	CAD	COORD.	DIR.
CEDA	1.00	**	**	**
CAD	.338	1.00		
COORD.	.462	.121	1.00	*
DIR.	.413	.056	.311	1.00

* Significativa a un Alpha = 0.01

** Significativa a un Alpha = 0.05

• El CEDA (Alumnos) correlaciona significativamente con otras 3 fuentes de valoración.

• No hay correlaciones entre las valoraciones del Profesor, Coordinador y Director.

• El CAD (como fuente) correlaciona levemente con la valoración del alumno (CEDA).

• Las mayores correlaciones corresponden al CEDA, con COOR. y DIR., lo que permite apoyar la validez convergente del CEDA en la medición de este rasgo (Planificación-Programación) y, en menor medida, las valoraciones del Coordinador y del Director en este rasgo.

b. Núcleo-rasgo de Aspectos Didácticos

Tabla 8: Intercorrelaciones entre fuentes de Valoración del rasgo Aspectos Didácticos

Núcleo: Aspec. Didac.	CEDA	CAD	COORD.	DIR.
CEDA	1.00		**	
CAD	.119	1.00		
COORD.	.569	.051	1.00	**
DIR.	.246	.037	.446	1.00

- * Significativa a un Alpha=0.01
- ** Significativa a un Alpha=0.05

El CEDA (Alumno) correlaciona de forma significativa y elevada con el coordinador, lo que apoya la validez convergente de ambas fuentes, en la valoración de este rasgo.

No hay correlaciones significativas del CAD (profesor) con ninguna otra fuente.

La validez Convergente de la fuente Coordinador, se ve reforzada por la correlación con la fuente Director.

c. Núcleo-rasgo de Aspectos Personales y Motivacionales

Tabla 9: Intercorrelaciones entre fuentes de Valoración del rasgo Aspectos Personales Motivacionales

Núcleo: Aspe. Pers- Motiv.	CEDA	CAD	COORD.	DIR.
CEDA	1.00	**	**	
CAD	.383	1.00		
COORD.	.477	.089	1.00	
DIR.	.128	-.034	.180	1.00

- * Significativa a un Alpha=0.01
- ** Significativa a un Alpha=0.05

El CEDA (Alumno) correlaciona de forma significativa con el CAD y con COORD., lo que apoya la validez convergente del CEDA.

El CAD, sólo correlaciona con el CEDA.

El DIR., no correlaciona con ninguna otra fuente.

d. Núcleo-rasgo de Evaluación-Exámenes

Tabla 10: Intercorrelaciones entre fuentes de Valoración del rasgo Evaluación-Exámenes

Núcleo: Evaluación- Exámenes	CEDA	CAD	COORD.	DIR.
CEDA	1.00	**		**
CAD	.490	1.00		
COORD.	.108	.041	1.00	
DIR.	.482	.225	.251	1.00

- * Significativa a un Alpha=0.01
- ** Significativa a un Alpha=0.05

El CEDA (Alumno) correlaciona de forma significativa y elevada con CAD y DIR., lo que refuerza la validez convergente del CEDA.

El CAD (profesor) correlaciona sólo con el CEDA.

El COORD., no correlaciona con ninguna otra fuente de valoración.

La conclusión básica de los estudios correlacionales heterofuente-mononúcleo (heterométodo-monorrasgo) es la consolidación de la validez

convergente del CEDA en la valoración de los diferentes núcleos teóricos valorativos de la eficacia docente del profesor. En todos los estudios mononúcleo ó monorrasgo, la validez de la valoración realizada a través del CEDA (alumno con fuente) recibe el refuerzo de correlaciones significativas y elevadas con otras fuentes: 3, en el primer núcleo (planificación), 1, en el segundo núcleo (aspectos didácticos) y 2 en los núcleos Aspectos personales y motivacionales y Evaluación-Exámenes. Todo ello apoya la validez del constructo Competencia Docente, especialmente al ser valorado a través del CEDA.

Las otras fuentes de valoración (profesor, coordinador, director) alcanzan menos apoyos (y menos importantes), a la validez convergente de sus instrumentos, como medidas alternativas del mismo constructo.

3.2. Estudios de Validez Divergente.

a) Análisis Hetero-rasgo/Mono-método.

En función de las correlaciones obtenidas entre las distintas variables podemos estudiar las correlaciones entre rasgos distintos, utilizando el mismo método o fuente.

Dado que las subdimensiones del Constructo Competencia Docente se hallan correlacionados (tal como señala la amplia literatura consultada sobre el tema y los análisis anteriormente realizados), se esperan correlaciones importantes entre los rasgos, dentro de cada fuente o método. Sólo correlaciones próximas a 1.00 cuestionarían la validez discriminante (Marsh y Hocevar, 1983; Marsh, Byrne y Shavelson, 1988; Martínez Arias, 1995). Las siguientes tablas nos muestran las correlaciones obtenidas entre rasgos, para cada método:

Tabla 11 Intercorrelación entre					. Correlaciones elevadas, no próximas a 1.00
Fuente: CEDA	Planificación	Aspectos Didácticos	Aspectos Personales y Motivacionales	Evaluación Exámenes	
Planificación	1.00	**	**	**	
Aspectos Didácticos	.783	1.00	**	**	
Aspectos Personales y Motivacionales	.669	.839	1.00	**	
Evaluación-Exámenes	.738	.813	.831	1.00	

Tabla 12: Intercorrelaciones entre rasgos del CAD

Fuente: CAD	Planificación	Aspectos Didácticos	Aspectos Personales y Motivacionales	Evaluación-Exámenes
Planificación	1.00	**	**	**
Aspectos Didácticos	.677	1.00	**	**
Aspectos Personales y Motivacionales	.424	.480	1.00	
Evaluación-Exámenes	.550	.730	.363	1.00

Correlaciones elevadas, no próximas a 1.00

Tabla 13: Intercorrelaciones entre rasgo Coordinador

Fuente: COORD.	Planificación	Aspectos Didácticos	Aspectos Personales y Motivacionales	Evaluación-Exámenes
Planificación	1.00	**	**	**
Aspectos Didácticos	.597	1.00	**	
Aspectos Personales y Motivacionales	.590	.382	1.00	
Evaluación-Exámenes	.433	.038	.024	1.00

Correlaciones moderadas, no próximas a 1.00

Tabla 14: Intercorrelaciones entre rasgos (Director)

Fuente: DIR.	Planificación	Aspectos Didácticos	Aspectos Personales y Motivacionales	Evaluación-Exámenes
Planificación	1.00	**	**	**
Aspectos Didácticos	.455	1.00	**	*
Aspectos Personales y Motivacionales	.828	.699	1.00	**
Evaluación-Exámenes	.569	.387	.656	1.00

Correlaciones moderadas, no próximas a 1.00

Tal como era de esperar, las correlaciones entre rasgos, en cada fuente, son significativas —en la mayoría de los casos— y elevadas (no próximas a la unidad) en algunos casos. Por tanto, estos resultados no parecen cuestionar gravemente la validez discriminante del constructo.

B. Análisis Heterorrasgo - Heterométrodo.

Se esperan bajas correlaciones entre rasgos distintos medidos con fuentes distintas y en todo caso, que sean sustancialmente inferiores a las correlaciones monorrasgo - heterométrodo (indicadores de validez convergente) y a las correlaciones monométrodo - heterorrasgo (indicadores también de validez divergente o discriminante).

Tabla 15: Matriz de Correlaciones Multi-Método / Multi-Rasgo

	Método 1 (CEDA)				Método 2 (CAD)				Método 3 (COORD)				Método 4 (DIR)				
	A1	B1	C1	D1	A2	B2	C2	D2	A3	B3	C3	D3	A4	B4	C4	D4	
Método Fuente 1 (CEDA)																	
A1 Planificación	(.84)																
B1 Asp. Didácticos	.783	(.87)															
C1 Asp. Pers. y Motiv.	.669	.839	(.91)														
D1 Evaluación Exámenes	.738	.813	.831	(.86)													
Método Fuente 2 (CAD)																	
A2 Planificación	.338	.130	.186	.356	(.76)												
B2 Asp. Didácticos	.244	.119	.283	.441	.667	(.79)											
C2 Asp. Pers. y Motiv.	.046	.173	.383	.306	.424	.480	(.73)										
D2 Evaluación Exámenes	.206	.150	.279	.490	.550	.730	.363	(.78)									
Método Fuente 3 (COORD)																	
A3 Planificación	.462	.130	.186	.356	.127	.227	.077	.291	(.81)								
B3 Asp. Didácticos	.346	.569	.283	.441	.192	.050	.081	.021	.597	(.60)							
C3 Asp. Pers. y Motiv.	.350	.339	.383	.306	.049	.174	.089	.212	.590	.382	(.56)						
D3 Evaluación Exámenes	.433	.221	.279	.490	.193	.103	.041	.433	.038	.024	(.42)						
Método Fuente 4 (DIR)																	
A4 Planificación	.413	.142	.007	.138	.096	.079	.205	.231	.314	.126	.012	.424	(.68)				
B4 Asp. Didácticos	.261	.246	.378	.209	.124	.037	.143	.124	.347	.446	.414	.067	.455	(.58)			
C4 Asp. Pers. y Motiv.	.423	.215	.128	.205	.198	.159	.034	.245	.385	.241	.180	.392	.828	.699	(.80)		
D4 Evaluación Exámenes	.587	.524	.389	.482	.077	.122	.121	.225	.444	.497	.275	.251	.569	.387	.656	(.73)	

La tabla anterior nos muestra todas las correlaciones e indicadores de validez convergente y discriminante:

- Entre paréntesis (), aparecen los índices de fiabilidad de las subescalas correspondientes a los distintos fuentes o métodos de medida.
- En negrita, aparecen los indicadores de Validez Convergente correspondientes a las distintas subescalas y para los distintos métodos o fuentes (correlaciones monorrasgo-heterométrodo).
- Los coeficientes monométrodo-heterorrasgo aparecen en los triángulos formados por líneas de puntos.
- El resto de los valores, son indicadores de correlación heterorrasgo-heterométrodo.

Continuando con el análisis de la validez discriminante y observando el cuadro anterior podemos ver que:

- El CEDA presenta correlaciones heterorrasgo-heterométrodo satisfactorias: bajas, no significativas en su mayoría y casi siempre inferiores a los indicadores de validez convergente. Se observa alguna irregularidad en las correlaciones del CEDA con las medidas o valoraciones de COORDINADOR y DIRECTOR; en mi

opinión se debe a la menor fiabilidad de estos instrumentos en la medición del rasgo. Por tanto, los resultados, globalmente, apoyan la validez discriminante del CEDA, en la medición del constructo estudiado.

- b. Con el CAD ocurre algo similar. Las correlaciones heterorrasgo-heterométodo son satisfactorias: bajas, no significativas en su mayoría y casi siempre inferiores a los indicadores de validez convergente (aunque en menor grado que en el CEDA). También presenta algunos valores insatisfactorios, con COORDINADOR y DIRECTOR, probablemente debidos a los defectos de precisión en la medida, realizada a través de estos instrumentos y también, en parte, por la menor fiabilidad de las sub-escalas del propio CAD.
- c. Los instrumentos de valoración que realizan estimaciones a través de COORDINADOR y DIRECTOR presentan muchos más valores no satisfactorios, lo que confirma la baja idoneidad (validez) de estos instrumentos para apreciar las diferentes dimensiones del rasgo estudiado.

4. *Análisis Factorial Multirrasgo-Multimétodo*

Otra forma de aproximarnos, más globalmente, al estudio de la validez convergente y discriminante de los instrumentos en la medida del constructo Competencia Docente consiste en realizar A.F.E. o A.F.C. (Análisis Factorial Exploratorio y/o Confirmatorio) tomando como variables de medida las 16 variables resultantes de la medida de las 4 dimensiones a través de los 4 instrumentos utilizados.

Realizamos varios A.F.E., por Componentes Principales, con rotaciones Varimax y Oblimin. La solución Varimax resultó más satisfactoria. Por otra parte y tras la solución oblicua, se observó que los factores obtenidos no correlacionaban (o lo hacían, en algún caso, mínimamente). Ofrecemos a continuación la solución rotada Varimax.

Tabla 16: Solución rotada Varimax del A.F. E. multirrasgo-multimétodo

Variables empíricas		FACTORES			
		Factor I	Factor II	Factor III	Factor IV
1. CEDA III.	Aspectos Personales y Motivacionales.	.839			
2. VALDIR II.	Aspectos Didácticos.	.820			
3. VALCOOR 2.	Aspectos Didácticos.	.749			
4. CEDA II.	Aspectos Didácticos.	.635			
5. VALCOOR 3.	Aspectos Personales y Motivacionales.	.609			
6. CEDA IV.	Evaluación-Exámenes.	.573			
7. VALDIR I.	Planificación.		.819		
8. VALDIR IV.	Evaluación-Exámenes		.770		
9. CEDA I.	Planificación.		.703		
10. VALDIR III.	Aspectos Personales y Motivacionales.		.621		
11. CAD I.	Planificación.			.828	
12. CAD II.	Aspectos Didácticos.			.905	
13. CAD III.	Aspectos Personales y Motivacionales.			.634	
14. CAD IV.	Evaluación-Exámenes			.811	
15. VALCOOR 1.	Planificación.				.772
16. VALCOOR 4.	Evaluación-Exámenes.				.887

* Sólo aparecen las saturaciones superiores al Duplo del Valor de significación ($\text{Alpha} = 0.01$).

Como puede apreciarse:

1. Aparecen 2 factores (III y IV) sólo saturados por secciones o subescalas de un mismo instrumento; son Factores-Método (CAD completo o valoraciones del profesor, y Subescalas 1 y 4 del coordinador, ambas subescalas miden los aspectos más formales de la labor docente: Planificación y Evaluación).

2. El factor II es también, fundamentalmente, un Factor-Método, en este caso juicio del Director (aspectos formales de Planificación y Evaluación, al que se añade su juicio sobre elementos de relación profesor-alumno). En este factor satura también la Subescala de Planificación del CEDA (CEDA I).

3. El Factor Principal (el I) es un Factor-Rasgo. En el saturan casi todas las variables que miden Aspectos didácticos y Aspectos Personales y Motivacionales, tal como son apreciados por las distintas fuentes: Alumno (CEDA), COORDINADOR, DIRECTOR. Es por tanto un factor que apoya claramente la validez del constructo Competencia Docente del profesor universitario, apreciando fundamentalmente aspectos tales como: Motivación de Aprendizaje, Dominio de Contenidos, Interacción de Grupo, Atención Personal al alumno, Claridad Expositiva, Entusiasmo...

Estos resultados confirman las conclusiones obtenidas en los estudios anteriores, especialmente en los A.F.E. y en los A.F.C. realizados, pudiéndose proponer la existencia de un Constructo multidimensional de Competencia Docente del Profesor Universitario. Las dimensiones de dicho constructo se hallan correlacionadas, siendo la más definitoria la relativa a la interacción entre aspectos didácticos (dominio de contenidos, claridad expositiva, orden en la presentación de contenido, etc...) y aspectos personales-motivacionales del profesor (interacción con el grupo, motivación de aprendizaje, entusiasmo, atención personal al alumno, ...).

No obstante, habrá que realizar más investigación, en contextos educativos distintos, mejorando los instrumentos o escalas de estimación de Coordinación y Director y utilizando otros procedimientos y fuentes de observación y medida, al objeto de confirmar, por otras vías, estos resultados. Por otra parte, se necesita más investigación sobre las causas de las diferencias entre las valoraciones de los estudiantes; diferencias, por ejemplo, entre materias de un mismo profesor y entre profesores de una misma materia y similar trayectoria académico-profesional. Estas diferencias, probablemente estén vinculados a características personales del docente: personalidad, estilo cognitivo, actitudes, valores,..., a la naturaleza de las propias asignaturas (más o menos analíticas, más o menos aplicadas, etc...) y a características personales de los alumnos (similares las enunciadas para el profesor).

5. *Estudios diferenciales*

Muchos otros estudios complementarios podrían realizarse para acotar aún mejor el constructo estudiado y sus dimensiones. Entre ellos destacarían los que denominaríamos estudios diferenciales, estudios cuya finalidad sería analizar el comportamiento diferencial de las dimensiones del constructo en función de las principales variables cualitativas o clasificatorias: sexo, edad, tipo de carrera, experiencia docente del profesor, tipo de materias (por ejemplo, analíticas y molares o globales), personalidad del profesor, etc...

Todo ello podría llevarnos muy lejos, en cuanto que requeriría un desarrollo amplio y la medición de algunas otras variables no utilizadas en este estudio (personalidad del profesor, por ejemplo; actitudes hacia la formación universitaria, estilo cognitivo del profesor y del alumno; etc...). Por ello no lo trataremos en este momento con extensión, aunque aportaremos algunos análisis realizados que pueden servirnos de apoyo al capítulo final de conclusiones, recomendaciones y prospectiva.

Por otro lado, también sería interesante realizar A.F.E., A.F.C. y pro-

bar los modelos estructurales causales establecidos en submuestras de la muestra global analizada. Los modelos teóricos de relaciones estructurales entre variables pueden tener un mayor o menor poder explicativo en unas carreras respecto de otras, con un tipo de materias, con un tipo de profesor (sexo, mayor o menor edad y experiencia docente, etc...). Pensamos que estos estudios pueden ser útiles para ir avanzando en el estudio y acotación de este constructo, tan ligado a la verdadera calidad de la docencia universitaria, ya que nos permitirá concreciones superiores a las establecidas en los modelos generales, utilizados en la muestra total estudiada.

5.1. Diferencias entre carreras, sexo y otras variables clasificatorias, en la valoración de la Competencia Docente del Profesor.

Tomando como V.D. de análisis la Eficacia Docente Global del Profesor, medida a través del ítem-criterio n.º 43 del CEDA, los resultados obtenidos reflejan lo siguiente.

Tabla 17: Análisis de Varianza (VD: Ítem 43 / VI: Carrera)

F.V.	SC	Gl.	MC	F.	Prob.	Difer P = 0.05
Carrera	56.11	5	11.22	5.20	.000	Si
Error	8425.7	3905	2.16			
Total	8481.8	3910				

El contraste posterior a F, para estudiar entre qué carreras se dan las diferencias significativas, aparece en la siguiente tabla:

Tabla 18: Contrastes Posteriores Diferencias entre carreras

Carrera	N	Media	S	* Significatividad p = 0.05													
				LSD							Scheffé						
				1	3	4	5	6	7	1	3	4	5	6	7		
1. Derecho	854	4.38	1.50		*				*		*				*		
3. Economía (Plan Antiguo)	193	3.88	1.42	*		*	*	*	*	*					*	*	
4. Economía (Plan Nuevo)	187	4.24	1.32		*												
5. Empresas (Plan Antiguo)	451	4.27	1.39		*												
6. Empresas (Plan Nuevo)	923	4.27	1.37		*						*						
7. Periodismo	1303	4.14	1.57	*	*												
TOTAL:	3911	4.23	1.47														

Las valoraciones más bajas fueron dadas por los estudiantes de Economía (Plan Antiguo). Las más altas correspondieron a Derecho, no dándose diferencias entre esta carrera y Economía (P. Nuevo) o Empresas (P. Antiguo y P. Nuevo). Periodismo se mantiene en un valor intermedio, no difiriendo significativamente de Economía (P. Nuevo) y Empresas (Plan

Antiguo y Plan Nuevo), pero si de Derecho (mayor media en Derecho) y de Economía (Plan Antiguo mayor media).

Si incluimos una V.I. más en el análisis (Sexo) las diferencias entre carreras se hacen más claras, al igual que aparecen diferencias por sexo, pero no interacción entre sexo y carrera respecto de la valoración de Eficacia Docente Global del Profesor Universitario (ítem 43). La tabla siguiente nos muestra los resultados del ANOVA FACTORIAL realizado.

Tabla 19: Anova Factorial de la Eficacia Docente Global Sexo Carrera

F.V.		Gl.	MC	f	Prob.	Diferencias = 0.05
Carrera	63.90	5	12.78	5.95	.000	SI
Sexo	13.83		13.83	6.44	.011	SI
Car. x Sexo	5.1		1.01	.47	.796	
Error	8312.0	3869	2.15			
Total	8410.3	3880				

El mayor nivel de exigencia o de dureza en la valoración parece corresponder a los varones con medias inferiores en todas las carreras.

Se realizaron estudios de diferencias entre Cursos y entre Años de Experiencia del Profesorado (en este ítem 43), no encontrándose diferencias significativas a un nivel de probabilidad = 0.05.

También se realizaron Análisis de Varianza Factoriales, similares al anterior, con otros ítems, que no incluimos para no extendernos en lo que sólo son datos complementarios. En la mayoría de los ítems, especialmente en aquellos que definen cada subescala, se daban resultados semejantes a los aquí referidos al ítem 43 (Competencia Docente valorada globalmente por el Alumno); esto es, diferencias entre carreras y entre sexos, pero no debidas a la interacción (efectos aditivos no interactivos).

6. Valoración de los Instrumentos utilizados en la medida del Constructo Competencia Docente del Profesor Universitario (CEDA; CAD; COORD y DIR)

De una manera telegráfica resumimos aquí las principales conclusiones obtenidas sobre los instrumentos que tratan de medir el constructo estudiado (competencia docente del profesor universitario):

a) El mejor instrumento, en cuanto a sus características y para la medición del Constructo mencionado, es el CEDA (Cuestionario de Evaluación Docente por el Alumno).

— Fiabilidad total elevada y satisfactoria (Alpha = 0.96).

Subescalas, con indicadores de Fiabilidad elevados y satisfactorios. Con sólo 5 ítems en cada subescala, Rxx varía entre 0.84 y 0.88.

Los indicadores de Homogeneidad de los ítems son muy satisfactorios, tanto respecto a la Escala Total, como a la subescala propia.

Los diferentes estudios descriptivos, correlacionales y de validación realizados nos llevan a proponer la Escala CEDA-Mod. como instrumento para la medición del constructo Competencia Docente del Profesor Universitario, con uso para evaluación formativa.

Los estudios mediante Análisis Factorial Clásico o Exploratorio apoyan la propuesta mencionada, pudiendo afirmarse que —tal como lo mide el CEDA— el Constructo Competencia Docente del Profesor Universitario es un Constructo Multidimensional, con 6 dimensiones básicas, correlacionadas entre sí.

- b) El C.A.D., Cuestionario de Autoevaluación Docente, obtiene satisfactorios niveles de fiabilidad (0.84-0.87), con un menor número de ítems que el CEDA, similar estructura factorial al CEDA y una fuerte homogeneidad interna entre sus subescalas.
- c) Los instrumentos de Medida del Constructo, utilizados a partir de las valoraciones de COORDINADOR y DIRECTOR presentan índices de fiabilidad no satisfactorios en su conjunto, lo que confirma la baja idoneidad de ambos instrumentos o fuentes de valoración para apreciar el constructo.
- d) Los estudios de Validez Convergente, utilizando los cuatro instrumentos, muestran el mayor grado de validez en las valoraciones realizadas a través del CEDA, que obtienen bastantes correlaciones significativas y elevadas mononúcleo-heterofuente (monorrasgo-heterométrico). Las otras fuentes de valoración (CAD, COORD. y DIR.), obtienen menos apoyos y de menor magnitud a la validez convergente de sus subescalas, como medidas alternativas del constructo analizado.
- e) Los estudios de Validez Divergente, realizados a partir de la matriz multirrasgo-multimétodo y de análisis heterorrasgo-monométrico permiten afirmar que el CEDA presenta satisfactorias correlaciones heterorrasgo-monométrico que apoyan la validez divergente del CEDA como medida de constructo analizado. El CAD presenta valores, o correlaciones también favorables en apoyo de su validez divergente, pero en menor grado. COORD. y DIR., presentan muchos más valores insatisfactorios de los de-

seables, lo que confirma aún más la baja idoneidad (validez) de estos instrumentos o fuentes de valoración para apreciar las diferentes dimensiones del constructo Competencia Docente.

- f) El Análisis Factorial Multirrasgo-Multimétodo presenta 2-3 factores-método (CAD., COOR. y DIR. aproximadamente), y un Factor Principal (el primero y más consistente) claramente de Rasgo; en él saturan subescalas de distintas fuentes (CEDA-Alumno; Coordinador y Director), apoyando claramente la validez del constructo Competencia Docente. Dicho factor-rasgo se halla fundamentalmente definido por aspectos tales como:

Motivación de Logro-Entusiasmo.

Dominio de los Contenidos - Claridad Expositiva.

Comunicación con el alumno (Interacción con el grupo y Atención personal).

Tales aspectos parecen nuclear la naturaleza del Constructo estudiado, como constructo multidimensional, pero fuertemente definido por la correlación-convergencia de estas conductas-actitudes-destrezas del profesor universitario.

En todo caso, los resultados del presente trabajo parecen confirmar los resultados ya obtenidos en estudios anteriores (García Ramos, 1997 a y b y García Ramos, 1998).

Dirección del autor: José Manuel García Ramos. Facultad de Educación - Dpto. M.I.D.E., Universidad Complutense de Madrid, Edificio Almudena, Paseo Rector Royo Villanova, s/n, 28040 Madrid

Fecha de la versión definitiva de este artículo: 27.III.1999

BIBLIOGRAFÍA

- ABALDE, E.; DE SALVADOR, X.; GONZÁLEZ, R. Y MUÑOZ, J. M. (1995) *Análisis de la evaluación de la docencia universitaria por los alumnos en la universidad de la Coruña (1993-94)* (Valencia. VII Seminario Nacional de Modelos de Investigación Educativa).
- ABBOT, R. D.; WULFF, D. H.; NYQUIST, J. D.; ROPP, V. A. y HESS, C. W. (1990) Satisfaction with processes of Collecting Student Opinions about instruction: the student Perspective, *Journal of Educational Psychology*, 82:2, pp. 201-206.
- ABRAMI, P. C.; COHEN, P. A. y D'APOLLONIA, S. (1990) Validity of Student Ratings of Instruction: What we know and what do not, *Journal of Educational Psychology*, 82:2, pp. 219-231.
- ACHERMAN, H. (1992) *Evaluación de la calidad por el profesorado* (Madrid, Actas del Congreso Internacional de Universidades. Universidad Complutense de Madrid).
- APODAKA, P. y OTROS. (1990) *Experiencia evaluativa en la UPV/EHU: consideraciones en torno a la puesta en marcha de un proyecto*, **revista española de pedagogía**, 186, pp. 327-336.
- ARBIZU, F. (1994) *La función docente del profesor universitario* (Bizcaia, Tesis Doctoral, Universidad del País Vasco)
- ASENSIO, I. I. (1993) *La medida del clima en instituciones de educación superior* (Madrid, Tesis Doctoral, Universidad Complutense).
- BAXTER, E. P. (1991) The TEVAL Experience, 1983-1988: the impact of student evaluation of teaching scheme on university teachers, *Studies in Higher Education*, 16, pp. 151-178.
- BENEDITO, V.; CABRERA, F.; HERNÁNDEZ, F.; MERCADE, F. y RODRÍGUEZ ESPINAR, S. (1989) La evaluación del profesor universitario, *Revista de Educación*, 290, pp. 279-291.
- BOLLEN, K. A. Y SCOTT-LONG, J. (Eds.) *Testing structural equation models* (Sage Pb. Newbury Park).
- BORREGO, A. (1992) *Evaluación institucional universitaria: el nivel académico* (Madrid, Actas del Congreso Internacional de Universidades. Universidad Complutense de Madrid).
- CAJIDE, J. (1994) Análisis factorial confirmatorio de las características de la calidad docente universitaria, *Bordón* 46:4, pp. 389-405.
- CASTEJÓN, J. L.; CARDA, R. H. y VERA, M.ª Y. (1991) *Enseñanza universitaria: diseño y evaluación. Cuestiones teóricas y estudio aproximativo* (Universidad de Alicante).
- CENTRA, J. A. (1974) The relationship between student and alumni rating of teachers, *Educational and Psychological Measurement*, 34, pp. 321-326.
- CENTRA, J. A. (1977) Student ratings of instruction and the relationship to student learning, *American Educational Research Journal*, 14, pp. 17-24.
- CENTRA, J. A. (1980) *Determining faculty performance* (San Francisco, Jossey-Bass).
- COHEN, P. A. (1981) Student ratings of instruction and student achievement: a meta-analysis of multisection validity studies, *Review of Educational Research*, 51, pp. 281-309.

- COWAN, J. (1985) Effectiveness and efficiency in higher education, *Higher Education*, 14, pp. 235-239.
- DALHLLÖF, W. (1990a) *Evaluation de l'enseignement: pratiques et développement. Versus un nouveau modèle* (París, OCDE: Centre pour la recherche et l'innovation dans l'enseignement).
- DALHLLÖF, W. (1990b) *Practice and evidence in the evaluation of teaching* (París, Ed. por DALHLLÖF et al. en *Report of the IMHE study group on Evaluation in Higher Education*. OCDE, June 1990 cap. 5, pp. 139-161),
- DALHLLÖF, W. (1990c) *Towards a new model for the evaluation of teaching: an interactive process-centred approach* (París, Ed. por DAHLLÖF et al, en *Report of the IMHE study group on Evaluation in Higher Education*. OCDE, June, 1990, cap. 6, pp. 163-215).
- DUNKIN, M. Y BARNES, J. (1986) *Research on higher education* (New York, en Wittrock, P.: *Handbook of research on teaching*, McMillan, pp. 754-777).
- ENTWISTLE, N. J. y RAMSDEN (1983) *Understanding student learning* (London, Croom Helm).
- ENTWISTLE, N. J. Y TAIT, H. (1990) Approaches to learning, evaluations of teaching, preferences for contrasting academic environments, *Higher Education*, 19, pp. 169-194.
- ESCUDERO ESCORZA, T. (1989) Aproximación pragmática a la evaluación de la universidad, *Revista de Investigación Educativa*, 7:13, pp. 93-112.
- ESCUDERO ESCORZA, T. (1993) *Enfoques modélicos en la evaluación de la enseñanza universitaria* (Universidad de Gran Canaria, III jornadas nacionales de Didáctica Universitaria. Universidad de Gran Canaria, pp. 3-59).
- FERNÁNDEZ DÍAZ, M. J. (1986) *Investigación evaluativa de instituciones universitarias* (Madrid. Tesis Doctoral. Universidad Complutense).
- FERNÁNDEZ, J. (1995) Metaevaluación del profesorado de la calidad docente, *Gaceta Complutense*, 111, pp. 16-19.
- FERNÁNDEZ, M. R.; JORNET, J. M.; PÉREZ CARBONELL, A. Y GONZÁLEZ SUCH, J. (1995) *Estudio del cuestionario de evaluación de la docencia a partir de las opiniones de los estudiantes: I y II cuestionario planteado; estudio metodológico y estudio factorial* (Valencia, Comunicaciones al VII seminario de Modelos de Investigación Educativa, pp. 300 a 308).
- FRANKE-WIKBERG, S. (1990) Evaluating education quality on the instructional level, *Higher Education Management*, 2:3, pp. 271-292.
- FREY, P. W.; LEONARD, D. W. Y BEATTY, W. N. (1975) Student ratings of instruction: validation research American, *Educational Research Journal*, 12, pp. 327-336.
- GARCÍA GARRIDO, J. L. (1992) *El profesorado universitario: un análisis comparativo de los modelos de formación* (Madrid, Actas del Congreso Internacional de Universidades. Universidad Complutense).
- GARCÍA RAMOS, J. M. (1986) Validez de constructo en el ámbito pedagógico, *revista española de pedagogía*, 174, pp. 535-554.

- GARCÍA RAMOS, J. M. (1987) El análisis factorial confirmatorio aplicado a la investigación pedagógica no-experimental, *Bordón*, 267, pp. 245-267.
- GARCÍA RAMOS, J. M. (1988) Modelos exploratorios y confirmatorios en la investigación pedagógica no experimental, *Revista Ciencias de la Educación*, 136, pp. 423-450.
- GARCÍA RAMOS, J. M. (1989) Evaluación de la eficacia docente en PÉREZ JUSTE, R. Y GARCÍA RAMOS, J. M. *Diagnóstico. Evaluación y Toma de Decisiones*, Cap. 22. (Madrid, Rialp)
- GARCÍA RAMOS, J. M. (1991) Recursos metodológicos en la evaluación de programas, *Bordón*, 43:4, pp. 461-476.
- GARCÍA RAMOS, J. M.; SÁNCHEZ, F.; ALBERT, C.; SALORD, M. (1993) *Proyecto Educativo del Centro Universitario Francisco de Vitoria* (Madrid, Publicación del Centro Universitario Francisco de Vitoria).
- GARCÍA RAMOS, J. M. (1995) La Evaluación Institucional a través del departamento de Recursos Humanos de una Institución Universitaria Privada, *Bordón*, 47:1, pp. 17-30.
- GARCÍA RAMOS, J. M. Y CONGOSTO, E. (1995) *Un sistema de evaluación institucional en la universidad* (Valencia, Comunicación al VII Seminario Nacional del Modelos de Investigación Educativa, pp. 264-267).
- GARCÍA RAMOS, J. M.; FERNÁNDEZ DÍAZ, M. J. Y CONGOSTO LUNA, E. (1995) *La evaluación de un modelo de tutoría académica en la universidad* (Valencia, Comunicación al VII Seminario de Modelos de Investigación Educativa, pp. 268-271).
- GARCÍA RAMOS, J. M. (1997a) Valoración de la Competencia Docente del Profesor Universitario. Una Aproximación Empírica, *Revista Complutense de Educación*, 7:3.
- GARCÍA RAMOS, J. M. (1997b) Análisis factorial confirmatorio a la Valoración del Constructo Competencia Docente del Profesor Universitario, *Bordón*, 49:4, pp. 361-391.
- GARCÍA RAMOS, J. M. (1998) Análisis de Estructuras de Covarianza en el estudio de la Competencia Docente del profesor Universitario, *Revista de Investigación Educativa*, 16:1, pp. 155-184.
- HELMÁNTICA, (GRUPO de la Universidad de Salamanca) (1995) *Elaboración de un instrumento de evaluación de las actitudes de los universitarios hacia la universidad, los alumnos y los profesores* (Valencia, Comunicación al VII Seminario Nacional de Modelos de Investigación Educativa, pp. 313-318).
- HOWARD, R. D.; NICHOLS, J. O. Y GRACIE, L. W. (1987) *Institutional research support of the self study*, en MUFFO, J. A. y McLAUGHLIN, G. W. (1987) *A primer on Institutional research* (Florida University Press).
- JORDAN, T. E. (1989) *Measurement and evaluation in higher education* (London. Falmer Press).
- JÖRESKOG, K. G. (1974) Analyzing Psychological Data by structural analysis of Covariance Matrices, en D. H. KRAHTS, R. C. Y OTROS (ed.) *Contemporary Developments in Mathematical Psychology*. (vol. 2).

- JÖRESKOG, K. G. (1977) Structural equation model in the social sciences: specification, estimation and testing, en KRISHANAIAH, P. R. (ed.) *Applications of statistics* (Amsterdam, North-Holland).
- JÖRESKOG, K. G. Y SORBÖM, D. (1976) *C●FAMM. Confirmatory factor analysis with model modification, User's Guide*, (Chicago, Nacional Educational Resources).
- JÖRESKOG, K. G. Y SORBÖM, D. (1981) *LISREL V: Analysis of linear structural relationship by method of maximum likelihood: user's Guide*, (Chicago, Internacional Educational Services).
- JÖRESKOG, K. G. Y SORBÖM, D. (1983) *LISREL VI: Analysis of linear structural relationship by maximum likelihood and least square methods* (Chicago, International Educational services)
- JÖRESKOG, K. G. Y SORBÖM, D. (1988) *PRELIS: A program for multivariate data screening and data summarization: A preprocessor for LISREL*, (Suecia, Univ. of Uppsala) (2.ª Ed.).
- JÖRESKOG, K. G. Y SORBÖM, D. (1989) *LISREL 7: A Guide to the Program and applications. Jöreskog-Sorböm* (2.ª Ed.).
- JÖRESKOG, K. G. Y SORBÖM, D. (1991) *LISREL 7.20. A guide to the program and applications. Jöreskog-Sorböm-SPSS*.
- JORNET, J. M. Y SUÁREZ (1988) *Evaluación de la docencia en Facultades de la Universidad de Valencia, a partir de las opiniones de los estudiantes* (Valencia, Informes de investigación 3, 4, 5 y 6. Universidad de Valencia).
- JORNET, J. M. Y SUÁREZ, J. M. Y GONZÁLEZ SUCH, J. (1989) Estudio de validación de un cuestionario de valoración de la docencia universitaria para estudiantes en un conjunto homogéneo de centros de la universidad de Valencia, *Revista de Investigación Educativa*, 7:13, pp. 57-92.
- JORNET, J. M.; GONZÁLEZ SUCH, J.; SUÁREZ, J. M. Y PÉREZ CARBONELL, A. (1995) *Análisis de la consistencia de cuestionarios de opinión de estudiantes para evaluaciones docentes* (Valencia. Comunicación en VII Seminario Nacional de Modelos de Investigación Educativa, pp. 260-263).
- JORNET, J. M. (1993) *Enfoques de la evaluación universitaria* (Universidad de Gran Canaria, III Jornadas Nacionales de Didáctica Universitaria, pp. 63-83).
- L'HOMMEDIEU, R.; MENGES, R. J. Y BRINKO, K. T. (1988) *The effects of student ratings feedback to college teachers: a meta-analysis and review of research* (Evanston, Center for the teaching professions. Northwestern University).
- L'HOMMEDIEU, R.; MENGES, R. J. Y BRINKO, K. T. (1990) Methodological explanations for the modest effects of feedback from student ratings, *Journal of Educational Psychology*, 83:2, pp. 232-241.
- MARLIN, J.W. (1987) Student perception of end-of-course evaluation, *Journal of Higher Education*, 58, 6.
- MARSH, H. W. Y OVERALL, J. V. (1981) The relative influence of course level, course type, and instructor on students evaluations of college teaching, *American Educational Research Journal*, 18, pp. 103-112.
- MARSH, H. W. Y HOCEVAR, D. W. (1983) Confirmatory factor analysis of Multitrait-Multimethod matrices, *Journal of Educational Measurement*, 20, pp. 231-248.

- MARSH, H. W. (1982) SEEQ: a reliable, valid and useful instrument for collecting students evaluation of university teaching, *British Journal of Educational Psychology*, 52, pp. 485-497.
- MARSH, H. W. (1983) Multidimensional ratings of teaching effectiveness by student from different academic setting and their relation to student / course / instructor characteristics, *Journal of Educational Psychology*, 75, pp. 750-766.
- MARSH, H. W. (1984) Students evaluations of university teaching: dimensionality, reliability, validity, potential biases and utility. *Journal of Educational Psychology*, 76, pp. 707-754.
- MARSH, H. W. (1987) Students evaluations of university teaching. Research findings methodological issues, and directions for future research. *Internacional Journal of Educational Research*, 11.
- MARSH, H. W. (1993) The use of students evaluations and a individually structure intervention to enhance university teaching effectiveness. *Educational Research Journal*, 30:1, pp. 217-251.
- MARTÍNEZ ARIAS, R. (1995) *Psicometría: Teoría de los tests psicométricos y educativos* (Madrid, Síntesis).
- MATEO, J. (1989) *La evaluación del profesorado universitario. Algunas reflexiones en torno al estado de la cuestión* (Universidad de Valencia, en informes de investigación evaluativa n.º 1. Consideraciones metodológicas sobre la evaluación y mejora de la docencia universitaria, pp. 13-30).
- MATEO, M. A. Y FERNÁNDEZ, J. (1992) Análisis confirmatorio de la estructura dimensional de un cuestionario para la evaluación de la calidad de la enseñanza. *Revista de Investigaciones Psicológicas*, 11 pp. 73-82.
- MCKEACHIE, W. J. (1990) Research on College Teaching: the historical background. *Journal of Educational Psychology*, 80:2, pp. 189-200.
- MIGUEL, M. de (1988) *Modelos de investigación sobre organizaciones educativas. IV Seminario de Modelos de Investigación Educativa* (Santiago de Compostela, Publicado en R.I.E. (1989), pp. 21-56).
- MIGUEL, M. de (1990) *Indicadores de calidad en la docencia universitaria* (Cádiz, Congreso sobre calidad de la Educación Universitaria).
- MIGUEL, M. de (1991) *Utilización de indicadores en la evaluación docente universitaria*, en MIGUEL, M. DE; MORA, J.G. Y RODRÍGUEZ ESPINAR, S. (1991). *La evaluación de las Instituciones Universitarias* (Madrid, Secretaría General del Consejo de Universidades).
- MILLER, R. I. (1987) *Evaluating faculty for promotion and tenure* (San Francisco, Jossey Bass).
- MILLMAN, J. (ed.) (1981) *Handbook of teacher evaluation* (Beverly Hills, Sage).
- MORA, J. G. (1991) *Calidad y rendimiento en las instituciones universitarias* (Madrid, Consejo de Universidades).
- MULAİK, S. A. Y OTROS (1989) An evaluation of goodness of fit indices for structural equation models, *Psychological Bulletin*, 10, pp. 430-445.
- MURRAY, H.G. (1984) The impact of formative and summative evaluation of teaching

- in North American Universities. *Assesment and Evaluation in Higher Education*, 9, pp. 117-132.
- NEAVE, G. (1992) *Hacia el gusto por la calidad: evaluación y universidad de masas* (Madrid, Congreso Internacional de Universidades. Universidad Complutense).
- ORDEN, A. de la (1985) Modelos de evaluación universitaria, **revista española de pedagogía**, pp. 169-170 y 521-537.
- ORDEN, A. de la (1988) Calidad de los Centros Educativos. Asunto para un Congreso, *Bordón*, 40:2.
- ORDEN, A. de la (1992) *Calidad y Evaluación de la Enseñanza Universitaria* (Madrid, Actas del Congreso Internacional de Universidades. Universidad Complutense).
- OSORO, J. M. (1995) *Los indicadores de rendimiento en la evaluación institucional universitaria* (Zaragoza, ICE. Universidad de Zaragoza).
- PÉREZ JUSTE, R. (1986) Análisis y valoración del modelo español de educación superior a distancia, *Revista de Investigación Educativa*, 4:7.
- PÉREZ JUSTE, R. Y GARCÍA RAMOS, J. M. (1989) *Diagnóstico, Evaluación y Toma de decisiones* (Madrid, Rialp).
- RAMSDEN, P. (1991) A performance indicator of teaching quality in higher education. The experience questionnaire, *Studies in Higher Education*, 16, pp. 129-150.
- RAMSDEN, P. Y ENTWISTLE, N. J. (1981) Effects of academic departament on students approaches to studying, *British Journal of Educational Psychology*, 51, pp. 368-383.
- RODRÍGUEZ DIÉGUEZ, J. L. (1992) *Evaluación y formación del profesorado universitario: el caso español* (Madrid, Actas del Congreso Internacional de Universidades. Universidad Complutense).
- RODRÍGUEZ ESPINAR, S. (1986) *Predicción y explicación del rendimiento relativo en función de los factores de personalidad del H.S.P.Q.* (Barcelona, Investigación presentada al Concurso para la Cátedra de la Universidad de Barcelona).
- RODRÍGUEZ ESPINAR, S. (1989) *La evaluación del profesorado universitario. Análisis de la opinión del profesorado de la universidad de Barcelona* (Valencia, en informes de Investigación Evaluativa n.º 1. Consideraciones metodológicas sobre la evaluación y mejora de la docencia universitaria. Universidad de Valencia, pp. 211-235).
- RODRÍGUEZ ESPINAR, S. (1990) *Evaluación e innovación universitaria. ¿Por qué y para qué?* (Barcelona, en I. Coloquio Inernacional: La Pedagogía Universitaria. Un reto en la Enseñanza Superior. Universidad de Barcelona).
- RODRÍGUEZ ESPINAR, S. (1991a) (Cádiz, *Dimensiones de la calidad universitaria*. Ponencia presentada en el I Congreso Internacional sobre la calidad de la educación universitaria. 6-8 de marzo).
- RODRÍGUEZ ESPINAR, S. (1991b) *Calidad universitaria: un enfoque institucional y multidimensional*. En: Miguel, M. de; Mora, J. G. y Rodríguez Espinar, S. (1991). *La Evaluación de las Instituciones Universitarias* (Madrid, Secretaría General del Consejo de Universidades).
- RODRÍGUEZ ESPINAR, S. (1993) *Experiencias españolas de evaluación de la enseñanza universitaria y nuevas perspectivas. Otro punto de vista* (Las III Jornadas Nacionales de Didáctica Universitaria. Universidad de Las Palmas de Gran Canaria, pp. 111-132).

- RODRÍGUEZ ESPINAR, S. (1996) *Evaluación institucional y planificación universitaria*. Ponencia en I (Madrid, Congreso Internacional de Evaluación de la Calidad de los Servicios en la Empresa, febrero 1996).
- SALVADOR, L. Y SANZ, J. J. (1988) *Evaluación de la docencia 1987-88. Opiniones del Alumno* (Cantabria, Vicerrectorado de Ordenación Académica. Universidad de Cantabria).
- SÁNCHEZ CARRIÓN, J. J. (1984) *Introducción a las técnicas de análisis multivariante aplicadas a las CC. Sociales* (Madrid, C.I.S).
- SCOTT LONG, J. (1987) *Confirmatory factor analysis*, Sage University. Paper. 33. Beverly Hills. California.
- SCOTT LONG, J. (1987) *Covariance Structure Models*, Sage University Paper, 34.
- TANAKA, J. S. (1993) *Multifaceted conceptions of fit in structural equation models*, en BULLEN, K. A. Y SCOTT-LONG, J. (ed.) *Testing structural equation models*. (Sage Pb. News Bury Park).
- TEJEDOR, F. J. (1985) Problemática de la enseñanza universitaria, *Revista de Investigación Educativa*, 3:6, pp. 322-337.
- TEJEDOR, F. J., CASTRO, C. Y MÍNGUEZ, C. (1988). Evaluación del profesorado universitario por los alumnos, *Studia Pedagogica*, 20, pp. 73-134.
- TEJEDOR, F. J. (1990) La evaluación del profesorado en la universidad de Santiago, **revista española de pedagogía**, 186, pp. 337-362.
- TEJEDOR, F. J. Y MONTERO, L. (1990) Indicadores de la calidad docente para la evaluación del profesor universitario, **revista española de pedagogía**, 186, pp. 260-279.
- TEJEDOR, F. J. (1993) *Experiencias españolas de evaluación de la enseñanza universitaria y nuevas perspectivas* (Gran Canaria, III Jornadas Nacionales de Didáctica Universitaria. Universidad de las Palmas de Gran Canaria, pp. 85-109).
- TOURÓN, J. (1989) *La validación de constructo: su aplicación al CEED*, (cuestionario de evaluación de la eficacia docente), *Bordón*, pp. 735-756.
- WESTERHEIJDEN, D. F. (1991) *La evaluación de la universidad y su contexto político: gestión de la calidad y toma de decisiones en la educación superior*, en MIGUEL, M. DE; MORA, J. G. Y RODRÍGUEZ, S. (1991). Evaluación de las Instituciones Universitarias. (Madrid, Secretaria General del Consejo de Universidades)
- WHITELEY, S. Y DAYLE, K.O. (1976) Implicit theories in student rating, *American Educational Research Journal*, 13, pp. 241-253.
- VILLA, A. y MORALES, P. (1993) *La evaluación del profesor. Una visión de los principales problemas y enfoques en diversos contextos*. (N.º 14 de Estudios y Documentos. Departamento de Educación, Universidades e Investigación. Gobierno Vasco. Vitoria).
- WOLF, R.L. (1992) *La realidad actual de la universidad: un análisis de la calidad de la educación universitaria*. (Actas del Congreso Internacional de Universidades. Universidad Complutense Madrid).

SUMMARY: MULTIFACE AND MULTIMETHOD ANALYSIS IN THE VALIDATION OF THE INSTRUMENTS FOR THE EVALUATION OF TEACHING QUALITY IN UNIVERSITIES.

The present article presents the principal results of a study made by using different statistical multivariate methods, principally studies multiface-multimethod and studies of convergent validity and discriminant, in the validation of a model of Formative Evaluation of University Professor core, with samples of different majors, using multiple instruments of data recruiting (Questionnaires of students, professors and major directors).

KEY WORDS: Evaluation of University Quality, Construct Validity, Validation of measure instruments.