

## EL ANÁLISIS FACTORIAL COMO HERRAMIENTA PARA LA INFERENCIA DE MODELOS DE EVALUACIÓN DEL APRENDIZAJE UNIVERSITARIO, DESDE LA OPINIÓN DEL ALUMNADO

*Clemente Rodríguez Sabiote<sup>1</sup>  
José Gutiérrez Pérez<sup>2</sup>  
Juana Rienda Díaz  
Ana M<sup>a</sup> Díaz León*

### RESUMEN

La evaluación del aprendizaje en el nivel universitario ha sido una línea de investigación en el que no se ha indagado excesivamente. Esa razón junto con otras de mayor calado, nos condujeron a la deducción de modelos de evaluación del aprendizaje universitario, que en opinión de los estudiantes deberían ser practicados. Para esta labor utilizamos una escala de actitudes, donde fueron propuestas diez formas de evaluación. A los datos obtenidos les fue aplicado un análisis factorial, del cual deducimos determinados componentes, que identificamos con diferentes modelos de evaluación del aprendizaje.

### ABSTRACT

The assessment of the learning in the university level has been a research line that hasn't been enough investigated. That reason, with other reasons of greater concern they guide us to deducing the assessment models of the university learning, that in opinion of the students would have to be practiced. For this labour, we use an attitudes scale where were proposed ten forms of assessment. To the obtained data was applied them a factorial analysis, of wich we deduce determined components, that we identify with different assessment models of the learning.

### 1. INTRODUCCIÓN

Inspirados en la convicción de que la evaluación del aprendizaje es el elemento clave en el conjunto de procesos que conforman la enseñanza-aprendizaje, y por tanto es un elemento modulador en la calidad de la enseñanza universitaria (De la Orden, 1969 y 1981), decidimos emprender el presente estudio. En el pretendemos mostrar los posibles modelos de evaluación del aprendizaje que pueden ser inferidos a partir de la opinión del alumnado, utilizando como instrumento de apoyo la técnica estadística multivariada del Análisis Factorial.

Si bien es verdad que el resto de elementos del proceso de enseñanza-aprendizaje (objetivos, contenidos, metodologías...) determinan en gran medida la naturaleza y alcance del modelo evaluativo, no deja de ser menos cierto, que a su vez este último condiciona al resto. Bajo esta premisa podemos entonces situar la evaluación en una posición de privilegio dentro del proceso instructivo, dotándola de un poder de modulación que aparentemente no parecen tener los demás elementos. Baste recordar a este respecto, los trabajos de Miller y Parlett (1974), Dearden (1976), Marton y Saljo (1976), Elton y Lanvillard (1979), Fuchs y otros

---

<sup>1</sup> Dpto. de Pedagogía (Área M.I.D.E.). Universidad de Granada

<sup>2</sup> Idem

(1984) o los más recientes de De la Orden (1993) y Abraira (1995) en el contexto español. Todos estos trabajos consolidan la idea de que existe un influjo predominante de los modelos de evaluación sobre los elementos del proceso de enseñanza-aprendizaje. Así, no es de extrañar que sea un hecho constatado que por ejemplo el alumno estudie para aprobar, o estudie de diferente forma, según la prueba a que tenga que hacer frente, o que el profesor planifique y desarrolle la enseñanza en función de los criterios de las pruebas de evaluación que usualmente emplee. Debemos tener en cuenta, si mantenemos la hipótesis anterior, que los criterios y modos de evaluación sean coherentes con los objetivos formalmente establecidos, o sobre los que existe un consenso implícito, y a través de ellos con los fines generales de la educación (dimensión teleológica) y el sistema de valores del que derivan (dimensión axiológica), puesto que si no lo son, la evaluación constituye uno de los mayores obstáculos para una educación de calidad.

## 2. OBJETIVOS

El objetivo fundamental de la presente investigación ha sido tratar de inferir diversos modelos de evaluación del aprendizaje. Dicha inferencia se ha hecho a partir de la opinión que los alumnos de la Facultad de Ciencias de la Educación de la Universidad de Granada han mostrado mediante una escala likert, donde se le proponían diversas formas de evaluación. Otro objetivo no menos importante, ha sido también tratar de comprobar si existía desacuerdo entre los modelos de evaluación inferidos desde la opinión estudiantil y los que dicen practicar los profesores.

## 3. METODOLOGÍA

### 3.1. MUESTRA

Las características específicas de los sujetos participantes en el estudio (sujetos pertenecientes a diferentes especialidades y cursos), así como la finalidad del mismo (ánimo de extrapolar los resultados a la población de referencia), hicieron que para la selección de la muestra optásemos por el muestreo probabilístico estratificado proporcional según especialidad y curso (2 estratos). Esta modalidad de muestreo es la que genera un menor error muestral, y en consecuencia una mayor representatividad de la muestra (Cea D'Ancona, 1996:174). El número de sujetos a los que les fue administrado el instrumento asciende a 1300 (aproximadamente 1/3 de la población de alumnos de la Facultad de Educación de la Universidad de Granada, curso 95-96), siendo sus características y composición, las que a continuación detallamos:

VARIABLE	NIVELES	NÚMERO DE SUJETOS
SEXO	HOMBRE	373
	MUJER	927
	PRIMERO	566
	SEGUNDO	319

CURSO	TERCERO	334
	CUARTO	52
	QUINTO	29
ESPECIALIDAD	PEDAGOGÍA	291
	PSICOPEDAGOGÍA	26
	E. INFANTIL	136
	E.MUSICAL	78
	E.PRIMARIA	192
	E. FÍSICA	210
	LENGUA EXTRANJERA	75
	AUDICIÓN Y LENGUAJE	21
	E. ESPECIAL	55
	PREESCOLAR	84
	FILOLOGÍAS	47
	CIENCIAS	59
	HUMANAS	26

TABLA 1. Características y Composición de la Muestra

### 3.2. INSTRUMENTO

El instrumento utilizado en la recogida de datos fue una escala likert de actitudes. En dicha escala los alumnos expresaron su opinión sobre diversas formas de evaluación del aprendizaje (hasta un total de diez) que les fueron propuestas. Constaba pues, el instrumento de una primera parte donde se podía constatar el sexo, especialidad y curso del encuestado/a, y una segunda parte donde se le presentaban la colección de formas de evaluación. El alumno/a expresó su opinión en una escala de 1 a 5, en la que el 1 era muy en desacuerdo... y el 5 muy de acuerdo.

	1	2	3	4	5
1. TRABAJOS					
2. LECTURA-RESUMEN DE PUBLICACIONES					
3. ENSAYOS					
4. AUTOEVALUACIÓN					

5. OBSERVACIONES DEL PROFESOR					
6. CUESTIONARIOS					
7. ENTREVISTAS					
8. EXÁMENES					
9. PRÁCTICAS					
10. PARTICIPACIÓN Y ASISTENCIA					

TABLA 2. Escala de Actitud hacia diversas formas de evaluación del aprendizaje

### 3.3. PARÁMETROS DE FIABILIDAD Y VALIDEZ DEL INSTRUMENTO

Quizás con demasiada frecuencia elaboramos informes de investigación en los que se pasan por alto aspectos tan importantes como la fiabilidad y validez del instrumento utilizado en el estudio. Por ello y para emitir un informe de investigación lo más completo posible presentamos a continuación los apartados de fiabilidad y validez de nuestra escala.

#### FIABILIDAD

Para la estimación de la fiabilidad del instrumento, hemos empleado tres procedimientos diferentes basados en la estimación de la consistencia interna del mismo, dado que sólo disponemos de una aplicación. Los procedimientos utilizados se basan en la corrección de la correlación momento-producto de Pearson ( $r_{xy}$ ) y para su cálculo hemos dividido la prueba en dos mitades (items pares vs. impares). Los procedimientos, su ecuación y los resultados obtenidos en cada uno de ellos son los siguientes:

Procedimiento	Spearman-Brown	Rulon	Guttman
Ecuación	$R_{xx} = 2r_{xx}/1+r_{xx}$	$r_{xx} = 1 - S^2d/S^2t$	$r_{xx} = 2(1 - S^2i + S^2p/S^2t)$
Resultado	6	6	66

TABLA 3. Procedimientos de Fiabilidad del Instrumento

Debemos incluir también la verificación del supuesto de homocedasticidad, como condición previa al cálculo del procedimiento Spearman-Brown. Dicho supuesto ha sido comprobado mediante la *Prueba de Hartley*, habiéndose obtenido para los dos casos posibles (niveles de significación: 0.05 y 0.01) valores que inducen a pensar que las varianzas de ambas mitades son homogéneas y por tanto el procedimiento viable. A continuación detallamos algunos datos de la prueba de verificación.

F.empírica :  $S^2_{mayor}/S^2_{menor} = 1.21$   
 F.critica:  $gl(1,1299)$  con un  $\alpha 0.01 \Rightarrow 6.63 / \alpha 0.05 \Rightarrow 3.84$   
 Como  $F_e = 1.21 < \text{que } 6.63 \text{ y } 3.84$  entonces aceptamos  $H_0$   
 es decir  $S^2_i = S^2_p \exists$  Homocedasticidad.

En conclusión, el análisis de fiabilidad de nuestro instrumento ofrece unos resultados medios. Se puede considerar que se ha conseguido una fiabilidad aceptable dadas las características del formato de la prueba (sólo diez ítems).

**VALIDEZ**

Validez de Constructo: Tal validez ha sido estimada a partir del análisis factorial a que hemos sometido los resultados de la prueba. Más adelante, el lector puede consultar los resultados de dicho análisis, así como la estructura del constructo evaluación de los aprendizajes.

Validez de Contenido: Para asegurar este tipo de validez, se sometieron a un análisis de contenido una amplia muestra de programas de asignaturas de la propia Facultad objeto del estudio (112 programas). Los resultados de tal análisis dieron como resultado, la presencia de determinadas unidades de registro, (Bardin, 1986) y (Krippendorff, 1990) (exactamente 10), que coinciden con cada uno de los ítems de la escala.

**4. ANÁLISIS DE DATOS**

Para el análisis de datos hemos utilizado el paquete estadístico SPSS V. 6.0

**ANÁLISIS FACTORIAL**

Para tratar de inferir modelos de evaluación del aprendizaje a partir de la opinión del alumnado y dotar de validez de constructo a nuestro instrumento, hemos recurrido al análisis factorial. El procedimiento específico utilizado ha sido el método de componentes principales, siguiendo el criterio de Kaiser para la determinación del número de factores (supresión de aquellos factores con  $\lambda < 1$  ó  $\%S^2 < 10\%$ ), utilizando rotación ortogonal VARIMAX.

Previo al comentario de la solución factorial resultante, creemos adecuado hacer mención de algunos aspectos acerca de la matriz de correlaciones obtenida:

1º El determinante de la matriz de correlaciones es de .54

2º El valor obtenido en el test de esfericidad de Bartlett es de  $\chi^2 = 784.47$ , habiéndose obtenido un valor  $p = .000$ , con lo cual rechazamos la  $H_0$  (Hipótesis nula), lo que supone aceptar que la matriz de correlaciones no es una matriz de identidad, sino que existe intercorrelación significativa dado que el valor hallado en la prueba de esfericidad es alto.

3º Los valores obtenidos en la  $H^2$  (comunalidad) son en general altos, lo cual indica que las variables que conforman la solución factorial están bien representadas.

TABLA 4. Resultados del Análisis Factorial de Matriz rotada de componentes principales, método VARIMAX.

Valor mínimo saturación en cada factor  $a \geq 0.30$ . Punto de corte  $\lambda \geq 1$ .  $\% S^2 \geq 10\%$

Ítems	Factor I	Factor II	Factor III	Factor IV	H <sup>2</sup>
Trabajos	37	-63	22	15	61
Lectura-Resumen Publicac.	49	-55	-2	8	52
Ensayos	46	-28	-22	5	55
Autoevaluación	31	11	-6	-75	34

Observac. Del Profesor	43	51	12	-18	68
Cuestionario	49	32	-38	34	50
Entrevista	54	27	-45	20	61
Examen	-6	34	45	51	61
Prácticas	43	14	49	3	59
Participación y Asistencia	48	6	51	-14	46
LAMBDA	185	139	119	107	$\sum \lambda = 5.5$
% Varianza explicada por cada factor	186	139	119	108	$\sum \%S^2 = 55.2$

#### FACTOR I

El primer componente explica el 18.6% de la varianza. Es un factor unipolar en el que saturan todos los ítems (formas de evaluación), excepto el examen. Por eso siguiendo la terminología empleada en anteriores trabajos (Rodríguez y Gutiérrez, 1997 y Rodríguez, 1996) hemos creído conveniente denominar a este factor como un Modelo de evaluación Crédito-formativo *Mixto*. Saturan pues, aquellas formas de evaluación que pudieran contribuir al desarrollo de una evaluación más formativa que sumativa en la que se combinan formas que dependen de los criterios del profesor (p.ej. observaciones del profesor, cuestionarios...), con otras que dependen más del desempeño/criterio de los alumnos (p.ej. trabajos, ensayos...).

#### FACTOR II

El segundo componente explica el 13.9% de la varianza. Es un factor bipolar en el que saturan los ítems trabajos (-), lectura-resúmenes de publicaciones (-), observaciones del profesor, examen y cuestionarios, lo que puede inducir a denominar a este factor como un Modelo de evaluación Crédito-sumativo *mixto*, en el que el examen se combina con las anteriores formas de evaluación. La polaridad del factor nos lleva a pensar que a mayor competencia del alumno en la gestión de los procesos de evaluación, menos incidencia del profesor, y viceversa.

#### FACTOR III

El tercer componente explica el 11.9% de la varianza. Es al igual que el primer factor, un componente unipolar en el que saturan los ítems, cuestionarios, entrevistas, exámenes, prácticas y participación y asistencia. Es también un Modelo Crédito-sumativo, pero centrado en el *profesor*, ya que las formas de evaluación que saturan dependen en un mayor grado de los criterios del docente.

#### FACTOR IV

El cuarto factor explica el 10.8% de la varianza. Es un componente unipolar en el que saturan los ítems examen (considerable saturación), autoevaluación (alta saturación) y el cuestionario (baja saturación). Se trata sin duda del Modelo de evaluación más duro, en el que la evaluación del aprendizaje casi recae exclusivamente en dos formas de evaluación (examen

y autoevaluación). Puede considerarse pues, si seguimos con la terminología empleada como un Modelo Crédito-sumativo *mixto*, ya que lo conforman formas de evaluación que dependen del criterio del profesor, así como del criterio del alumno.

## 5. HALLAZGOS

Un estudio anterior de Rodríguez (1996), reveló que entre los modelos de evaluación del aprendizaje que los profesores decían utilizar en su práctica docente en la misma Facultad objeto de estudio, predominan aquéllos que utilizan el examen en combinación con otras formas de evaluación. Eran los denominados Modelos de evaluación del aprendizaje Crédito-Sumativos.

El presente trabajo revela, que si bien los alumnos desean un modelo en donde no se contemple el examen (factor I de la solución factorial), también parecen encontrar factible la práctica de modelos, donde el examen está sopesado con otras formas de evaluación (factores II, III, y IV de la solución factorial). No parece en consecuencia, haber un distanciamiento acusado entre el deseo de los estudiantes y la presumible práctica evaluadora que dicen practicar los docentes en el ámbito universitario objeto de la experiencia. Cabe preguntarse si los resultados hubiesen sido otros, en el caso de que la investigación se hubiese desarrollado en un ámbito diferente al de las Humanidades. A pesar de todo, no nos llevemos al engaño, ya que ambos estudios también nos han revelado ciertas contradicciones como p. ej., que la forma de evaluación peor valorada por los alumnos con una media de 2.95 (entre el desacuerdo y la indiferencia), haya sido el examen, precisamente la forma de evaluación que dicen utilizar con más asiduidad los profesores (fue utilizada en 104 asignaturas, de las 112 analizadas). No cabe duda en suma, de que nos encontramos ante un campo de investigación actual y en expansión, al que sería deseable dedicar más esfuerzo investigador.

## BIBLIOGRAFÍA UTILIZADA

- ABRAIRA FERNÁNDEZ, C. (1995). ¿La evaluación formativa ayuda a los alumnos a aprender?. *Bordón*, 47 (1), pp.53-65.
- BARDIN, L. (1986). El análisis de contenido. Madrid. Akal.
- CEA D'ANCONA, M.A. (1996). Metodología cuantitativa. Estrategias y técnicas de investigación social. Madrid. Síntesis.
- DEARDEN, M. (1976). Assessment in a electrical and electronic engineering course.
- DE LA ORDEN, A. (1969). Evaluación del rendimiento educativo y la calidad de la enseñanza. *Revista de Educación*, XV III, nº206.
- DE LA ORDEN, A. (1981). Evaluación del aprendizaje y calidad de la educación. En VV.AA.. La calidad de la educación. Madrid. CSIC.
- DE LA ORDEN, A. (1985). Modelos de evaluación universitarios. *R.E.P.*, nº 169, pp.521-537
- DE LA ORDEN, A. (1993). Influencia de la evaluación del aprendizaje en la eficacia de la enseñanza. *R.I.E.*, nº22, pp.7-41.
- ELTON, L. y LANVILLARD, D.M. (1979). Trends in research on students learning. *Studies in Higher Education*, nº4, pp.87-182.
- FUCHS, L.S. y OTROS (1984). The effects of frequent curriculum-based measurement and evaluation in pedagogy, student achievement and student awareness of learning.

- KRIPPENDORF, K. (1990). Metodología del análisis de contenido. Teoría y Práctica. Barcelona. Paidós.
- MARTON, F. y SALJO, R.L (1976). On qualitative difference on learning outcome and process. *British Journal of Educational Psychology*.
- RODRÍGUEZ SABIOTE, C. (1996). Evaluación del aprendizaje en la Universidad. Tesis en curso. Universidad de Granada.
- RODRÍGUEZ, C. y GUTIÉRREZ, J. (1997). Estudio descriptivo sobre modos de evaluación del aprendizaje universitario (En prensa).