

CONCEPCIONES Y CREENCIAS DEL PROFESORADO DE SECUNDARIA SOBRE EVALUACIÓN EN MATEMÁTICAS¹

Francisco Gil Cuadra

Departamento de Didáctica de la Matemática y de las Ciencias Experimentales
Universidad de Almería

Luis Rico Romero

Departamento de Didáctica de la Matemática
Universidad de Granada

Antonio Fernández Cano

Departamento de Métodos de Investigación y Diagnóstico en Educación
Universidad de Granada

RESUMEN

La contribución de este trabajo consiste en una descripción y caracterización de las concepciones y creencias que sobre evaluación en matemáticas mantienen los profesores de secundaria andaluces. Se trata de un estudio exploratorio que utiliza la técnica de encuesta (survey), por medio de la administración de un cuestionario cerrado a modo de escala de valoración sobre una muestra de la población que se estudia (N= 163).

La elaboración del cuestionario cerrado de escala de valoración se apoya en la identificación empírica de los juicios de los profesores, la generación inductiva de un sistema de categorías teóricamente fundamentado para clasificar tales juicios y el control del proceso por

E-mail: fgil@ualm.es - lrico@ugr.es - afcano@ugr.es

1 Este trabajo ha recibido financiación en el Programa Sectorial de Promoción General del Conocimiento de la Dirección General de Investigación Científica y Técnica, en el Proyecto *Evaluación de conocimientos, procesos y actitudes en matemáticas* (PS93-0195 y PS96-1442).

expertos. Un estudio descriptivo de las valoraciones de los profesores indica el grado de aceptación de cada categoría. El análisis factorial de los datos permite detectar un factor general que establece la concepción global sustentada por el profesorado sobre este tópico. El factor general se articula en 13 factores específicos, que muestran diversas creencias de los profesores. El análisis de conglomerados (clusters) de las puntuaciones factoriales de los profesores permite inferir tendencias de pensamiento respecto de las creencias detectadas para la evaluación.

Palabras clave: Pensamiento del profesor, Concepciones y creencias, Evaluación en Matemáticas, Educación Matemática, Educación Secundaria, Encuestas, Análisis Factorial, Análisis Clúster.

ABSTRACT

This work describes and characterizes the conceptions and beliefs on mathematics' assessment sustained by Andalusian secondary teachers'. It is an exploratory study. Survey technique is used by means of a closed multiple-choice scale questionnaire applied to a sample of the studied population (N = 163).

The design and production of the closed multiple-choice scale questionnaire are based on the empirical identification of teachers' opinions, the inductive generation of a theoretically founded system of categories useful to classify such opinions, and the process control done by experts. A descriptive study of teachers' valuations establishes the level of acceptance for each category. The factor analysis of the data let us to detect a general factor that establishes the global conception sustained by the teachers on this topic. The general factor is structured by means of 13 partial factors, which show several teachers' beliefs. The cluster analysis of the teachers' factor scores allows to establish thought tendencies with regard to the beliefs in assessment already detected.

Key words: Teachers' Thinking, Conceptions and beliefs, Mathematics assessment, Mathematics Education, Secondary education, Survey, Factorial Analysis, Cluster Analysis.

I. DESCRIPCIÓN GENERAL

Este estudio describe las concepciones y creencias que sobre evaluación tienen los profesores de matemáticas de educación secundaria obligatoria de la Comunidad Autónoma de Andalucía. Es una contribución a los trabajos que se vienen realizando para comprender y caracterizar los modos que tienen de interpretar la evaluación los profesores de matemáticas de secundaria.

Se trata de un estudio exploratorio que utiliza la técnica de encuesta (*survey*) por medio de un cuestionario. El trabajo forma parte de una investigación más amplia: *Marco Conceptual y Creencias de los Profesores sobre la Evaluación en Matemáticas*, (Gil, 2000), donde se establece el contexto, marco conceptual y la línea general de investigación de la que deriva este informe. En el presente artículo vamos a centrarnos en el proceso de investigación, limitándonos a dar ejemplos de los resultados obtenidos.

1.1. Términos clave

En el presente estudio entendemos por:

Creencias: las verdades personales indiscutibles llevadas por cada uno, derivadas de la experiencia o de la fantasía, teniendo una fuerte componente evaluativa y afectiva (Pajares, 1992).

Concepciones: los marcos organizadores implícitos de conceptos, con naturaleza esencialmente cognitiva y que condicionan la forma en que afrontamos las tareas (Ponte, 1994).

1.2. Objetivos

Este estudio se propone describir y caracterizar el estado de opinión y las diversas creencias que sobre evaluación tienen los profesores de matemáticas, detectando las ideas y conceptos en los que se sustentan y las tendencias de pensamiento que se comparten.

La investigación realizada es un estudio muestral de tipo transversal y se ha llevado a cabo mediante aplicación de un cuestionario de escala valorativa: *Creencias de los Profesores sobre Evaluación en Matemáticas*, durante los años 1994 a 1998, en los que se ha producido la implantación de la Educación Secundaria Obligatoria.

El propósito central específico del estudio que aquí presentamos consiste en establecer el estado de opinión, es decir, los conceptos, ideas y valoraciones usuales de los profesores de matemáticas sobre evaluación así como la concepción general y las creencias parciales que sobre este tópico sustentan y caracterizar las tendencias de pensamiento mediante la delimitación de ideas y conceptos compartidos por grupos de profesores, este estado de opinión se quiere determinar de modo sistemático. Para ello es preciso delimitar de manera contrastable las distintas valoraciones que los profesores de matemáticas asignan a los conceptos e ideas sobre evaluación y caracterizar, si resulta posible, el factor que estructura esa concepción general, o sea describir el sistema conceptual en que se encuadran. A partir de este sistema conceptual delimitaremos los grupos de profesores que comparten concepciones similares y determinaremos tendencias de pensamiento.

Este propósito se ha desglosado en los siguientes objetivos parciales:

1. Establecer el estado de opinión relativo a una serie de conceptos e ideas sobre evaluación y su grado de aceptación, interpretando las valoraciones asignadas a dichos conceptos e ideas.
2. Detectar y caracterizar factores en el sistema de ideas inferido sobre evaluación, teniendo en cuenta las valoraciones hechas por los profesores de la muestra. Establecer las concepciones y creencias de los profesores de secundaria andaluces sobre la evaluación en matemáticas.
3. Establecer y estudiar las distintas tendencias o tipos de pensamiento que grupos significativos de profesores sostienen sobre evaluación en matemáticas, en función de los diferentes sistemas de concepciones y creencias detectados.

1.3. Antecedentes y estado de la cuestión

A partir de mediados de los 80 comenzó en España un proceso de renovación curricular que culminó con la Ley de Organización General del Sistema Educativo (LOGSE) de 1990. La estructura derivada de la ley ha implicado una serie de cambios considerables para el currículo de Educación Secundaria (Rico y Sierra, 1997). Dentro de los cambios impuestos por la LOGSE en Educación Secundaria destacan los relativos a nuevos criterios y prácticas sobre evaluación, que actualmente son objeto de revisión en muchos aspectos.

En la tradición de las innovaciones curriculares en matemáticas, los cambios en evaluación han sido, por lo general, limitados y escasamente difundidos (Howson, Keitel y Kilpatrick, 1981; Romberg, 1989). Ha sido usual que planteamientos de reforma e innovación hayan tenido escasa o nula incidencia sobre la evaluación; esto ha sucedido en gran medida por la inercia del sistema, sostenida en las prácticas del profesorado. En el caso de la reforma del Sistema Educativo Español se ha detectado un considerable desconcierto y un rechazo apreciable hacia algunas de las propuestas que sobre evaluación se han marcado para la Educación Secundaria (Castro y otros, 1996; Hidalgo, 1996; Buendía y otros, 1999). Durante el periodo de implantación general de la LOGSE en Andalucía (1994-1997), se produjeron frecuentes manifestaciones de disconformidad por parte del profesorado. Sin embargo, es muy poco lo que se conoce sobre los argumentos en que se funda tal rechazo y los estudios sobre la opinión de los profesores en ejercicio son muy escasos (Gil, 2000).

La investigación educativa ha centrado uno de sus focos de interés en el pensamiento del profesor y, más concretamente, en la investigación sobre el conocimiento, las concepciones y las creencias de los profesores como factores determinantes de su práctica profesional y de sus acciones en el aula (Houston, 1990; Thompson, 1992; Llinares, 1998). En particular, los estudios e investigaciones sobre el pensamiento del profesor de matemáticas y su conocimiento profesional han experimentado un desarrollo considerable en los últimos años (Ponte y otros, 1996; García, 1997).

El estudio del conocimiento del profesor de matemáticas, del que las concepciones y creencias sobre evaluación forman parte, parece un tema de interés tanto por su actualidad curricular como por sus conexiones con corrientes actuales de investigación en Educación Matemática. En este contexto hemos diseñado la investigación que nos ocupa y que aquí presentamos.

Otros objetivos parciales han sido presentados en trabajos previos (Rico y otros, 1995a; Rico y otros, 1995b; Rico y Gil, 1996; Gil y otros, 1997; Gil, 2000).

2. METODOLOGÍA

Nuestro estudio está enmarcado en un paradigma integrativo, que utiliza sucesivamente métodos complementarios, aunque predomina una orientación exploratoria y una metodología descriptiva de tipo encuesta, que se lleva a cabo mediante la administración de un cuestionario a una muestra de la población en estudio, y trata de combinar lo descriptivo con lo interpretativo. Los resultados obtenidos se interpretan tras

un estudio descriptivo de respuestas, en primer lugar, después, a partir de un análisis factorial, y, finalmente, a partir de un análisis clúster.

2.1. Instrumento

El procedimiento seguido para inferir un sistema de conceptos e ideas mediante el cual estructurar y comprender las concepciones y creencias del profesor de secundaria sobre evaluación y evaluación en matemáticas, ha sido descrito con detalle en un trabajo anterior (Rico y otros, 1995a). Partimos de un cuestionario abierto mediante el cual recogemos una amplia muestra de juicios y opiniones de los profesores en ejercicio sobre cuestiones básicas relativas a la evaluación en matemáticas: contenido, finalidades, instrumentos, agentes y dificultades que presenta la evaluación. La diversidad de juicios recogidos dan lugar a un sistema de categorías, conceptualmente fundado, que permiten su clasificación; llegamos así a establecer una familia de conceptos e ideas que estructuran la diversidad de los juicios emitidos por los profesores de matemáticas. Una vez determinado este sistema de conceptos e ideas, y basado en él, construimos un nuevo cuestionario cerrado a modo de escala de valoración.

El proceso general seguido para llegar a la elaboración y análisis formal del instrumento utilizado en este estudio (Gil y otros, 1997), ha sido el siguiente:

- a) Elaboración de un cuestionario abierto:
 - a.1) Creación de un cuestionario piloto.
 - a.2) Revisión de ese cuestionario.
 - a.3) Elaboración de un segundo cuestionario.
 - a.4) Versión definitiva del cuestionario abierto.

- b) Aplicaciones sucesivas del cuestionario abierto:
 - b.1) Primera aplicación para recoger los juicios de los profesores sobre evaluación.
 - b.2) Segunda aplicación para completar los tipos y diversidad de juicios.
 - b.3) Tercera aplicación, en la que se comprueba la saturación de los juicios emitidos.

- c) Clasificación de las respuestas obtenidas:
 - c.1) Reducción de datos.
 - c.2) Determinación de criterios de clasificación.
 - c.3) Validación del criterio de clasificación.
 - c.4) Reconsideración de aquellos criterios no validados.
 - c.5) Sistema final de categorías.

- d) Elaboración de un cuestionario cerrado a modo de escala de valoración multicategorial, *rating scale* (Wright y Masters, 1982; Spector, 1992; Fernández Cano, 1995):
 - d.1) Creación de un cuestionario piloto sobre las categorías anteriormente establecidas.

- d.2) Revisión de ese cuestionario.
- d.3) Elaboración de un segundo cuestionario.
- d.4) Versión definitiva del cuestionario.

En el Anexo aparece la versión final de la escala de valoración, con 44 enunciados y puntuaciones de 1 a 9, que se obtuvo al culminar el proceso descrito.

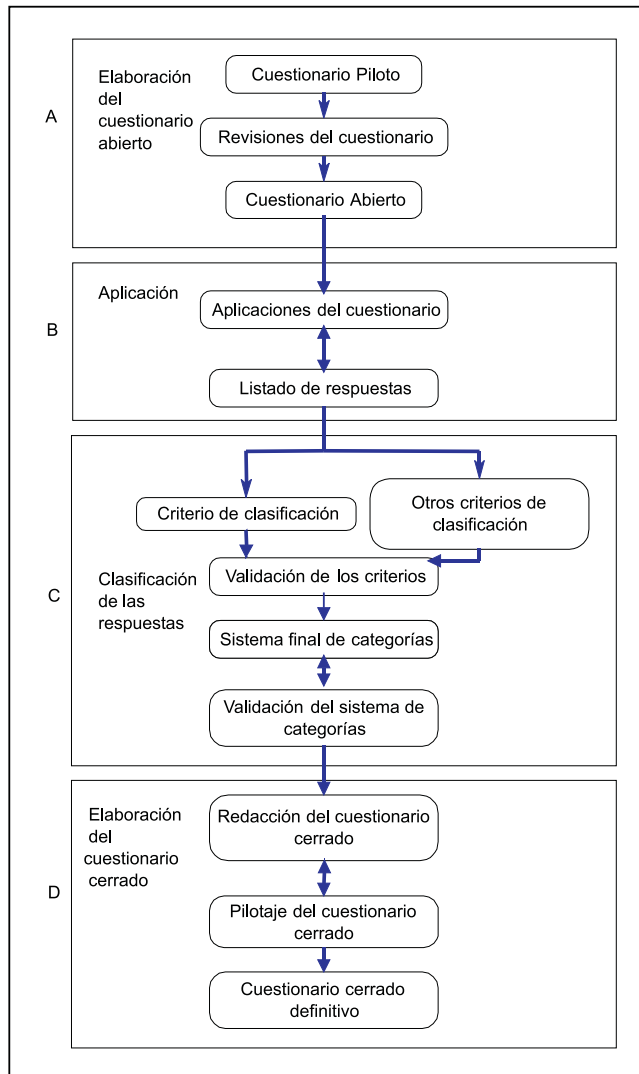


Ilustración 1

Proceso de elaboración del instrumento escala de valoración.

2.2. Contenidos del cuestionario

La versión final del cuestionario cerrado está organizada en diez preguntas:

- Las preguntas 1^a a 5^a plantean cuestiones relativas a la evaluación en general, donde la 1^a y la 2^a se refieren a los objetivos y fines de la evaluación; las preguntas 3^a, 4^a y 5^a se refieren a aspectos prácticos y técnicos.
- Las preguntas 6^a a 10^a son cuestiones específicas sobre la evaluación en matemáticas, donde las preguntas 6^a y 7^a se refieren a objetivos y dificultades de la evaluación en matemáticas; las preguntas 8^a, 9^a y 10^a proponen evaluar otros elementos del currículo de matemáticas.

Cada una de las preguntas presenta diversas opciones de respuesta; así, la pregunta 1^a presenta 12 opciones diferentes, mientras que otras preguntas como la 3^a, 4^a y 8^a presentan sólo dos opciones de respuesta. Las opciones en cada caso no son alternativas pero sí expresan diferentes conceptualizaciones o modos de interpretar la cuestión general que las precede (ver Anexo). En total el cuestionario está constituido por 44 enunciados; cada uno de ellos expresa, junto con la pregunta que lo origina, una idea o concepto relativo a la evaluación. El profesor encuestado debe valorar uno a uno los enunciados propuestos para las cuestiones generales; de este modo manifiesta su juicio sobre el tema. Hay que considerar que los profesores asignan las puntuaciones expresando sus preferencias relativas ante las diversas opciones a cada cuestión.

2.3. Administración del cuestionario

El cuestionario cerrado se administró en el transcurso de una sesión de los cursos de actualización que los profesores de matemáticas tenían que realizar para incorporarse a la reforma. La contestación se realizó tras una breve introducción que consistió en presentar los fines de la investigación, identificar los autores, solicitar y agradecer la colaboración a los encuestados e indicar la valoración asignada a cada una de las cuestiones en una escala de 1 a 9, donde: 1 indica muy en desacuerdo, 9 muy de acuerdo, y el 5 expresa indiferente.

Durante el proceso de administración del cuestionario siempre estuvo presente uno de los investigadores, quién constató la seriedad con que los profesores contestaron a los cuestionarios. El tiempo para responder el cuestionario fue libre y giró en torno a una hora.

2.4. Muestra

La población objeto de nuestro estudio está formada por los profesores de matemáticas del Segundo Ciclo de la Enseñanza Secundaria Obligatoria de Andalucía en el período de implantación de la reforma derivada de la LOGSE. Debido a las dificultades que planteaba el proceso de implantación de la reforma y la incorporación escalonada de los profesores de secundaria al nuevo sistema, optamos por trabajar con una mues-

tra estratificada o por cuotas, seleccionada en dos etapas según los criterios que se presentan a continuación.

Puesto que la incorporación del profesorado a la reforma estaba previsto que se realizase a lo largo de tres cursos (95-96, 96-97 y 97-98), consideramos la población dividida en tres cohortes, cada una de ellas formada por los profesores que en un mismo curso se incorporan a la reforma. Para determinar la muestra seleccionamos los profesores de matemáticas de la segunda cohorte, es decir, los profesores de matemáticas de los centros que en el curso 96-97 comienzan la implantación del Segundo Ciclo de la Educación Secundaria Obligatoria; en total 412 profesores.

Como, por otra parte, existe una división territorial y administrativa del profesorado por provincias, esta población se considera compuesta por ocho estratos que se corresponden con cada una de las provincias andaluzas. Por tanto, encontramos una población dividida en tres cohortes correspondientes a los cursos 95-96, 96-97 y 97-98 y organizada en ocho estratos, correspondientes a las provincias de Almería, Cádiz, Córdoba, Granada, Huelva, Jaén, Málaga y Sevilla. Nuestra muestra la hemos elegido de la segunda cohorte, cuya distribución por estratos era la siguiente:

AL	CA	CO	GR	HU	JA	MA	SE	Total
6	65	31	73	37	43	60	97	412

Por razones de disponibilidad, la aplicación de los cuestionarios se realizó finalmente sobre cuatro de los estratos de la cohorte seleccionada: Almería, Granada, Málaga y Sevilla. La aplicación se hizo mediante un censo, pues se encuestó a todos los sujetos de cuatro de los ocho estratos de la cohorte central (curso 96-97) de la población. Esto supone 236 sujetos de un total de 412. No obstante, se ha producido una mortalidad muestral considerable de 73 sujetos: 55 por falta de asistencia a la sesión donde se encuestó y 18 porque se negaron a responderla o la entregaron en blanco. Todas estas circunstancias implican que hemos recogido datos de un total de 163 sujetos, los cuales representan un porcentaje del 40% sobre el total de la cohorte.

3. ANÁLISIS DE DATOS

El análisis de datos lo vamos a estructurar en tres partes: una primera parte dedicada al estudio descriptivo e interpretación de las respuestas, una segunda donde se presentaran los resultados del análisis factorial efectuado, y una tercera donde se presentarán los resultados del análisis clúster.

3.1. Análisis descriptivo

En este apartado se presentan los resultados de la aplicación del cuestionario a los 163 profesores de matemáticas de secundaria de Andalucía finalmente encuestados.

En cada una de las preguntas de las encuestas se presentaron sus respuestas alternativas ordenadas según orden decreciente de sus medias y, cuando hay coincidencia, según su dispersión, de menor a mayor desviación típica.

Para tratar de responder a las cuestiones: ¿qué diferencias entre valoraciones podemos considerar apreciables? y ¿cuáles podemos considerar estadísticamente significativas?, hemos calculado el *tamaño del efecto* entre cada pareja de enunciados de una misma pregunta (Cohen, 1988; Hedges y Olkin, 1985) y un test inferencial de igualdad de medias para comparar parejas de enunciados utilizando el paquete estadístico BMDP (Dixon, 1990). El cálculo del tamaño del efecto (TE) permite determinar qué diferencias consideraremos como apreciables y cuáles no. Hemos considerado tamaños de efecto sustanciales, los superiores a 0.5, y los asociados a *niveles de significación* (valores p) menores que 0.05.

A modo de ejemplo, presentamos una de las preguntas generales (la primera) con las distintas opciones de respuesta propuestas. Cada respuesta va precedida de un código que indica el número de la pregunta (P1, pregunta 1) y seguida de otro número que indica el enunciado; posteriormente aparece el enunciado, la media y la desviación típica obtenidas. Cada tabla presenta los distintos agrupamientos de respuestas que pueden formarse en función de los tamaños del efecto y de los test de significación de diferencias de medias. Finalmente realizamos un estudio descriptivo global con todas las respuestas.

3.1.1. Resultados según ítem/pregunta

A modo de ejemplo vamos a presentar los resultados de la primera pregunta, la totalidad del análisis para el resto de las preguntas puede verse en Gil y otros (2000).

Primera pregunta

¿Qué debe ser objeto de evaluación? En evaluación es prioritario valorar:

		\bar{X}	S	TE	P
P1.1	el conocimiento adquirido por los alumnos	8.01	1.07		
P1.2	el trabajo realizado por los alumnos	7.78	1.08		
P1.10	los logros alcanzados respecto de los objetivos	7.74	1.40		
				0.27	0.00
P1.3	la actitud y el interés de los alumnos	7.39	1.22		
P1.8	la madurez y formación del alumno	7.15	1.52		
P1.4	las capacidades de los alumnos	7.07	1.56		
P1.7	la labor del profesor	6.87	1.74		
P1.9	los contenidos	6.57	1.58		
				0.32	0.00
P1.5	la conducta de los alumnos	6.04	1.72		
P1.6	el currículo	5.91	1.87		
P1.11	los medios y materiales	5.89	1.98		
P1.12	las instituciones y el sistema educativo	5.51	2.47		

De las valoraciones asignadas conjeturamos que, en primer lugar, el objeto prioritario de evaluación debe ser *el alumno*. Se consideran diversos criterios para su evaluación corroborando la idea de que la evaluación, prioritariamente, debe recaer sobre el alumno. En primer lugar, los conocimientos adquiridos y el trabajo realizado. Se manifiesta prioritariamente una noción de evaluación en la que priman los resultados (conocimientos adquiridos, trabajo realizado) por encima del logro de unos objetivos, concepción de Tyler de la evaluación por objetivos, que tan familiar ha sido en nuestro sistema educativo años atrás.

Después de evaluar al alumno también se deben evaluar otros agentes y componentes del sistema educativo, y se comienza por la labor del profesor.

Al realizar comparaciones intra-ítems aparecen tres subgrupos de variables², debido a que entre el último ítem de un grupo y el primero del siguiente surgen tamaños del efecto medianos (entre 0.25 y 0.30). El primero está formado por aquellos conceptos que han recibido una alta valoración en la encuesta; son ideas que tradicionalmente se consideran pertinentes al tratar de la evaluación de los alumnos: son objetivos de la evaluación el conocimiento adquirido, el trabajo realizado y los logros alcanzados. El segundo subgrupo está formado por un conjunto de enunciados que reciben una valoración intermedia; se trata de aspectos colaterales con los enunciados anteriores y parecen ocupar un segundo plano en la evaluación; éstos son: la actitud e interés de los alumnos, la valoración de su madurez, sus capacidades, la labor del profesor y los contenidos. El tercer subgrupo está formado por un núcleo de conceptos distantes de la conceptualización clásica sobre evaluación; los profesores parecen considerar estas ideas poco importantes en relación con la evaluación. Se contempla aquí la conducta de los alumnos (de la conducta del alumno al profesor interesa la parte que tiene que ver con su trabajo), conceptos más abstractos como son el currículo y las instituciones, o bien otros pocos usuales como son los medios y materiales.

El hallazgo central es que los profesores establecen prioridades que van de lo más inmediato, próximo y concreto a lo mediato, remoto y abstracto. Hay preferencia por centrar la evaluación sobre el trabajo de los alumnos.

3.1.2. Categorías de enunciados. Análisis de clúster

Se trata de clasificar la totalidad de los conceptos en función de las valoraciones que han recibido de los profesores, de modo que un mismo grupo esté formado por enunciados que han recibido, estadísticamente, puntuaciones similares.

Esta clasificación permite determinar una serie de intervalos para caracterizar el grado de aceptación de la totalidad de los enunciados, es decir, determinar qué conceptos tienen aceptación similar y expresar un orden de aceptación entre estos grupos, que se concretaría en una categoría nominal: su *nivel o grado de aceptación*. También queremos estudiar qué juicios expresan un alto consenso, cuáles otros son aceptados comúnmente y cuáles producen disensión entre los profesores.

2 Hablamos de grupos de ítems cuando el tamaño del efecto entre ellos es grande, superior a 0.5, y de subgrupos cuando es inferior.

Para ello hemos construido una matriz de distancias que recoge la distancia entre cada par de enunciados; por distancia entre dos enunciados tomamos el valor del tamaño del efecto correspondiente. Sobre esta matriz de distancias hemos realizado un análisis *clúster* con todas las variables del cuestionario, cuyos resultados presentamos a continuación.

Aparecen cinco conglomerados o clusters. Las medias de los enunciados de cada clúster están entre un valor mínimo y un valor máximo; tomando ambos valores extremos establecemos un intervalo, en el cual están las medias de todos los enunciados del clúster. Establecemos así una categoría nominal para cada clúster, que llamamos su *grado de aceptación*, como sigue:

- * Primer clúster: conceptos cuya media se encuentra comprendida entre los valores 4.23 y 4.86, o en el intervalo [4.23, 4.86]. Consideramos que el *grado de aceptación* de estos enunciados es *muy bajo*. Los enunciados de este clúster suelen tener una desviación típica alta, que indica un alto grado de disensión entre los profesores con porcentajes de rechazos considerable. Estaría conformado por 5 enunciados.
- * Segundo clúster: conceptos cuya media se encuentra comprendida en el intervalo [5.51, 6.12], son enunciados que tienen un *bajo grado de aceptación*; expresan juicios que merecen poca atención de los profesores, sobre ellos hay una valoración positiva muy escasa o cierta disensión. Estaría conformado por 8 enunciados.
- * Tercer clúster: conceptos cuya media se encuentra comprendida en el intervalo [6.57, 7.45], representan un *grado de aceptación medio*; expresan juicios que merecen cierta atención de los profesores, sobre ellos hay valoración positiva poco pronunciada. Estaría conformado por 18 enunciados.
- * Cuarto clúster: conceptos cuya media se encuentra comprendida en el intervalo [7.70, 7.80], representan un *grado de aceptación alto*; expresan juicios que merecen bastante atención de los profesores, sobre ellos hay una valoración positiva pronunciada. Estaría conformado por 5 enunciados.
- * Quinto clúster: conceptos cuya media se encuentra comprendida en el intervalo [8.01, 8.48], les corresponde un *grado de aceptación muy alto*; expresan juicios que tienen el consenso de los profesores, sobre ellos hay una valoración bastante positiva. Estaría conformado por 9 enunciados.

3.2. Solución factorial y constructo subyacente

Con el fin de obtener un modelo explicativo más sencillo que permita interpretar la estructura de los conceptos presentados realizamos un análisis factorial, complementario del estudio descriptivo, teniendo en cuenta los datos obtenidos de la aplicación del cuestionario.

Para ello consideramos cada uno de los conceptos e ideas enunciados como una variable que toma valores en un conjunto discreto, y trataremos de reducir el espacio

de estas variables, es decir, descubrir las dimensiones de variabilidad común existentes en este campo.

En concreto intentamos responder la pregunta ¿existe un factor general que aglutine la mayoría de estas variables de modo que de él pueda deducirse que los profesores de matemáticas tienen una *concepción global* de la evaluación?

Para contestar esta cuestión sometimos los datos a un análisis factorial utilizando el programa 4M del paquete estadístico BMDP (Dixon, 1990).

Los resultados se exponen en la siguiente tabla 1:

TABLA 1
SOLUCIÓN FACTORIAL POR COMPONENTES PRINCIPALES A DATOS DE
UNA ESCALA DE VALORACIÓN SOBRE CONCEPCIONES/CREENCIAS EN
EVALUACIÓN POR PROFESORES DE SECUNDARIA

ÍTEMS	R ²	h ²	F. General	ÍTEMS	R ²	h ²	F. General
P1.1	0.76	0.78	0.29	P5.4	0.52	0.69	0.28
P1.2	0.67	0.71	0.53	P6.1	0.79	0.81	0.37
P1.3	0.75	0.80	0.57	P6.2	0.71	0.72	0.62
P1.4	0.72	0.79	0.52	P6.3	0.71	0.73	0.64
P1.5	0.79	0.81	0.64	P6.4	0.74	0.82	0.53
P1.6	0.58	0.63	0.65	P6.5	0.76	0.75	0.65
P1.7	0.59	0.67	0.61	P6.6	0.61	0.64	0.55
P1.8	0.50	0.54	0.62	P6.7	0.70	0.70	0.57
P1.9	0.60	0.64	0.56	P6.8	0.71	0.70	0.59
P1.10	0.73	0.76	0.62	P7.1	0.37	0.65	0.28
P1.11	0.78	0.82	0.65	P7.2	0.22	0.55	0.14
P1.12	0.60	0.71	0.58	P7.3	0.45	0.65	0.35
P2.1	0.41	0.53	0.35	P7.4	0.32	0.70	0.19
P2.2	0.43	0.57	0.46	P8.1	0.38	0.50	0.34
P2.3	0.50	0.64	0.48	P8.2	0.31	0.58	0.28
P3.1	0.56	0.70	0.33	P9.1	0.37	0.59	0.36
P3.2	0.50	0.77	0.18	P9.2	0.45	0.65	0.42
P4.1	0.53	0.69	0.33	P9.3	0.39	0.56	0.42
P4.2	0.41	0.63	0.18	P10.1	0.58	0.70	0.46
P5.1	0.37	0.64	0.24	P10.2	0.57	0.70	0.47
P5.2	0.47	0.71	0.36	P10.3	0.46	0.63	0.38
P5.3	0.42	0.59	0.44	P10.4	0.47	0.73	0.22

La solución factorial obtenida permite establecer las siguientes consideraciones:

Primera. Todas las variables, menos una, tienen un cuadrado de correlación múltiple (R^2) superior a 0.25 lo que justifica su inclusión en el análisis factorial, este valor mide el cuadrado de la correlación de cada variable con las restantes. La única variable que presenta un SMC inferior a 0.25 es la P7.2 cuyo valor es de 0.22, pero al estar próximo al valor de corte hemos decidido no eliminarla del análisis factorial.

Segunda: Todas las variables presentan una comunalidad alta (h^2), superior a 0.60 en todos los casos salvo en nueve que está próxima (superior a 0.50).

Tercera: El valor de la Theta de Carmines obtenido es $\theta = 0.91$, valor muy estimable. Este dato mide la consistencia interna de los datos y controla la fiabilidad del instrumento utilizado.

Cuarta: Todas las variables, excepto cinco, cargan en el primer factor; las variables que no cargan son: P3.2, P4.2 (estas dos han recibido una valoración negativa por el profesorado), P7.2 (que tiene un valor R^2 bajo), P7.4 y P10.4.

Las cargas en este factor de las restantes variables son buenas, lo cual permite *afirmar la existencia de un factor general sobre la concepción de la evaluación en matemáticas*. Este factor explica un porcentaje de varianza del 22% en el espacio de datos y de un 32% en el de factores.

Quinta: El análisis realizado proporciona 13 factores que explican un porcentaje de varianza del 68%. Para facilitar su interpretación hemos realizado una rotación ortogonal de los factores y hemos considerado despreciables aquellas cargas que en valor absoluto eran inferiores a 0.32.

3.2.1. Factor general

A la vista de estos datos, sostenemos que *el profesorado de matemáticas encuestado presenta una concepción global compartida sobre la evaluación en matemáticas*, en el sentido de un marco organizador implícito de conceptos y juicios de naturaleza esencialmente cognitiva, ya que se produce una valoración coordinada entre los docentes encuestados sobre conocimientos relevantes relativos a la evaluación.

En este sentido *postulamos el constructo concepción de los profesores sobre evaluación en matemáticas*, con las siguientes características generales obtenidas del factor general:

- * El factor general muestra una preocupación diversificada por el alumno y el currículo como sujetos de la evaluación y de la evaluación en matemáticas, sobre los que se pueden establecer distintos criterios para determinar el objeto de la evaluación. También muestra preocupación por la valoración del profesor y la valoración del centro.
- * Los fines de la evaluación tienen distintas cargas en el factor general, siempre bajas. Tomar decisiones y controlar el proceso son fines de la evaluación que caracterizan al factor general.
- * El factor general contempla las variables técnicas con cargas muy bajas y rechaza los evaluadores externos y pruebas estandarizadas. Tampoco contempla las dificultades de la evaluación y la valoración del libro de texto.

- * Las prioridades del constructo general sobre evaluación están en los criterios para establecer su objeto. Alumnos y componentes del currículo reciben cargas altas; profesores y centro reciben cargas bajas.
- * En el factor general se excluye que la valoración del conocimiento de los alumnos sea objetivo prioritario de la evaluación.

Este factor general determina la concepción predominante en la población encuestada sobre evaluación. Si repasamos los posicionamientos que recoge este factor general vemos que sólo uno de ellos choca con los planteamientos de la reforma, el que estipula que la finalidad de la evaluación es tomar decisiones sobre la promoción de los alumnos, ya que los nuevos currículos la promoción de los alumnos se produce de manera casi automática.

3.2.2. Factores específicos

A continuación se establecen una serie de factores específicos, obtenidos mediante solución rotada, determinando para cada factor las variables que cargan, los enunciados correspondientes, el grado de aceptación y el peso de cada variable. Después de caracterizar cada factor llevamos a cabo su interpretación.

A modo de ejemplo presentamos el tratamiento general de los dos primeros factores, de los restantes nos limitamos a dar su denominación.

Primer factor

FACTOR 1

P1.3	En evaluación es prioritario valorar la actitud y el interés de los alumnos	0.87
P6.3	En matemáticas es prioritario evaluar la actitud y el interés de los alumnos hacia la asignatura	0.76
P1.2	En evaluación es prioritario valorar el trabajo realizado por los alumnos	0.73
P6.2	En matemáticas es prioritario evaluar el trabajo realizado por los alumnos	0.67
P1.7	En evaluación es prioritario valorar la labor del profesor	0.53
P1.5	En evaluación es prioritario valorar la conducta de los alumnos	0.46
P9.2	El profesor de matemáticas se valora por su formación científica y didáctica	0.38

Este factor lo determinan siete variables; cinco de ellas hacen referencia al alumno y dos al profesor. Del alumno se valora su actitud, interés y participación, el trabajo realizado y su conducta, tanto en términos generales como de manera específica en matemáticas. Del profesor se valora su labor docente y su formación científica y didáctica. Las variables que mayor carga presentan son las relativas a la actitud, interés y trabajo del alumno.

La idea que transmiten conjuntamente estas variables es la prioridad que debe darse en la evaluación a las actitudes y esfuerzos personales de alumnos, principalmente, pero también de los profesores, lo cual se pone de manifiesto en el trabajo realizado y en la formación alcanzada.

En el apartado 3.1 vimos que las variables corresponden, por su grado de aceptación, a distintos clusters. Así, la variable P1.5 tiene un bajo grado de aceptación en la

muestra, sin embargo tiene una carga media en el factor, lo que la hace significativa y característica en el factor.

Denominamos al primer factor: *Prioridad de la actitud, el interés y la conducta para evaluar a los alumnos y del trabajo personal para evaluar a alumnos y profesores.*

Segundo factor

FACTOR 2

P1.11	En evaluación es prioritario valorar los medios y materiales	0.81
P1.12	En evaluación es prioritario valorar las instituciones y el sistema educativo	0.79
P6.8	En matemáticas es prioritario evaluar medios y materiales	0.72
P1.6	En evaluación es prioritario valorar el currículo	0.56
P1.7	En evaluación es prioritario valorar la labor del profesor	0.45
P1.9	En evaluación es prioritario valorar los contenidos	0.34
P1.8	En evaluación es prioritario valorar la madurez y formación del alumno	0.33
P1.5	En evaluación es prioritario valorar la conducta de los alumnos	0.32
P1.10	En evaluación es prioritario valorar los logros alcanzados respecto de los objetivos	0.32
P6.6	En matemáticas es prioritario evaluar los contenidos	0.32

Este factor lo integran diez variables, que proceden de las cuestiones primera y sexta, y se ajustan al enunciado general: *¿Qué evaluar/ qué evaluar en matemáticas?*

La idea que transmiten conjuntamente estas variables es complementaria al primer factor respecto del factor general, y se centran en considerar el objeto de la evaluación en las componentes del currículo; destacan la importancia de evaluar las variables sobre *Aspectos generales del currículo* en el proceso de clasificación. También consideran la madurez, formación y conducta de los alumnos, pero con carga baja.

Las variables P1.11, P1.12, P6.8, P1.6 y P1.5 tienen un grado de aceptación bajo en la muestra, lo cual hace que determinen mejor al factor; además, las cuatro primeras variables tienen carga alta o muy alta, con lo cual establecen sus rasgos principales.

Denominamos al segundo factor: *Prioridad de los medios y materiales, de la institución escolar y de otros componentes curriculares como objeto de la evaluación.*

Denominamos al tercer factor: *Evaluación cuantitativa del conocimiento del alumno sobre contenidos. Para su promoción y orientación.* Se trata de una evaluación clásica y convencional, cercana y próxima a los intereses formativos del alumno, poco contemplada en el factor general.

Denominamos al cuarto factor: *Valorar el logro de los objetivos con el fin de controlar el proceso y el resultado.*

Denominamos al quinto factor: *Objeto de la evaluación son la organización y proyecto del centro, en primer lugar, y la conducta de los alumnos. En segundo lugar.*

Denominamos al sexto factor: *Dificultades de la evaluación debidas al profesor y a los instrumentos.*

Denominamos al séptimo factor: *Evaluación estandarizada mediante agentes externos.*

Denominamos al octavo: *Prioridad para evaluar las capacidades de los alumnos, distancia con la evaluación del aprendizaje de sus alumnos.*

Denominamos al noveno factor: *Evaluación del centro por la formación de sus profesores y el nivel de sus alumnos.*

Denominamos al décimo factor: *Evaluación convencional y comunicación por escrito.*

Denominamos al undécimo factor: *Evaluación de las cualidades del profesor e informes orales cualitativos.*

Denominamos el duodécimo factor: *Evaluar el contenido de los libros y la profesionalidad de los profesores.*

Denominamos el decimotercer factor: *Complejidad del proceso de evaluación, sin responsabilidad de los alumnos.*

3.2.3. Hallazgos del estudio factorial

El constructo o factor general sobre evaluación en matemáticas, queda organizado y diversificado en 13 factores parciales. Hay 1 variable que interviene en 4 factores distintos (P1.5), 4 variables que intervienen en 3 factores (P1.9, P4.1, P6.6 y P9.2). Las 39 variables restantes intervienen en 1 ó 2 de los factores. No hay variables excluidas.

Destacamos que, en 10 ocasiones, las versiones alternativas correspondientes a las preguntas 1ª y 6ª cargan en el mismo factor y con cargas muy similares, cosa que también ocurre en el factor general y en el estudio sobre el grado de aceptación de las variables.

Los 13 factores expresan diversas opciones para unas pocas relaciones centrales.

En primer lugar, hay una serie de factores que relacionan un sujeto con un objeto de evaluación. Estos factores marcan diversas prioridades entre las posibilidades de evaluar a los alumnos y, en segundo lugar, de evaluar a los profesores. De este modo se establecen distintas creencias relativas a las criterios para establecer el objeto de la evaluación.

En un segundo término aparecen factores que consideran diversas opciones sobre los fines de la evaluación. Sólo 3 factores consideran las finalidades (F3, F4 y F8). Por otra parte, los factores F5, F6 y F13 establecen diferentes responsabilidades sobre la dificultad del proceso de evaluación.

En tercer lugar hay 3 factores (F7, F10 y F11) que se manifiestan de manera precisa sobre el modo de evaluar y alguna de sus características técnicas. Estos factores inciden en aspectos técnicos de la evaluación a practicar en el aula con los alumnos.

Todos los factores ponen de manifiesto las creencias sobre evaluación en matemáticas de los profesores de la muestra que, a nuestro juicio, son:

- * El sujeto principal de la evaluación es el alumno, pero el objeto de la evaluación varía entre unas creencias y otras. Los indicadores utilizados para las distintas creencias son: Interés y trabajo (F1 y F5); Conocimientos y nivel (F3 y F9); Capacidades (F8); Logro de objetivos (F4 y F11); Actitudes y conducta (F5 y F12 negativamente).

Detectamos, pues, cinco creencias diferentes sobre el objeto de la evaluación en los alumnos.

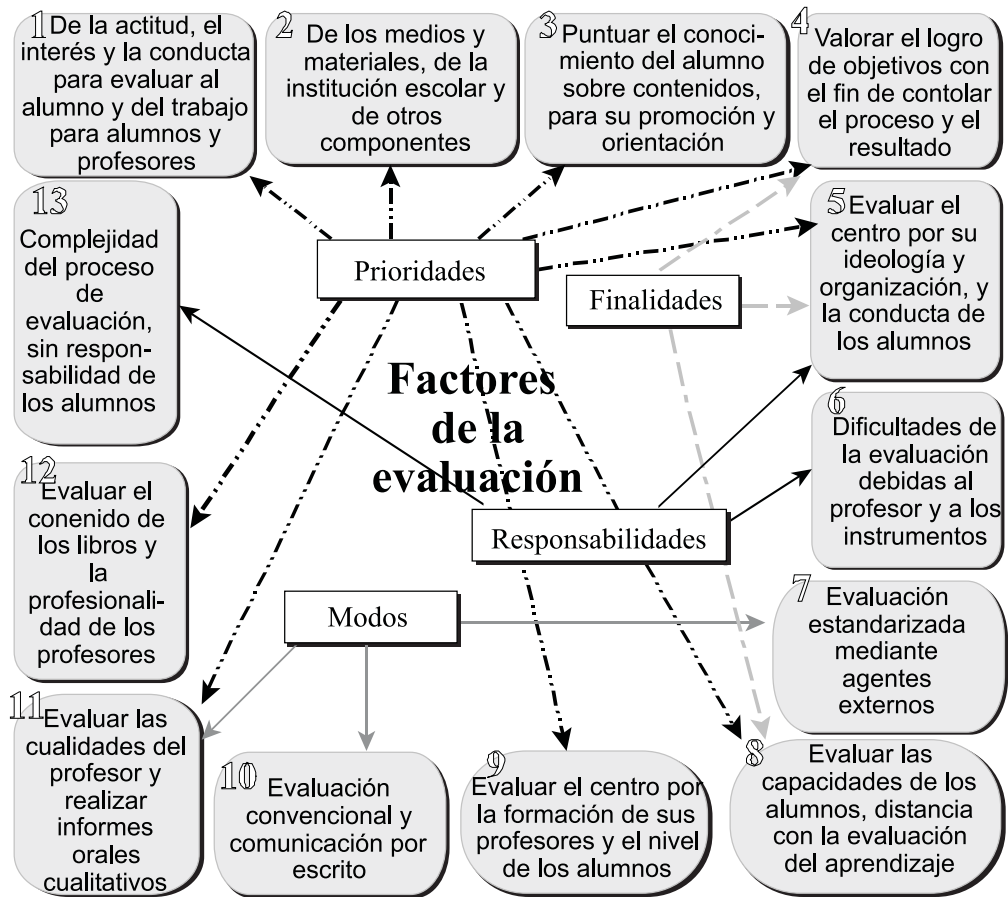


Figura 2
Estructura conceptual de los factores (creencias) sobre evaluación.

- * El profesor también es sujeto de evaluación pero, en muchas ocasiones, en relación con la evaluación del centro o con la complejidad del proceso. Detectamos las siguientes creencias: El profesor se debe evaluar por su labor docente y su formación profesional (F1, F6 y F12); El profesor se debe evaluar por su formación científica (F9); El profesor se debe evaluar por sus cualidades personales (F11), en relación con la valoración del centro.
- * Hay dos creencias diferentes sobre la evaluación del centro. El centro se puede evaluar por su proyecto y organización (F5); también por la preparación de sus alumnos y profesores (F9).
- * Una creencia sostiene que la evaluación del libro debe hacerse por sus contenidos (F12).

- * Sobre el objeto general de la evaluación hay una creencia fuerte sobre la prioridad de evaluar los aspectos institucionales y curriculares, que se manifiesta en el factor F2.
- * En relación con las finalidades de la evaluación hemos detectado dos creencias. Hay una creencia explícita relativa a que el fin prioritario está en el control del proceso y de los resultados (F4); una segunda creencia contempla, conjuntamente, obtener información sobre el aprendizaje y tomar decisiones sobre la promoción y orientación (F3 y F8).
- * Desde un punto de vista técnico hemos detectado cuatro creencias contrapuestas: F3 y F11; el primero prima la evaluación cuantitativa y el segundo prima la cualitativa; F7 y F10; el primer factor prima la evaluación estandarizada y externa frente al segundo, que incide en la evaluación convencional con comunicación escrita.
- * También hemos detectado creencias relativas a las causas de las dificultades de la evaluación; estas creencias hacen referencia siempre a un agente. Así, en un primer caso (F6), las dificultades son debidas al profesor y a los instrumentos. En un segundo caso (F13 y F5 débilmente) las dificultades son debidas a la complejidad del proceso, pero no a los alumnos.

3.3. Análisis clúster de respondientes (profesores)

Una vez caracterizadas las distintas creencias de los sujetos de la muestra sobre nuestro foco de estudio, establecemos y estudiamos las distintas tendencias o tipos de pensamiento de los profesores que se pueden determinar por los diferentes sistemas de concepciones y creencias que grupos significativos de profesores sostienen sobre evaluación en matemáticas.

Se trata de mostrar que las diversas creencias de los profesores de matemáticas sobre evaluación no son individuales, si no que son grupos de profesores, más o menos amplios, quienes las sostienen de manera coherente y las comparten; en este sentido es posible hablar de tendencias en el pensamiento de los profesores.

Descendiendo a un nivel más concreto, planteamos la pregunta: ¿cómo poner de manifiesto que un grupo de profesores mantiene, esencialmente, las mismas posiciones sobre las cuestiones claves que han determinado las diferentes creencias establecidas en el apartado anterior?

Los sujetos en este caso son los profesores, no las variables, y los datos que consideramos en esta fase son las puntuaciones factoriales obtenidas por cada sujeto en los factores. Puesto que son los factores establecidos los que permiten caracterizar las diferentes creencias, es por lo que consideramos como valores para cada sujeto sus puntuaciones factoriales, obtenidas en cada uno de esos factores. De este modo conectamos los sujetos con las creencias, ya que tales creencias se han determinado en base a los factores.

Para dar respuesta a la cuestión anteriormente planteada, el procedimiento utilizado en la mayoría de las investigaciones revisadas consiste en realizar un análisis clúster de casos siguiendo algún método aglomerativo (Bisquerra, 1989; Everit, 1993, Aldenderfer y Blashfield, 1984), como el de las distancias mínimas, el de las distancias máximas o el del centroide, entre otros.

En nuestro caso comprobamos que todos estos métodos usuales y convencionales producían un efecto de encadenamiento (Aldenderfer y Blashfield, 1984) y originaban un único clúster al que se iban añadiendo uno a uno todos los casos. Esta situación impide determinar clúster, ya que si se decide aceptar dos clúster, éstos estarían formados uno de ellos por un sujeto y el otro por todos los restantes; si se decide tomar tres, dos de ellos estarían formados por un sujeto cada uno, y el tercer clúster por los restantes, etc.

A continuación presentamos el procedimiento seguido para minimizar las dificultades mencionadas. Por su interés metodológico vamos a reproducir las consideraciones que fundamentan cada una de nuestras decisiones.

3.3.1. Proceso de clusterización

En este procedimiento tomamos las puntuaciones factoriales de cada individuo para cada constructo y realizamos sucesivos análisis clúster para determinar la distancia a utilizar, el número apropiado de clúster y los sujetos que corresponden a cada clúster.

Podemos distinguir tres pasos en el procedimiento que vamos a describir:

- * primero, determinar el tipo de distancia;
- * segundo, determinar el número de clústers;
- * tercero, determinar qué sujetos pertenecen a cada clúster.

Pasamos a describir cada uno de ellos.

Primer paso

Hemos realizado tres análisis clúster jerárquicos por medio del *método de Ward*, puesto que es el que mejor minimiza los problemas de encadenamiento³.

Se han tomado las tres distancias (euclídea, euclídea al cuadrado y la de ciudad) que tiene el paquete estadístico Statgraphics Plus (versión 3.0 para Windows) para comprobar si se producen soluciones estables (lo cual indicaría la presencia de una estructura bien definida en los datos).

Así pues, el proceso consiste en determinar, en primer lugar, si las soluciones proporcionadas usando cada una de las tres distancias son parecidas. Ello lo hacemos con el coeficiente R_g (Jolliffe y otros 1986)⁴ que marca la proporción de *parecido* o *similitud* (porcentaje si se expresa como tal) entre dos clasificaciones.

3 Este método calcula la media de todas las variables de cada clúster, luego calcula la distancia entre cada individuo y la media de su grupo, sumando después las distancias de todos los casos. En cada paso, los clúster que se forman son aquéllos que resultan con el menor incremento en la suma total de las distancias intra-clúster (Bisquerra, 1989; Everit, 1993, Aldenderfer y Blashfield, 1984).

4 R_g es un coeficiente que mide el grado de correspondencia entre dos particiones de una misma población (con igual número de clúster k). Toma valores entre 0 y 1, y cuando la correspondencia entre las dos particiones es perfecta toma el valor 1. Entre sus propiedades se encuentra la de tender rápidamente a 1 cuando se hace crecer k (incluso para particiones no relacionadas), con lo que su utilidad se reduce a valores de k pequeños.

En nuestro caso, el nivel de parecido es bastante grande entre las soluciones cuando se considera distinto número de clúster, lo cual es indicativo de la presencia de una estructura bien definida de los datos. Por ello debemos ahora fijar una solución de las tres ensayadas, es decir, tenemos que optar por una de las tres distancias y, además, señalar posibles soluciones, esto es, posibles niveles de corte. La decisión corresponde en este caso a los investigadores y, por las características de nuestro trabajo, hemos considerado apropiado establecer entre 5 y 10 clúster.

En primer lugar seleccionamos la distancia que consideramos, sin olvidar tampoco los posibles niveles de corte de los dendogramas. Observamos que, tomando la distancia euclídea (L2) se consigue que el nivel de parecido con las otras dos distancias sea del 84.37% y 86.77% para 10 clúster (conforme aumenta el número es obvio que el parecido es mayor). Ello nos conduce a considerar adecuada la distancia euclídea.

Segundo paso.

Una vez que hemos decidido emplear el método jerárquico de Ward y la distancia euclídea, pasamos a señalar el punto de corte. Para ello, nos ayudamos del método de Beale (Milligan y Cooper, 1985) que representa los contrastes que se realizan para determinar si, por ejemplo, es mejor la partición en 8 clúster que en 7. No es tan importante el valor concreto del Pvalor como el hecho de que se observe un decrecimiento y/o aumento en su magnitud entre varias comparaciones de niveles. Observamos que los Pvalores no son significativos al principio pero van disminuyendo, lo cual es indicio de que va mejorando en sucesivas particiones. Señalemos que, estrictamente hablando, no apreciamos mejoría entre 5 y 9 particiones pero se ve un cambio sustancial de magnitud con 10 (de orden 10^{-2}). Así, ese primer cambio brusco nos *determina a tomar 10 clúster* en lugar de 9. Si no incluyéramos el 10º observamos que 9 no es preferible a 8 y así sucesivamente hacia arriba, lo cual causaría serios problemas de interpretación.

Para reafirmar esta opción sobre el nivel de corte hemos obtenido también los valores de los estadísticos empleados por Calinsky (Milligan y Cooper, 1985). En esencia, según este método, se toma entre las alternativas planteadas por nosotros, la partición con coeficiente C mayor. En éste caso es la de 10, en el rango 5-10.

Además hemos obtenido también estos resultados para las otras dos distancias. Esto lo hemos hecho puesto que, como las tres distancias aportan soluciones parecidas (véase el porcentaje anterior), sería bueno que los tres métodos (aunque ya nos hayamos decantado por uno) condujeran a iguales niveles de corte. Con los resultados obtenidos, podemos comprobar que la solución de 10 clúster es la óptima en el rango 5-10.

Así pues, todos estos análisis nos llevan a afirmar que, con las puntuaciones factoriales obtenidas por los sujetos, *tenemos mediante un método jerárquico (el de Ward) y la distancia euclídea una solución en 10 clúster.*

Tercer paso.

En esta etapa se realiza un análisis clúster no jerárquico mediante el método de las KMedias (Bisquerra, 1989; Everit, 1993, Aldenderfer y Blashfield, 1984), del paquete estadístico BMDP (Dixon, 1990). Este es un método bastante más interesante que los anteriores puesto que da directamente la partición de la población según el número de

clúster propuesto de antemano. Además, no está sujeto a estructura jerárquica como los anteriores; esto es, se trata de un método iterativo que puede intercambiar individuos entre clúster en las sucesivas etapas siempre y cuando ello ayude a mejorar la partición obtenida. Por otra parte, se basa en las distancias al centroide (vector de medias) al igual que el método de Ward. El problema en ese caso es que requiere que se fije de antemano un número de clúster concreto. Cuando no se dispone de ese número se suelen hacer ensayos con distintos valores y se compara *de visu* la estabilización de la partición.

Este no es nuestro caso puesto que disponemos de información anterior que nos dice que una división en 10 clúster puede ser óptima. Además disponemos en cada caso de una solución inicial.

El procedimiento que se sigue para el análisis no jerárquico es el siguiente:

Para cada cuestionario se hacen dos ejecuciones. En la primera se deja al programa que seleccione la partición inicial, y en la segunda se dice que tome como solución inicial la proporcionada en el apartado anterior. La idea que perseguimos es ver si las soluciones en ambos casos son parecidas, puesto que así tenemos una partición bastante estable. Las comparaciones se realizan de nuevo con el coeficiente Rg y son bastante altas.

Recogemos las configuraciones de los clúster, así como una tabla de medias de cada variable (en este caso cada factor) para cada clúster. Las tipologías las da la gráfica del final.

Nos decantamos por la opción sin solución inicial porque en ella el número de clúster claramente interpretables es muy superior (9 frente a 3).

Esta gráfica da la *tipología de los clúster obtenidos por este método*.

CLUSTER PROFILES - VARIABLES ARE ORDERED BY F-RATIO SIZE					
	*	*	*	*	*
FACTOR1	-1	-2	-3	-4	-5
FACTOR3	-1-	-2	-3-	-4-	5-
FACTOR13	-1-	-2	-3-	-4-	5-
FACTOR6	1-	2-	-3-	-4-	--5-
FACTOR4	-1-	-2-	-3-	-4	-5-
FACTOR9	1-	-2	-3-	-4-	-5-
FACTOR2	-1-	-2-	-3-	-4-	-5-
FACTOR12	-1-	-2-	-3-	-4-	--5-
FACTOR7	1-	-2	-3-	-4-	--5--
FACTOR11	--1-	-2-	-3-	-4-	--5---
FACTOR5	--1-	--2-	--3--	--4-	--5--
FACTOR10	-1-	---2--	-3-	-4-	----5----
FACTOR8	--1-	-2--	-3--	--4--	---5---
	*	*	*	*	*
	*	*	*	*	*
FACTOR1	-6	-7-	-8-	9-	A-
FACTOR3	-6-	7-	-8-	9-	A-
FACTOR13	6-	7	-8-	-9-	A-
FACTOR6	-6	-7-	-8-	-9	A-
FACTOR4	-6	7	-8-	-9	A-
FACTOR9	-6-	-7	-8-	-9-	A-
FACTOR2	-6-	7	-8-	9-	A-
FACTOR12	-6-	---7---	-8-	9-	A-
FACTOR7	-6-	-7--	-8-	-9	A-
FACTOR11	--6-	-----7-----	-8-	-9-	--A-
FACTOR5	--6--	---7---	-8-	-9--	--A-
FACTOR10	--6-	---7-	--8--	-9-	A-
FACTOR8	--6--	7	-8-	-9-	A-
	*	*	*	*	*

EACH COLUMN DESCRIBES A CLUSTER .

3.3.2. Descripción de los clusters

Primer Clúster CE1

Agrupa a 45 profesores, y es el clúster que incluye mayor número de sujetos (con el 27.6% de la muestra). Las medias de los sujetos que lo integran no presentan diferencias significativas en cada uno de los factores respecto de la media global de la muestra. Podemos interpretar que se trata de un grupo de profesores promedio de la población, es decir, un grupo cuya valoración de los factores coincide con las valoraciones medias de la población encuestada y, en este sentido, los sujetos que lo componen representan al factor general.

Estos profesores mantienen una preocupación prioritaria por los alumnos y las componentes curriculares como sujetos de evaluación, atendiendo a una diversidad de criterios, entre los que no destaca la valoración del conocimiento de los alumnos. También muestran cierta preocupación por la valoración del profesor y del centro. Establecer el criterio adecuado para evaluar a los distintos sujetos es su preocupación central.

Entienden que las finalidades prioritarias de la evaluación son tomar decisiones y controlar el proceso, pero sin que ello sea determinante. También sostienen unas determinadas creencias relativas a los aspectos técnicos: evaluadores internos, pruebas estandarizadas e informes cualitativos, pero sin que estas creencias sean determinantes para el clúster.

Los profesores del primer clúster muestran un prototipo de *profesor de matemáticas estándar respecto a la evaluación y mantienen la posición del constructo general*.

Segundo Clúster CE2

Este clúster lo conforman 16 profesores (9.8% de la muestra) que se distinguen por dar una valoración inferior, respecto del total de la población encuestada, al primer factor que habíamos interpretado como *Prioridad de la actitud, el interés y la conducta para evaluar a los alumnos y del trabajo personal para evaluar a alumnos y profesores*. Se trata pues de un grupo de profesores que da menos valor al interés y al trabajo personal como datos prioritarios para evaluar a los alumnos y profesores en matemáticas pero, en especial, consideran escasamente la conducta como indicador para evaluar a los alumnos. Son profesores para los que la conducta de los alumnos, su actitud e interés no son prioritarios como objetos de evaluación; igualmente parecen dar escaso valor al trabajo personal. En este caso, son las capacidades y el logro de los objetivos lo que parece establecer el criterio para la evaluación de los alumnos. En el resto de las cuestiones, los profesores de este clúster no parecen tener una caracterización distintiva.

Podemos describir así la tendencia que muestran los profesores del segundo clúster: *sostienen la creencia de que las actitudes y la conducta de los alumnos y el trabajo de los profesores no son un criterio importante para su evaluación, en especial en matemáticas. Consecuentemente, los alumnos se evalúan por sus capacidades y los logros alcanzados en la consecución de objetivos; los profesores por su profesionalidad*.

Tercer Clúster CE3

Está formado por 17 profesores (10.4% de la muestra) que sustentan la creencia: *Los alumnos tienen responsabilidad en las dificultades del proceso de evaluación, que no se considera de complejidad especial.*

Cuarto Clúster CE4

Este clúster lo constituyen 20 profesores (12.3% de la muestra) y su rasgo más característico es que muestran una *tendencia a conceder escaso interés a la formación de los profesores y al nivel de los alumnos para evaluar al centro.*

Quinto Clúster CE5

Este clúster lo forman 6 profesores (3.7% de la muestra), que se distinguen por compartir una *tendencia a valorar el interés y el trabajo de los alumnos, como objeto prioritario y considerar la orientación y promoción como sus principales finalidades.*

Sexto Clúster CE6

Agrupar 15 profesores (9.2% de la muestra), que se caracterizan por *centrar el objeto de evaluación en el logro de los objetivos y en las capacidades de los estudiantes y, secundariamente, en sus conocimientos pero no en sus actitudes. No consideran los aspectos curriculares e institucionales, aunque sí parecen dar un valor medio a la valoración del profesor, del libro de texto y el centro.*

Séptimo Clúster CE7

Este clúster agrupa a 2 sujetos (1.2% de la muestra) que se destacan por compartir que *los criterios para la evaluación del alumno se orientan hacia sus actitudes, capacidades y logro de objetivos; se considera al alumno responsable en el proceso de evaluación. Sobre el profesor parecen tener prioridad su profesionalidad y cualidades personales. Son profesores con creencias no convencionales sobre evaluación.*

Octavo Clúster CE8

Está formado por 14 profesores (8.6% de la muestra) y el rasgo que los *se orienta a considerar la conducta de los alumnos como criterio prioritario para su valoración. La profesionalidad del profesor y los contenidos del libro de texto parecen tener menor importancia.*

Noveno Clúster CE9

En este clúster quedan encuadrados 15 profesores (9% de la muestra), que presentan como rasgo *la tendencia que se muestra parece situar las dificultades sobre la complejidad del proceso y reafirman la creencia de que los evaluadores deben ser internos y las pruebas convencionales.*

Décimo Clúster CE10

Este clúster recoge a 13 profesores (8% de la muestra) y *la tendencia que se muestra parece situar como criterio de evaluación de los alumnos el logro de objetivos y como criterio para valorar a los profesores su formación científico-didáctica.*

Los diez clúster obtenidos presentan ciertas regularidades. Así, en relación con sus tamaños, podemos distinguir tres tipos: grande, mediano y pequeño:

	Grande	Mediano	Pequeño
Intervalo	≥ 25%	8% - 12%	≤ 5%
Clúster en cada tipo	CE1	CE2, CE3, CE4, CE6, CE8, CE9, CE10	CE5, CE7

3.3.3. Hallazgos del estudio clúster

El primer clúster, CE1, que comprende al 27% de la muestra, parece configurar la conceptualización global sobre evaluación. Esta tendencia la ponen de manifiesto un grupo de profesores que sostienen la concepción general.

Hay seis clúster cuya tendencia parece centrarse en diversas opciones sobre el criterio prioritario y las finalidades para evaluar a los alumnos. Así, apreciamos en ellos las siguientes creencias:

- CE2: los alumnos se evalúan por sus capacidades y los logros alcanzados en la consecución de objetivos;
- CE5: valorar el interés y el trabajo de los alumnos, como objeto prioritario y considerar la orientación y promoción como sus principales finalidades;
- CE6: centrar el objeto de evaluación en el logro de los objetivos y en las capacidades de los estudiantes, secundariamente en sus conocimientos pero no en sus actitudes;
- CE7: los criterios para la evaluación del alumno se orientan hacia sus actitudes, capacidades y logro de objetivos; se considera al alumno responsable en el proceso de evaluación;
- CE8: considerar la conducta de los alumnos como criterio prioritario para su valoración;
- CE10: sitúa como criterio de evaluación de los alumnos el logro de objetivos.

Hay tres clúster cuyas tendencias muestran diversas opciones sobre el criterio prioritario para evaluar a los profesores. Así, apreciamos las siguientes creencias:

- CE2: los profesores se evalúan por su profesionalidad;
- CE7: tienen prioridad la profesionalidad y cualidades personales del profesor;
- CE10: la formación científico-didáctica de los profesores es el criterio para su valoración.

Hay un clúster cuya tendencia se sostiene en una determinada creencia sobre la evaluación del centro, CE4: escaso interés en la formación de los profesores y en el nivel de los alumnos para evaluar al centro.

Hay un clúster cuya tendencia de pensamiento se sostiene sobre una determinada creencia sobre los aspectos técnicos de la evaluación, CE9: reafirman la creencia de que los evaluadores deben ser internos y las pruebas convencionales.

Hay dos clúster cuya tendencia se sostiene en una determinada creencia sobre las dificultades del proceso de evaluación:

- CE3: los alumnos tienen responsabilidad en las dificultades del proceso de evaluación, que no se considera de complejidad especial;
- CE9: sitúa las dificultades sobre la complejidad del proceso.

Los clúster considerados muestran diversas tendencias de pensamiento de los profesores que se articulan en torno a las creencias sobre evaluación.

4. CONCLUSIONES

Hemos constatado que todos los enunciados propuestos en el cuestionario son aceptados por el profesorado y valorados como alternativas posibles a las cuestiones planteadas.

Se consigue así establecer un sistema de ideas comunes sobre evaluación, compartidas por los profesores de matemáticas, es decir, un estado de opinión. También hemos determinado un conjunto de enunciados en los que el profesorado manifiesta un cierto disenso o desacuerdo; son ideas sobre las que los profesores mantienen posiciones encontradas, y que suelen referirse a nuevos planteamientos sobre la evaluación.

El empleo de distintas técnicas estadísticas para el análisis de las respuestas al cuestionario de escala de valoración cerrado ha servido para establecer y estudiar el estado de opinión de los profesores y sus concepciones y creencias. La selección, coordinación y empleo progresivo de estas técnicas ha sido uno de los logros de la investigación. Como consecuencia de la aplicación de estas técnicas hemos realizado una interpretación de los factores que surgen del cuestionario de escala de valoración; se establecen así concepciones y creencias de los profesores sobre evaluación en matemáticas.

En nuestro estudio de los factores que estructuran las ideas y juicios expresados sobre evaluación por los profesores de matemáticas de secundaria, hemos encontrado un factor general sobre evaluación que caracteriza tales ideas y juicios, lo cual indica que existe una dimensión global en este constructo. El estudio permite establecer una concepción sobre evaluación en matemáticas, que expresa las ideas generales que comparten los profesores de la muestra. Esta concepción se sostiene sobre una visión convencional de la evaluación, que mantienen la generalidad de profesores de matemáticas, pero también en ideas que surgen de nuevos planteamientos curriculares con distinto grado de aceptación entre el profesorado.

La concepción general viene matizada por distintas creencias, que muestran diferentes criterios a la hora de establecer el contenido y las finalidades de la evaluación. Se detectan diversos sistemas de ideas coherentes, sostenidos por grupos reducidos de profesores. No podemos hablar de un conocimiento homogéneo y organizado de los profesores de matemáticas sobre evaluación. Aunque se sostiene y puede justificarse

sobre un marco conceptual elaborado, los distintos sujetos no lo suelen presentar estructurado formalmente, ni basarlo en conceptos y definiciones bien establecidos. En sus manifestaciones singulares, se trata de un conocimiento parcial, influenciado por opiniones y experiencias personales; es decir, también hemos logrado poner de manifiesto algunas creencias.

De los agrupamientos de profesores detectados y de las caracterizaciones realizadas parece desprenderse que distintos grupos de profesores sostienen diferentes creencias sobre los mismos constructos; esto se muestra porque se han detectado diferentes tendencias de pensamiento en relación con las concepciones establecidas sobre evaluación.

De esta manera se pone de manifiesto que las diversas creencias de los profesores de matemáticas sobre evaluación no son individuales; son grupos de profesores, más o menos amplios, quienes las sostienen de manera coherente y las comparten. En este sentido es posible hablar de tendencias en el pensamiento de los profesores.

REFERENCIAS

- Aldenderfer, A. y Azen, S. (1991). *Statistical analysis. A computer oriented approach*. London: Academic Press.
- Begle, E. (1968). Curriculum Research in Mathematics. En R. Ashlock y W. Jerman (Edts.) *Current Research in Elementary School Mathematics*. New York: Mcmillan.
- Buendía, L.; González, D. y Carmona, M. (1999). «Procedimientos e instrumentos de evaluación en educación secundaria». *Revista de Investigación Educativa*, 17 (1), 215-236.
- Bisquerra, R. (1989). *Introducción conceptual al análisis multivariante. Un enfoque informático de los paquetes SPSS-X, BMDP, LISREL y SPAD*. Barcelona: PPU.
- Castro, E.; Rico, L.; Gutiérrez, J., Castro, E.; Segovia, I.; Morcillo, N.; Fernández, F.; González, E. y Tortosa, A. (1996). «Evaluación de la resolución de problemas aritméticos en Primaria». *Revista de Investigación Educativa*, 14 (2), 121-140.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Hillsdale, N.J.: Lawrence Erlbaum.
- Dixon, J. (1990). *BMDP Statistical Software* (Vols. 1 y 2). Berkeley CA: University of California Press.
- Everit, B. (1980). *Cluster Analysis*. New York: Halsted Press.
- Fernández Cano, A. (1995). *Métodos para evaluar la investigación en Psicopedagogía*. Madrid: Síntesis.
- García, M. (1997). *Análisis del Conocimiento Profesional del Profesor de Matemáticas de Enseñanza Secundaria y el Concepto de Función como objeto de enseñanza-aprendizaje. Aportaciones metodológicas*. Sevilla: Universidad de Sevilla.
- Gil, F. (2000). *Marco Conceptual y Creencias de los Profesores sobre la Evaluación en Matemáticas*. Almería: Universidad de Almería.
- Gil, F.; Moreno, M. F.; Olmo, M.A. y Fernández, A. (1997). Elaboración de cuestionarios para determinar las creencias de los profesores. *UNO Revista de Didáctica de la Matemática*. 11, 43-54.
- Hedges, L.V. y Olkin, I. (1985). *Statistical methods for meta-analysis*. Londres: Academic Press.

- Hidalgo, E. (1996). La evaluación en la Educación Secundaria Obligatoria en el nuevo sistema educativo en Andalucía. *Revista de Investigación Educativa*, 14 (2), 51-72.
- Houston, R. (Edt.) (1990). *Handbook of Research on Teacher Education*. New York: Mcmillan.
- Howson, G.; Keitel, C. y Kilpatrick, J. (1981). *Curriculum Development in Mathematics*. Cambridge: Cambridge University Press.
- Jolliffe, I.; Jones, B. y Morgan, B. (1986). «Comparison of Cluster Análisis of the English Personal Social Services Authorities». *Journal S. of Statistical Society A.*, 149 (3), 253-270.
- Llinares, S. (1998). «La investigación «sobre» el profesor de matemáticas: aprendizaje del profesor y práctica profesional». *Aula*, 10, 153-179.
- Pajares, M. (1992). «Teachers' Beliefs and Educational Research: Cleaning up a Messy Construct». *Review of Educational Research*, 62 (39), 307-332.
- Ponte, J.P. (1994). «Mathematics Teachers' professional knowledge». En J. Ponte y J. Matos (eds.) *Proceedings of the Eighteenth International Conference for the Psychology of Mathematics Education*. Lisboa: International Group for the Psychology of Mathematics Education.
- Ponte, J.; Monteiro, L.; Maia, M.; Serrazina, L.; Loureiro, C. (1996). *Desenvolvimento Profissional dos Professores de Matemática, Que Formação?* Lisboa: Sociedade Portuguesa de Ciências de Educação.
- Rico, L.; Castro, E.; Castro, E.; Fernández, F.; Gil, F.; Moreno, M.F.; Olmo, M.A. y Segovia, I. (1995a). Conceptualizaciones sobre evaluación del Profesorado de Matemáticas. En L. Blanco y V. Mellado (Edts.) *La Formación del Pofesorado de Ciencias y Matemáticas en España y Portugal*. Badajoz: Diputación Provincial.
- Rico, L.; Castro, E.; Castro, E.; Fernández, F.; Gil, F.; Moreno, M.F.; Olmo, M.A. y Segovia, I. (1995b). Teacher's Conceptual Framework on Mathematics Assessment. *Proceedings of the Nineteenth International Conference for the Psychology of Mathematics Education*. Vol. II (130-137) Recife (Brasil): International Group for the Psychology of Mathematics Education.
- Rico, L. y Gil, F. (1997). Teachers' beliefs and implicit theories about mathematics assessment. *First Mediterranean Conference Mathematics Education and Applications* (255-268). Nicosia (Chipre).
- Rico, L. y Sierra, M. (1997). Antecedentes del Currículo de Matemáticas. En L. Rico (coord.) *Bases Teóricas del Currículo de Matemáticas en Educación Secundaria*. Madrid: Síntesis
- Romberg, T. (1989). Evaluation: A coat of many colours. En D. Robitaille (Edt.) *Evaluation and Assessment in Mathematics Education*. París: Unesco.
- Spector, P.E. (1992) *Summated Rating Scale Construction: An Introduction*. Newbury Parks CA: Sage.
- Thompson, A. (1992): Teachers' Beliefs and Conceptions: A Synthesis of the Research. En D. Grouws (Edt.) *Handbook of Research on Mathematics Teaching and Learning*. New York: Macmillan.
- Webb, N. (1992): Assessment of Student's Knowledge of Mathematics: Steps toward a Theory. En D. Grouws (Edt.) *Handbook of Research on Mathematics Teaching and Learning*. New York: Macmillan.
- Wright, B.D. y Masters, G.N. (1982): *Rating Scale Analysis*. Chicago: Mesa.

Anexo

Cuestionario cerrado de escala de valoración:

1. ¿Qué debe ser objeto de evaluación?	
En evaluación es prioritario:	
valorar el conocimiento adquirido por los alumnos	1 2 3 4 5 6 7 8 9
valorar el trabajo realizado por los alumnos	1 2 3 4 5 6 7 8 9
valorar la actitud y el interés de los alumnos	1 2 3 4 5 6 7 8 9
valorar las capacidades de los alumnos	1 2 3 4 5 6 7 8 9
valorar la conducta de los alumnos	1 2 3 4 5 6 7 8 9
valorar el currículo	1 2 3 4 5 6 7 8 9
valorar la labor del profesor	1 2 3 4 5 6 7 8 9
valorar la madurez y formación del alumno	1 2 3 4 5 6 7 8 9
valorar los contenidos	1 2 3 4 5 6 7 8 9
valorar los logros alcanzados respecto de los objetivos	1 2 3 4 5 6 7 8 9
valorar los medios y materiales	1 2 3 4 5 6 7 8 9
valorar las instituciones y el sistema educativo	1 2 3 4 5 6 7 8 9
2. ¿Por qué evaluar?	
Se evalúa para:	
obtener información sobre los alumnos y su aprendizaje	1 2 3 4 5 6 7 8 9
tomar decisiones sobre la promoción y orientación de los alumnos	1 2 3 4 5 6 7 8 9
controlar en relación con el proceso y el resultado	1 2 3 4 5 6 7 8 9
3. ¿Quién debe evaluar a los alumnos?	
La evaluación debe ser realizada por:	
evaluadores internos al aula	1 2 3 4 5 6 7 8 9
evaluadores externos al aula	1 2 3 4 5 6 7 8 9
4. ¿Qué instrumentos se deben utilizar?	
Para evaluar hay que:	
utilizar exámenes, observaciones y actividades de aula	1 2 3 4 5 6 7 8 9
utilizar tests estandarizados y pruebas generales	1 2 3 4 5 6 7 8 9
5. ¿Cómo deben expresarse los resultados de la evaluación?	
Para transmitir los resultados de la evaluación hay que dar prioridad	
a la comunicación oral	1 2 3 4 5 6 7 8 9
a la comunicación escrita	1 2 3 4 5 6 7 8 9
al informe de tipo cualitativo	1 2 3 4 5 6 7 8 9
al informe de tipo cuantitativo	1 2 3 4 5 6 7 8 9
6. ¿Qué aspectos deben evaluarse en matemáticas?	
En matemáticas es prioritario:	
evaluar el conocimiento adquirido por los alumnos	1 2 3 4 5 6 7 8 9
evaluar el trabajo realizado por los alumnos	1 2 3 4 5 6 7 8 9

evaluar la actitud y el interés de los alumnos hacia la asignatura	1 2 3 4 5 6 7 8 9
evaluar las capacidades de los alumnos	1 2 3 4 5 6 7 8 9
evaluar la conducta de los alumnos	1 2 3 4 5 6 7 8 9
evaluar los contenidos	1 2 3 4 5 6 7 8 9
evaluar los logros alcanzados respecto de los objetivos	1 2 3 4 5 6 7 8 9
evaluar medios y materiales	1 2 3 4 5 6 7 8 9
7. ¿Qué dificultades plantea la evaluación en matemáticas?	
En matemáticas, las dificultades de la evaluación:	
son debidas a la insuficiente preparación del profesor	1 2 3 4 5 6 7 8 9
son debidas al alumno	1 2 3 4 5 6 7 8 9
son debidas a los instrumentos utilizados	1 2 3 4 5 6 7 8 9
son debidas a la complejidad del proceso	1 2 3 4 5 6 7 8 9
8. ¿Qué criterios consideras importantes para valorar el libro de matemáticas?	
Para valorar el libro de matemáticas:	
el criterio prioritario es la presentación	1 2 3 4 5 6 7 8 9
el criterio prioritario es el desarrollo de los contenidos	1 2 3 4 5 6 7 8 9
9. ¿Qué aspectos deben evaluarse en un profesor de matemáticas?	
El profesor de matemáticas:	
se valora por sus cualidades personales	1 2 3 4 5 6 7 8 9
se valora por su formación científica y didáctica	1 2 3 4 5 6 7 8 9
se valora por su profesionalidad	1 2 3 4 5 6 7 8 9
10. ¿Qué interesa evaluar sobre los centros en relación con la educación matemática?	
Respecto a la educación matemática, el centro:	
se valora por su organización	1 2 3 4 5 6 7 8 9
se valora por su proyecto	1 2 3 4 5 6 7 8 9
se valora por su equipo de profesores	1 2 3 4 5 6 7 8 9
se valora por el nivel de sus alumnos	1 2 3 4 5 6 7 8 9

Fecha de recepción: 6 de marzo de 2000.

Fecha de aceptación: 6 de septiembre de 2001.