

Usos y formas de la evaluación por parte del profesorado de la ESO. Un modelo multinivel¹

Jose F. Lukas, Karlos Santiago, Luis Joaristi , Luis Lizasoain

Universidad del País Vasco-Euskal Herriko Unibertsitatea²

Resumen

En este trabajo se analizan los usos y las formas de evaluar de los profesores de la Educación Secundaria Obligatoria (ESO) de la Comunidad Autónoma del País Vasco (CAV). Durante cuatro cursos académicos (1999-00/2002-03) a una muestra de docentes de la ESO se les ha aplicado un cuestionario sobre procesos de centro y aula. Con las respuestas aportadas se describen los procesos evaluativos de los docentes para a continuación realizar un estudio diferencial que, dada la estructura anidada de los datos, se lleva a cabo mediante modelización multinivel. Los resultados obtenidos apuntan a que, en general, los profesores asumen y ponen en práctica las directrices de la LOGSE en lo que a la evaluación se refiere, aunque usos más académicos son igualmente bastante frecuentes (evaluación normativa, el profesor como único agente evaluador). Los modelos propuestos en el estudio diferencial apuntan a que las covariables más importantes son la antigüedad del profesorado y, sobre todo, el ciclo de la ESO en que ejerce.

Palabras clave: educación secundaria, profesorado, evaluación de los alumnos, tipos de evaluación, modelos multinivel.

Abstract: *Types of assessment and their uses by compulsory Secondary Education teachers. A multilevel model*

This article deals with the analysis of the different types of assessment and their uses by compulsory Secondary Education teacher in the autonomous community of the Basque

⁽¹⁾ Esta investigación se ha podido realizar gracias a la financiación del Instituto de Evaluación y Asesoramiento Educativo (IDEA), del Departamento de Educación del Gobierno Vasco y de la Universidad del País Vasco-Euskal Herriko Unibertsitatea (proyecto UPV 0218.230-H-15288).

⁽²⁾ Este cuestionario ha sido construido por el Instituto de Evaluación y Asesoramiento Educativo (IDEA) (REDES, 1999).

Country. A sample of Secondary Education teacher filled in a questionnaire on school and classroom processes throughout four school years (1999-00/2002-03). Then, teacher's assessment processes are described according to the resulting to the data. Afterwards, taking into account the interrelated structure of the data, a differential analysis is carried out by means of a multilevel modelation. The outcomes come to show that teachers usually adopt and put into practice LOGSE guidelines concerning assessments, although more academic uses are also quite frequent (for instance, normative assessment and the teacher as the only evaluator agent). The models proposed in the differential study show that the most significant covariables are teacher's seniority and, specially, the cycle of compulsory Secondary Education in which s/he teaches.

Key words: secondary education, teachers, student evaluation, evaluation types, multilevel models.

Introducción

Durante los últimos años hemos asistido a una serie de cambios educativos amparados en diferentes leyes que se han ido aprobando y que han tenido repercusiones directas en los usos y formas de evaluar de los profesores.

Hasta la aprobación en 1970 de la Ley General de Educación (LGE), la evaluación de los alumnos no era una responsabilidad directa del profesor. Con la aprobación de esa Ley esta tarea pasa a recaer en el docente. Pero además, la LGE supuso un avance significativo en el quehacer evaluativo debido entre otras razones a:

- La aparición del término evaluación por primera vez en la legislación educativa española.
- La orientación de la evaluación hacia la mejora, incorporando la recuperación como estrategia para la misma.
- La ampliación del objeto evaluativo también a otros ámbitos educativos.
- La consideración de otros aspectos a evaluar además del rendimiento académico del alumnado.
- La necesidad de la evaluación continua y orientadora.

Sin embargo, esa Ley, basada en muchos aspectos en lo que a la evaluación se refiere en la ideas de Tyler (1950), se centraba en la evaluación por objetivos, dándose

excesiva importancia a los resultados finales y olvidándose del proceso. Por ello, aunque en la LGE se hable de la función formativa, ésta es difícil que se dé si sólo se evalúa al final (Lukas y Santiago, 2004a).

Durante los 20 años de vigencia de la misma se constató que la evaluación se centró exclusivamente en el rendimiento académico de los alumnos y, tal y como señalan Sáez Brezmes y Carretero (1995), los conceptos novedosos de la evaluación se convirtieron en términos de una retórica oficial asumida, pero nunca utilizada en su profundidad. Quizás, las carencias formativas del profesorado para garantizar una evaluación continua y formativa también contribuyeron a que las potencialidades de la LGE no se desarrollaran.

Con la aprobación en 1990 de Ley de Ordenación General del Sistema Educativo (LOGSE) se planteó un cambio educativo sin precedentes en España. Este cambio fue acompañado de variaciones en lo que a la evaluación se refiere:

- Se subraya el carácter formativo de la evaluación.
- Se amplía el agente dejando de ser el profesor el único evaluador.
- Se amplía el objeto evaluado tomándose en consideración además del alumno, el profesor, los procesos, el proyecto curricular, el funcionamiento del centro, etc.
- Se amplían los contenidos a evaluar, asumiéndose como importantes los conceptos, los procedimientos y las actitudes.
- Se pasa de una evaluación normativa a una evaluación criterial e idiográfica.
- Se amplían las estrategias utilizadas para la recogida de información desechándose el examen tradicional como único instrumento de evaluación.

Por último, en el año 2002 se aprobó la Ley de Calidad (LOCE) en la que se define la evaluación como la identificación de los errores y de los aciertos como un factor básico de calidad que constituye un instrumento para hacer políticas educativas inteligentes y para incrementar su oportunidad y su adecuación a los cambios. Esta ley quiere introducir una serie de cambios en lo que se refiere a los propósitos de la evaluación y la utilización de ésta de una manera diferente a la propuesta en la LOGSE (Lukas y Santiago, 2004a). En el segundo eje de medidas que propone la Ley se declara claramente «en orientar más abiertamente el sistema educativo hacia los resultados, pues la consolidación de la cultura del esfuerzo y la mejora de la calidad están vinculadas a la intensificación de los procesos de evaluación de los alumnos, de los profesores, de los centros y del sistema en su conjunto, de modo que unos y otros

puedan orientar convenientemente los procesos de mejora». Es decir, si en la LOGSE se acentuaba la importancia de los procesos, en esta Ley se acentúa la importancia de los resultados. En este sentido, la evaluación idiográfica o personalizada que se aconsejaba en la LOGSE desaparece por completo dado que el único criterio utilizado para evaluar a los alumnos tanto en Educación Primaria como en Secundaria son los objetivos específicos establecidos en el currículo.

Los alumnos accederán al ciclo siguiente si han alcanzado los objetivos correspondientes establecidos en el currículo. Cuando un alumno no haya alcanzado los objetivos, podrá permanecer un curso más en el mismo ciclo (Educación Primaria).

La evaluación del aprendizaje de los alumnos en la Educación Secundaria Obligatoria será continua y diferenciada según las distintas asignaturas del currículo... cuando el número de asignaturas no aprobadas sea superior a dos, el alumno deberá permanecer otro año en el mismo curso (Educación Secundaria).

No obstante, en el momento en que se escribe este texto, la situación de esta Ley se encuentra a expensas de la anunciada suspensión de su aplicación por parte del nuevo Gobierno.

Ante esta situación, este trabajo se plantea como finalidad analizar cuáles son los usos y las formas de evaluar de los profesores de la Educación Secundaria Obligatoria (ESO) del País Vasco. Hace unos años (Lukas et al., 2001) presentamos algunos datos preliminares sobre este estudio que ahora se ofrece completado.

Pensamos que ésta es una cuestión importante pues la forma en que los profesores evalúan a sus alumnos y el uso que de los resultados de la misma hacen es uno de los indicadores más claros de la acción educativa de los docentes.

En esta misma IDEA abundan Coll, Martín y Onrubia (2001) cuando plantean que el conocimiento de las prácticas evaluativas de los docentes es una buena vía para conocer la manera de entender la educación que los mismos tienen.

En la evaluación de Secundaria realizada por el equipo REDES (Marchesi y Martín, 2002) ésta es una de las cuestiones que se abordan a la hora de examinar los procesos de aula (Martín, 2002). Struyf y otros (2001) estudiaron las prácticas evaluadoras de 132 profesores belgas en 18 centros de secundaria y concluyeron caracterizando a los profesores en tres grupos según su práctica estuviera orientada preferentemente hacia el juicio, la solución de los problemas o los procesos.

En nuestro caso, en la evaluación externa de centros de la eso que se ha llevado a cabo en la Comunidad Autónoma Vasca durante los cursos 1999/00 a 2002/03 (Lukas et al., 2000; Lukas y Santiago, 2004b) se ha aplicado un cuestionario a los profesores de los 55 centros participantes. Este cuestionario, denominado «Procesos de centro y aula»³ consta de 90 ítems tipo Likert de siete grados que pretenden medir diferentes dimensiones.

Coincidimos con Martín (2002) en que la observación, la entrevista o el análisis de documentos serían procedimientos adecuados para comprender en profundidad la actividad evaluadora de los docentes. Pero en este caso se trata más bien de tratar de caracterizar los diferentes tipos y usos de dicha actividad. Y para ello, y para obtener información de un conjunto numeroso de sujetos, el cuestionario resulta la herramienta adecuada.

Somos conscientes de que este tipo de herramientas tiene sus limitaciones entre las que no queremos dejar de citar el efecto de la deseabilidad social máxime si consideramos que a los docentes se les solicita que expresen su grado de acuerdo o desacuerdo con una serie de conductas o afirmaciones. Como veremos en las conclusiones, es plausible pensar que esta cuestión afecte a las respuestas que se dan a determinados ítems y que, en consecuencia, en algunos casos parezca que determinados discursos están teóricamente asumidos mientras que la praxis cotidiana va por otros derroteros.

De cualquier forma, en el instrumento que hemos empleado hay 22 ítems dedicados a medir las percepciones y usos que hacen los profesores de la eso con respecto a la evaluación de los alumnos (Lukas et al., 2001). Estos 22 ítems a su vez, están agrupados en 4 dimensiones:

- **Funcionalidad/temporalidad de la evaluación.** Se refiere al uso que hacen los profesores de los distintos momentos evaluativos. A saber, si realizan una evaluación inicial, procesual, continua, final, etc. De la misma manera, se les interroga sobre la función que cumple la evaluación que ellos realizan sobre los alumnos. Es decir, si utilizan la evaluación con carácter diagnóstico (conocimiento del punto de partida para adecuar el proceso de enseñanza-aprendizaje), con carácter formativo (ir recogiendo información sobre el proceso de enseñanza-aprendizaje para ir mejorando el propio proceso) o con carácter sumativo (recoger información al final del proceso para ratificar y sancionar los logros obtenidos).

³⁾ En el curso 2002/03 no se aplicaron algunos de los ítems por lo que en la casilla correspondiente no aparece la media. Lógicamente, en tales casos la ponderación se ha ajustado.

- **Normotipo utilizado.** En esta dimensión se pregunta por el estándar, norma o referente utilizado para discriminar las realizaciones aceptables de las no aceptables. En función de este criterio se suele hablar de tres tipos de evaluación, aunque en ocasiones la terminología utilizada por diversos autores suele variar. Así se distingue entre evaluación nomotética (cuando el referente es externo al propio sujeto) y evaluación idiográfica o personalizada (cuando el referente es el propio sujeto evaluado, es decir, su situación anterior). Dentro de la evaluación nomotética se puede establecer la diferencia entre la evaluación normativa (cuando un sujeto es evaluado positiva o negativamente en función de la posición relativa que ocupa dentro de su grupo de referencia) y la evaluación criterial (cuando se tiene en cuenta la realización absoluta del sujeto, es decir, se establecen de antemano cuáles son los estándares que debe alcanzar).
- **El agente evaluador.** En estos ítems el profesorado responde sobre quién es el encargado de evaluar a los alumnos. Como se mencionó previamente, si bien hasta la entrada en vigor de la logse la responsabilidad era única y exclusivamente del profesor (por lo tanto estaríamos hablando de la heteroevaluación), a partir de ésta, se aboga por la utilización también de la autoevaluación y de la coevaluación (evaluación entre iguales o de forma conjunta).
- **Contenidos de la evaluación.** Aquí se pregunta a los profesores sobre la importancia que, a la hora de ser evaluados, tienen los distintos contenidos (conceptos, procedimientos y actitudes) así como los temas transversales.

En su conjunto la prueba es reflejo del enfoque constructivista aunque dentro de los 22 ítems considerados hay algunos que presentan una forma más tradicional de plantear la evaluación. En concreto son las cuestiones 6, 9, 10, 11 y 19.

Metodología

Cuestiones de investigación

En la presente investigación se ha pretendido dar respuesta a las siguientes cuestiones:

- ¿Cuáles son los procesos de evaluación utilizados por el profesorado de la ESO durante los cuatro años de la evaluación? Hemos querido determinar cuáles son los usos que da dicho profesorado a la evaluación que realiza de sus alumnos.
- ¿Ha habido alguna evolución a lo largo de estos cuatro cursos en los usos dados a la evaluación por parte del profesorado de la ESO de la Comunidad Autónoma Vasca?
- ¿Qué diferencias se dan en los usos de la evaluación que hace el profesorado de la ESO en función de las siguientes variables?:
 - El ciclo de la ESO en el que imparte docencia el profesorado.
 - La antigüedad docente del profesorado.
 - El contexto socioeconómico y cultural de los centros.
 - La titularidad del centro.

Se trata de tres cuestiones que responden a tres enfoques distintos. El primero tiene una pretensión fundamentalmente descriptiva; el segundo es un enfoque longitudinal que, como luego veremos, sólo vamos a apuntar; y, por último, el tercero es un estudio multinivel en el que distinguimos dos niveles, el primero (N-1) el de los profesores donde se incorporan las variables ciclo y antigüedad, y el segundo (N-2) el de los centros escolares donde se incluyen el contexto y la titularidad.

Variables

En función de lo antedicho, las variables con las que se va a operar son las siguientes:

Variable Dependiente: en este caso planteamos la existencia de una variable dependiente que denominaremos «Usos y formas de la Evaluación por parte del profesorado de la ESO de la Comunidad Autónoma Vasca». Esta variable está compuesta por 22 ítems que responden a formas y usos diferentes de la evaluación del alumnado y que son analizados independientemente (cada ítem es una puntuación) sin que tenga sentido una puntuación total o única para toda la variable ni subtotales correspondientes a las agrupamientos de ítems originados de la aplicación de un análisis factorial pues no hay una solución factorial coherente.

Sin embargo y de cara a la interpretación de los resultados se han hecho agrupaciones de ítems considerando las cuatro dimensiones que antes apuntábamos:

- Funcionalidad/temporalidad de la evaluación (ítems 1-9).
- Normotipo utilizado (ítems 10-12).
- Agente evaluador (ítems 13-18).
- Contenidos evaluados (ítems 19-22).

VARIABLES INDEPENDIENTES. Asumiendo el enfoque multinivel que anteriormente hemos señalado, las variables independientes consideradas en esta investigación han sido las siguientes:

Nivel 1: de los docentes

- Ciclo de la ESO. Aquí distinguimos entre el profesorado que imparte docencia en el primer ciclo de la ESO y aquel que lo hace en el segundo. Los primeros están mayoritariamente diplomados en Magisterio y se incorporaron a la secundaria con la implantación de la LOGSE. El segundo grupo está formado por profesorado licenciado.
- Antigüedad docente. En función de los años de docencia del profesorado, éste se ha clasificado en tres grupos: profesorado de menos de cuatro años de docencia, entre cuatro y diez años y profesorado con más de diez años de docencia.

Nivel 2: de los centros

- Contexto socioeconómico y cultural del centro. A partir de un cuestionario construido al efecto, se distingue entre centros de contexto alto, medio-alto, medio-bajo y bajo. El criterio básico es el del nivel socioeconómico y cultural de las familias.
- Titularidad del centro. Los 55 centros que han participado en esta evaluación están clasificados en dos grupos: centros públicos y centros privados concertados.

Muestra

En la Tabla I puede observarse el número de docentes y de centros que han respondido al cuestionario en cada uno de los cursos académicos. Lo primero que hay que dejar claro es que no es una muestra representativa pues no ha sido elegida al azar ya que la incorporación de los centros al proceso de evaluación era voluntaria. Una de las limitaciones de este estudio está relacionada con la variabilidad entre los diferentes cursos académicos en cuanto al número de profesores y centros que contestaron.

TABLA I. Muestra total de profesores y centros

	1999/2000	2000/2001	2001/2002	2002/2003
Nº docentes	1266	529	869	900
Nº centros	55	26	45	42

TABLA II. Distribución de la muestra en función de las distintas variables

			1999/2000	2000/2001	2001/2002	2002/2003
Nivel I (Profesores)	Titulación	Magisterio	664	304	472	496
		Licenciado	591	225	397	404
	Antigüedad docente	Menos de 4	199	75	84	87
		4-10	287	107	199	196
		Más de 10	674	308	531	561
Nivel 2 (Centros)	Titularidad	Público	482	234	346	348
		Privado	784	295	523	552
	Contexto	Alto	243	107	118	213
		Medio-alto	306	123	254	181
		Medio-bajo	433	193	287	285
		Bajo	284	106	210	220

En cuanto a la distribución de la muestra teniendo en cuenta las distintas variables independientes y los cuatro cursos académicos, los datos se encuentran en la Tabla II presentada a continuación.

Aunque no se puede considerar representativa desde el punto de vista estrictamente estadístico, esta muestra nos permite obtener una visión bastante fidedigna de las formas y usos de la evaluación por parte de las y los docentes de Secundaria de la Comunidad Autónoma Vasca ya que, en términos globales, se dispone de alrededor de 3.500 respuestas.

Pero aquí nos topamos también con otra limitación. Dado que el cuestionario aplicado era sobre procesos de centro y aula, un requisito importante era garantizar el anonimato en las respuestas por lo que no se dispone de identificadores de los casos.

Es altamente probable que a lo largo de estos cuatro cursos haya docentes que hayan respondido más de una vez por lo que el supuesto de independencia de los casos no se verifica. Por ello el total de 3.500 debe ser considerado de forma muy matizada y nunca como total de casos sino de respuestas. Por esto los resultados se van a presentar habitualmente curso por curso de forma que entonces sí se pueda considerar que los mismos se refieren a docentes.

Análisis estadísticos

A la hora de seleccionar las técnicas estadísticas más apropiadas para dar respuesta a las cuestiones de investigación que han sido planteadas hay que considerar aspectos como el diseño empleado y la información disponible.

Con respecto a la primera cuestión, el enfoque va a ser fundamentalmente descriptivo empleando los recuentos de frecuencias y los índices habituales.

El estudio diferencial lo vamos a abordar mediante las técnicas estadísticas de modelización multinivel o jerárquica dado que los docentes están anidados en los centros y que se conoce el centro de pertenencia de cada uno.

Con respecto al estudio estrictamente longitudinal, no es posible el planteamiento de un modelo de estas características ya que no se dispone de un sistema de identificación de los docentes que permita fusionar los registros correspondientes a los distintos docentes a lo largo de los cursos. Por ello resulta imposible estudiar el cambio individual, el de los docentes.

En un trabajo posterior esta cuestión será planteada a nivel de centro, de manera que se podrá concluir si se ha producido o no una evolución global y en qué tipo de centros en función de su Titularidad y Contexto, así como en características globales de su profesorado referidas a Antigüedad docente y Ciclo de la ESO en que imparten docencia.

- Como anteriormente apuntamos, las covariables o variables independientes del N-1, es decir las correspondientes a los docentes, son:
 - Ciclo de la ESO, tratada como binaria (1 = el docente ejerce en el primer ciclo de la ESO; 0, no ejerce) con el fin de comparar de forma sencilla los modelos de los que ejercen y de los que no.
 - Antigüedad docente, centrada en la media del Centro, pues interesa interpretar la intersección de la ecuación de N-1 (β_0) como el valor medio de los ítems de los docentes de cada centro, controlando la Antigüedad docente (Singer, 1998).
- Las covariables correspondientes a las unidades del N-2, los Centros, son:
 - Contexto socioeconómico y cultural del Centro, la cual se centra en la media global o gran media, con el fin de interpretar la intersección de su ecuación como la puntuación en los distintos ítems en un centro de Contexto medio.
 - Titularidad del Centro, tratada como binaria (1 = Centros privados) para comparar los modelos de los centros públicos y privados.

- A estas dos hemos añadido la Antigüedad media de cada Centro, con el fin de corregir los efectos de una incorrecta especificación del modelo de N-1 (Raudenbush, Bryk, 2002).

La estrategia de obtención del modelo se basa en formular inicialmente el modelo incondicional, que proporciona la referencia con que comparar el resto de modelos, para posteriormente obtener el modelo completo, es decir, sin restricciones, y por lo tanto, incluyendo también las interacciones, tanto dentro de cada nivel como inter-niveles. El modelo definitivo, en ocasiones no único, se obtiene restringiendo a cero los parámetros no significativos (al 5%) del modelo completo. El método de estimación es el de Máxima Verosimilitud Completa (FML) para poder comparar entre modelos de diferentes efectos, tanto fijos como aleatorios (Singer, 1998). Esto permite utilizar el criterio AIC de comparación de modelos no anidados (*Akaike's Information Criteria*) (Akaike, 1987). No se ha creído necesario utilizar el criterio AIC condicional dado el elevado número de casos.

El programa informático utilizado es el HLM para dos niveles y con varianza residual homogénea, debido a que la existencia de heterogeneidad implicaría una incorrecta especificación del modelo del N-1 (Raudenbush, Bryk, 2002) y a que la finalidad de este trabajo es estudiar las diferencias únicamente en función de las cuatro covariables especificadas y no de ninguna más.

Discusión de los resultados

Estudio descriptivo

Vamos en primer lugar a presentar los datos descriptivos globales que pretenden dar respuesta a la primera cuestión de investigación: ¿Cuáles son los procesos de evaluación utilizados por el profesorado de ESO durante los cuatro años de la evaluación?

Tal y como hemos señalado en el apartado dedicado a la muestra, los resultados se presentan desagregados por curso.

La Tabla III está basada en las respuestas a los 22 ítems formuladas por los docentes de la muestra en el curso 1999-2000. En la misma aparecen las frecuencias agrupadas en tres intervalos, la media y la desviación típica.

TABLA III. Curso 1999-00

Ítems		Nada de acuerdo 1 2 %	Indiferente 3 4 5 %	Muy de acuerdo 6 7 %	\bar{X}	S
Funcionalidad Temporalidad	1. Ajusto la programación de un tema a partir de los conocimientos previos de los alumnos	3,1	43,9	53,0	5,37	1,27
	2. Voy conociendo y revisando el trabajo de los alumnos mientras ellos lo realizan	3,0	41,2	55,8	5,46	1,25
	3. Obtengo información durante el proceso de aprendizaje acerca de lo que van aprendiendo mis alumnos	2,0	33,0	65,0	5,70	1,18
	4. Comento con el alumno los resultados de las pruebas y de sus trabajos para que sepa cómo mejorar su aprendizaje	1,3	21,7	77,0	6,03	1,03
	5. Doy importancia a los trabajos que realizan los alumnos	1,6	41,7	56,7	5,53	1,15
	6. Utilizo la calificación como instrumento para resolver los problemas de disciplina	64,5	29,7	5,8	2,43	1,59
	7. Doy importancia al trabajo continuado del alumno	1,2	24,9	73,9	6,01	1,09
	8. Dedico un tiempo en cada tema a asegurar que los conocimientos básicos han quedado claros para los alumnos	1,3	26,1	72,6	5,92	1,06
	9.-Doy importancia a las pruebas al final de cada tema	2,2	35,8	62,0	5,63	1,17
Normotipo	10. A la hora de evaluar doy importancia al nivel medio de cada grupo	5,6	60,6	33,7	4,88	1,31
	11. A la hora de evaluar doy importancia a los objetivos establecidos en la unidad didáctica	0,9	34,0	65,1	5,74	1,03
	12. A la hora de evaluar doy importancia al nivel de partida del alumno	5,1	47,0	48,0	5,20	1,41
Agente evaluador	13. Mis alumnos podrían corregir sus pruebas y trabajos porque tienen claros los criterios de evaluación	11,3	61,5	27,2	4,47	1,48
	14. Procuero que mis alumnos preparen parte de las pruebas de evaluación	54,2	34,6	11,2	2,84	1,79
	15. Mis alumnos corrigen sus pruebas y ejercicios	36,2	43,9	19,9	3,58	1,89
	16. Doy importancia a la autoevaluación de los alumnos	33,7	53,5	12,8	3,48	1,68
	17. Al acabar un tema o unidad didáctica la reviso para modificarla en el futuro	8,3	51,0	40,7	4,94	1,46
	18. Mis alumnos corrigen los trabajos y pruebas de sus compañeros	60,0	33,5	6,5	2,58	1,59
Contenidos	19. A la hora de evaluar doy importancia a los conceptos	1,5	33,8	64,7	5,73	1,13

Continúa en pág. siguiente

TABLA III. Curso 1999-00

Ítems	Nada de acuerdo		Indiferente			Muy de acuerdo		\bar{X}	S
	1	2	3	4	5	6	7		
	%		%			%			
Contenidos	20. A la hora de evaluar doy importancia a los procedimientos								
	0,3		30,2			69,5		5,86	0,99
	21. A la hora de evaluar doy importancia a los valores y las actitudes								
	4,5		39,9			55,6		5,46	1,39
	22. A la hora de evaluar doy importancia a los temas transversales								
	19,3		60,1			20,5		4,15	1,61

Dado que las tablas e índices de los otros tres cursos son muy similares, vamos ahora a examinar brevemente esta tabla y luego presentaremos resultados más agregados.

Como puede observarse, el patrón de respuesta es bastante similar y se caracteriza en primer lugar por una dispersión muy baja (la media de las desviaciones típicas es de 1,31) por lo que podemos afirmar que, globalmente, los profesores responden de manera muy similar a cada ítem. La mayoría de las respuestas se sitúan en los valores altos de la escala (6 y 7) y en los medios (3, 4 y 5) y bajos (1 y 2) el número de efectivos es tanto menor cuanto más bajo es el grado.

Este patrón de distribución de frecuencias se da en 13 de los 22 ítems (1, 2, 3, 4, 5, 7, 8, 9, 11, 12, 19, 20 y 21).

En los 10 ítems restantes nos encontramos bien con que la categoría central es la más frecuente o bien que incluso la distribución es inversa disminuyendo las frecuencias conforme se incrementa el grado de acuerdo (ítems 6, 14, 18).

En la Tabla IV tenemos las medias aritméticas para cada ítem y curso. A la vista de la misma lo primero que se puede concluir es que las mismas son bastantes altas en todos los ítems (la escala va de 1 a 7) con algunas excepciones que luego comentaremos.

Los primeros nueve ítems hacen referencia a la funcionalidad y temporalidad de la evaluación, es decir, del uso que hacen los profesores de la evaluación inicial, procesual y final y de la función que cumple la evaluación en dichos momentos (formativa, diagnóstica o sumativa). De la observación de ambas tablas podemos concluir que la mayoría de los docentes realiza una evaluación inicial/diagnóstica de los conocimientos del alumnado para ajustarlo a su programación. De la misma manera, en las respuestas a los ítems 2, 3, 4 y 7, aseguran que realizan una evaluación del proceso de enseñanza-aprendizaje de cara a mejorar el propio proceso y el rendimiento de sus

TABLA IV.

Ítems	1999/2000	2000/2001	2001/2002	2002/2003	Media promediada
1. Ajusto la programación de un tema a partir de los conocimientos previos de los alumnos	5,37	5,22	5,29	5,33	5,32
2. Voy conociendo y revisando el trabajo de los alumnos mientras ellos lo realizan	5,46	5,37	5,37	5,51	5,43
3. Obtengo información durante el proceso de aprendizaje acerca de lo que van aprendiendo mis alumnos	5,70	5,50	5,61	5,68	5,64
4. Comento con el alumno los resultados de las pruebas y de sus trabajos para que sepa cómo mejorar su aprendizaje	6,03	5,94	5,93	6,02	5,99
5. Doy importancia a los trabajos que realizan los alumnos	5,53	5,39	5,38	- ^a	4,07
6. Utilizo la calificación como instrumento para resolver los problemas de disciplina	2,43	2,58	2,49	2,83	2,57
7. Doy importancia al trabajo continuado del alumno	6,01	5,87	5,85	-	5,93
8. Dedico un tiempo en cada tema a asegurar que los conocimientos básicos han quedado claros para los alumnos	5,92	5,73	5,81	5,84	5,84
9. Doy importancia a las pruebas al final de cada tema	5,63	5,55	5,67	-	5,62
10. A la hora de evaluar doy importancia al nivel medio de cada grupo	4,88	4,67	4,75	4,73	4,78
11. A la hora de evaluar doy importancia a los objetivos establecidos en la unidad didáctica	5,74	5,55	5,65	5,69	5,67
12. A la hora de evaluar doy importancia al nivel de partida del alumno	5,2	4,99	4,98	5,18	5,11
13. Mis alumnos podrían corregir sus pruebas y trabajos porque tienen claros los criterios de evaluación	4,47	4,45	4,54	4,53	4,50
14. Procuro que mis alumnos preparen parte de las pruebas de evaluación	2,84	2,97	2,93	3,14	2,96
15. Mis alumnos corrijen sus pruebas y ejercicios	3,58	3,55	3,59	3,67	3,60
16. Doy importancia a la autoevaluación de los alumnos	3,48	3,57	3,45	-	3,49
17. Al acabar un tema o unidad didáctica la reviso para modificarla en el futuro	4,94	4,85	5	5,08	4,97
18. Mis alumnos corrijen los trabajos y pruebas de sus compañeros	2,58	2,62	2,65	2,91	2,68
19. A la hora de evaluar doy importancia a los conceptos	5,73	5,64	5,75	-	5,71
20. A la hora de evaluar doy importancia a los procedimientos	5,86	5,74	5,73	-	5,79
21. A la hora de evaluar doy importancia a los valores y las actitudes	5,46	5,23	5,34	-	5,37
22. A la hora de evaluar doy importancia a los temas transversales	4,15	4,09	4,08	-	4,12

^{a)} En el curso 2002/03 no se aplicaron algunos de los ítems por lo que en la casilla correspondiente no aparece la media. Lógicamente, en tales casos la ponderación se ha ajustado.

alumnos. Además, la utilización de la evaluación/calificación como instrumento para resolver los problemas de disciplina es mínima.

Por otra parte, el profesorado mayoritariamente dedica un tiempo a asegurar que los conocimientos básicos han quedado claros para los alumnos. No obstante, todavía el profesorado de secundaria concede una gran importancia a las pruebas o exámenes del final de cada tema.

En cuanto al normotipo o estándar utilizado (ítems 10 a 12) puede observarse que en consonancia con la LOGSE que abogó por la desaparición de la evaluación nomotética normativa, la mayoría del profesorado tiene más en cuenta la situación de partida del alumno, es decir la evaluación idiográfica y sobre todo los objetivos preestablecidos o la evaluación criterial. La evaluación normativa es la que obtiene una media más baja.

Los ítems 13 a 18 hacen referencia al agente evaluador, es decir, al uso que hacen los profesores de la autoevaluación, coevaluación y heteroevaluación. Aunque en la LOGSE se señala la importancia de la autoevaluación del alumnado, todavía el profesorado no la ha incluido de forma mayoritaria en los procesos evaluativos. Esto sucede tanto en la preparación por parte del alumnado de las pruebas de evaluación, corrección por parte del alumnado de sus pruebas y ejercicios, como en la importancia que dan a la autoevaluación del alumnado. En cuanto a la autoevaluación del profesor (ítem 17) la media obtenida es también moderadamente baja. La coevaluación (ítem 19) tampoco parece ser una práctica asumida en la ESO pues es reducida la proporción del profesorado que declara utilizar esa estrategia.

Por último, los ítems 19 a 22 se refieren a la importancia que dan a los distintos tipos de contenido a la hora de evaluar. Los contenidos más importantes para el profesorado son los procedimentales y los conceptuales, seguidos por los valores y las actitudes, siendo los menos importantes los temas transversales.

La segunda pregunta que nos planteábamos en este estudio era si ha habido alguna evolución a lo largo de estos cuatro cursos en los usos dados a la evaluación por parte del profesorado de la ESO de la Comunidad Autónoma Vasca. En un apartado anterior ya hemos comentado cuál es el enfoque que en un próximo trabajo pretendemos plantear para abordar el estudio del cambio. De cualquier forma, del examen de la Tabla IV podemos concluir que las medias obtenidas durante estos cuatro años en los distintos ítems se han mantenido bastante estables. Cabe, en consecuencia suponer que durante estos cuatro años las respuestas de los profesores se hayan mantenido constantes.

Estudio diferencial

Como ha quedado dicho, en esta parte del trabajo pretendemos examinar las posibles diferencias que puedan darse en el uso de la evaluación por parte del profesorado de la ESO. Dada la estructura anidada de los datos (profesores/centros) este estudio diferencial se lleva a cabo mediante modelización jerárquica.

Tras utilizar las técnicas apropiadas de análisis exploratorio de datos para detectar violaciones a las condiciones de aplicación, no se han encontrado anomalías considerables. En la totalidad de los modelos definitivos las interacciones entre covariables intranivel como la Antigüedad media de cada Centro han resultado no significativas.

En su momento ya hemos apuntado cómo no es posible utilizar ningún tipo de puntuación compuesta y cómo tampoco las técnicas de reducción de la información nos han aportado una estructura factorial clara. En consecuencia, en este punto nos encontramos con la necesidad de formular un modelo para cada ítem y curso por lo que al final son 80 los modelos estudiados (hay ocho ítems no medidos en el curso 2002-03).

Ante esto, y con el fin de simplificar en lo posible la presentación de los resultados, vamos en primer lugar a examinar -a modo ilustrativo- uno cualquiera de los modelos con cierto detalle. A continuación y para la totalidad de los restantes modelos nos limitaremos a tablas esquemáticas con la información más relevante.

Examen detallado de un modelo (ítem I I/curso 1999-00)

El modelo definido resultante es:

$$Y_{ij} = \beta_{0j} + \beta_{1j} \text{ESO} I_{ij} + r_{ij} \quad r_{ij} \sim N(0, \sigma^2)$$

$$\beta_{0j} = \gamma_{00} + g_{01} \text{Tipo}_j + u_{0j}$$

$$\beta_{1j} = u_{1j} \begin{pmatrix} u_{0j} \\ u_{1j} \end{pmatrix} \sim N \left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} t_{00} & \\ & t_{10} \ t_{11} \end{pmatrix} \right]$$

Resultados:

- Fiabilidad (β_{0j}) = 0,257
- Fiabilidad (β_{1j}) = 0,224
- $\sigma^2 = 0,97$ (0,04)
- $\tau_{00} = 0,03$ (0,02) $p_{50} = 0,003$
- $\tau_{10} = 0,000$ (0,024)
- $\tau_{11} = 0,057$ (0,040) $p_{52} = 0,013$
- $\gamma_{00} = 5,57$ (0,060) $p_{52} = 0,000$

- $\gamma_{01} = 0,24$ (0,080) $p_{52} = 0,005$
- $-2\ln\Delta = 3562,97_6$
- AIC = 3574,97

Nota: Entre paréntesis se encuentra el error típico de estimación de cada parámetro y como subíndice, los grados de libertad asociados a la prueba de hipótesis de nulidad del parámetro correspondiente.

Los valores de AIC para el modelo incondicional y el completo son 3580,82 y 3590,31 respectivamente.

Por tanto, se concluye que:

Ni la variable Contexto del Centro ni la Antigüedad de los docentes tiene efecto sobre el ítem 11 (Y_{ij}),

El modelo combinado es:

$$Y_{ij} = \gamma_{00} + \gamma_{01}\text{Tipo}_j + (u_{0j} + u_{1j}\text{ESO}_{ij} + r_{ij}) = 5,57 + 0,24\text{Tipo}_j + (u_{0j} + u_{1j}\text{ESO}_{1ij} + r_{ij}).$$

$$r_{ij} \sim N(0, 0,97) \quad \begin{pmatrix} u_{0j} \\ u_{1j} \end{pmatrix} \sim N \left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0,030 & \\ & 0,057 \end{pmatrix} \right]$$

El modelo, teniendo presente que tanto ESO como Tipo están codificadas como:

ESO: 0 = no ejerce en el primer ciclo de la ESO,

1 = sí ejerce en el primer ciclo de la ESO,

Tipo: 0 = público,

1 = privado,

Público - no 1er ciclo ESO $Y_{ij} = 5,57 + (u_{0j} + r_{ij})$.

Público - sí 1er ciclo ESO $Y_{ij} = 5,57 + (u_{0j} + u_{1j} + r_{ij})$.

Privado - no 1er ciclo ESO $Y_{ij} = 5,81 + (u_{0j} + r_{ij})$.

Privado - sí 1er ciclo ESO $Y_{ij} = 5,81 + (u_{0j} + u_{1j} + r_{ij})$.

Un modelo así formulado puede ser interpretado como sigue:

$t_{00} \neq 0 \rightarrow$ la intersección varía en los distintos centros, controlando el Tipo de Centro, siendo mayor la puntuación en el ítem 11 en los profesores de los Privados.

$t_{11} \neq 0 \rightarrow$ la influencia de la ESO 1 varía entre los centros ($t_{11} = 0,057$ (0,04) $p_{50} = 0,013$). Son más heterogéneos los profesores que sí imparten en el primer ciclo de la ESO (la diferencia es u_{1j}).

En resumen, las diferencias detectadas en el ítem 11 son debidas a la titularidad, y en lo que se refiere a la heterogeneidad de los docentes, a la variable «ciclo».

A continuación, en las tablas 5, 6, 7 y 8 se presentan esquemáticamente los modelos de los distintos ítems agrupados según las cuatro dimensiones mencionadas. En estas cuatro tablas es preciso hacer notar que:

- Aunque en las tablas esquemáticas no se represente, la intersección es significativa en todos los modelos y representa en general la puntuación media en cada ítem correspondiente a un docente que no ejerce en el primer ciclo de la ESO y una Antigüedad docente igual a la media de su Centro.
- La presencia del efecto aleatorio u_0 , bastante frecuente en las distintas tablas, indica que, aún controlando la Titularidad del Centro y su Contexto, hay diferencias entre los Centros.
- En cursiva se representan las covariables con coeficientes negativos. Téngase presente que el Contexto del Centro está medido según una escala inversa, es decir, cuanto menor es su valor, mayor es el Contexto.
- Los códigos para las covariables son los siguientes:
 - E: El docente no ejerce (0) o sí (1) en el primer ciclo de la ESO,
 - A: Antigüedad docente, centrada en la media del Centro,
 - C: Contexto socioeconómico y cultural del Centro, centrado en la media global de todos los Centros,
 - T: Titularidad del centro: público (0) y privado (1).
- El resto de términos, u_1 y u_2 , se refieren a efectos aleatorios del N-2.
- Si se denomina genéricamente como X y W a las covariables del N-1 y N-2 respectivamente, el término $X*W$ indica la interacción entre tales covariables.
- En las cuatro columnas de la derecha, encabezadas por *oneway*, se encuentran los resultados para los cuatro cursos. Han sido obtenidos aplicando el Análisis de la varianza de un factor, considerando que las covariables anteriormente mencionadas son factores y omitiendo el esquema anidado. Se presentan aquí a modo orientativo.
- En las casillas que aparecen vacías no hay ningún efecto significativo.

Modelización de la dimensión funcionalidad/temporalidad de la evaluación

En general, podemos anticipar que del examen de las cuatro tablas no se desprende una estructura común clara, ni por ítems ni por cursos. El aspecto que, en general, más diferencia a los docentes es si imparten o no docencia en el primer ciclo de la ESO, siendo más favorables aquéllos que sí la imparten. Además la Antigüedad docente presenta una influencia negativa, en el sentido del signo del correspondiente coeficiente, bastante generalizada. Prácticamente en la totalidad de los ítems las diferencias entre Centros no se explican por la Titularidad ni por el Contexto, si bien aparecen ciertas interacciones entre estas variables y las propias de los docentes, indicando que la influencia del ciclo de la ESO en que el docente ejerce y la Antigüedad en el ejercicio profesional varían en función de la Titularidad y el Contexto del Centro. Por último, conviene aclarar que ciertas influencias pueden resultar no significativas en el curso 2000-01 dado el reducido número de Centros implicados, 26.

1. Ajusto la programación de un tema a partir de los conocimientos previos de los alumnos.

No hay diferencias entre los docentes de cada centro, exceptuando el curso 2001-02, en que los docentes que ejercen en el primer ciclo de la ESO son más favorables a tener en cuenta los conocimientos previos de los alumnos.

TABLA V. Esquema de los modelos de los ítems de Funcionalidad/temporalidad de la evaluación

Ítems	1999/2000	2000/2001	2001/2002	2002/2003	Oneway		
1	U_0		$E + u_1$	u_1	E	E	E
2	A		A + E	A	A E	A E	A
3	$E + E^*u_1 + u_1$	$E + u_1$	$E + u_1$		E	E	E
4			u_1	$A + A^*T$			
5	$A + E + C^*E + E^*u_1$	$E + u_1$	$E + u_1$	$A + E$	E	E	
6	$A + A^*T + U_1$	$A + u_1$	$A + u_1$		A	A	A
7	$A + E + T + E^*T + U_1$	$E + u_1$	$A + E + A^*T + u_1$		E	E	E
8	A^*C		$A + A^*T$		C	C	C
9	$E + C + E^*u_1$	E^*C	$E + E^*T$		T	T	T

2. Voy conociendo y revisando el trabajo de los alumnos mientras ellos lo realizan.
La Antigüedad docente incide de forma negativa en este aspecto y son más favorables los docentes que ejercen en el primer ciclo de la eso. Las diferencias entre Centros no son significativas.
3. Obtengo información durante el proceso de aprendizaje acerca de lo que van aprendiendo mis alumnos.
Es ejercer en el primer ciclo de la eso lo que influye positivamente; en el curso 1999-00 los docentes que ejercen en el primer ciclo de la eso son más heterogéneos.
4. Comento con el alumno los resultados de las pruebas y de sus trabajos para que sepa cómo mejorar su aprendizaje.
Hasta el curso 2002-03 no hay diferencias entre los docentes, y es en tal curso cuando la Antigüedad docente influye de forma positiva en los centros públicos y ligeramente negativa en los privados.
5. Doy importancia a los trabajos que realizan los alumnos.
Ejercer en el primer ciclo de la eso influye positivamente y las diferencias entre Centros no son explicadas por el Contexto y la Titularidad. En el curso 1999-00 la situación es más compleja, apareciendo la Antigüedad docente con su influencia negativa y la interacción entre Contexto y el primer ciclo de la eso que hace que los docentes de primer ciclo de la eso sean más favorables cuanto menor sea el Contexto del Centro. Además los docentes del primer ciclo de la eso son más heterogéneos.
6. Utilizo la calificación como instrumento para resolver los problemas de disciplina.
Nuevamente la Antigüedad docente se presenta bajo una influencia inversa en todos los cursos, aunque en el curso 1999-00 ésta se ve aminorada en los Centros privados, y en el curso 2002-03 en los docentes que ejercen en el primer ciclo de la eso.
7. Doy importancia al trabajo continuado del alumno.
Los docentes que ejercen en el primer ciclo de la eso así como los de Centros privados son más favorables, pero en el curso 1999-00 los docentes de Centros privados que ejercen en el primer ciclo de la eso lo son menos. La Antigüedad docente influye negativamente, aunque en el curso 2001-02 tal influencia disminuye en los Centros privados.
8. Dedico un tiempo en cada tema a asegurar que los conocimientos básicos han quedado claros para los alumnos.

En este caso hay homogeneidad entre los Centros. La Antigüedad docente influye positivamente, si bien en el curso 1999-00 la interacción con el Contexto del Centro hace que tal influencia se reduzca para los Centros de Contextos más altos; en el curso 2001-02 la influencia de la Antigüedad docente es negativa en los Centros privados.

9. Doy importancia a las pruebas al final de cada tema.

Los docentes del primer ciclo de la ESO son menos favorables; además en los cursos 1999-00 y 2000-01 se presenta una influencia positiva del Contexto del Centro. En el curso 2001-02 las mayores diferencias se dan entre los docentes de Centros privados y públicos que imparten en primer ciclo de la ESO, a favor de los primeros. Los docentes del primer ciclo de la ESO son más heterogéneos.

Modelización de la dimensión normotipo utilizado

En la Tabla 6 se refleja que las covariables de más clara influencia son si el docente ejerce en el primer ciclo de la ESO, la Titularidad y de forma bastante puntual la Antigüedad docente. Las diferencias entre Centros apenas resultan explicadas por sus correspondientes covariables.

10. A la hora de evaluar doy importancia al nivel medio de cada grupo.

En los dos primeros cursos los docentes del primer ciclo de la ESO son más favorables. En los dos últimos la influencia corresponde a la Antigüedad docente, siendo su influencia positiva y aumentando en el último curso con el Contexto del Centro. En el curso 1999-00 los docentes que sí imparten en el primer ciclo de la ESO son más heterogéneos.

11. A la hora de evaluar doy importancia a los objetivos establecidos en la unidad didáctica.

TABLA VI. Esquema de los modelos de los ítems de Normotipo utilizado

Ítems	1999/2000	2000/2001	2001/2002	2002/2003	Oneway			
I0	$E + E^*u_i + u_i$	E	A	$A^*C + u_i$				
I1	$T + u_i$	$T + u_i$	u_i	T	T	ECT	ET	CT
I2	$E + A^*u_i + E^*u_i$	$E + A^*u_i$	E + A	E	E	E	E	E

Los docentes de los Centros privados son más favorables.

12. A la hora de evaluar doy importancia al nivel de partida del alumno.

Los docentes del primer ciclo de la ESO son más favorables; en los dos primeros cursos los docentes de más Antigüedad docente son más heterogéneos; en el 2001-02 la Antigüedad docente incide negativamente.

Modelización de la dimensión agente evaluador

En la Tabla VII se puede observar que globalmente las covariables que más frecuentemente influyen en los distintos ítems y cursos son la Titularidad y si el docente ejerce en el primer ciclo de la ESO. Las diferencias entre Centros quedan en muchas ocasiones inexplicadas.

13. Mis alumnos podrían corregir sus pruebas y trabajos porque tienen claros los criterios de evaluación.

En los tres primeros cursos los docentes de los centros privados son más favorables. En el último la interacción entre Antigüedad docente y Contexto indica que la influencia de la Antigüedad docente aumenta con el Contexto del centro.

14. Procuro que mis alumnos preparen parte de las pruebas de evaluación.

El Contexto del centro incide negativamente en el curso 1999-00 y en el 2002-03 lo hace la Antigüedad docente.

15. Mis alumnos corrigen sus pruebas y ejercicios.

En los dos primeros cursos los Centros privados son más favorables.

TABLA VII. Esquema de los modelos de los ítems de Agente evaluador

Ítems	1999/2000	2000/2001	2001/2002	2002/2003	Oneway			
13	$T + u_0$	T	T	$A^*C + u_0$	AT	T	AT	AT
14	$C + u_0$	u_0	u_0	A				
15	$E + u_0$	E		u_0	EC	E	E	C
16	$E + A + E^*T + A^*T + u_0$	$E + u_0$	$E + A + A^*T + u_0$		EC	EC	EC	
17		$A + A^*C + A^*T + u_0$						
18	$E + E^*C + A^*C + u_0$	$E + u_0$	$E + u_0$	$E + u_0$	E	E	E	

16. Doy importancia a la autoevaluación de los alumnos.

En general los docentes del primer ciclo de la ESO son más favorables, aunque en el curso 1999-00 en los Centros privados la diferencia se reduce. En ese mismo curso y en el 2001-02 la Antigüedad docente incide negativamente si bien con menos intensidad en los centros privados.

17. Al acabar un tema o unidad didáctica la reviso para modificarla en el futuro.

En general no hay diferencias. Sin embargo en el curso en que menos docentes han contestado al cuestionario, el 2000-01, la Antigüedad docente influye positivamente, menos en los Centros privados y más cuanto mayor es el Contexto del Centro.

18. Mis alumnos corrigen los trabajos y pruebas de sus compañeros.

Los docentes del primer ciclo de la ESO son más favorables. En el curso 1999-00 las diferencias aumentan según disminuye el Contexto, así como la influencia de la Antigüedad docente para los Contexto más bajos.

Modelización de la dimensión contenidos evaluados

El esquema de los modelos se encuentra en la Tabla VIII, y se puede observar que para este grupo de de ítems resultan más complejos. Se puede decir que en general el ejercer en el primer ciclo de la ESO es favorable, aunque haya bastantes excepciones. Algo análogo ocurre con la Antigüedad docente, cuya influencia es, con más frecuencia, negativa. Por otra parte se presentan múltiples interacciones y las diferencias entre los Centros son escasamente explicadas por la Titularidad y el Contexto.

19. A la hora de evaluar doy importancia a los conceptos.

En los dos primeros cursos los docentes de Centros públicos que ejercen en el

TABLA VIII. Esquema de los modelos de los ítems de Contenidos evaluados

Ítems	1999/2000	2000/2001	2001/2002	2002/2003	Oneway		
19	$E + A + E^*T + u_i$	$E + A^*C + E^*T + u_i$	$u_i + E^*u_i$		ET	ET	ET
20	$E + u_i + E^*u_i$	$E + A + E^*C + u_i$	E		E	E	
21	$A + E + T + A^*T + E^*T + = u_i$	$A + E + A^*C + u_i$	$A + E + u_i$		E	AE	AE
22	$E + C + T + E^*T + u_i$	A + E	$A + E + u_i$		E	E	E

primer ciclo de la ESO son menos favorables, siendo contraria la tendencia en los privados, aunque no de forma clara en el curso 2000-01; además la Antigüedad docente incide negativamente sobre todo en los centros de Contexto bajos. En el curso 2001-02 los docentes que ejercen en el primer ciclo de la ESO son más heterogéneos.

20. A la hora de evaluar doy importancia a los procedimientos.

Los docentes del primer ciclo de la ESO son más favorables, resultando en el curso 1999-00 más heterogéneos, y en el 2000-01, en que la Antigüedad docente influye positivamente, también más favorables en los Centros de Contexto más altos.

21. A la hora de evaluar doy importancia a los valores y las actitudes.

El que ejerzan en el primer ciclo de la ESO hace que los docentes sean más favorables, aunque en el curso 1999-00 en los Centros privados, en que en general son más favorables, esta tendencia disminuye. Por su parte la Antigüedad docente incide negativamente, aunque menos intensamente en los Centros privados en el curso 1999-00, así como con el aumento del Contexto del centro en el curso 2000-01.

22. A la hora de evaluar doy importancia a los temas transversales.

Los docentes que ejercen en el primer ciclo de la ESO son más favorables, siéndolo algo menos los docentes de Centros privados en el curso 1999-00. En el curso 1999-00 el Contexto influye negativamente, y en los otros dos presenta la misma tendencia la Antigüedad docente.

En resumen, uno de los objetivos implícitos era analizar si se produce estabilidad, sobre todo estructural, en los modelos tanto entre cursos como entre agrupaciones de ítems, es decir, dirigirnos a la búsqueda de un modelo lo más general posible. Pero nada hacía presagiar que este fin fuese alcanzable. Así, ni la exploración a través del Análisis de Componentes Principales para encontrar clases de ítems, ni la Clasificación, bajo sus distintas técnicas (la más ilustrativa a este respecto es la clasificación en dos fases) para encontrar clases de sujetos (docentes) han dado fruto alguno.

Así pues, tras decidir que había que explorar sobre modelos para cada ítem y cada curso académico, teniendo presente que los docentes están anidados en los Centros, y a la vista de los resultados, parece que la casuística es más bien variada. Y esto obliga a una síntesis difícil y, en el mejor de los casos, muy general.

Por otra parte, la sospecha sobre la causa de la dificultad para la generalización podría recaer en los ítems que no siguen claramente una distribución normal, pero ha

resultado no ser así. Se ha apuntado también que en el curso 2000-01 el número de Centros es bajo; sin embargo no hay nada claro, según los distintos autores, sobre el número mínimo de unidades en ambos niveles (Raudenbush, Bryk, 2002).

Por ello sólo nos podemos sustentar en la frecuencia de aparición de las distintas covariables en los distintos modelos:

- Ejercer en el primer ciclo de la ESO da resultados favorables, excepto en los ítems 9 y 19.
- La Antigüedad influye negativamente.
- Los docentes de Centros privados son más favorables.
- En general hay variabilidad entre centros, lo cual nos indica que el no contemplar un enfoque multinivel dará lugar a estimaciones no eficientes.

Conclusiones

A nivel general los aspectos de la evaluación que destacan relacionados con el total de los profesores son los siguientes:

- La utilización de la evaluación como medio de mejora del proceso de enseñanza-aprendizaje.
- La no utilización de la evaluación como un instrumento para mantener la disciplina.
- La utilización de la evaluación personalizada y sobre todo criterial, aunque hay un porcentaje no desdeñable de profesores que utiliza la evaluación normativa.
- El uso masivo de la heteroevaluación, es decir, el profesor es prácticamente el único que evalúa a los alumnos, desechándose otras prácticas evaluativas como la coevaluación y la autoevaluación.
- La importancia dada a los contenidos conceptuales, procedimentales y actitudinales (aunque en menor medida) y la poca importancia dada a los temas transversales.

Es en cuanto al Ciclo de la ESO en el que se imparte la docencia donde se dan el mayor número de diferencias en los usos de la evaluación en las cuatro dimensiones

estudiadas. A partir de esto se puede concluir que es el profesorado del primer ciclo de la ESO, que mayoritariamente son maestros y maestras que provienen de la antigua egb quienes han adoptado de una forma más mayoritaria las directrices marcadas por la LOGSE en cuanto a la evaluación. Frente al profesorado licenciado y/o proveniente de los antiguos BUP y FP, el profesorado de primer ciclo tiene más en cuenta la evaluación inicial y continua, tiene en cuenta los objetivos establecidos y la situación de partida del alumnado, da una mayor importancia a la autoevaluación (aunque ambos colectivos arrojan valores bajos), y a la evaluación de procedimientos, actitudes, valores y temas transversales.

La importancia dada a la evaluación de los objetivos establecidos en la unidad didáctica es otra característica en la que también se dan diferencias en función del tipo de centro y del contexto socioeconómico-cultural del mismo.

Como conclusión final se puede señalar que el profesorado de la ESO de la CAV muestra una tendencia a asumir las directrices marcadas por la LOGSE, es decir que poco a poco la evaluación se va haciendo más procesual y formativa, más criterial e idiográfica y cada vez va teniendo más en cuenta otros aspectos además de los contenidos conceptuales. Como lagunas en este sentido habría que mencionar la práctica poco habitual de la autoevaluación y la coevaluación. Esta última situación es más acusada en el profesorado del segundo ciclo de la ESO.

Por último, y como evidente limitación de este estudio, queda pendiente el estudio de la evolución a través de los cuatro cursos, es decir, el análisis longitudinal. Al no estar identificados los distintos docentes a lo largo de los cursos, es imposible estudiar el cambio individual, el de los docentes. Ello implica que tal cuestión sólo será resoluble a nivel de centro, de manera que se podrá concluir si se ha producido o no una evolución global y en qué tipo de centros en función de su Titularidad y Contexto, así como en características globales de su profesorado referidas a Antigüedad docente y Ciclo de la ESO en que imparten docencia.

Referencias bibliográficas

- AKAIKE, H. (1987): «Factor analysis and AIC». en *Psychometrika*, 52, pp. 317-332.
- COLL, C.; MARTÍN, E.; Y ONRUBIA, J. (2001): « La evaluación del aprendizaje escolar: dimensiones psicológicas, pedagógicas y sociales», en C. COLL; J. PALACIOS; A. MARCHESI

- (coords.): *Desarrollo psicológico y educación, vol II, Psicología de la educación escolar*. Madrid, Alianza.
- LUKAS, J.F., SANTIAGO, C., MUNARRIZ, B., MOYANO, N.; SEDANO, M. (2000): «Euskal Autonomia Erkidegoko ikastetxeen kanpoebaluazioa». *Tantak*. 24., pp. 3963.
- LUKAS, J.F., SANTIAGO, C., SEDANO, M., MOYANO, N., JOARISTI, L.; LIZASOAIN, L. (2001): «Usos de la evaluación en profesores/as de la eso de la Comunidad Autónoma Vasca», en AIDIPE (comp.): *Actas del X Congreso Nacional de Modelos de Investigación Educativa*. A Coruña, AIDIPE., pp. 427433.
- LUKAS, J.F.; SANTIAGO, C. (2004a): *Evaluación Educativa*. Madrid, Alianza Editorial. (En prensa.).
- LUKAS, J.F.; SANTIAGO, C. (2004b): «Evaluación de centros de Secundaria del País Vasco», en *Revista Electrónica de Investigación Educativa* (redie). V6. N-2.
- MARCHESI, A. (2002): «El proyecto redes», en A. MARCHESI; E. MARTÍN (comp.): *Evaluación de la educación secundaria. Fotografía de una etapa polémica*. Madrid, Fundación Santa María, pp. 21-33.
- MARCHESI, A.; MARTÍN, E. (2002) (comp.): *Evaluación de la educación secundaria. Fotografía de una etapa polémica*. Madrid, Fundación Santa María.
- MARTÍN, E. (2002): «Los procesos de aula y su influencia en la calidad de la enseñanza», en A. MARCHESI; E. MARTÍN (comp.): *Evaluación de la educación secundaria. Fotografía de una etapa polémica*. Madrid, Fundación Santa María, pp. 95-118.
- RAUDENBUSH S. W., BRYK A. S. (2002): *Hierarchical Linear Models. Applications and Data Analysis Methods*. California, Sage Publications.
- REDES (1999): «Una red de evaluación de centros de Educación Secundaria». *Infancia y Aprendizaje*. 85, pp. 59-73.
- SÁEZ BREZMES, M.J.; CARRETERO, A.J. (1995): «Un modelo posible: La evaluación en la España actual». En M.J. SÁEZ BREZMES (ed.): *Conceptualizando la evaluación en España*. Madrid, Servicio de Publicaciones de la Universidad de Alcalá, pp. 131-141.
- SINGER J. D. (1998): «Using SAS PROC MIXED to Fit Multilevel Models, Hierarchical Models, and Individual Growth Models», en *Journal of Educational and Behavioral Statistics*, 24(4), pp. 323-355.
- STRUYE, E.; LANDENBERGHE, R.; LENS, W. (2001): «The Evaluation Practice of Teachers as a Learning Opportunity for Students», en *Studies in Educational Evaluation*. Vol. 27. 3, pp. 215-238.
- TYLER, R. (1950): *Basic principles of curriculum and instruction*. Chicago, University of Chicago Press.