

Capítulo 3

ATRIBUCIÓN DE IMPORTANCIA A LAS DIMENSIONES COMPETENCIALES EN LOS EXÁMENES Y RENDIMIENTO EN INGLÉS COMO PRIMERA LENGUA EXTRANJERA EN EUROPA

Eva Exposito Casas*

Esther López Martín*

Enrique Navarro Asencio**

José Luis Gaviria***

**Universidad Nacional de Educación a Distancia.*

***Universidad Internacional de La Rioja.*

****Universidad Complutense de Madrid.*

■ Resumen

La evaluación es un elemento fundamental en el proceso de enseñanza y aprendizaje. Si las dimensiones competenciales que un profesor incluye en su examen definen operativamente sus opciones didáctico-metodológicas, estas deberían tener una repercusión en los resultados de aprendizaje de sus alumnos.

En este estudio se han definido dos modelos de actitud respecto de la evaluación en los profesores. Por una parte, hablamos del modelo *exhaustivo* cuando nos referimos a profesores que dan la máxima importancia a todas y cada una de las ocho

dimensiones competenciales analizadas en el EECL. Por otra, denominamos modelo *orientado* a la opción de aquellos profesores que consideran algún subconjunto de las mencionadas dimensiones como más relevantes que las otras, independientemente del contenido del subconjunto seleccionado.

La hipótesis de este trabajo es que los alumnos de los profesores que se atienen al modelo aquí denominado *exhaustivo* obtienen resultados medios superiores en Comprensión lectora y oral en lengua inglesa a los de los alumnos de los profesores del modelo denominado *orientado*.

Los resultados muestran que siempre que se observan diferencias estadísticamente significativas, estas son a favor de las medias de los alumnos que se encuentran en centros con modelo *exhaustivo* predominante.

Es especialmente interesante el caso de España. En nuestro país el 85% de los alumnos están en centros en los que el modelo predominante es el *orientado*, y solo un 15% de los casos corresponden a centros de modelo *exhaustivo*. Las diferencias son estadísticamente significativas a favor de este último modelo tanto en Comprensión lectora como en Comprensión oral.

En España, el 15% de los alumnos que se encuentra en centros en los que el modelo predominante es el aquí denominado *exhaustivo* su rendimiento es, términos de tamaño del efecto, 0,4 desviaciones típicas en Comprensión lectora y 0,32 desviaciones típicas en Comprensión oral mayor que el 85% restante.

Palabras Clave:

Inglés como lengua extranjera, Evaluación de rendimiento, Creencias de los profesores, Modelos de evaluación, Evaluación a gran escala

■ Abstract

Evaluation is a key element in the process of teaching and learning. If the dimensions of competence that a teacher includes in his/her exams, operationally define his/her didactical-methodological choices, then they should have an impact on the learning outcomes of his/her pupils.

In this study we have defined two models of teachers attitude toward evaluation. First, we define the *exhaustive* model when referring to teachers who give the highest importance to each and every one of the eight dimensions of competence analyzed in the EECL study. Additionally, we call *oriented* model to the choice of those teachers

who consider some subset of the above dimensions as more relevant than others, regardless of the content of the selected subset.

The hypothesis of this study is that students of teachers who stick to the model here called *exhaustive* obtain higher average results in Reading and Listening Comprehension than those of the students of teachers linked to the so called *oriented* model.

The results show that whenever a statistically significant difference exists, it is in favor of the students who are in schools where the predominant model is the *exhaustive* one.

Especially interesting is the case of Spain. In this country 85% of students are in schools where the predominant model is the *oriented* one and only 15% of cases are in the alternative centers. The differences are statistically significant in favor of this last *exhaustive* model both in Reading and Listening Comprehension.

In Spain, the 15% of students who are in schools where the predominant model is the *exhaustive* one, have an average achievement whose corresponding effect size is of 0.4 standard deviation in Reading and 0.32 standard deviation in Listening higher than the remaining 85% students assigned to the *oriented* model.

Key-words:

English as Second Language, Performance Assessment, Teachers believes, Large Scale Assessment

INTRODUCCIÓN

El estudio de los factores que intervienen en el proceso de enseñanza-aprendizaje de una segunda lengua, así como la interacción de estos con el desempeño de los estudiantes, resulta indispensable para identificar qué procesos o ambientes pueden llegar a favorecer la calidad de la enseñanza en idiomas. Al igual que ocurre con otros indicadores de resultado del sistema educativo, el estudio del logro de los estudiantes en la adquisición de otro idioma, ha despertado el interés de numerosos autores por analizar el efecto que las variables relativas a las características individuales y familiares de los alumnos, el profesorado, o, incluso, el contexto, ejercen en los niveles de rendimiento que obtienen los alumnos (véase De la Rica y González, 2012; Vez, Martínez y Lorenzo, 2012, y Anghel y Guell, 2012, todos ellos en este mismo volumen).

Y este es el caso de este artículo, que toma como referencia las medidas de rendimiento del Estudio Europeo de Competencia Lingüística (EECL).

Junto con las características individuales y familiares de los alumnos, es importante señalar el papel que juegan las características de los profesores sobre el desempeño de los estudiantes. Este es el caso de "sus conocimientos y su experiencia docente, sus creencias, sus metas personales y docentes, sus expectativas y sus atribuciones sobre el rendimiento de los alumnos y la efectividad de sus propias actuaciones" (García Legazpe, 2008, p.42). Estas características de los profesores pueden influir en su práctica docente, creando ambientes que favorezcan el aprendizaje de un segundo idioma y que, a su vez, repercutan sobre los resultados que consiguen los alumnos. De esta forma, son las pautas de actuación del profesor en el aula las que influyen de manera más directa sobre el aprendizaje (García Legazpe, 2008).

La investigación sobre la enseñanza de las lenguas extranjeras comienza en la década de los 60 centrándose en los métodos y procedimientos que se muestran eficaces para el aprendizaje, teniendo claramente un afán prescriptivo con respecto al método que aún hoy se refleja en los materiales de enseñanza. Poco a poco comienza a moverse la investigación sobre la enseñanza del Inglés como segunda lengua desde el campo de la lingüística aplicada hacia el campo de las creencias cognitivas de los profesores (Woods, 1996; Rodger, 2001; Richards y Rodgers, 2001). Este cambio implicaría que para que la enseñanza de un segundo idioma sea eficaz no solo hay que definir qué necesita conocer el profesor, sino también hay que analizar cuáles son sus concepciones y creencias sobre la enseñanza, los procesos de aprendizaje, los contextos de enseñanza y sus prácticas pedagógicas (Freeman y Johnson, 1998; 2005).

Las creencias cognitivas de los profesores lógicamente afectan a qué y a cómo enseñan. Además están influenciadas por su formación, por el contexto de enseñanza y por las experiencias docentes (Tsui, 2011). Sin embargo, no se encuentran estudios sobre cómo las creencias de los profesores influyen en lo que hacen en su labor docente ni en cómo estas creencias impactan en los resultados de sus alumnos (Kiely, 2001).

La forma más habitual de estudiar el pensamiento del profesor es analizar los autoinformes sobre los objetivos que plantean para sus clases, los principios pedagógicos en los que se apoyan y las actividades docentes que desarrollan en el aula. Sin embargo, probablemente donde mejor se expresan las opciones metodológicas y didácticas de los profesores es en la evaluación que estos llevan a cabo. Aquello de lo que un profesor examina es aquello que considera realmente importante. Por tanto, el análisis de las opciones evaluativas de los profesores proporciona una información de primera mano respecto a cómo estos conciben el proceso de enseñanza y de

aprendizaje. Es en los contenidos y criterios de la evaluación donde el profesor establece de forma operativa lo que considera que es más valioso o relevante para sus estudiantes. La valoración de lo que considera muy importante evaluar es también un referente de lo que ha enseñado en su aula y puede considerarse un indicador indirecto de las expectativas respecto al rendimiento de sus alumnos.

La evaluación constituye una vía de comunicación indirecta entre el profesor y los alumnos. El contenido de la evaluación que realiza un profesor constituye la guía más eficaz para orientar el trabajo individual de los alumnos. La primera información que un alumno trata de obtener cuando accede a una nueva materia es cómo examina ese profesor. La evaluación constituye un medio en el que el profesor expresa operativamente sus creencias sobre la enseñanza y sus opciones metodológicas y didácticas, más allá de su propio discurso.

En el Estudio Europeo de Competencia Lingüística (EECL), además de la medida del rendimiento de los estudiantes en lenguas extranjeras, se ha obtenido información a través del Cuestionario del Profesor. En este cuestionario se pide al profesor que valore ocho dimensiones competenciales y el grado en el que considera importante incluirlas en sus evaluaciones. Estas ocho dimensiones son: a) escribir bien en Inglés, b) hablar bien en Inglés, c) entender bien el Inglés hablado, d) conocer bien la gramática inglesa, e) leer bien en Inglés, f) pronunciar correctamente el Inglés, g) conocer bien el vocabulario Inglés y h) tener conocimientos sobre la cultura y la literatura de países de habla inglesa.

Como puede apreciarse, estas dimensiones hacen referencia a las distintas destrezas que el profesor piensa exigir a sus estudiantes. Partimos del hecho de que los profesores siempre tienden a pensar que todo lo que enseñan es muy importante. Cierta tipo de profesor puede dar mucha importancia a escribir bien en Inglés, a conocer bien la gramática y a conocer el vocabulario, por encima de las otras cinco dimensiones. La orientación de este profesor podría calificarse como formal y académica de la enseñanza del Inglés, centrada en la gramática, dando menos importancia a los elementos comunicativos. Alternativamente, podemos concebir un perfil de profesor que valora sobre las demás dimensiones hablar y entender, conocer el vocabulario y tener conocimientos sobre la cultura y la literatura. Se trataría de un perfil de enseñanza con claro énfasis en las funciones comunicativas del lenguaje. También es posible concebir un tipo de profesor que valore mucho todos los elementos relacionados en el párrafo anterior. Estaríamos ante un perfil de profesor exigente y exhaustivo que demanda todo a todos sus alumnos. A diferencia de los dos anteriores, este perfil implica simultáneamente una enseñanza centrada tanto en los elementos formales del lenguaje como en las dimensiones comunicativas.

PROBLEMA DE INVESTIGACIÓN

El objetivo fundamental en cualquier investigación relacionada con rendimiento de los alumnos en el aprendizaje de la lengua extranjera es la identificación de aquellos factores que permiten explicar las diferencias en los niveles de desempeño exhibidos por los estudiantes. Un dato que aparece de forma recurrente tanto en la investigación como en la experiencia cotidiana de cualquier ciudadano que se desplaza por los distintos países de la Unión Europea es el muy distinto nivel de dominio de la lengua inglesa, considerada esta como *lingua franca*, que presentan los ciudadanos de los distintos países. Esto no sería un tema relevante si no fuese por la gran cantidad de recursos y esfuerzos, públicos y privados, que en todos ellos se dedican al fomento del aprendizaje del Inglés.

Como ya se ha señalado, la evaluación es un elemento fundamental en el proceso de enseñanza y aprendizaje. Si las dimensiones competenciales que un profesor incluye en su examen definen operativamente sus opciones didáctico-metodológicas, estas deberían tener una repercusión en los resultados de aprendizaje de sus alumnos.

Si un profesor incluye un amplio elenco de dimensiones competenciales en su evaluación, sus estudiantes tendrán más probabilidades de superar con éxito aquellas pruebas adaptadas a ese espectro competencial. Por otra parte, los alumnos de los profesores que orientan su evaluación selectivamente a algunos aspectos específicos del lenguaje, bien comunicativos, bien formales, tendrán más dificultad en superar con éxito una prueba de amplio espectro competencial.

A efectos operativos de este estudio, se han definido dos modelos de actitud respecto de la evaluación en los profesores. Por una parte, hablamos del modelo *exhaustivo* cuando nos referimos a profesores que dan la máxima importancia a todas y cada una de las ocho dimensiones competenciales analizadas. Por otra, denominamos modelo *orientado* a la opción de aquellos profesores que consideran algún subconjunto de las mencionadas dimensiones como más relevantes que las otras, independientemente del contenido del subconjunto seleccionado.

Nuestra hipótesis de partida es que los alumnos de los profesores que se atienen al modelo aquí denominado *exhaustivo* obtienen resultados medios superiores en Comprensión lectora y oral en lengua inglesa a los de los alumnos de los profesores del modelo denominado *orientado*.

METODOLOGÍA

La metodología utilizada en este estudio está condicionada por su propia naturaleza, dado que se trata de un análisis secundario de los datos procedentes del EECL. El diseño de EECL responde a un patrón complejo de medida, por lo que se ha hecho uso de la metodología de valores plausibles (Wu, 2005, 2010) basada en los trabajos previos de Rubin (1976, 1987, 1996) sobre imputación múltiple y un diseño de muestreo adaptado a la estructura de la población, lo que para el cálculo de los errores típicos de algunos estadísticos exige procedimientos de remuestreo adecuados al objeto del estudio. Los detalles técnicos completos se pueden consultar en ESCL (2012).

En este trabajo se han utilizado dos bases de datos distintas: una referida a estudiantes y otra referida a profesores. Aunque el EECL incluye datos de rendimiento referidos tanto a la primera como a la segunda lengua extranjera, aquí se han seleccionado las puntuaciones obtenidas por los estudiantes para los que el Inglés es la primera lengua extranjera. Consecuentemente, se ha trabajado solo con las respuestas de los profesores que enseñan Inglés como primera lengua extranjera.

De entre todos los países evaluados, los incluidos en este estudio son solo aquellos en los que el Inglés puede estudiarse como lengua extranjera en primera opción. Estos países son: Bélgica (comunidad francófona), Bulgaria, Estonia, Grecia, España, Francia, Croacia, Países Bajos, Polonia, Portugal, Suecia y Eslovenia. Se ha excluido Malta intencionalmente de esta muestra al considerar que hasta 1979 ha estado bajo influencia británica, y el Inglés es lengua oficial junto con el maltés.

El diseño muestral de sujetos e ítems supone la utilización de los procedimientos de remuestreo adecuados para obtener estimaciones insesgadas de los errores asociados a los estadísticos estimados, y de valores plausibles obtenidos de las distribuciones a posteriori para cada sujeto, garantizando de ese modo que no se subestima el tamaño del error de medida. En este estudio se ha hecho uso de las réplicas generadas por el procedimiento JRR y particularmente de las variables de ponderación correspondientes a la Comprensión lectora y oral (para más información véase el ESCL *Technical Report*, 2012).

VARIABLES

Son cuatro las variables implicadas en este análisis secundario de los datos de EECL: la variable independiente que está relacionada con la valoración de los elementos

importantes en la evaluación por parte de los profesores; las variables de respuesta que son las puntuaciones obtenidas en las pruebas de Comprensión lectora y de Comprensión oral, y dos variables de control como son el país al que pertenece el estudiante y el nivel socio-cultural de la familia que es un factor clásico en los estudios de eficacia y que contribuye a explicar una parte de la varianza entre los resultados de los estudiantes.

■ ***Variable Independiente: Opinión sobre los elementos importantes en la evaluación del Inglés como segunda lengua***

El Cuestionario del Profesor se ha aplicado a los docentes de un mismo centro que voluntariamente han querido responder al cuestionario. Sin embargo, debido al diseño del EECL no es posible vincular las respuestas de un profesor específico con los resultados de los estudiantes. Esto supone que aunque las respuestas de los alumnos no pueden ser vinculadas directamente con los profesores que ese año académico están impartiendo clases, se pueden ligar las respuestas de los alumnos a características globales del centro al que está adscrito. La agregación de las respuestas de los profesores de un centro se convierte en una característica del mismo. Por tanto, en el nivel del alumno los valores del rendimiento individual pueden relacionarse con el conjunto de respuestas proporcionadas por los profesores del centro al que asiste durante la secundaria. Esta decisión de agregación tiene además una justificación sustantiva. El estudio EECL evalúa a los estudiantes con una edad aproximada de 15 años, estos llevan en el sistema educativo una media superior a doce años de escolaridad, por tanto no tendría sentido vincular sus rendimientos exclusivamente a la acción de un profesor específico. Después de haber cursado toda la Educación Secundaria Obligatoria tiene más sentido relacionar el rendimiento de los alumnos a la intervención concertada del centro escolar en su conjunto. Naturalmente, algunos alumnos habrán pasado por más de un centro escolar en su vida académica. Pero, además de tratarse de datos aislados, no es una información de la que se disponga en la base de datos. Por tanto, como se verá más adelante por su forma de construcción, la variable independiente es una variable que caracteriza el modelo evaluativo predominante del centro donde el alumno está estudiando.

El tratamiento de las respuestas del profesor se ha hecho de la siguiente forma:

- En primer lugar se ha calculado la media de las ocho variables que incluye el cuestionario sobre elementos importantes en su evaluación para obtener la valoración media de cada profesor.
 - En el Cuestionario del Profesor, las ocho cuestiones sobre evaluación se valoran en una escala tipo Likert de cuatro

alternativas. Los valores de la escala son: “Nada importante” (0), “Poco importante” (1), “Bastante importante” (2) y “Muy importante” (3).

- El rango de la variable promedio de cada profesor varía entre 0 y 3.
- Para obtener el perfil de la escuela sobre los elementos que consideran importantes sus profesores en la evaluación de la lengua inglesa se calcula la puntuación media de forma agregada para cada centro. De esta forma se obtiene una puntuación media de importancia de la evaluación para cada escuela.
- Esta media es un indicador indirecto del estilo evaluador del centro. Con la intención de obtener dos modelos evaluativos diferenciados, se han dicotomizado las respuestas medias de centro en función del valor promedio de esta variable. Como resultado de un análisis exploratorio de conglomerados, dada la distribución de los valores observados se tomó como punto de corte el valor 2,50.
 - Los centros con un valor promedio igual o inferior al punto de corte fueron asignados al denominado modelo *orientado*. En ese caso se asignó el código cero en la variable independiente.
 - Los centros con un valor promedio superior al punto de corte fueron asignados al denominado modelo *exhaustivo*. En ese caso se asignó el código uno (1) en la variable independiente.

■ **Variable Dependiente: Puntuaciones obtenidas en las pruebas de Comprensión lectora y oral en Inglés**

Se utilizan dos variables de respuesta: la comprensión lectora y la Comprensión oral. Como en EECL cada alumno contestó solo una parte de todas las preguntas utilizadas, la asignación de puntuaciones individuales se ha llevado a cabo mediante la imputación de Valores Plausibles (Wu, 2005, 2010; ESCL, 2012).

En esencia, en lugar de obtener para cada alumno un estimador puntual de la destreza medida, se obtiene para cada uno una distribución a posteriori, de la que luego se extraen al azar cinco valores, denominados valores plausibles. Este procedimiento tiene la ventaja de que permite una mejor estimación de la varianza del error de medida, por lo que se disminuye la probabilidad de los errores de tipo I al hacer inferencias respecto de la media y otros valores poblacionales. Como contrapartida, no disponemos de un solo valor para cada individuo, y de hecho, dos alumnos con el mismo conjunto de respuestas a cada ítem pueden tener distintos conjuntos de valores plausibles, con medias diferentes. Por este motivo todos los análisis

estadísticos realizados con estas variables difieren de los procedimientos estándar que son habituales en los paquetes estadísticos convencionales.

PLAN DE ANÁLISIS DE DATOS

La base de datos utilizada finalmente ha vinculado los valores de cada centro correspondientes a la variable independiente construida del modo que anteriormente se explicó, con la base de datos de los estudiantes. A cada estudiante se le ha asignado el valor medio calculado para su centro en la variable independiente.

El muestreo corresponde a un diseño estratificado de dos etapas, con las escuelas como PSU's (*Primary Sampling Unit*), con probabilidad proporcional al tamaño del centro y con igual tamaño de muestra para todos los centros en la segunda etapa. A este diseño corresponde una variable de ponderación para poder calcular adecuadamente los estimadores puntuales de cada variable.

Por otra parte, para la estimación de la varianza de error el proyecto EECL utiliza un procedimiento de resmuestro, en este caso concreto una variante del método *Jackknife* conocido como JRR o JK2. Para la lengua extranjera de primera opción este procedimiento utiliza 40 réplicas. Esto supone que se generan 40 variables de ponderación que nos permiten obtener una mejor estimación de la varianza muestral de cada estimador.

Como para cada registro de la base de datos a cada variable dependiente corresponden cinco valores plausibles, la estimación de los errores de medida y de muestreo alcanza cierto nivel de complejidad máxime considerando un contexto multinivel como el definido. Por este motivo se han adaptado al caso las macros de SPSS desarrolladas para PISA (2009, p. 203 y 217) y por Eveline Gebhardt (Australian Council for Educational Research)¹.

Se ha utilizado un modelo jerárquico lineal de dos niveles –alumnos (*i*) y centros (*j*)–para cada país (Snijders y Bosker, 1999; Gaviria y Castro, 2005). El modelo presentado se ha aplicado a cada una de las dos variables dependientes (y_{ij}) incluidas en este trabajo. Este modelo incluye como predictores el nivel socio-cultural de la familia (*ESCS*) y el modelo de evaluación del centro (*M_Evaluación*), conforme a la *ecuación 1*:

$$y_{ij} = \beta_0 + \beta_1 ESCS_{ij} + \beta_2 M_Evaluación_{ij} + (\mu_{1j} ESCS_{ij} + \mu_{2j} M_Evaluación_{ij} + \mu_{0j} + \varepsilon_{ij})$$

¹ Disponibles en <http://www.oecd.org/dataoecd/52/24/42628268.zip>.

Los errores aleatorios de los alumnos y las escuelas se distribuyen conforme a una distribución normal con media cero y varianzas constantes.

DISCUSIÓN DE RESULTADOS

El *Gráfico 3.1* muestra las respuestas brutas al Cuestionario del Profesor en las ocho variables relacionadas con la valoración de las dimensiones competenciales en la evaluación por parte del docente.

Se observa una distribución bastante asimétrica, en la que la práctica totalidad de las respuestas se concentran en las valoraciones “Bastante importante” y “Muy importante”. Sobre este hecho caben al menos dos consideraciones.

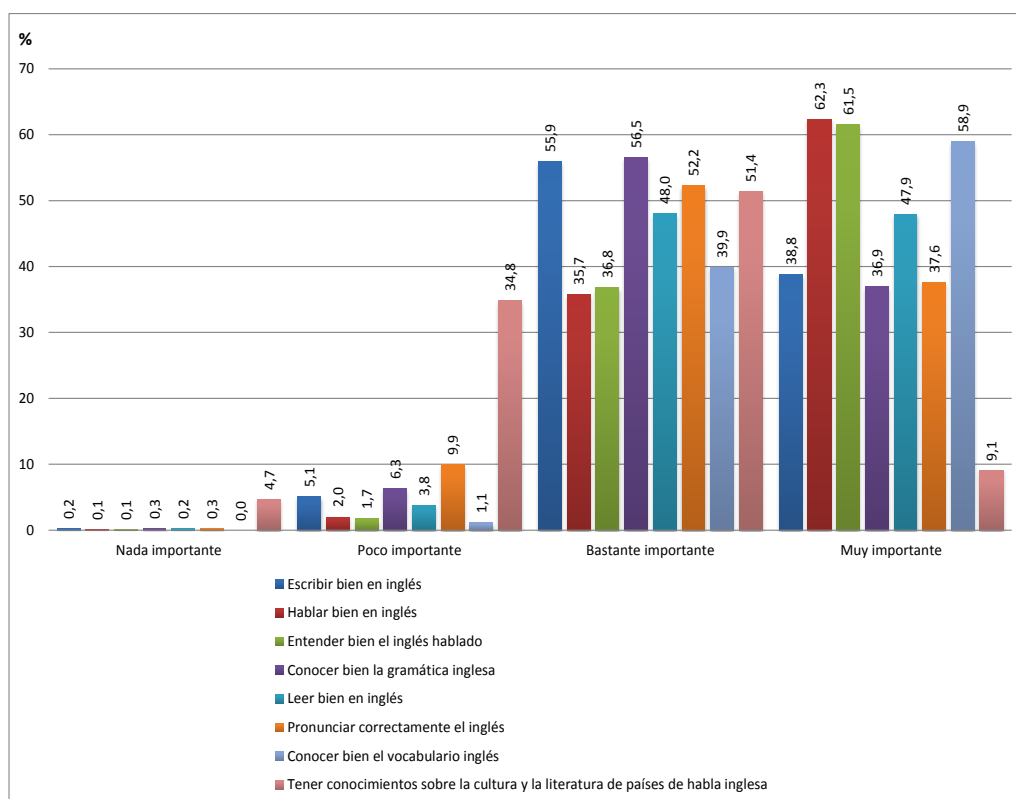
La primera hace referencia a la posibilidad más que real de que las respuestas estén influidas por la deseabilidad social. Es muy difícil que un profesor considere poco importante cualquiera de las dimensiones investigadas, especialmente en una materia como la enseñanza de las lenguas en la que es especialmente importante el trabajo departamental, y, en consecuencia, donde está más presente la presión del grupo.

La segunda se refiere al planteamiento de la pregunta, pues lo que muestra esta distribución es la poca capacidad de discriminación de las preguntas. Hay muy poca varianza en las respuestas de los profesores, lo que equivale a decir que estas variables son poco informativas.

Pero lo que es más importante, la forma en la que las preguntas están planteadas no permite obtener una imagen precisa de cómo es la evaluación que llevan a cabo los profesores, ya que difícilmente las respuestas obtenidas pueden considerarse como un fiel reflejo de la evaluación que los profesores llevan a cabo. No es que todos los profesores hagan lo mismo, es que todos los profesores contestan de la misma forma. No es un problema de actividad uniforme, sino de falta de potencia discriminadora del instrumento de medida.

Lo interesante de esta variable está en que incluso con esta distribución tan concentrada en solo dos valores, aparentemente muy similares, muestra diferencias en los niveles de logro de los alumnos, como veremos más adelante.

Gráfico 3.1. Valoraciones de los profesores sobre la importancia de cada elemento en la evaluación de los alumnos



Efectivamente, una cosa es lo que ideológicamente considera el profesor respecto a las dimensiones a evaluar, el discurso que explícita o implícitamente asume y otra distinta lo que operativamente incluye como contenido de su evaluación. El tiempo para la evaluación es un recurso escaso. Y necesariamente todos los profesores deben hacer un uso y un reparto intencional que responde a sus preferencias personales, preferencias de las que muchas veces incluso el propio profesor no es completamente consciente y que responden a modelos interiorizados durante toda su etapa de formación y de práctica profesional.

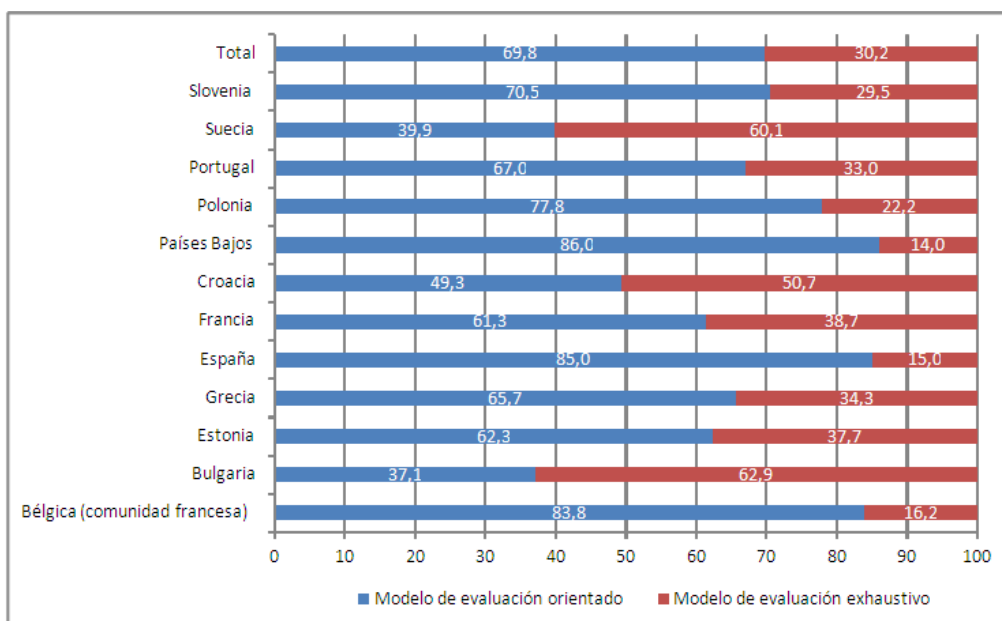
Es posible que un profesor considere que todas las competencias son igualmente importantes pero que a la hora de la verdad asigne doble tiempo, por ejemplo, a la traducción de textos que al examen de la Comprensión oral, o que en un examen una de las partes relacionadas con vocabulario tenga menos puntuación asignada que otra relacionada con la correcta construcción de oraciones gramaticales. Y la forma en que estas preguntas están planteadas no permite inferir cuál es la conducta real de los profesores en la evaluación. Y esto es muy importante, pues como se ha sugerido

anteriormente, son precisamente los exámenes que los alumnos tienen que responder los que establecen a sus ojos cuáles son las verdaderas prioridades del profesor en el aprendizaje de la lengua extranjera. Y es precisamente a esos mensajes implícitos en la evaluación a los que más fuertemente reaccionan los alumnos, adaptando su conducta del modo que garantice mayores probabilidades de éxito escolar.

Como consecuencia de estas consideraciones, cabe apuntar la posibilidad de que en el futuro estas cuestiones se presenten en formatos que sean homogéneos con la realidad sobre la que pretenden recabar información. En efecto, al planificar su evaluación, el profesor dispone de ciertos recursos limitados, como son el tiempo dedicado a la evaluación y la puntuación relativa que asigna a cada una de las dimensiones competenciales. Pues las preguntas debieran ser isomórficas con esa realidad. Por ejemplo, podría pedirse a cada profesor que distribuyese cien puntos entre las distintas dimensiones del mismo modo que puntuaría su examen. Otra pregunta podría pedir que distribuyese un tiempo fijo de evaluación entre las dimensiones analizadas. Otra alternativa sería que los profesores ordenasen todas las competencias en función de varios criterios, por ejemplo de mayor a menor importancia en la puntuación, en el tiempo que le asignarían, en las tareas que pedirían cada día o a la semana, etc... Todas estas formas de evaluación tienen en común la particularidad de que piden al profesor que reparta un recurso fijo entre las dimensiones competenciales. Esto permitiría diferenciar mejor los distintos estilos evaluadores de los docentes. Naturalmente aun así nos enfrentaríamos con el problema de la correspondencia entre lo afirmado por los profesores y la realidad del aula, algo que siempre ocurre cuando se toma a los individuos como fuente de información respecto de realidades objetivas. Pero eliminaríamos el efecto homogeneizador que se produce con las preguntas con formato de respuesta irrestricta, como es el caso del cuestionario que nos ocupa.

En el *Gráfico 3.2* podemos ver la distribución de los dos modelos construidos por países. En el conjunto de países estudiados se observa que un 70% de los centros de la muestra se asignan al modelo *orientado* mientras que el 30% restante es asignado al modelo *exhaustivo*. Esta tendencia se mantiene en la descripción del conjunto de centros por países, salvo en el caso de Suecia, Croacia y Bulgaria, lo que como veremos más adelante tiene cierta importancia.

Gráfico 3.2. Distribución por países de los dos modelos evaluativos



La *Tabla 3.1* muestran los resultados de los principales coeficientes del modelo multinivel definido en la *ecuación 1* para las variables dependientes de Comprensión lectora y Comprensión oral en función del modelo evaluativo del centro (*orientado-exhaustivo*) y del nivel socio-cultural de la familia.

La presentación de los resultados va a estar centrada en los coeficientes fijos del modelo, pues nos va a permitir cuantificar la magnitud de la influencia de cada uno de los dos predictores incluidos y su significación estadística. Nos centraremos especialmente en el modelo evaluativo. Hay que tener en cuenta que el punto de corte representa el rendimiento medio de un alumno (*i*) en el centro (*j*) de un determinado país cuando el nivel socio-cultural de su familia está en la media y el modelo evaluativo del centro en el que estudia es el denominado *orientado*. Como en cualquier modelo de regresión, el coeficiente del predictor indica el incremento en el rendimiento con respecto al punto de corte cuando aumenta un punto el nivel socio-cultural de la familia, por un lado, y por otro, cuando el modelo evaluativo del centro pasa a ser *exhaustivo*.

Tal y como muestran todos los trabajos que estudian la influencia del nivel socio-cultural de las familias en el rendimiento académico, en términos generales, la magnitud de este coeficiente es mayor que el coeficiente del modelo evaluativo en las

dos variables dependientes y en todos los países, salvo como veremos a continuación en un par de significativas excepciones que son Bulgaria y España.

El estudio del comportamiento del predictor Modelo de Evaluación por países muestra que en Comprensión Lectora (Tabla 3.1) tiene una influencia significativa en esta competencia a favor del modelo *exhaustivo* en Bulgaria, España, Portugal y Eslovenia ($p < 0,05$). Además es destacable que la magnitud de este coeficiente es superior al mostrado por el nivel socio-cultural de la familia en Bulgaria (0,344 para el Nivel Socio-económico frente a 0,691) y España (0,349 para el Nivel Socio-económico frente a 0,373). Esto es un fuerte indicador de que en estos dos países el modelo evaluativo denominado *exhaustivo* es especialmente conveniente para el desarrollo de la Comprensión lectora. En el resto de países, donde el modelo evaluativo no tiene efecto significativo, la magnitud del coeficiente tiende a cero.

Tabla 3.1. Modelo multinivel para Comprensión lectora y Comprensión oral en función del nivel socio-cultural de la familia y del modelo evaluativo del centro por país

| PAÍS | | C. ORAL | | C. LECTORA | |
|-----------------------|-------------------------------------|---------|----------|------------|----------|
| | | COEF. | E.TÍPICO | COEF. | E.TÍPICO |
| BÉLGICA (C. FRANCESA) | P. Corte (β_0) | 0,330 | (0,027) | 0,372 | (0,035) |
| | Nivel Socio-económico (β_1) | 0,165 | (0,034) | 0,299 | (0,064) |
| | Modelo Evaluativo (β_2) | 0,000 | (0,047) | 0,073 | (0,049) |
| | Varianza entre escuelas | 0,152 | (0,013) | 0,187 | (0,038) |
| | Varianza intra escuelas | 0,017 | (0,024) | 0,787 | (1,127) |
| BULGARIA | P. Corte (β_0) | 0,238 | (0,036) | -0,030 | (0,046) |
| | Nivel Socio-económico (β_1) | 0,348 | (0,034) | 0,344 | (0,056) |
| | Modelo Evaluativo (β_2) | 0,734* | (0,060) | 0,691* | (0,057) |
| | Varianza entre escuelas | 0,485 | (0,051) | 0,664 | (0,088) |
| | Varianza intra escuelas | 0,018 | (0,025) | 1,082 | (1,552) |
| ESTONIA | P. Corte (β_0) | 1,418 | (0,032) | 1,298 | (0,044) |
| | Nivel Socio-económico (β_1) | 0,368 | (0,069) | 0,388 | (0,258) |
| | Modelo Evaluativo (β_2) | 0,003 | (0,049) | 0,015 | (0,095) |
| | Varianza entre escuelas | 0,502 | (0,047) | 0,539 | (0,205) |
| | Varianza intra escuelas | 0,096 | (0,127) | 1,334 | (1,904) |
| GRECIA | P. Corte (β_0) | 0,766 | (0,039) | 0,680 | (0,051) |
| | Nivel Socio-económico (β_1) | 0,312 | (0,037) | 0,383 | (0,045) |
| | Modelo Evaluativo (β_2) | -0,032 | (0,063) | 0,013 | (0,090) |
| | Varianza entre escuelas | 0,228 | (0,037) | 0,073 | (0,029) |
| | Varianza intra escuelas | 0,013 | (0,018) | 1,576 | (2,259) |
| ESPAÑA | P. Corte (β_0) | 0,162 | (0,016) | 0,390 | (0,048) |
| | Nivel Socio-económico (β_1) | 0,267 | (0,026) | 0,349 | (0,038) |
| | Modelo Evaluativo (β_2) | 0,257* | (0,052) | 0,373* | (0,103) |
| | Varianza entre escuelas | 0,291 | (0,038) | 0,252 | (0,047) |
| | Varianza intra escuelas | 0,002 | (0,003) | 1,001 | (1,441) |

| | | | | | |
|--------------|-------------------------------------|---------|---------|--------|---------|
| FRANCIA | P. Corte (β_0) | -0,042 | (0,053) | -0,155 | (0,079) |
| | Nivel Socio-económico (β_1) | 0,281 | (0,033) | 0,359 | (0,025) |
| | Modelo Evaluativo (β_2) | -0,029 | (0,061) | 0,036 | (0,101) |
| | Varianza entre escuelas | 0,215 | (0,027) | 0,083 | (0,030) |
| | Varianza intra escuelas | 0,001 | (0,001) | 0,838 | (1,206) |
| CROACIA | P. Corte (β_0) | 1,155 | (0,024) | 0,609 | (0,033) |
| | Nivel Socio-económico (β_1) | 0,365 | (0,031) | 0,507 | (0,041) |
| | Modelo Evaluativo (β_2) | -0,165* | (0,033) | 0,019 | (0,056) |
| | Varianza entre escuelas | 0,171 | (0,010) | 0,168 | (0,043) |
| | Varianza intra escuelas | 0,024 | (0,033) | 1,260 | (1,807) |
| PAÍSES BAJOS | P. Corte (β_0) | 1,855 | (0,069) | 1,198 | (0,065) |
| | Nivel Socio-económico (β_1) | 0,153 | (0,063) | 0,222 | (0,066) |
| | Modelo Evaluativo (β_2) | -0,041 | (0,126) | 0,034 | (0,153) |
| | Varianza entre escuelas | 0,414 | (0,093) | 0,445 | (0,037) |
| | Varianza intra escuelas | 0,007 | (0,009) | 0,836 | (1,199) |
| POLONIA | P. Corte (β_0) | 0,483 | (0,106) | 0,355 | (0,142) |
| | Nivel Socio-económico (β_1) | 0,348 | (0,054) | 0,578 | (0,082) |
| | Modelo Evaluativo (β_2) | 0,158* | (0,054) | 0,092 | (0,099) |
| | Varianza entre escuelas | 0,208 | (0,137) | 0,192 | (0,106) |
| | Varianza intra escuelas | 0,003 | (0,004) | 0,945 | (1,357) |
| PORTUGAL | P. Corte (β_0) | 0,787 | (0,018) | 0,430 | (0,052) |
| | Nivel Socio-económico (β_1) | 0,380 | (0,021) | 0,363 | (0,024) |
| | Modelo Evaluativo (β_2) | 0,166* | (0,056) | 0,155* | (0,068) |
| | Varianza entre escuelas | 0,109 | (0,015) | 0,164 | (0,034) |
| | Varianza intra escuelas | 0,012 | (0,017) | 1,149 | (1,650) |
| SUECIA | P. Corte (β_0) | 2,110 | (0,034) | 1,858 | (0,040) |
| | Nivel Socio-económico (β_1) | 0,347 | (0,018) | 0,451 | (0,027) |
| | Modelo Evaluativo (β_2) | 0,141* | (0,054) | 0,035 | (0,043) |
| | Varianza entre escuelas | 0,101 | (0,014) | 0,161 | (0,062) |
| | Varianza intra escuelas | 0,010 | (0,014) | 1,389 | (1,992) |
| ESLOVENIA | P. Corte (β_0) | 1,452 | (0,029) | 0,886 | (0,055) |
| | Nivel Socio-económico (β_1) | 0,388 | (0,036) | 0,434 | (0,025) |
| | Modelo Evaluativo (β_2) | 0,102* | (0,038) | 0,229* | (0,075) |
| | Varianza entre escuelas | 0,122 | (0,010) | 0,127 | (0,041) |
| | Varianza intra escuelas | 0,060 | (0,081) | 1,339 | (1,912) |

* Coeficiente vinculado al modelo evaluativo significativo con $p < 0,05$.

El estudio del comportamiento del predictor Modelo de Evaluación por países en Comprensión Oral (*Tabla 3.1*) muestra una influencia significativa en esta competencia a favor del modelo *exhaustivo* en Bulgaria, España, Polonia, Portugal, Suecia y Eslovenia. Se observa también una influencia significativa y a favor del modelo *orientado* en Croacia. Al igual que en Comprensión lectora, el coeficiente en Bulgaria (0,348 para el Nivel Socio-económico frente a 0,734) y en España (0,267 para el Nivel Socio-económico frente a 0,257) tiene un peso similar o incluso superior al peso del Nivel Socio-cultural de la familia. Se puede decir por tanto que en Bulgaria y España es especialmente significativo el modelo *exhaustivo* en la Comprensión oral. En este caso,

solo en cinco países de los estudiados el modelo evaluativo no tiene efecto significativo, siendo la magnitud del coeficiente tendente a cero.

Es importante resaltar el hecho de que siempre que se observan diferencias estadísticamente significativas son a favor de las medias de los alumnos que se encuentran en centros con modelo *exhaustivo* predominante, salvo en el caso de Croacia en Comprensión oral.

El caso de España es especialmente interesante. Los alumnos españoles no destacan por su desempeño, pues España ocupa la penúltima posición entre todos los países participantes en las pruebas de Inglés como primera lengua extranjera. En nuestro país el 85% de los alumnos están en centros en los que el modelo predominante es el *orientado*, y solo un 15% de los casos corresponden a centros de modelo *exhaustivo*. Pero la influencia del modelo evaluativo *exhaustivo* es significativa tanto en Comprensión lectora como en Comprensión oral, teniendo tanta influencia sobre el rendimiento como en el Nivel Socio-cultural de la familia.

Se ha realizado también un análisis de diferencia de medias con los datos españoles en función del modelo evaluativo. Se ha comprobado la existencia de diferencias estadísticamente significativa a favor del modelo *exhaustivo* tanto en Comprensión lectora ($t = 3,01$, $gl = 1001$) como en Comprensión oral ($t = 3,25$, $gl = 1026$). Estos análisis nos permiten también informar de la magnitud del efecto pues habla del impacto que tiene un tratamiento sobre otro, de la eficacia cuantificada entre los distintos niveles de la variable independiente. Estadísticamente es una estandarización de las diferencias encontradas entre los grupos de tratamiento. Desde este punto de vista, complementa la información ofrecida por los contrastes de diferencia de medias, pues además de confirmar la existencia de estas, también informa de la magnitud estandarizada de las diferencias de un grupo sobre otro.

El tamaño del efecto en Comprensión lectora es 0,4 desviaciones típicas a favor del modelo *exhaustivo*. Esto significa que el 65% de los estudiantes escolarizados en centros con modelo *orientado* están por debajo de la media de los estudiantes de los centros con modelo *exhaustivo* predominante. De modo análogo, el tamaño del efecto en Comprensión oral es 0,32 desviaciones típicas a favor del modelo *exhaustivo*, mostrando que el 63% de los estudiantes escolarizados en centros con modelo *orientado* están por debajo de la media de los alumnos escolarizados en centros con modelo *exhaustivo*.

Resulta muy sugerente comprobar que en la variable modelo evaluativo tiene influencia sobre el desempeño en Inglés como primera lengua extranjera, pues es una variable que es manipulable por el profesor y el centro educativo. Cuando se

encuentran resultados significativos, el modelo evaluativo *exhaustivo* tiene una influencia positiva (salvo en el caso de Croacia en Comprensión oral). Esto quiere decir que la media de rendimiento de los alumnos que están en centros en los que predomina el modelo *exhaustivo* es superior a la de los alumnos que están en centros en los que predomina el modelo *orientado*.

CONCLUSIONES

Aunque los resultados no son estadísticamente concluyentes, apuntan en la dirección de que existe una relación sistemática entre la forma de evaluación predominante en un centro y el resultado promedio obtenido por sus alumnos en las dimensiones de Comprensión oral y Comprensión lectora, incluso después de haber controlado una variable con acreditada influencia sobre el rendimiento como es el Nivel socio-cultural de la familia. Estas diferencias apuntadas se inclinan mayoritariamente del lado del modelo denominado *exhaustivo*. Es importante señalar que aunque parece razonable que exista una relación entre el modo de evaluar de los profesores y el rendimiento de sus alumnos, en este estudio este rendimiento no es el medido por los propios profesores, sino el obtenido en una evaluación independiente y objetiva, lo que realza la importancia del resultado.

Las limitaciones del estudio en lo que a los instrumentos de medida utilizados, especialmente los relacionados con las informaciones proporcionadas por los propios profesores, pueden explicar, parcialmente, el hecho de que las diferencias observadas, aunque casi siempre apuntando en una misma dirección, no siempre lleguen a ser estadísticamente significativas.










Cabe legítimamente preguntarse por el verdadero significado de la variable independiente que aquí se ha construido. Que refleja alguna diferencia importante entre unos y otros profesores y entre centros, parece indudable, especialmente después de comprobar que estas diferencias se reflejan parcialmente en los rendimientos de los alumnos en unas pruebas que son independientes de las percepciones de los profesores.













Ahora bien, ¿es posible que los profesores que asignan a todas o casi todas las dimensiones competenciales la máxima importancia sin hacer distinciones, tengan una actitud más exigente en general hacia los alumnos y que esto se refleje en su rendimiento? Si así fuese, es evidente que la consecuencia más lógica de esto sería la promoción de una actitud de mayor rigor académico entre los profesores. También es

posible que esta variable independiente esté asociada a otros elementos causales que en este momento no son evidentes. Para aclarar este punto sería necesario llevar a cabo algunos cambios en los instrumentos de medida utilizados y en el propio diseño de las evaluaciones como las que ahora nos ocupa.

Y no se escapa a nuestra consideración la enorme importancia de una variable independiente que en última instancia puede ser modificable, aunque como todo en relación a las actitudes de los profesores, no sin grandes esfuerzos por parte de todo el sistema, administradores, directores e incluso los propios profesores.

Referencias bibliográficas

-  Anghel, B. y Güell, M. (2012). Los determinantes del aprendizaje del Francés en España.
-  De la Rica, S. y González de San Román, A. (2012). Determinantes del rendimiento académico en competencia inglesa en España: claves para la mejora.
-  European Survey on Competence Language (ESCL) (2012). First European Survey on Language Competences: Technical Report. Survey Lang.
-  Freeman, D. y Johnson, K.E. (1998). Re-conceptualizing the knowledge-base of language teacher education. *TESOL Quarterly*, 32 (3), 397-417.
-  Freeman, D. y Johnson, K.E. (2005). Toward linking teacher knowledge and student learning. En D. Tedick (Ed.) *Second language teacher education: International perspectives* (pp. 73-96). Mahwah, NJ: Lawrence Erlbaum Associates.
-  García Legazpe, F. (2008). *Motivar para el aprendizaje desde la actividad orientadora*. Madrid: Centro de Investigación y Documentación Educativa.
-  Gaviria, J. L. y Castro, M. (2005). *Modelos Jerárquicos Lineales*. Madrid: La Muralla-Hespérides.
-  Kieli, R. (2001). Classroom evaluation, values, interest and teacher development. *Language Teaching Research*, 5 (3), 241-261.
-  Little, R.J.A. y Rubin, D.B. (1987). *Statistical Analysis with Missing Data*. J. Wiley & Sons, New York.

-  OCDE (2009). PISA Data Analysis Manual. SPSS Second Edition.
-  Richards, J., & Rodgers, T. (2001). Approaches and methods in language Teaching (2nd ed.). Cambridge: Cambridge University Press.
-  Rodgers, T. S. (2001). Language Teaching Methodology. Issue Paper. Center for Applied Linguistics. Online Resources: Digests.
-  Rubin, D.B. (1976). Inference and missing data. *Biometrika*, 63, 581-592.
-  Rubin, D.B. (1987). Multiple Imputation for Nonresponse in Surveys. J. Wiley & Sons, New York.
-  Rubin, D.B. (1996). Multiple imputation after 18+ years (with discussion). *Journal of the American Statistical Association*, 91, 473-489.
-  Snijders, T.A.B. y Bosker, R.J. (1999). Multilevel Analysis. An Introduction to basic and advanced multilevel modeling. London: Sage Publications.
-  Tsui, A.M.B. (2011). Teacher education and teacher development. En Eli Hinkel (2011). *Handbook of research in second language and learning* (vol. II). NY: Routledge, (pp. 21-39).
-  Vez, J. M., Martínez, E. y Lorenzo, A. (2012). La exposición y uso ambiental de la lengua extranjera en contextos no formales mejora su Comprensión oral.
-  Woods, D. (1996). Teacher cognition in language teaching: beliefs, decision-making and classroom practice. Cambridge: Cambridge University Press.
-  Wu, M. (2005). The Role of Plausible Values in Large-Scale Surveys. *Studies in Educational Evaluation*, 31, 114-128.
-  Wu, M. (2010). Measurement, Sampling, and Equating Errors in Large-Scale Assessments. *Educational Measurement: Issues and Practice*, 29, (4), 15–27.