

EL DESARROLLO DE LAS ESPECIFICACIONES DE LAS SUBPRUEBAS DE COMPRENSIÓN DE LECTURA Y DE EXPRESIÓN E INTERACCIÓN ESCRITAS EN UNA PRUEBA DE APROVECHAMIENTO ESTANDARIZADA PARA LA EVALUACIÓN DE LA COMPETENCIA EN LENGUA INGLESA EN UN ENTORNO ESCOLAR

JOSÉ CUADRADO MORENO*

Universidad Pablo de Olavide, Sevilla

MARÍA REYES FERNÁNDEZ

Universidad Pablo de Olavide, Sevilla

RESUMEN. *Se presenta un procedimiento para desarrollar la estructura de las subpruebas de comprensión de lectura y de expresión e interacción escritas de una prueba de aprovechamiento estandarizada de inglés para el alumnado del segundo curso de Educación Secundaria Obligatoria de Sevilla (España). Este procedimiento está basado en un estudio observacional de las actividades de aprendizaje en las clases de inglés del segundo curso de Educación Secundaria Obligatoria en Sevilla y en una revisión del currículo de inglés para este curso y en pruebas similares de lengua extranjera. Al contrario de los métodos tradicionales de recoger evidencias basadas en el contenido para las pruebas educativas, este procedimiento es realizado antes de escribir los ítems y las tareas de la prueba.*

PALABRAS CLAVE. *Validez del contenido, prueba de inglés, primer ciclo de educación secundaria, prueba estandarizada.*

ABSTRACT. *A procedure is presented to develop the structure of the subtests of written reception (reading) and written production and interaction (writing) in a standardised achievement test of English for Year-8 pupils in Seville (Spain). This procedure is based on an observational study of the learning activities in the Year-8 classes of English in Seville and on a review of both the Year-8 curriculum of English and similar foreign language tests. As opposed to the usual ways of collecting content-based evidence for educational tests, this procedure is carried out before both test items and tasks are written.*

KEY WORDS. *Content validity, test of English, lower secondary education, standardized test.*

1. INTRODUCCIÓN

El grupo de investigación EALIS está desarrollando la prueba sumativa estandarizada EALIS-ESO-2, cuya finalidad es evaluar, de acuerdo con lo establecido en el R.D. 1631/2006 (art. 3, España 2007: 679), las capacidades de comprensión y expresión en la lengua inglesa de manera apropiada del alumnado de la ciudad de Sevilla al final del segundo curso de la etapa de la educación secundaria obligatoria (E.S.O.). En nuestra opinión, el contenido de esta prueba, es decir, “the themes, wording, and format of the items, tasks, or questions on a test, as well as the guidelines for procedures regarding administration and scoring” (American Educational Research Association, American Psychological Association y National Council on Measurement in Education 1999: 11), debe tener en cuenta:

- 1) Los objetivos y los contenidos del área de lengua extranjera para el segundo curso de la E.S.O., que en la Comunidad Autónoma de Andalucía están regulados por el Real Decreto 1631/2006 (España 2007) y por la Orden de 10 de agosto de 2007 (Andalucía 2007).
- 2) El contenido de las actividades de aprendizaje realizadas en el aula. Uno de los principios que debe regir la elaboración de una prueba sumativa es que la prueba debe incluir tareas que traten los mismos temas y que utilicen el mismo tipo de expresiones, instrucciones y formato que las actividades de aprendizaje realizadas en el aula. Si el contenido de una prueba sumativa consigue ajustarse al contenido de las actividades de aprendizaje realizadas en el aula, esto contribuirá, indudablemente, a la validación de la prueba (American Educational Research Association, American Psychological Association y National Council on Measurement in Education 1999: 184).
- 3) El hecho de que la investigación en evaluación lingüística parece que ha demostrado que la actuación de los examinandos puede estar influida por el tipo de ítem o tarea que tiene que realizar. Shohamy (1983: 532-533, 537), por ejemplo, sugería que las variables “tipo de actividad lingüística comunicativa” y “tema de la actividad” influían en las puntuaciones asignadas a los participantes en un experimento para evaluar sus competencias de expresión e interacción orales (cf. también, por ejemplo, Bachman y Palmer 1982: 461-462). Este hecho viene a apoyar el principio de que la prueba sumativa debe incluir tareas similares a las actividades de aprendizaje realizadas en el aula.

En los estudios acerca de los métodos para obtener evidencias basadas en el contenido de la prueba (por ejemplo, Crocker, Miller y Franks 1989; Osterlind 1989; Oltman, Stricker y Barrows 1990; Shavelson, Gao y Baxter 1995; Sireci 1998), estas evidencias se suelen obtener a partir del análisis de los juicios de una serie de jueces o mediante análisis estadísticos *después* de la redacción de los ítems o tareas. Así, Hambleton (1980) desarrolló un índice de congruencia ítem-objetivo (*item-objective congruence index*) para los ítems de pruebas criterioles que evalúan un único objetivo, cuyo valor dependía de la valoración realizada por una serie de especialistas en la materia del grado en que cada ítem

medía cada objetivo. En este artículo se presenta un procedimiento para ajustar parte del contenido de las especificaciones de la prueba EALIS-ESO-2 a los objetivos y contenidos curriculares, es decir, *antes* de la redacción de los ítems y las tareas. Este procedimiento, que fue ideado por Cuadrado Moreno (2009), consta de dos fases: (i) estudio observacional de clases de lengua inglesa del segundo curso de E.S.O. en la ciudad de Sevilla y (ii) elaboración de la especificación de la prueba. Con este procedimiento se pretende, en última instancia, solucionar un problema sobre el cual Popham (2000: 98) llamó la atención: la escasa calidad de las evidencias basadas en el contenido de la prueba que se proporcionan durante el proceso de validación de la prueba.

2. ESTUDIO OBSERVACIONAL DE CLASES DE LENGUA INGLESA DEL SEGUNDO CURSO DE E.S.O. EN LA CIUDAD DE SEVILLA

El estudio observacional estuvo guiado por la siguiente pregunta de investigación: ¿Qué actividades de aprendizaje realizan durante la clase de lengua inglesa los alumnos sevillanos del segundo curso de E.S.O.? Este estudio constó de las siguientes fases (Brown 2001: 8-10):

- 1) Planificación del proyecto de observación.
- 2) Diseño de un instrumento para la observación.
- 3) Recogida y compilación de los datos de la observación.
- 4) Análisis estadístico de los datos.

2.1. *Planificación del proyecto de observación*

La conducta que debía ser observada era la actividad de aprendizaje que el alumnado sevillano del segundo curso de E.S.O. realiza durante la clase de lengua inglesa. El estudio observacional comenzó con una revisión de la literatura sobre la observación en el aula de segundas lenguas (Jarvis 1968: 335-338; Clark 1969: 392; Munby 1978: 154-189; Johnson y Morrow 1981; Malamah-Thomas 1987; Allwright 1988; Chaudron 1988; Lier 1988; Nunan 1989; Allwright y Bailey 1991; Wragg 1994; Spada and Frohlich 1995; Nikolov 1999; Seedhouse 2004).

La cuestión de cuántas categorías de la conducta observada debían distinguirse condujo al problema de determinar cuántas subpruebas debían componer la prueba EALIS-ESO-2. Partiendo de la declaración explícita del legislador español, según la cual éste había tomado como base para la elaboración del currículo del área de lengua extranjera en la E.S.O. el *Marco Común Europeo de Referencia para las Lenguas: Aprendizaje, Enseñanza, Evaluación* (M.C.E.R.) (“Estas pautas [del M.C.E.R.] han sido un referente clave en este currículo [del área de lengua extranjera para la etapa de la E.S.O.]”, España 2007: 742), se decidió que las subpruebas debían tener en cuenta las categorías del M.C.E.R. El problema que se debía solucionar a continuación era elegir como base de la prueba o el constructo de competencia lingüística comunicativa del M.C.E.R. (Trim, North y Coste

2002: 106-127) o el constructo de actividad lingüística comunicativa del M.C.E.R. (Trim, North y Coste 2002: 60-88). Según Trim, North y Coste (2002), las competencias comunicativas de la lengua están compuestas de (i) las competencias lingüísticas, (ii) las competencias sociolingüísticas y (iii) las competencias pragmáticas (p. 106), mientras que para estos mismos autores existen los siguientes tipos de actividades lingüísticas comunicativas: (i) actividades de expresión, (ii) actividades receptivas, (iii) actividades interactivas, (iv) actividades de mediación, y (v) comunicación no verbal (pp. 60-88). En la traducción española del M.C.E.R., por cierto, se tradujeron los términos ingleses *production activities* y *receptive activities* (Trim, North and Coste 2001: 57 y 65, respectivamente) como *actividades de expresión* y *actividades de comprensión*, respectivamente (Trim, North y Coste 2002: 61 y 68, respectivamente). Puesto que en la redacción del objetivo *i* de la etapa de la E.S.O. (“Comprender y expresarse en una o más lenguas extranjeras de manera apropiada”, España 2007: 679), que está relacionado con la enseñanza/aprendizaje de las lenguas extranjeras, aparecen los mismos términos que aparecen en el M.C.E.R. para referirse a tipos de actividades lingüísticas comunicativas, se concluyó que el legislador español había decidido basar la organización del currículo en la taxonomía de actividades lingüísticas comunicativas del M.C.E.R. y, por tanto, se decidió usar esta taxonomía como criterio para la organización de la prueba EALIS-ESO-2.

Se acordó, asimismo, utilizar como modelo para la prueba EALIS-ESO-2 el examen *Flyers* (University of Cambridge ESOL Examinations 2007: 16-26), puesto que *Flyers* (i) está dirigido a un grupo de edad similar (edad media 11.9 años, Marshall y Gutteridge 2002: 5) y (ii) está compuesto de tres subpruebas (*Listening*, *Speaking* y *Reading & Writing*), que, en realidad, son lo que en la terminología del M.C.E.R. se denominan actividades de comprensión auditiva, comprensión lectora, producción e interacción orales y producción e interacción escritas. Por esta razón se decidió observar tres tipos diferentes de actividades de aprendizaje en el aula: (i) actividades de comprensión auditiva, (ii) actividades de comprensión de lectura y de expresión e interacción escritas y (iii) actividades de expresión e interacción orales.

2.2. Diseño de un instrumento para la observación

El instrumento utilizado para registrar las observaciones está basado en el modelo de las facetas de los métodos de una prueba (Bachman 1990: 116-152) y en su desarrollo posterior, el modelo para la descripción de las tareas de una prueba de Bachman y Palmer (1996: 47-58). En la didáctica de segundas lenguas y en la evaluación lingüística se han elaborado diferentes instrumentos que se pueden utilizar para observar las actividades de aprendizaje en un aula de enseñanza de segundas lenguas y para describir las tareas de una prueba (Jarvis 1968: 335-338; Clark 1969: 392; Munby 1978: 154-189; Allwright 1988; Brown 1989: 51-52; Spada and Frohlich 1995; Norris, Brown, Hudson y Yoshioka 1998: 69-97; Freedle y Kostin 1999: 26-31; Nikolov 1999; Weigle 2002: 60-64; Carr 2006; North, Figueras, Takala, Van Avermaet y Verhelst 2009: 25-34). Sin embargo, hasta donde hemos podido saber, ninguno de ellos ha utilizado para la obser-

vación un número tan elevado de variables como el modelo de Bachman y Palmer (1996), que, además, puede ser utilizado para la descripción tanto de actividades de aprendizaje como de tareas en una prueba. Este modelo ha sido utilizado en diferentes estudios de evaluación lingüística, como Bachman, Davidson, Ryan y Choi (1995) o Bachman, Davidson y Milanovic (1996).

2.3. *Recogida y compilación de los datos de la observación*

Los observadores coordinaron los procedimientos para la observación antes de realizar la observación en una serie de reuniones en las que se analizaron el modelo para la descripción de las tareas de una prueba de Bachman y Palmer (1996: 47-58) y el instrumento para la observación. A continuación, se observaron a lo largo del primer trimestre del curso 2007-2008 las actividades de aprendizaje realizadas en la clase de lengua inglesa (que suele tener una duración aproximada de una hora) en una muestra aleatoria simple proporcional estratificada de diecinueve grupos, seleccionados entre los trescientos grupos del segundo curso de E.S.O. que, según la Delegación Provincial de la Consejería de Educación en Sevilla, existían en la ciudad de Sevilla durante el primer trimestre del curso académico 2007-2008 (Kish 1965: 75-86). Los estratos utilizados fueron: (i) grupos no bilingües de centros públicos, (ii) grupos bilingües de centros públicos y (iii) grupos de centros privados. Por tanto, la forma de la conducta observada fue no secuencial (Kitsantas, Ware y Kitsantas 2005: 913-914).

2.4. *Análisis estadístico de los datos recogidos y presentación de los datos*

Los datos recogidos durante el estudio observacional han sido publicados en Sánchez, Cuadrado, Reyes, Fernández, Hernández, Fielden y Ocaña (2009), Reyes, Fernández, Hernández, Fielden, Ocaña y Sánchez (2009) Cuadrado Moreno, Reyes Fernández, Fernández Amaya, Sánchez Fernández, Fielden, Ocaña Escolar y Hernández López (en prensa).

3. ELABORACIÓN DE LAS TABLAS DE ESPECIFICACIONES DE LAS SUBPRUEBAS DE COMPRENSIÓN DE LECTURA Y DE EXPRESIÓN Y DE INTERACCIÓN ESCRITAS

La elaboración de las tablas de especificaciones de la subpruebas estuvo guiada por el procedimiento de Linn y Gronlund (2000: 143), que consta de tres fases: (i) elaboración de una lista de los objetivos de la enseñanza, (ii) elaboración de una lista con los contenidos del curso y (iii) elaboración de la tabla de las especificaciones.

Aunque durante el estudio observacional se habían utilizado tres categorías de actividades de aprendizaje, la revisión de la literatura acerca de las pruebas de aprovechamiento similares destinadas a un grupo de edad semejante y basadas en el M.C.E.R. (por ejemplo, France 2007; Rupp, Vock, Harsch y Köller 2007; Tuokko 2007; Gassner, Mewald

y Sigott 2008; Gassner, Mewald y Sigott s.d.) condujo a revisar la decisión de diseñar una prueba con tres subpruebas. Finalmente, se acordó que la prueba constaría de las siguientes subpruebas: (i) subprueba de comprensión escrita, (ii) subprueba de expresión e interacción escritas, (iii) subprueba de comprensión auditiva, y (iv) subprueba de expresión e interacción orales.

3.1. *Elaboración de una lista de los objetivos de la enseñanza*

Los objetivos relacionados con las competencias lingüísticas comunicativas para la materia de lengua extranjera en el segundo curso de la E.S.O. aparecen en el Anexo II del Real Decreto 1631/2006, de 29 de diciembre, por el que se establecen las enseñanzas mínimas correspondientes a la Educación Secundaria Obligatoria (España 2006: 743), que establece, entre otras cosas, que

la enseñanza de la lengua extranjera en la etapa de la educación secundaria obligatoria tendrá como finalidad el desarrollo de las siguientes capacidades [...]

3. Leer y comprender textos diversos de un nivel adecuado a las capacidades e intereses del alumnado con el fin de extraer información general y específica, y utilizar la lectura como fuente de placer y de enriquecimiento personal.
4. Escribir textos sencillos con finalidades diversas sobre distintos temas utilizando recursos adecuados de cohesión y coherencia.

Estos objetivos son desarrollados para el segundo curso de E.S.O. en la sección “Criterios de evaluación” (España 2007: 746):

1. Comprender la información general y la específica de diferentes textos escritos, adaptados y auténticos, de extensión variada, y adecuados a la edad, demostrando la comprensión a través de una actividad específica.
2. Redactar de forma guiada textos diversos en diferentes soportes, utilizando estructuras, conectores sencillos y léxico adecuados, cuidando los aspectos formales y respetando las reglas elementales de ortografía y de puntuación para que sean comprensibles al lector y presenten una corrección aceptable.

Desafortunadamente, el R.D. 1631/2006 no establece un objetivo de nivel de competencia lingüística que el alumno debería alcanzar al final de la etapa de la E.S.O. A partir de la revisión de otras pruebas dirigidas a alumnos del mismo grupo de edad (por ejemplo, France 2007; Rupp, Vock, Harsch and Köller 2007: 86-96; University of Cambridge ESOL 2007: 3; Gassner, Mewald y Sigott 2008: 20; Gassner, Mewald y Sigott s.d.: 1), los autores decidieron establecer como objetivo para el alumnado al final de la etapa de E.S.O. un nivel de competencia lingüística comunicativa en la lengua extranjera correspondiente al nivel de referencia común A2 (Trim, North and Coste 2001: 21-42). Esto conlleva que (i) el alumnado del segundo curso de E.S.O. debería tener una competencia lingüística en la lengua extranjera situada entre los niveles comunes de referencia A1 y A2

y (ii) la prueba de aprovechamiento EALIS-ESO-2 debe tener como objetivo determinar el grado de competencia lingüística comunicativa de los alumnos del segundo curso de E.S.O., que se debe encontrar entre los niveles comunes de referencia A1 y A2.

3.2. *Elaboración de una lista con los contenidos del curso*

En el R.D. 1631/2006 los contenidos de la materia de lengua extranjera para el segundo curso de E.S.O aparecen distribuidos en “bloques”, entre los que se encuentra el “Bloque 2. Leer y escribir” (España 2007: 745):

Bloque 2. Leer y escribir.

- Anticipación del contenido antes y durante la lectura de textos sencillos.
- Comprensión de la información general y específica en diferentes textos, en soporte papel y digital, auténticos y adaptados, sobre asuntos familiares y relacionados con contenidos de otras materias del currículo.
- Iniciativa para leer de forma autónoma textos de cierta extensión.
- Uso de estrategias de comprensión lectora: identificación del tema de un texto, con ayuda de elementos textuales y no textuales; utilización de los conocimientos previos sobre el tema; inferencia de significados por el contexto, por elementos visuales, por comparación de palabras o frases similares en las lenguas que conocen.
- Reconocimiento e iniciación en el uso de algunas fórmulas que diferencian el lenguaje formal e informal en las comunicaciones escritas.
- Composición de distintos textos con ayuda de modelos, atendiendo a elementos básicos de cohesión y utilizando estrategias elementales en el proceso de composición escrita (planificación, textualización y revisión).
- Comunicación personal con hablantes de la lengua extranjera a través de correspondencia postal o utilizando medios informáticos.
- Uso de reglas básicas de ortografía y puntuación y valoración de su importancia en las comunicaciones escritas.
- Interés por la presentación cuidada de los textos escritos, en soporte papel y digital.

3.3. *Elaboración de la tabla de las especificaciones*

El primer paso en la elaboración de las estructuras de las tablas de las especificaciones para las subpruebas consistió en establecer los objetivos de cada subprueba. Esta tarea fue realizada teniendo en cuenta (i) la distinción de Gronlund (2004: 13) entre objetivos generales de la enseñanza y resultados específicos del aprendizaje, (ii) los contenidos que el R.D. 1631/2006 asigna al área de lengua extranjera en el segundo curso de E.S.O., y (iii) que el alumno debía alcanzar al final del segundo curso de E.S.O. un nivel común de referencia entre los niveles A1 y A2.

Gronlund (2004: 13) presentó una serie de definiciones relacionadas con los objetivos de la enseñanza:

1. *Objetivo general de la enseñanza (general instructional objective):* “an intended outcome of instruction that has been stated in general enough terms to

- encompass a domain of student performance (e.g., “Comprehends the literal meaning of written material”)
2. *Resultado específico del aprendizaje (specific learning outcome)*: “an intended outcome of instruction that has been stated in terms of specific and observable student performance (e.g., “Identifies details that are explicitly stated in the passage”)
 3. *Actuación del estudiante (student performance)*: “any measurable or observable student response that is the result of learning”.

Para la selección de los objetivos para la subprueba de comprensión de lectura se consideró que los contenidos del Bloque 2 relacionados con la lectura podían ser asociados a diferentes momentos del proceso lector, cuyo resultado es el segundo contenido del Bloque 2 (“Comprensión de la información general y específica en diferentes textos, en soporte papel y digital, auténticos y adaptados, sobre asuntos familiares y relacionados con contenidos de otras materias del currículo”) (España 2007: 745), que viene a coincidir parcialmente con el objetivo 3 del área de lengua extranjera (“Leer y comprender textos diversos [...] con el fin de extraer información general y específica”) (España 2007: 743) y con el criterio de evaluación 3 en el área de lengua extranjera para el segundo curso de E.S.O. (“Comprender la información general y la específica de diferentes textos escritos”) (España 2007: 746). La redacción de estos objetivos demuestra que el legislador español ha abrazado un concepto multidivisible de la lectura, que no ha sido confirmado por la investigación empírica. En realidad, en la investigación de la lectura no existe un consenso acerca del concepto de lectura y por ello se han propuesto básicamente tres conceptos diferentes de lectura (Alderson 2000: 95): (i) la lectura es una destreza unitaria (Rost 1993), (ii) la lectura es multidivisible (Alderson 1990: 478) y (iii) la lectura consta de dos subdestrezas (Gough, Juel and Griffith 1992). Los investigadores que se basan en un concepto multidivisible de la lectura, por cierto, tampoco se ponen de acuerdo en el número de las destrezas lectoras que se pueden distinguir empíricamente. El hecho de que la prueba EALIS-ESO-2-Reading deba seguir lo dispuesto en el R.D. 1631/2006 nos obligó a establecer como resultados específicos del aprendizaje la identificación de la información general y de la información específica expresadas en un texto escrito, a pesar de ser conscientes de que estos dos objetivos que impone el R.D. 1631/2006 no están fundados en resultados empíricos, sino en la especulación. Por tanto, la redacción de los resultados específicos del aprendizaje evaluados en la subprueba de comprensión de lectura quedó como sigue:

1. El alumno identifica la información general en un texto escrito.
2. El alumno identifica la información específica expresada explícitamente en un texto escrito.

Con respecto a la subprueba de expresión e interacción escritas, se tuvo en cuenta que parece que la redacción de los contenidos del Bloque 2 *Leer y escribir* tiene en cuen-

ta los dos tipos de actividades lingüísticas comunicativas relacionadas con la destreza de la escritura que el M.C.E.R. distingue: (i) la actividad de expresión escrita y (ii) la actividad de interacción escrita (Trim, North and Coste 2001: 61-63 y 82-84, respectivamente). De hecho, esta hipótesis es reforzada por el hecho de que el contenido séptimo del Bloque 2 (“Comunicación personal con hablantes de la lengua extranjera a través de correspondencia postal o utilizando medios informáticos”) incluye (i) uno de los ejemplos de actividades de interacción escrita de la traducción del M.C.E.R. al español (“correspondencia por carta, fax, correo electrónico, etc.”, Trim, North y Coste 2002: 81) y (ii) el título de una de las escalas ilustrativas para valorar las actividades de interacción escrita, *Escribir cartas* (Trim, North y Coste 2002: 82). Si ese contenido séptimo del Bloque 2 se refiere a las actividades lingüísticas comunicativas de interacción escrita, entonces parece lógico concluir que el contenido sexto del Bloque 2 (“Composición de distintos textos con ayuda de modelos, atendiendo a elementos básicos de cohesión y utilizando estrategias elementales en el proceso de composición escrita (planificación, textualización y revisión)”, España 2007: 745) debe estar referido al otro tipo de actividades lingüísticas comunicativas que se distingue en el M.C.E.R., es decir, las actividades de expresión escrita. Además, el verbo principal utilizado en el criterio de evaluación 4 para el segundo curso de E.S.O., *redactar* (España 2007: 746), indica que el legislador optó claramente por obligar a que los alumnos realicen actividades de actuación, es decir, actividades en las que una situación, que funciona como criterio, es simulada en un grado relativamente elevado (Fitzpatrick y Morrison 1971: 268; Hamp-Lyons 1991: 5). Por tanto, se acordó que la subprueba de expresión e interacción escritas evaluara los siguientes resultados específicos del aprendizaje:

1. El estudiante realiza una actividad de expresión escrita.
2. El estudiante realiza una actividad de interacción escrita.

3.3.1. Duración de las subpruebas

En primer lugar, aparecen una serie de tablas que presentan la duración de subpruebas similares a las que estamos diseñando:

Prueba	Duración (minutos)
<i>DELTA A-2 Version Scolaire et Junior – Compréhension des écrits</i> (Centre International d’Études Pédagogiques s.d.: 1)	30
<i>Flyers – Reading & Writing</i> (University of Cambridge ESOL Examinations 2007: 29)	50
<i>E8-Standards Reading Test</i> (Gassner, Mewald and Sigott s.d.: 4)	35

Tabla 1. Duración de diversas subpruebas similares de comprensión de lectura.

Prueba	Duración (minutos)
<i>DELFA-2 Version Scolaire et Junior – Production écrite</i> (Centre International d'Études Pédagogiques s.d.: 1)	30
<i>Flyers – Reading & Writing</i> (University of Cambridge ESOL Examinations 2007: 29)	50
<i>E8-Standards Writing Test</i> (Gassner, Mewald and Sigott 2008: 4)	35

Tabla 2. *Duración de diversas subpruebas similares de expresión e interacción escritas.*

La duración de una prueba está muy determinada, entre otras cosas, por el tiempo disponible para la convocatoria (Schmeiser and Welch 2006: 311) y el grado necesario de fiabilidad (Hambleton, Swaminathan, Algina y Coulson 1978: 23-24), que, a su vez, es influido por el grado de importancia de la decisión que se toma a partir de la puntuación en la prueba (Linn y Gronlund 2000: 130). Si se acepta que el grado de importancia de la decisión es elevado y que, como máximo, se dispondría de dos horas para realizar las subpruebas de comprensión escrita, producción e interacción escritas y comprensión auditiva, esto obliga a repartir dos horas entre estas subpruebas. Las características prácticas de los centros que imparten E.S.O., con clases con una duración aproximada de una hora, hacen aconsejable no dedicar más de 30-40 minutos a cualquier subprueba. A partir de las modas de los valores de la variable *Duración*, se acordó asignar las siguientes duraciones a las subpruebas EALIS-ESO-2:

Subprueba	Duración (minutos)
Comprensión de lectura	30
Expresión e interacción escritas	30

Tabla 3. *Duración de las subpruebas de comprensión de lectura y de expresión e interacción escritas EALIS-ESO-2.*

3.3.2. Estructura de las subpruebas

Para cada subprueba se presenta, en primer lugar, una tabla con la estructura de subpruebas similares y, a continuación, se justifica y se aporta la propuesta de especificación para esa subprueba.

La Tabla 4 muestra la estructura de subpruebas similares de comprensión de lectura:

Subprueba	Número de tareas	Número de ítems	Media de ítems por tarea
<i>DELTA A-2 Version Scolaire et Junior – Compréhension des écrits</i> (Centre International d'Études Pédagogiques s.d.: 1)	3	18	6
<i>Flyers – Reading & Writing</i> (University of Cambridge ESOL Examinations 2007: 29)	7	50	7
<i>E8-Standards Reading Test</i> (Gassner, Mewald and Sigott s.d.: 4)	3	20	6

Tabla 4. Estructura de diversas subpruebas similares de comprensión de lectura.

En la Tabla 4 las modas en las variables *Número de tareas* y *Media de ítems por tarea* son 3 y 6, respectivamente. Esto llevó a plantear una subprueba de comprensión de lectura compuesta de tres tareas, cada una de las cuales incluiría seis ítems. Para poder evaluar los resultados específicos del aprendizaje que deben ser evaluados en la subprueba de comprensión de lectura (el alumno identifica la información general en un texto escrito y el alumno identifica la información específica expresada explícitamente en un texto escrito), se acordó que en cada tarea un ítem evaluaría el objetivo de identificar la información general y que los otros cinco ítems evaluarían el objetivo de identificar la información específica. Por tanto, la estructura de la especificación de la subprueba EALIS-ESO-2-Reading quedó como sigue (Tabla 5):

Tarea	Objetivo específico	Número de ítems	Duración (minutos)
1	El estudiante identifica información general en un texto escrito.	1	10
	El estudiante identifica información específica en un texto escrito.	5	
2	El estudiante identifica información general en un texto escrito.	1	10
	El estudiante identifica información específica en un texto escrito.	5	
3	El estudiante identifica información general en un texto escrito.	1	10
	El estudiante identifica información específica en un texto escrito.	5	

Tabla 5. Estructura de la subprueba de comprensión de lectura EALIS-ESO-2-Reading.

La Tabla 6 muestra la estructura de subpruebas similares a la subprueba de expresión e interacción escritas que se está diseñando:

Prueba	Número de tareas
<i>DELFA-2 Version Scolaire et Junior – Production écrite</i> (Centre International d'Études Pédagogiques s.d.: 1)	2
<i>Flyers – Reading & Writing</i> (University of Cambridge ESOL Examinations 2007: 29)	7
<i>E8-Standards Writing Test</i> (Gassner, Mewald and Sigott 2008: 4)	2

Tabla 6. Estructura de diversas subpruebas similares de expresión e interacción escritas.

En la Tabla 6 la moda en la variable *número de tareas* es 2, lo cual condujo a proponer una subprueba de expresión e interacción escritas con dos tareas. Por tanto, la estructura de la subprueba EALIS-ESO-2-*Writing* quedó en la especificación de la prueba como sigue (Tabla 7):

Tarea	Objetivo	Duración (minutos)
1	El estudiante realiza una actividad de expresión escrita	15
2	El estudiante realiza una actividad de interacción escrita	15

Tabla 7. Estructura de la subprueba de expresión e interacción escritas EALIS-ESO-2-*Writing*.

4. CONCLUSIÓN

Este estudio ha presentado un procedimiento para el desarrollo de una parte de las especificaciones de las subpruebas de comprensión de lectura y de expresión e interacción escritas de una prueba de aprovechamiento de lengua inglesa (EALIS-ESO-2) para el alumnado del segundo curso de E.S.O. de la ciudad de Sevilla. El diseño está basado en (i) un estudio observacional de las actividades de aprendizaje que el alumnado del segundo curso de E.S.O. de la ciudad de Sevilla realizaba durante la clase de lengua inglesa y (ii) un examen del currículo de la materia de lengua extranjera para el segundo curso de E.S.O. (España 2007) y de los presupuestos que subyacen a este currículo, lo cual constituye una alternativa a los métodos tradicionales para obtener evidencias basadas en el contenido de la prueba y una solución al problema de la escasa calidad de este tipo de evidencias.

NOTAS

* Correspondencia a: María Reyes Fernández. *Universidad Pablo de Olavide*. Área de Filología Inglesa. Departamento de Filología y Traducción. Edificio 14, 1ª planta, despacho 50. Autovía A-376, Km. 1. 41013 Sevilla. E-mail: mreyfer@upo.es.

REFERENCIAS

- Alderson, J. C. 1990. "Testing reading comprehension skills (Part two): Getting students to talk about taking a reading test (A pilot study)". *Reading in a Foreign Language* 7: 465-503.
- Alderson, J. C. 2000. *Assessing Reading*. Cambridge: Cambridge University Press.
- Allwright, D. 1988. *Observation in the Language Classroom*. London y New York: Longman.
- Allwright, R. y K. M. Bailey. 1991. *Focus on the Language Classroom: An Introduction to Classroom Research for Language Teachers*. Cambridge: Cambridge University Press.
- American Educational Research Association, American Psychological Association y National Council on Measurement in Education. 1999. *Standards for Educational and Psychological Testing*. Washington, D.C.: American Educational Research Association.
- Andalucía. 2007. Orden de 10 de agosto de 2007, por la que se desarrolla el currículo correspondiente a la Educación Secundaria Obligatoria en Andalucía. *Boletín Oficial de la Junta de Andalucía* 171: 23-65.
- Bachman, L. F. 1990. *Fundamental Considerations in Language Testing*. Oxford: Oxford University Press.
- Bachman, L. F., F. Davidson y M. Milanovic. 1996. "The use of test method characteristics in the content analysis and design of EFL proficiency tests". *Language Testing* 13: 125-150.
- Bachman, L. F., F. Davidson, K. Ryan y I. Choi. 1995. *An Investigation into the Comparability of Two Tests of English as a Foreign Language*. Cambridge: Cambridge University Press.
- Bachman, L. F. y A. S. Palmer. 1982. "The construct validation of some components of communicative proficiency". *TESOL Quarterly* 16: 449-465.
- Bachman, L. F. y A. S. Palmer. 1996. *Language Testing in Practice*. Oxford: Oxford University Press.
- Brown, J. D. 1989. "Cloze item difficulty". *JALT Journal* 11: 46-67.
- Brown, J. D. 2001. *Using Surveys in Language Programs*. Cambridge: Cambridge University Press.
- Carr, N. T. 2006. "The factor structure of test task characteristics and examinee performance". *Language Testing* 23: 269-289.

- Centre International d'Études Pédagogiques. s.d. *Diplôme d'Études en Langue Française DELF A-2: Version Scolaire et Junior*. Paris: Centre International d'Études Pédagogiques.
- Chaudron, C. 1988. *Second Language Classrooms: Research on Teaching and Learning*. Cambridge: Cambridge University Press.
- Clark, J. L. D. 1969. "The Pennsylvania Project and the 'audio-lingual vs. traditional' question". *The Modern Language Journal* 53: 388-396.
- Crocker, L. M., D. Miller y E. A. Franks. 1989. "Quantitative methods for assessing the fit between test and curriculum". *Applied Measurement in Education* 2: 179-194.
- Cuadrado Moreno, J. 2009. "Evidencias basadas en el contenido para las pruebas estandarizadas en primeras y segundas lenguas en entornos escolares formales". *Applied Linguistics Now: Understanding Language and Mind / La Lingüística Aplicada Actual: Comprendiendo el Lenguaje y la Mente*. Eds. C. M. Bretones Callejas et al. Almería: Universidad de Almería: 361-372.
- Cuadrado Moreno, J., M. Reyes Fernández, L. Fernández Amaya, B. Sánchez Fernández, L. V. Fielden, P. Ocaña Escolar y M. O. Hernández López. 2010. "La validación del contenido de la subprueba de producción e interacción orales de la prueba de aprovechamiento EALIS-ESO-2 para la evaluación de la competencia en lengua inglesa del alumnado del segundo curso de la Educación Secundaria Obligatoria". *Interacció comunicativa i ensenyament de lengües*. Valencia: Universitat de Valencia: 495-500.
- España. 2007. Real Decreto 1631/2006 de 29 de diciembre, por el que se establecen las enseñanzas mínimas correspondientes a la educación secundaria obligatoria. *Boletín Oficial del Estado* 5: 677-773.
- Fitzpatrick, R. y E. J. Morrison. 1971. "Performance and product evaluation". *Educational Measurement*. Ed. R. L. Thorndike. Washington, D.C.: American Council on Education: 237-270.
- France. 2007. Arrêté du 15-5-2007: Modalités d'attribution du diplôme national du brevet. *Bulletin Officielle*, 22. [Documento de Internet disponible en <http://www.education.gouv.fr/bo/2007/22/MENE0753209A.htm>].
- Freedle, R. y I. Kostin. 1996. *The Prediction of TOEFL Listening Comprehension Item Difficulty for Mini-talk Passages: Implications for Construct Validity*. Princeton, New Jersey: Educational Testing Service.
- Freedle, R. y I. Kostin. 1999. "Does the text matter in a multiple-choice test of comprehension? The case for the construct validity of TOEFL's minitalks". *Language Testing* 16: 2-32.
- Gassner, O., C. Mewald y G. Sigott. 2008. *Testing Writing: Specifications for the E8-Writing Tests*. Klagenfurt: Language Testing Centre, Alpen-Adria-Universität.
- Gassner, O., C. Mewald y G. Sigott. s.d. *E8 Reading Test Specifications Version 03*. Klagenfurt: Language Testing Centre, Alpen-Adria-Universität.

- Gough, P., C. Juel y P. Griffith. 1992. "Reading, speaking and the orthographic cipher". *Reading Acquisition*. Eds. P. Gough, L. Ehri y R. Treiman. Hillsdale, New Jersey: Lawrence Erlbaum: 35-48.
- Gronlund, N. E. 2004. *Writing Instructional Objectives for Teaching and Assessment*. Upper Saddle River, New Jersey y Columbus, Ohio: Pearson Education.
- Hambleton, R. K. 1980. "Test score validity and standard-setting methods". *Criterion-referenced Measurement: The State of the Art*. Ed. R. A. Berk. Baltimore: John Hopkins University Press: 80-123.
- Hambleton, R. K., H. Swaminathan, J. Algina y D. B. Coulson. 1978. "Criterion-referenced measurement and testing: a review of technical issues and developments". *Review of Educational Research* 48: 1-47.
- Hamp-Lyons, L. 1991. "Basic concepts". *Assessing Second Language Writing in Academic Contexts*. Ed. L. Hamp-Lyons. Norwood, New Jersey: Ablex: 5-15.
- Jarvis, G. A. 1968. "A behavioral observation system for classroom foreign language skill acquisition activities". *The Modern Language Journal* 52: 335-341.
- Johnson, K. y K. Morrow. 1981. *Communication in the Classroom: Applications and Methods for a Communicative Approach*. London: Longman.
- Kish, L. 1965. *Survey Sampling*. New York: John Wiley and Sons.
- Kitsantas, A., H. W. Ware y P. Kitsantas. 2005. "Observational studies". *Encyclopedia of Social Measurement: Vol. 2 F-O*. Ed. K. Kempf-Leonard. Kidlington, Oxford: Elsevier: 913-918.
- Lier, Leo van. 1988. *The Classroom and the Language Learner: Ethnography and Second Language Classroom Research*. London, New York: Longman.
- Linn, R. L. y N. E. Gronlund. 2000. *Measurement and Assessment in Teaching*. Saddle River, New Jersey: Prentice-Hall.
- Malamah-Thomas, A. 1987. *Language Teaching: A Scheme for Teacher Education: Classroom Interaction*. Oxford: Oxford University Press.
- Marshall, H. y M. Gutteridge. 2002. "Candidate performance in the Young Learner English Tests in 2000". *Research Notes* 7: 5-8.
- Munby, J. 1978. *Communicative Syllabus Design: A Sociolinguistic Model for Defining the Content of Purpose-specific Language Programmes*. Cambridge: Cambridge University Press.
- Nikolov, M. 1999. "Classroom observation project". *English Language Education in Hungary: A Baseline Study*. Eds. H. Fekete, E. Major y M. Nikolov. Budapest: The British Council: 221-246.
- Norris, J. M., J. D. Brown, T. D. Hudson y J. Yoshioka. 1998. *Designing second language performance assessments*. Honolulu, Hawai'i: Second Language Teaching and Curriculum Center, University of Hawai'i Press.
- North, B., N. Figueras, S. Takala, P. Van Avermaet y N. Verhelst. 2009. *Relating Language Examinations to the Common European Framework of Reference for Languages: Learning, Teaching, Assessment (CEFR): A Manual*. Strasbourg: Council of Europe, Language Policy Division.

- Nunan, D. 1989. *Understanding Language Classrooms: A Guide for Teacher-Initiated Action*. London: Prentice-Hall.
- Oltman, P. K., L. J. Stricker y T. S. Barrows. 1990. "Analyzing test structure by multi-dimensional scaling". *Journal of Applied Psychology* 75: 21-27.
- Osterlind, S. J. 1989. *Constructing Test Items*. Hingham, Massachusetts: Kluwer.
- Popham, W. J. 2000. *Modern Educational Measurement: Practical Guidelines for Educational Leaders*. Boston: Allyn and Bacon.
- Reyes, M., L. Fernández, M. O. Hernández, L. Fielden, P. Ocaña. y B. Sánchez. 2009. "La validación del contenido de la subprueba de recepción oral de la prueba de aprovechamiento EALIS-ESO-1 para la evaluación de la competencia en lengua inglesa del alumnado del primer curso de la Educación Secundaria Obligatoria". *Applied Linguistics Now: Understanding Language and Mind / La Lingüística Aplicada Actual: Comprendiendo el Lenguaje y la Mente*. Eds. C. M. Bretones Callejas et al. Almería: Universidad de Almería: 543-560.
- Rost, D. 1993. "Assessing the different components of reading comprehension: fact or fantasy?". *Language Testing* 10: 79-92.
- Rupp, A. A., M. Vock, C. Harsch y O. Köller. 2007. *The Development, Calibration, and Validation of Standards-based Tests for English as a First Foreign Language at the IQB: Part I –The Context for and Processes of Task Development*. Berlin: Institut zur Qualitätsentwicklung im Bildungswesen.
- Sánchez B., J. Cuadrado, M. Reyes, L. Fernández, M. O. Hernández, L. V. Fielden y P. Ocaña. 2009. "La validación del contenido de la subprueba de recepción y producción escritas de la prueba de aprovechamiento EALIS-ESO-1 para la evaluación de la competencia en lengua inglesa del alumnado del primer curso de la Educación Secundaria Obligatoria". *Applied Linguistics Now: Understanding Language and Mind / La Lingüística Aplicada Actual: Comprendiendo el Lenguaje y la Mente*. Eds. C. M. Bretones Callejas et al. Almería: Universidad de Almería: 581-598.
- Schmeiser, C. B. y C. J. Welch. 2006. "Test development". *Educational Measurement*. Ed. R.L. Brennan. Westport, Connecticut: Praeger Publishers: 307-353.
- Seedhouse, P. 2004. *The Interactional Architecture of the Language Classroom: A Conversation Analysis Perspective*. Malden, Massachusetts: Blackwell.
- Shavelson, R. J., X. Gao y G. P. Baxter. 1995. "On the content validity of performance assessments: Centrality of domain specification". *Alternatives in Assessment of Achievements, Learning Process and Prior Knowledge*. Eds. M. Birenbaum y F. Douchy. Boston: Kluwer Academic: 131-141.
- Shohamy, E. 1983. "The stability of oral proficiency assessment on the oral interview testing procedure". *Language Learning* 33: 527-540.
- Sireci, S. G. 1998. "The construct of content validity". *Social Indicators Research* 45: 83-117.

- Spada, N. y M. Frohlich. 1995. *Communicative Orientation of Language Teaching Observation Scheme: Coding Conventions and Applications*. Sydney: National Centre for English Language Teaching and Research.
- Trim, J. L. M., B. North y D. Coste. 2001. *Common European Framework of Reference for Languages*. Strasbourg: Council of Europe. [Documento de Internet disponible en http://www.coe.int/T/DG4/Linguistic/Source/Framework_EN.pdf].
- Trim, J. L. M., B. North y D. Coste. 2002. *Marco Común Europeo de Referencia: Aprendizaje, Enseñanza, Evaluación*. Madrid: Secretaría General Técnica del Ministerio de Educación, Cultura y Deporte y Grupo Anaya. [Documento de Internet disponible en http://cvc.cervantes.es/obref/marco/cvc_mer.pdf].
- Tuokko, E. 2007. Mille tasolle perusopetuksen englannin opiskelussa päästään? Perusopetuksen päättövaiheen kansallisen arvioinnin 1999 Eurooppalaisen viitekehyksen taitotasoihin linkitetty tulokset [What Levels do pupils reach in English at the end of the Comprehensive School? National Assessment Results linked to the Common European Framework]. Jyväskylä: Jyväskylä University Publishing House.
- University of Cambridge ESOL Examinations. 2007. *Cambridge Young Learners English Tests: Starters/Movers/Flyers: Handbook for Teachers*. Cambridge: University of Cambridge ESOL Examinations.
- Weigle, S. C. 2002. *Assessing Writing*. Cambridge: Cambridge University Press.
- Wragg, E. C. 1999. *An Introduction to Classroom Observation*. London: Routledge.

