

An Exercise in Course-book Evaluation: Strengths, Weaknesses, and Recommendations regarding *New English File: Elementary*

Un Ejercicio en Evaluación de Libros de Estudio: Fortalezas, Debilidades, y Recomendaciones Acerca de *New English File: Elementary*

Julián Felipe Gutiérrez Bermúdez
Secretaría de Educación Distrital
(Bogotá, Colombia)

Abstract

Often overlooked and reduced to a quick and shallow 'flick test', materials evaluation is a necessary skill for English Language teachers to acquire. It serves the logistical purpose of making an effective selection of materials that cater to their students' needs as well as their own, and puts them in contact with opportunities for further development of their practices based upon the contents and activities found throughout the evaluation process of a given teaching material. The purpose of this article will be then, to conduct an exercise in materials evaluation using Oxford University Press' New English File series. By making a review of the different theoretical stances and experiences on the matter, it will be possible to develop a series of criteria with which to conduct a structured, qualitative evaluation. Based on the results of the application of the aforementioned criteria, it will be possible to conclude on the suitability of this teaching material on a particular context.

Key Words: second language learning; materials evaluation; course-book evaluation; English language teaching; materials authenticity.

Resumen

Constantemente ignorada y reducida a una rápida y superficial 'prueba de hojear', la evaluación de materiales es una habilidad de necesaria adquisición para los profesores de lengua inglesa. Contribuye al propósito logístico de hacer una selección efectiva de materiales que se adapten tanto a las necesidades de sus estudiantes como a las propias, y los pone en contacto con oportunidades para el desarrollo de sus prácticas basados en los contenidos y actividades que aparecen en el proceso evaluativo de un material determinado. El propósito de este artículo será entonces, conducir un ejercicio en evaluación de materiales usando la serie New English File de la editorial Oxford University Press. Al hacer una revisión de las diferentes perspectivas teóricas y experiencias en el tema, se posibilita el desarrollo de una serie de criterios con los cuales realizar una evaluación estructurada y cualitativa. Basándose en los resultados de la aplicación de los ya mencionados criterios, será posible elaborar conclusiones en lo apropiado de este material de enseñanza en un contexto particular.

Palabras Claves: aprendizaje de segunda lengua; evaluación de materiales; evaluación de libros de texto; enseñanza de lengua inglesa; autenticidad de los materiales.



WHYS OF MATERIAL EVALUATION

The requirements of lesson planning, grading and off-classroom attention make teaching a very time-intensive job with a good number of associated tasks. This situation brings the unintended consequence of teachers prioritizing certain tasks before others in order to effectively conduct their job, and leading them, in some cases, to over rely on the teaching materials they have at hand for evaluation and lesson planning purposes, as evidenced by the multiple uses Zohrabi, Sabouri and Behroozian (2012) list for English course-books in the classroom. In their view, course-books can become tools for learners' self-study, showcases for the latest theories, syllabus planners and overall guidelines for teachers. Kayapinar (2009) establishes the usefulness of course-books as their included syllabi provide the framework for courses given their inherent structure, providing teachers with a proved and tested sequence for the contents they aim to develop with their students. However, this situation might result problematic in the long run as prolonged dependence from teachers on course-books might increase their authority within the classroom and the teaching process, giving them an unnecessary and dangerous power over the learning experience, a highly dangerous power since course-books 'Are not perfect and they are only as [sic] simple devices for teachers' (Zohrabi et. al, 2012, p. 90).

In Sheldon's view (1988), course-books are somewhat of a necessary evil since they provide a wealth of evaluation tools and tasks, but the publishers' need for a profit might in some cases overweight the pedagogical aims and needs of the materials. The long writing process might also mean that, in some cases, the books are outdated regarding curricular content just as they are released into the market which, due to its globalized nature, forces course-book authors to make a series of political concessions that have consequences on their topical content, affecting negatively the learners' ability to relate and effectively engage with the materials. McGrath (2013) also recognizes profitability as a decisive factor on course-book writing for publishers, as it often gets in the way of the elaboration of relevant or even effective materials.

Another factor that has affected the course-book writing process throughout the last years has to do with the growing need for immediate results in learners' language acquisition process. Whether for business, immigration purposes or another, obtaining immediate results has become a must for English Language teachers; putting them in the position of having to use course-books



even though they are dissatisfied with the ones they use (Tomlinson and Masuhara, 2013)

In view of this risk of over reliance and the circumstances affecting the materials' production process, the need for a conscious and thorough evaluation process becomes evident, calling for an evaluation beyond the so-called 'flick test' of a course-book, thus leading 'the way to the exploration of course-book evaluation by teachers' (Kayapinar, 2009, p. 70).

It must be also noted that, along with the practical reasons that call for the evaluation of materials, comes a reason that while not as practical, also holds great importance, as a thorough and structured evaluation will turn the empirical knowledge developed by teachers through their practice into a technical or theoretical one, thus bridging the gap that exists between the research in Second Language Acquisition and the second language classrooms (Ellis, 2010).

The question is then, how should teachers evaluate course-books? Littlejohn (2011) sheds an initial light on the matter by establishing two dimensions or levels of the evaluation process: A theoretical level, in which materials are evaluated according to their curricular construction, and a practical level, related to their usability in the classroom environment. When discussing the practical evaluation, Ellis (1997) establishes two further classifications when talking about an empirical evaluation of materials, differentiating whether the evaluation is of a predictive nature (Is the material going to work?) or a prospective nature (Did the material work?), and suggesting a permanent nature of the evaluating process.

Roberts (1996) also sees evaluation as a continuous process with pre and post publication evaluation stages. The former, conducted by the publisher with a balance between content and profitability in mind, and the latter, conducted by teachers concerned with the usability of the book in their particular teaching contexts. Evaluation then, becomes a 'total evaluation process' (Roberts, 1996, p. 377) that happens at various stages, meaning the text on teachers' hands has already been tested and evaluated, so the book is not just the expression of the author's inputs and also that it represents 'a certain way of looking at the world' (Zohrabi *et. al*, 2012, p. 91) that has an effect on the books' degree of usability and effectiveness.

McGrath (2013), on another hand, sees materials evaluation and the need for it as a practical necessity that has massive implications for those involved. In



the case of teachers, it allows them to be agents of change within their academic communities, while at the same time raising the standards of professionalism within faculties.

The method of evaluation also changes according to the type of material. When discussing the evaluation of task-based learning, Ellis (2011) proposes a division of the process at micro and macro levels. Understanding tasks as activities that are focused on learning, have a knowledge goal, rely on learners' resources and have defined outcomes, the macro evaluation of tasks seeks to establish their effectiveness and room for improvement, taking into account administrative and curricular matters; micro-evaluations on the other hand, are in-depth reviews of one aspect covered in the macro-evaluations.

These evaluations require a combination of qualitative and quantitative methods, providing different types of insight on the learning process. While macro-evaluations aim to understand the functioning of the learning environment as a whole, micro-evaluations focus more on specific aspects, such as the materials themselves, providing information on the learners' views on the task and how it is performed.

The discussion on micro- and macro-evaluation can link to questions regarding the scope of the evaluation process. Barnard and Randall's (1995) recall of the evaluation experience of English course-books at Oman is a good example of how, large-scale evaluation processes, that involve nation-wide material distribution and information collection, guarantee a highly representative sample, yet also lead to a series of logistic inconveniences that end hampering the reliability of the evaluation, making a case for small-sized evaluation processes that allow an immediate flow of information, suggesting a higher validity of small-scale evaluations conducted by teams of teachers within English language departments at schools. A small-scale evaluation that nevertheless, must be, preferably, conducted through the combination of a series of subjective viewpoints, a viewpoint that Chambers (1997) argues is more effective in reaching a compromised decision that satisfies as many needs as possible. The Omani experience and Chambers' viewpoint stand out as elements to be taken into consideration by policy makers when piloting English language teaching materials in order to establish what process is the most effective in order to obtain relevant and usable data with which make decisions regarding what materials are to be used in, for example a country or a city's public schools.



Having stated why there is a need for an evaluation process, comes the time to establish both how to conduct it. The complexity of the evaluation subject makes it difficult to apply a quantitative logic to the evaluation process, but at the same time, a heavily impressionistic and unstructured evaluation will be as effective as a quick flick test. This dichotomy calls for the establishment of evaluation frameworks that allow teachers to develop their own criteria for evaluation, taking into account their particular teaching contexts.

In this definition of what an evaluation framework is, Littlejohn (2011) considers its need to show the materials' aspects to examine, how can they be examined and how these findings determine the materials' effectiveness in a given context. These three requirements are helpful in narrowing down the scope of the evaluation, resulting in an analysis of the materials' features *as they relate* to the teaching context's requirements.

However, this narrowing down must not be regarded in a restrictive manner; it cannot be reduced to a checklist to be filled out in a automatic manner for it will also become a non-effective evaluation that says nothing about the material. This structuring stage is one of the reasons that has turned materials evaluation into a bogeyman of sorts for teachers as they dread the idea of mindlessly filling a random checklist, but, as Roberts (1996) states, checklists must be seen in an advisory manner, providing elements that teachers use to devise their own evaluation criteria, and means of evaluation, since not all materials can be evaluated in the same manner.

Devising evaluation criteria for materials evaluation is a decisive step in establishing what makes effective materials. Through the analysis of a long-standing and highly popular English language series in Hungary, Illés (2009) marks a difference between effective materials and popular materials, establishing as a goal of evaluation accounting what makes course-books effective, not popular. This is determined by the teaching context and the goals. For example, a communicative focus in the English language curriculum will favour materials that, regardless of the authenticity of contents, foster the development of communicative skills on learners with a structure that goes beyond the focus on forms approach.

In this view through the theory behind material evaluation, we can see how it is a response to the pre-eminence of course-books in the English language-teaching environment that can swing between two opposite



approaches, being either fully quantitative or qualitative. The role played by the teaching context in determining the needs of both learners and teachers turns the different theoretical guidelines on the matter into non-definitive viewpoints that teachers must take into account and combine with their own particular needs in order to define their evaluation criteria. With this as a starting point; the following section will present the evaluation criteria developed for the evaluation of a course-book in the *New English File Series*.

This evaluation process is made taking into mind the considerations and needs of teachers and school officials. As these two are tasked with the selection of materials for students in schools, they must be provided or develop a set of criteria with which to conduct an evaluation process that is effective and suits their needs. Moreover, this specific material was selected in a scenario In what context is this evaluation process being conducted? The evaluation is conducted taking into consideration the needs and context of a particular population who would see the course-book as a prospective material. In this case, this population is composed of Colombian high school students, aged sixteen to seventeen years old (Tenth grade of secondary instruction), studying English as a foreign language in Bogotá, Colombia, with Spanish as their first language.

In terms of their access to L2 materials, their socioeconomic situation somewhat restricts this access outside the school environment, leaving the school materials (course-books, multimedia and online access) as the main provider of quality L2 materials relevant to their acquisition process. This situation puts the school in the position to be the main provider for the L2 materials that will allow the students' acquisition of the foreign language, giving school officials the responsibility of authorizing acquisition of materials based on the judgments made by teachers.

EVALUATION CRITERIA

Following Sheldon's (1988) and Roberts' (1996) posture on the development by teachers of evaluation criteria based on their own curricular needs instead of reliance on external checklists, the following evaluation criteria have been formulated. These criteria aim to conduct a structured, qualitative evaluation of a given material in order to get a clear picture of its effectiveness in a given teaching context that will be exposed in the criteria application stage.



The advantage of structured qualitative evaluations resides on having a descriptive nature but at the same time, being defined by specific guidelines, providing an evaluation suited to the particular needs of both teachers and learners. In this particular scenario, each criterion is made up by an evaluation question (or two), which is to be answered through a series of evaluators.

Table 1. Book design.

Evaluation Question	Evaluators	Description
<i>Does the book design allow for its effective use?</i>	Design layout. Typographical design. Content/Activity separation. Physical Quality.	Taking into account the ultimate nature of the course-book as a printed, tangible product, the book design evaluators describe how the physical attributes of the course-book affect its effectiveness either in a positive or negative way, through a description of the book's design and its effect on the book's usability.

Table 2. Curricular and material design.

Evaluation Question	Evaluators	Description
<i>Can the book's curricular content be easily distinguished and examined?</i>	Course-book syllabus, correspondence with proficiency standards (Common European Framework)	These evaluators relate to the book's construction from a curricular viewpoint. The establishment of a Common European Framework of Reference for languages, or CEFR (Council of Europe, 2011) has turned into a valuable yardstick for publishers to grade their course-books according to the guidelines established by the CEFR. This criterion compares the course-book's syllabus against its CEFR claims and the CEFR's standards in order to see if there is a true correspondence. This framework, devised among other things to standardize the measurement of learners' proficiency in the acquisition of foreign languages, establishes clear goals and evaluation criteria that define basic, intermediate and proficient users. Since its adoption, publishers have been able to tailor the syllabi and activities of their course-books according to these proficiency markers.



Table 3. Task design.

Evaluation Question	Evaluators	Description
<i>Do the book's activities allow for a development of proficiency within set standards?</i>	Task–Ability correspondence. Task–Syllabus correspondence. Task–Proficiency correspondence.	Very much related to curricular and material design, this evaluator takes a look at a sample of tasks from the course-book and compares them to the syllabus and its CEFR claims in order to establish the degree of correspondence between them.

Table 4. Usability.

Evaluation Question	Evaluators	Description
<i>Are the book's contents appropriate in terms of sourcing?</i>	Content authenticity.	The issue of authenticity is one well explored and discussed in the field of Materials Evaluation (Rinvolucrí, 1999; Thornbury, 1999), and it is an issue that goes beyond the sourcing of it. In this evaluation context, authenticity also refers to the particular vocabulary used in it and its similarity to real-life situations.

Table 5. Versatility.

Evaluation Question	Evaluators	Description
<i>Are there any other resources used by the course-book besides printed text?</i>	Multimedia (Online resources, CDs, DVDs). Book-Multimedia connection.	The development of online and multimedia technologies have broadened course-books' possibilities beyond the printed page both for learners and teachers. However, in some cases these possibilities might go untapped by publishers and authors. These evaluators describe the course-book's online and multimedia resources, should they exist, and their articulation with its curricular design.

EVALUATION

After defining evaluation criteria, it is now possible to apply them to a given piece of English language learning material. For this case, the Elementary level course-book in the *New English File* series has been selected (Oxenden, Latham-Koenig, & Seligson, 2009), conducting the evaluation on a sample composed of



the book's fifth and sixth chapter (Pages fifty-two to seventy-four in the student's book).

Aimed at learners from levels A1 to A2 (Basic) of the CEFR, the book includes a teacher's edition and a workbook, which were also included in the evaluation process. According to the guideline document of the CEFR, (Council of Europe, 2011, p.24), basic users are able to understand simple sentences, communicating in simple and routine tasks that require routine and concrete exchanges of information, and providing basic description regarding their immediate environment.

As these learners started their study of English at the end of their primary schooling (Age 10) they are in the transition from A1 to A2 as defined by the CEFR's standards, which puts them in the proficiency range aimed by the course-book.

Having the aforementioned descriptors in mind, each evaluation question was scored in a scale from one to three, where one corresponded to 'Yes', two corresponded to 'Partly', and three corresponded to 'No'.

Table 6. Book design.

Evaluation Question	Score	Comments
<i>Does the book design allow for its effective use?</i>	1 - Yes	The two-column design of the course-book allows for a maximization of the page space. It also uses different fonts and colors that allow for a differentiation of activity prompts and contents. Its magazine-like nature in the design of the contents is reinforced by the use of glossy paper stock.



Table 7. Curricular design.

Evaluation Question	Score	Comments
<i>Can the book's curricular content be easily distinguished and examined?</i>	2 - Partly	While the student's book syllabus appears basic in its contents, the teacher's book syllabus is far more developed, including pronunciation, speaking, listening and reading goals and contents, and corresponding with the development of basic communicative competences that correspond with the CEFR's goals for A1-A2 learners. The activity prompts are written in a simple, yet non-telegraphic manner, using distinct sentence structures without coming off as drill-like or repetitive.

Table 8. Task design.

Evaluation Question	Score	Comments
<i>Do the book's activities allow for a development of proficiency within set standards?</i>	2 - Partly	There is a sense of continuity and progressive difficulty from one chapter to another. In the reviewed sample (Chapters 5-6), this was shown in the movement from the contents of past simple and were/was forms in chapter 5 to the contents of present continuous and there were/there was forms in chapter 6, thus corresponding to the contents established in the course-book and teacher's book syllabi and seeking the development of a communicative competence at a written and spoken level as shown by the pronunciation exercises in the workbook (Oxenden, Latham-Koenig and Seligson, 2009c, p.81) However, these exercises are more of a reinforcement to the course-book than a complement, intensifying already worked upon areas rather than covering possible holes. The proficiency level of the book itself does not allow for a great depth in certain areas such as reading, aiming to develop basic comprehension and not a critical use of this skill.



Table 9. Usability.

Evaluation Question	Score	Comments
<i>Are the book's contents appropriate in terms of sourcing?</i>	2 - Partly	While the content relates to the CEFR's standards for communication regarding basic, everyday activities and immediate environments, and strides to have an authentic look and feel, it does not feel authentic enough to give learners the impression they're looking into another culture; in fact, in some cases it feels quite generic with vague sourcing such as being 'adapted from a British newspaper' (Oxenden, Latham-Koenig and Seligson, 2009a, p.75). It also requires a level of familiarity with British and American culture that might not be present in all teaching contexts

Table 10. Versatility and multimedia.

Evaluation Question	Score	Comments
<i>Are there any other resources used by the course-book besides printed text?</i>	2 - Partly	<p>The book's online resources provide the complement in terms of activity types that the workbook does not. These are organized clearly according to the course-book's units, and the pronunciation exercises are divided according to segmental and supra-segmental phonological features.</p> <p>The CD-ROM included with the teacher's book includes a massive wealth of teacher resources such as evaluation formats and documentation that establishes a clear link between the book's contents and the CEFR's goals and standards. The format of the included tests allows teachers to create new evaluations from them, giving the course-book a high degree of flexibility in this area.</p> <p>The audio exercises however, present a certain degree of disorganization and it is sometimes difficult to match the activities with the appropriate audio companion.</p>



RECOMMENDATIONS

After conducting the evaluation of *New English File: Elementary* using the aforementioned criteria, it can be stated that the book, has a very solid and verifiable curricular foundation, as shown by the book's syllabi and documentation included in the teacher resources, something that guarantees a correspondence between the book's contents and stated goals with the CEFR's standards for proficiency levels A1-A2 (Basic language user).

However, certain elements of its content in terms of authenticity and sourcing prevent it to be effective in certain learning contexts. As stated during the review, the contents used in the activities require a level of general knowledge that might be available only to older learners and also, a degree of familiarity with British and American culture that might not be available in all contexts such as the one taken in reference for this evaluation, of tenth grade students in a public school in Bogotá Colombia, aged 16-17 years old.

It can be stated that, although highly appropriate for A1/A2 learners given its curricular organisation and articulation, the accessibility and authenticity of its topical content limit its effectiveness, making it most appropriate to young adults who are starting to study English as a second language. It is interesting to note how, in a way, the findings of this analysis mirror those of the one conducted by Tomlinson and Masuhara (2013), where their analysis of multiple materials reveals an underlying assumption made by publishers: namely, that their materials are used by 'aspirational, middle-class, well-educated, westernized computer users' (Tomlinson & Masuhara, 2013, p. 248). These results, also update the discussion on materials' authenticity and the worldview they sponsor started by Thornbury (1999) and Rinvolutri (1999), and discussed earlier in the paper. Furthermore, it could be seen as a consequence of the globalized nature of materials writing, where course-books must be written in order to tailor an audience as wide as possible, posing the question on the balance between profitability and effectiveness, and what is being regarded as more important for publishers.

Taking this into consideration, and given the context of the course-book's potential users, this material might not be wholly suitable for their use, as evidenced partly through the scores obtained in the evaluation.



The results obtained in this evaluation underscore the need for the establishment of solid evaluation criteria to be established in a case to case basis, taking into account the needs of the particular context and parties involved, whether it be students, instructors or policy makers at a school, district or even national level. Materials evaluation must be understood as the most effective way to close the often-mentioned gap between the theory that feeds the materials elaborated by publishers and the reality of teachers, thus creating more effective materials that will ultimately result in advancement for learners.

REFERENCES

- Barnard, R., & Randall, M. (1995). Evaluating course materials: A contrastive study in text book trialling. *System*, 23(3), 337–346.
- Chambers, F. (1997). Seeking consensus in coursebook evaluations. *ELT Journal*, 51(1), 29–35
- Council of Europe (2011). *Common European framework of reference for languages*. Strasbourg, France: Language Policy Unit. Retrieved from http://www.coe.int/t/dg4/linguistic/Source/Framework_en.pdf
- Ellis, R. (1997). The empirical evaluation of language teaching materials. *ELT Journal*, 51(1), 36-42
- Ellis, R. (2011). Macro- and micro- evaluations of task-based teaching. In Tomlinson, B. (ed.) (2011). *Materials development in language teaching* (pp. 212–235), Cambridge, England: Cambridge University Press.
- Ellis, R. (2010). Second language acquisition, teacher education, and language pedagogy. *Language Teaching*, 43(2), 182–201.
- Illés, E. (2009). What makes a coursebook series stand the test of time? *ELT Journal*, 63(2), 145-163.
- Kayapinar, U. (2009). Coursebook evaluation by English teachers. *Inonu University Journal of the Faculty of Education*, 10(1), 69–78.
- Littlejohn, A. (2011) The analysis of language teaching materials: Inside the Trojan horse. In Tomlinson, B. (ed.) (2011). *Materials development in language teaching* (pp. 212–235), Cambridge, England: Cambridge University Press.
- McGrath, I. (2013) *Teaching materials and the role of EFL/ESL teachers: Theory versus practice*. London, England: Bloomsbury.



- Oxenden, C., Latham-Koenig, C., & Seligson, P. (2009a) *New English file: Elementary student's book*. Oxford, England: Oxford University Press.
- Oxenden, C., Latham-Koenig, C., & Seligson, P. (2009b) *New English file: Elementary teacher's book*. Oxford, England: Oxford University Press
- Oxenden, C., Latham-Koenig, C., & Seligson, P. (2009c) *New English file: elementary workbook*. Oxford, England: Oxford University Press.
- Rinvolucrí, M. (1999). The UK, EFLese sub-culture and dialect. *Folio* 5(2), 12–14
- Roberts, J. T. (1996). Demystifying materials evaluation. *System*, 24(3), 378–389
- Sheldon, L. (1988). Evaluating ELT textbooks and materials. *ELT Journal*, 42(4), 237–246.
- Thornbury, S. (1999). Window-dressing vs. cross-dressing in the EFL sub-culture. *Folio* 5(2), 15–17.
- Tomlinson, B., & Masuhara, B. (2013). Adult coursebooks. *ELT Journal*, 6(2), 233–249. doi:10.1093/elt/cct007
- Zohrabi, M., Sabouri, H., & Behroozian, R. (2012). An assessment of strengths and weaknesses of Iranian first year high school coursebook using evaluation checklist. *English Language and Literature Studies*, 2(2), 89–99. doi:10.5539/elt.v4n2p213

BIODATA

Julián Felipe GUTIÉRREZ BERMÚDEZ graduated from the Universidad Nacional de Colombia in 2007 with a degree on psychology. Since 2010, he has worked as an elementary school teacher with Bogotá's Secretaría de Educación Distrital at the Débora Arango Pérez school, one of the public schools participating in the city's bilingual education programme. He has recently completed his studies in the MA in TESOL and Applied Linguistics programme at the University of Leicester (United Kingdom) and expects to receive his degree in summer 2014. His academic interests include the effect of cultural identity on bilingual education and the development of evaluation procedures for English-language teaching materials.

