The analysis and evaluation of language teaching materials

Andrew Littlejohn

Abstract

This chapter reviews some of the main approaches to the analysis and evaluation of language teaching materials. Three approaches to analysis are identified: *illustrative commentaries*, *theory-driven analyses* and *data-driven analyses*, with examples of each type. The chapter raises concerns about the subjective nature of some of these analyses, and recommends analysts explain the basis of their analysis and their motivation. With regard to materials evaluation, the chapter identifies two basic formats: *checklists* and *criteria-generating procedures* and, again, provides examples of each type. Some of the problems associated with checklists and with the counter-proposal of criteria-generating procedures are discussed. The chapter argues that materials evaluation should always include a separate prior stage of materials analysis, and that both activities need to be seen as highly specialised. Designers of analysis and evaluation tools need to specify clearly how they are using key terms and concepts, and the background knowledge and experience that is expected of users. The chapter concludes with an indication of some possible future directions for analysis and evaluation.

The analysis and evaluation of language teaching materials

Andrew Littlejohn

1 Introduction

The term 'language teaching materials' can, of course, refer to an enormous variety of objects including, amongst many other things, teacher-produced exercises, newspaper articles, radio programmes, online practice, games, dictionaries and full-length novels - as well as the traditional coursebook. For the purposes of this chapter, however, and to make it possible to operationalise a definition for research purposes, I will adopt a much more specific view which sees materials as exclusively relating to a pedagogic purpose. To this end, I will be using the term 'language teaching materials' to refer to *a blend of language content with instructions for how to interact with that content, with the aim of bringing about second language development*. This definition, then, will distinguish language teaching materials from not only free-standing samples of language use (e.g. newspapers and radio programmes), but also from reference materials, such as dictionaries and grammars, which do not include any teaching/learning procedure.

Despite a long history for language teaching materials, stretching back more than half a millennium (Howatt 2004), it is perhaps curious that it is only in very recent times, from the early 1990s onwards (Littlejohn 1992:3), that a concerted effort has been put into systematically examining the nature of these materials. Since then, however, a considerable amount of work has been done. Gray (2016:98) notes that there is now a substantial 'vitality of research' into materials, and lists many of the recent publications in this area, many of which are reviewed in this chapter. As this chapter will demonstrate, the net effect of this increased attention to language teaching materials is that we now have a much better idea of what, potentially, their use may imply in the classroom and a much better idea of how we may fine-tune the relationship between aims and implementation in materials design.

Language teaching materials may emerge from many sources. These may include teachers themselves, as they develop materials for their own classes. They may also include learners, as learners, too, may engage in materials production for themselves or their classroom peers (Brown et al, 2013). Most commonly, however, and increasingly so, materials emerge from specialist teams, such as those within ministries of education or language centres, or from author teams commissioned by commercial publishers. In this case, then, the materials which teachers and learners bring into their classrooms will have been conceived of by individuals remote from them, with whom they will probably never have direct contact, and who will thus have no direct knowledge of the particular context or users of the materials. It is this simple fact of remoteness, which Apple (2012:31) describes as the separation of the *conception* of plans for classroom work from the *execution* of those plans, which has probably been the main driving force in the rapid development of research into the nature of language teaching materials. It is useful then, before, I proceed with a discussion of this research, to set out some of the reasons why this has prompted such interest.

First and foremost is the fact that materials often form an *interface* between all participants in the classroom, setting out who is to say what to whom and when. How far materials do this will, of course, depend on the nature of their design and the

willingness of users to follow the instructions as given, but, potentially, materials may determine what it is 'legitimate' to say at any particular point. If, for example, an instruction indicates that learners are to scan a text for general meaning and at that point one individual learner attempts to ask about the meaning of particular words in the text, this may be deemed 'illegitimate' at that moment, and the request postponed for later. Secondly, and more broadly (as discussed below), a number of writers have also pointed to the potential existence of 'hidden outcomes' (whether positive or negative) in teaching materials as they will always encode particular views of what language knowledge *is*, how learning is to happen and the roles teachers and learners are to have (Littlejohn and Windeatt 1989; Littlejohn 1997; Canaragarah 1999; and Wallace 2006). These concerns have gained greater weight as the presence of published materials in the classroom has expanded enormously in recent years with the provision of not only the traditional package of a student's book, workbook and teacher's guide but also items such as online resources, computer assisted learning materials, electronic whiteboard materials, videos, test software, and more. The potential for published materials to effectively structure almost every moment of classroom time has thus increased significantly. Given this, it is perhaps not surprising that Littlejohn (1992:4, 2011:181) has described this situation as potentially a 'Trojan horse' in that the use of published materials may imply more than is immediately apparent.

In the study of published materials, we can distinguish two distinct perspectives (Tomlinson 2012; Gray 2016). One is the familiar activity of *materials evaluation*, an 'ends and means' perspective, which refers to examining and making a judgement on the suitability of a set of materials for a particular context and a particular teaching/learning purpose and which necessarily implies a view of how the materials *should* be. A related but entirely different perspective is *materials analysis*. This refers to examining materials 'as they are', to arrive at a description of those materials, most usually from the vantage point of a framework of analysis. Materials analysis is an activity in its own right, but it can also be seen as a preliminary step before materials evaluation (as further discussed below).

Given this key distinction between analysis and evaluation, this chapter will examine each of these activities in turn, before concluding with a discussion of future directions. It is important at the outset to point out that the focus will be on already published materials, not materials development, although both analysis and evaluation can offer much to writers during materials production. I will be viewing materials as *proposals* for what teachers and learners are to do together in the classroom. Following Breen (1987, 1989) I will be discussing materials as *workplans*, something quite distinct from materials *in action*, the point at which materials are actually used. Precisely what happens when materials are used in the classroom may be very different indeed from what was proposed by their creators, as both teachers and learners bring their own interpretations and purposes to bear on those materials. The latter is indeed a very interesting and fruitful area of research, but one which stands beyond the focus here.

Materials analysis
2 Critical issues and topics: approaches to analysis
As we shall see, many of the analytical approaches reviewed here do actually cross into 'evaluation', as it is clear that a particular view of how materials in general *should be* often motivates their study. The distinction here, however, is that what is prioritised first is an analysis of the materials and it is only after that is done that the outcome of the

analysis is discussed. In addition, the analysis will most usually be general, and not for a specific teaching/learning context.

Materials analysis has evolved considerably in both scale and approach since it first made a significant appearance in the language teaching literature. The earliest contributions to analysis can be termed ***illustrative commentaries***. These generally offered a particular perspective on materials and highlighted aspects for comment, usually supported by examples from published materials. They were, then, perspective-driven, most often reflecting a particular viewpoint from social commentary, rather than any view on language teaching pedagogy or language itself, and most often claiming to reveal a 'hidden curriculum' within the manifest language teaching curriculum. These commentaries had an important role in broadening our understanding of the role of teaching materials, and the kind of learning, in addition to language learning, that they may offer. One of the most well-known of these is Porreca's 1984 paper on pervasive sexism in ESL. Drawing on research on textbooks in other areas of the curriculum, Porreca extended this to ESL texts and showed how, for example, males and females were differently represented and how females were ascribed particular occupational roles. A similarly revealing illustrative commentary was provided by Auerbach and Burgess (1985) who examined textbooks in survival ESL for newly arrived immigrants, and found they reflected a 'hidden curriculum' which prepared students for 'subservient social roles' and which reinforced 'hierarchical relations within the classroom'. Littlejohn and Windeatt (1989) extended the range of commentary to show, with examples, how materials may advance learning in relation to many areas 'beyond language learning', such as the development of cognitive abilities, the learning of particular values and attitudes and the place of learners in classroom decision-making.

While illustrative commentaries have indeed raised our awareness of the ways in which teaching materials may be carrying a 'hidden curriculum', broader ***theory-driven analyses*** have emerged to offer a stronger basis for analysing materials. Similar to illustrative commentaries, theory-driven analyses usually provide a range of examples, rather than a detailed analysis of one particular text, to support the theoretical perspective on offer. Many of these analyses take as their starting point a Marxist or neo-Marxist argument that sees ideology and ideas as being socially and temporally located, that is, reflective of a particular society and a particular point in time. A good early example of this is Dendrinos' (1992) which shows how materials analysis can be situated within a macro-sociological and macro-sociolinguistic perspective, drawing in particular on concepts from critical discourse analysis, to analyse instructional texts and rubrics in teaching materials and reveal their underlying ideology.

Starting from a sociological viewpoint, Littlejohn (2012) argues that materials 'can be seen as potentially resonating in tune with social forces far beyond language teaching itself' (2012:284) and goes on to provide an historical perspective, with many examples from the 1950s to recent years, to show how wider social changes have been intimately reflected in the form and content of classroom materials. In a similar vein, a number of recent writers such as Gray (2012), Gray and Block (2014), Copely (2017), Babaii & Sheikhi (2017) and Bori (2018) have focussed in particular on the relationship between contemporary neoliberalism and language teaching materials. Copely (2017), for example, contrasts an analysis of UK ELT materials produced between 1975-1982 with those produced between 1998-2014 to show how there has been a marked shift away from materials which included references to social issues, such as divorce, homelessness and unemployment, towards neoliberalist concerns such as consumerism

and individual aspirations, with the erosion of any mention of the existence of hardship. Bori (2018), in a substantial book-length work, takes a similar perspective but sets out in more detail an underlying theoretical model derived from Marxist thinking. This allows him to situate language textbooks in an examination of today's capitalism, and to set out a framework for quantitative and qualitative analysis in relation to neoliberal values and practices.

Perhaps rather curiously, the number of theory-driven analyses which focus on issues directly related to language and language teaching seem to be relatively few on the ground. Those that do exist usually highlight aspects of how language use is represented in textbooks, typically from the perspective of pragmatics. Usó-Juan (2007), for example, focuses directly on how requests are presented in five popular textbooks, while Salazar Campillo (2007), in the same volume, examines transcripts from ten textbooks to see how the mitigation of requests are represented. Related analyses are also found in, inter alia, Alcón and Tricker (2000), Gilmore (2004), Boxer and Pickering (1995), and Usó-Juan and Salazar (2002). In the main, these analyses tend to focus on the presentation of specific speech acts or specific discourse markers in ELT textbook dialogues. A broader perspective, however, is offered by Sercu (2000) who sets out the theory and techniques of data collection and analysis for investigating how textbooks can contribute to the acquisition of intercultural communicative competence.

Rather than emphasising a particular theoretical perspective on materials, ***data-driven analyses***, in contrast, aim to reveal the nature of materials 'as they are', by the application of a framework which collects data from the materials in respect of certain categories. To the extent that any framework for analysis depends on a view of what is 'significant' in materials, data-driven analyses are of course also theory-driven. The distinction here, however, is that data-driven analyses aim to provide a minimal framework for data collection, with the bulk of the categories of analysis depending on what is found in the materials, and with the subjectivity of the analyst made explicit. One of the most frequently cited and used data-driven frameworks is that developed by Littlejohn (1998, 2011), sometimes referred to as the 'Three Levels Analysis' because of the way in which the framework moves from objective description, through subjective analysis to subjective inference, as shown in Figure X.1.

Figure X.1 *Levels of analysis of language teaching materials (Littlejohn 2011)*

1 'WHAT IS THERE'	'*objective description*'
• statements of description • physical aspects of the materials • main steps in instructional sections	

2 'WHAT IS REQUIRED OF USERS'	'*subjective analysis*'
• subdivision into constituent tasks • an analysis of tasks: What is the learner expected to do? Who with? With what content?	

3 'WHAT IS IMPLIED'	'*subjective inference*'
• deducing aims, principles of selection and sequence • deducing teacher and learner roles • deducing demands on learner's process competence	

The first level of the framework collects objective data for a description of the materials by setting out 'what is there', such as the way the material is divided up, the existence of different components, number of pages etc. The next level, Level 2, moves to subjective analysis and focuses on identifying 'what is required of users' by conducting

a detailed analysis of the tasks in the materials. The final level, Level 3, relies on subjective inference to determine 'what is implied' by using the materials. At this level, the analyst draws on findings from Levels 1 and 2 to set out how the materials view the roles of teachers and learners, and the underlying aims, principles of selection and sequence of the materials and their role in the classroom.

At the heart of the Three Levels Analysis, lies the task analysis in Level 2, which is the element of the framework most frequently used by researchers, evaluators and postgraduate students (see, inter alia, Nishiyama 2011, Humphries et al 2014, Aliakbari and Tarlani-Aliabadi 2017). To analyse a sample of teaching materials, the analyst first needs to divide it into constituent tasks. For this purpose, a task is identified as 'a task' when it contains three essential elements: 1) a *process* through which the learners are expected to go; 2) a mode of *classroom participation* concerning with whom (if anyone) they are to work; and 3) *content* upon which they are to engage. Focusing on these aspects, the framework then offers a schedule which seeks to identify how the learner is positioned in the learning discourse, the aspect of language they are to focus on (meaning/form), the cognitive process they are to engage, who (if anyone) they are to work with, and the content involved. While the framework offers some categories to select from, the bulk of the categories emerge from what is evident in the materials, as shown in Figure X.2. It is for this reason that the resulting analysis can be termed *data-driven,* allowing the materials to 'speak for themselves'.

Figure X.2 *A task analysis schedule (Littlejohn 2011)*

Task number:								
I What is the learner expected to do?								
A TURN-TAKE								
initiate / open response								
closed / scripted response								
not required								
B FOCUS on								
language system / form								
meaning								
meaning/system relationship								
C MENTAL OPERATION								
[detailed according to what is								
found in the materials]								
II Who with?								
[detailed according to what is								
found in the materials]								
III With what content?								
A INPUT TO LEARNERS								
form [detailed according to what is								
source found in the materials]								
Nature								
B OUTPUT FROM LEARNERS								
form [detailed according to what is								
source found in the materials]								
Nature								

While the Three Levels Analysis is intended to be neutral in terms of a view of what leads to successful language learning, Guilloteaux (2013) has modified the framework

so materials can be analysed and directly related to Second Language Acquisition Theory, from a psycholinguistic viewpoint. Drawing on principles contained in Ellis (2005) and other works, Guilloteaux sets out what she sees as 'universal SLA criteria' which can be related directly to the outcomes of applying Littlejohn's analytical framework. In this way, Guilloteaux has shown how it is possible to move from a data-driven analysis of materials towards evaluation of those materials against a set of desired characteristics that support language acquisition.

3 Implications, challenges and recommendations for materials analysis
A recurring assumption or claim in most approaches to analysis is that the outcome is 'objective', that is, it is not dependent on the subjective, personal judgements or biases of the analyst. This claim is underpinned by clear systematicity in the procedures of analysis, in such a way that similar findings would emerge whoever undertook the analysis. No framework for analysis is neutral, however. It is possible to analyse materials from any number of different perspectives, depending on the specific interests of the analyst. We have seen how gender roles, class representation, market orientation, and pedagogic aspects can be the focus of analysis. One can also imagine a limitless number of other bases for analysis – from content issues (diversity? age-roles? stereotypes?) to language issues (coverage of language forms? explanations of language rules? text density?) to production issues (use of colour? fonts? layout?) and beyond. By selecting a particular area, any framework of analysis is immediately not 'objective' as, by implication, it stresses the significance of that aspect.

A more difficult challenge is this regard lies in the analyst's decision-making during the process of analysis itself. Typically, in social science research, this problem is dealt with by asking two or more analysts to examine the same data, and then to produce a score showing the level of inter-rater reliability. This may certainly guard against a particular bias in analysis, but it does not avoid the basic fact that *all* analysis is subjective. Littlejohn's (2011) Three Levels Analysis tries to account for greater levels of subjectivity by setting out how the resulting description arises from increasing amounts of inference, yet, as noted earlier, even Level 1, the 'objective' level, is the result of a subjective selection of what it is significant to record.

There is probably no way to avoid these difficulties, but it does suggest that it will always be incumbent upon designers of analytical frameworks and all users of such frameworks to explain first and foremost *why* they have chosen a particular area to investigate and *why* it is to be considered significant, thereby revealing the basis for their subjective decisions.

Materials evaluation
4 Critical issues and topics: approaches to evaluation
Materials evaluation, as a defined activity, long predates the development of materials analysis. As noted earlier, the essential quality that distinguishes materials analysis from materials evaluation is that the latter is aimed at determining the 'fit' for a particular teaching/learning context and purpose, whereas materials analysis will not normally focus on the suitability or efficacy of materials. In common with materials analysis, however, the process of doing evaluation can have a consciousness-raising role for all language teaching professionals (including teachers, researchers and writers) by helping to identify key aspects of a teaching-learning relationship. As discussed later, it is this latter aspect which has shaped some of the more recent approaches to evaluation.

Materials evaluation can occur at one of three stages: *pre-use*, *in-use* (sometimes termed *whilst-use)*, and *post-use.* Most materials evaluation tools focus on the pre-use stage, with the development of in-use and post-use tools still relatively underexplored (but see Tomlinson and Masuhara 2018: 73-75 for ideas). In-use and post-use evaluation of materials are much more complicated endeavours than pre-use evaluation, for the simple reason that the number of factors involved are far more numerous (such as the unique aspects of the classroom) and it is impossible to determine which aspect of a particular outcome derives solely from the materials themselves. At the very least, additional methods of collecting and evaluating classroom observation data (in-use) and additional methods of collecting achievement and learner response data (post-use) will be required.

Pre-use evaluation systems appear in a variety of forms. The most common type is the **checklist** in which 'desirable features' of materials are listed, and which the evaluator uses to see if or how far those features are present in the materials under review. This may result in quantitative data, by using a rating system, or in qualitative data, in which evaluators consider their answers to open-ended questions. Two early examples of these different approaches are Williams (1983) and Cunningsworth (1995). Williams (1983) sets out a series of four 'basic assumptions' about the features materials should contain (for example, 'an up to date methodology') and then relates these assumptions to aspects of a language teaching syllabus such as 'speech', 'grammar', 'vocabulary' etc. Each area is then given a weighting by the evaluator, and a score (0-4), such that, according to Williams (1983:25) the numerical ratings can then 'be used for absolute or comparative evaluations of textbooks'. In contrast, a more qualitative approach is taken in Cunningsworth (1995), with a checklist of 45 questions which the evaluator needs to consider, covering areas such as aims, design, language content, skills coverage, and methodology. Rather than providing a rating scale, Cunningsworth provides a series of reflective questions such as 'Is the course book suited to the teaching/learning situation?' and 'Will the topics help expand the students' awareness and enrich their experience?'

Since the early days of materials evaluation tools, many more checklists have appeared (see Mukundan and Ahour 2010 for a good overview of four decades of such checklists). These have continued to include both quantitative checklists of features (such as Skierso 1991, Gearing 1999, Miekley 2005 and Nimehchisalem and Mukundan 2015) and qualitative checklists of questions (such as Richards 2001, McGrath 2002, and Rubdy 2003). While each of these checklists offers distinct points of emphasis (for example, authenticity of texts may be highly significant in one checklist but absent in another), Mukundan et al 2011 found, in their survey of evaluation tools, general agreement on what needs to be evaluated, which they used to develop their own checklist (2011, 2015):

> *'General' attributes*: relation to syllabus and curriculum, methodology, suitability to learners, physical and utilitarian attributes, and supplementary materials
> *'Learning-teaching content'*: general (i.e., task quality, cultural sensitivity, linguistic and situational realism), listening, speaking, reading, writing, vocabulary, grammar, pronunciation, and exercises.

While one may question some of the particular features that these instruments prioritise, there is no doubt that such checklists can aid in guiding the examination of teaching materials. That said, the checklist approach has been subject to a number of criticisms. Tomlinson (2018:55), for example, shows how many checklists slide uneasily back and

forth between providing analysis questions (with supposedly objective answers) and evaluation questions (with highly subjective answers) making it difficult for the evaluator to determine if judgements are based on fact or opinion. For this reason, and following Littlejohn 2011, Tomlinson argues that analysis and evaluation need to be kept as separate activities. Littlejohn (1992, 1998, 2011) further warns that checklist items can often involve 'general, impressionistic judgements … rather than examining in depth what the materials contain' (2011: 181). There is also the danger that checklists which utilise a numeric scoring system can give the false impression of a factual, objective outcome from what is essentially a highly personal, subjective process. Evaluation checklists will tend to reflect the personal experiences, reading, understandings and priorities of their designers. This means that they will often be reflective of issues and priorities in undisclosed contexts, and then offered as a global tool, when they are actually local in origin and relevance.

These problems have motivated a number of developments in materials evaluation. Firstly, many writers now insist that an analysis stage (as detailed in section 3) should precede an evaluation stage, so that the two activities can be clearly separated with the increasing levels of subjectivity made clear. A second major development is a movement away from providing any kind of checklists at all, by setting out a ***criteria-generating process*** which guides evaluators in identifying their own criteria, relevant to the local context and their own beliefs about successful language learning. To do this, Tomlinson (2003) distinguishes between what he terms 'universal criteria' (broadly referring to the conditions in which people most effectively acquire a second language) and 'local criteria' (such as factors relevant to a specific group, individual or culture). The procedure has strong echoes of the step-by-step questioning process outlined in Breen and Candlin (1987), in which evaluators are first asked to identify what *they* think are requirements for successful language learning and the particular requirements of their learners, before relating these to materials. Tomlinson and Masuhara (2018: 68-71) summarise a seven-stage procedure to do this, which takes evaluators from debating their own beliefs, creating a profile of the learning context to finally developing evaluation criteria. Initially developed as a procedure to adopt before beginning the development of any materials, Tomlinson and Masuhara (2018: 72) say that they 'strongly recommend, in any selection procedure for an institution or especially for a nation, that a formal criterion-referenced procedure is followed. In our view this means going through the seven-stage procedure … and comparing the grades and comments for each book evaluated before making a selection.'

5 Implications, challenges and recommendations for materials evaluation
In respect of the checklist approach to evaluation, and in addition to problems with the subjective nature of the categories used and their often loose, impressionistic terminology, commentators have also pointed to repeated problems in their design. Tomlinson and Masuhara (2004:7), for example, summarise a list of questions for evaluating evaluation checklists, highlighting five common errors (such as the lack of separation between analysis and evaluation questions, unanswerable questions, and an underlying dogma). Nimehchisalem and Mukundan (2015) further report that 'in spite of their importance, these instruments are often not tested for their validity or reliability…and are rarely tested for their practicality.' For this reason, Nimehchisalem and Mukundan (2015) developed their own checklist through multiple draft stages and with validation and reliability checks that included reviews by several ELT experts, checks for inter-rater reliability, and checks for a high degree of correlation with a

checklist of known validity. This, they argue, makes their scheme more time-efficient, and the outcome of the evaluation more objective, than comparable checklists,.

Criticisms of evaluation checklists and ideas for a more reliable design procedure certainly highlight the faulty nature of many evaluation tools, and suggest some pitfalls to avoid in the development or use of checklists. However, if a checklist is to be used by teacher-evaluators, it is perhaps of little relevance to hear that ELT experts agree on the interpretation and application of checklist categories. 'Experts', by definition, are members of a particular professional community, and will be immersed in the vocabulary of that community, be up to date with the literature, and be fully aware of the restricted meanings that terms can have. Teacher-evaluators are likely to be part of a different professional community, with different priorities, and so are likely to have quite different, perhaps more personal, interpretations of the same terms. It is therein that the essential problem with checklists lies: a tool devised by one professional community will always be subject to the reinterpretation of users from a different professional community.

It perhaps for this reason that Tomlinson and Masuhara (2018) appear to favour the abandonment of checklists all together. It is clear from the procedure that they advocate that they envisage a strong teacher development role in the process of materials evaluation. Rather than simply being a consumer who is selecting ready-made materials using a ready-made checklist, and who is dependent on outside 'expert' guidance, the teacher-evaluator is viewed as being their own 'expert' in their context. This places evaluators in a much more proactive, demanding relationship with language teaching materials, and immediately directs them towards adapting, modifying, supplementing, selecting or rejecting materials as they see fit. As a counterbalance to the worries, discussed earlier, about materials potentially being a 'Trojan horse' through their role as an interface between teachers and learners, and the dangers in the separation of the conception of work plans from the execution of those plans (Apple 2012:31), the procedure usefully presents a view of the teacher-evaluator as responsible for making curriculum decisions, not simply applying them.

An immediate issue that arises for designers of criteria-generating procedures, however, is the daunting and time-consuming process implied. For those involved in materials *development,* it is indeed imperative that they make clear their beliefs about how language learning effectively happens, particularly if the materials they will design are to be used by others. In this case, the procedure has much to offer. For those involved in materials *selection* however, the situation is not so clear, particularly for inexperienced teachers or teachers with limited exposure to different ways of working. In these cases, it seems quite likely that developing their own criteria, without guidance, may simply lead to them reproducing how they were taught, keeping the uninformed uninformed, in fact. The principal question in this case, then, is 'What *is* the role of 'expert' knowledge?' Checklists tend to discount local knowledge in favour of the expert. Criteria-generating processes do the opposite, potentially asking the teacher-evaluator to reinvent the wheel.

One of the key problems with checklists is that they suggest that evaluation is something anyone can do. I have yet to see a checklist that actually stipulates requirements lts of the user of the list (for example, familiarity with particular literature, years of experience, knowledge of other materials, and so on) and yet these attributes seem vital for an effective and meaningful application of a checklist. On the other hand,

criteria-generating processes seem to assume that the teacher-evaluator already has the required knowledge, and that it just needs to be made explicit. Neither assumption seems satisfactory to me. I think it would be wise for the designers of evaluation tools, and those who advocate criteria-generating procedures, to stress that evaluation is a highly specialised activity, which requires specialist knowledge, and that there are considerable dangers (in terms of coming to a wrong conclusion) in undertaking materials evaluation without the requisite background knowledge or training. Enlisting the involvement of more experienced teachers may help, as may suggestions for guided reading, but fundamentally we need to recognise that skills in *using* materials may be quite different from skills in *analysing* and *evaluating* those materials.

Whether a checklist or a criteria-generating approach is taken, it is clear that effective and reliable materials evaluation will always require materials *analysis* as a preliminary step, to ensure that the full nature of the materials is revealed and so that personalised judgements are not read into the materials. This suggests a three stage operation:

> 1 Analysis of the context of use and analysis of the materials
>
> 2 Match and evaluation of the materials in the light of the analyses
>
> 3 Decisions to adopt, adapt, supplement, critique or reject the materials.

It further suggests, as noted earlier, that designers need to stress that analysis and evaluation are both specialist activities, that require a certain level of teaching experience and background professional knowledge. Where this is not available, analytical and evaluation tools should, at the very least, be accompanied by clear definitions, with examples, of how particular terms and concepts are being used, readings on the key issues involved, and up to date guidance on where teacher-evaluators may find further professional knowledge.

6 Future directions for materials analysis and evaluation
It is likely that in the coming years, new priorities for materials analysis and evaluation will come to the fore. Two particular strands stand out for me, which point, in fact, in opposite directions.

The first derives from a currently emerging convergence of issues in language teaching with issues in educational research. One of the most obvious examples of this is in the development of sociocultural approaches to language teaching research (see, for example, Lantolf et al 2018), and in methodologies which take a more explicitly constructivist, rather than transmission-based approach (for example, negotiated syllabuses; see Breen and Littlejohn, 2000). It seems likely that, in years to come constructivism, already well-recognised in mainstream educational literature (see, for example, approaches to dialogic teaching in Alexander 2008) will come to take a more prominent role in language teaching methodology. This will present a major challenge to the current orthodoxy of language teaching materials (which tend to emphasise defined classroom scripts for teachers and learners to enact through mainly closed tasks), and we can expect to see materials developers experimenting with new forms of teaching and learning, whether in classrooms or online. For this, then, we will need new means of analysing and evaluating materials to take account of a much more fluid use of materials and the likely nature of classroom discourse, something which our current tools seem ill-suited for.

A second potential strand of development for materials analysis and evaluation derives, however, from a quite different direction for language teaching. If the thrust of the theory-driven analyses set out in section 2 is correct, we can expect to see, perhaps rather worryingly, the pressures of neoliberalist thinking begin to be reflected directly in the purposes of the analysis and evaluation of language teaching materials. Evidence of neoliberalist thinking is already present in many global language teaching materials in the packaged, commodified views of what language teaching and learning should focus on (as in, for example, the insistence on adhering to the prescriptions of the Common European Framework of Reference for Languages). It would seem entirely plausible to me that we will begin to see this directly reflected in systems of analysis and evaluation through, for example, specifications of employment/market-related 'skills' and employment/market-related 'language competencies'. This should cause us to pause, and reflect on what language teaching is *for* and how it relates to a wider picture of human growth.

7 Conclusion

This chapter has provided an overview of the development and current state of the distinct activities of materials analysis and evaluation. We have seen how analysis may offer *illustrative commentaries* (indicating how social themes are reflected in materials), *theory-driven analyses* (in which mainly critical theory is related directly to materials), and *data-driven analyses*, which aim to reveal materials 'as they are'. Evaluation, on the other hand, considers materials in the light of a particular context and purpose. Two main approaches to evaluation exist: *checklists* and *criteria-generating procedures*. We have seen some of the problems surrounding both these approaches, and the need to be absolutely precise and explicit in how evaluation is undertaken. The argument stressed in this chapter is that both analysis and evaluation are highly specialised activities which require particular background knowledge and experience, and which should not therefore be undertaken without detailed guidance.

8 Further reading

1. Tomlinson, B. and Masuhara, H. (2018). *The complete guide to the theory and practice of materials development for language learning*. Oxford: Wiley Blackwell
This book draws together a wide range of themes related to materials development generally, but specifically addresses (in Chapter 3) issues in materials analysis and evaluation. Drawing on many years of experience in, Tomlinson and Masuhara provide a grounded account of how their views on evaluation have changed.

2. Gray, J. (Ed). (2013). *Critical perspectives on language teaching materials*. Houndmills: Palgrave Macmillan.
This is a stimulating collection of papers, providing a wide range of theory-driven analyses and evaluations of materials. Topics covered include LGBT invisibility in ELT materials, neoliberalism in EAP materials, practitioners' perspectives on CLIL materials, and constructions of 'frenchness' in language coursebooks.

3. Nimehchisalem, V. and Mukundan, J. and others at the Universiti Putra Malaysia. Various publications freely available online.
An online search for these authors will generate a list of many useful papers which have come from the Materials Development and Evaluation Unit at the Universiti Putra Malaysia. The team has been particularly active in critiquing approaches to materials evaluation and in trialling and developing a grounded, reliable checklist.

References

Alcón E. and Tricker D. (2000). Teaching discourse competence in the classroom: An example based on the discourse marker well in EFL materials. In Gallardo, P., Llurdà, E. (eds). *Proceedings of the 22nd International Conference of Aedean*, Leida: Universitat de Lleida, 61–66

Alexander, R. 2008. Culture, dialogue and learning: Notes on an emerging pedagogy. In Mercer, N. and Hodgkinson, S. *Exploring talk in school*. London: SAGE.

Aliakbari, M. and Tarlani-Aliabadi, H. (2017). Categorization of writing tasks in Iranian senior high school English textbooks. *International Journal of Applied Linguistics*, 28(1), 119-146.

Apple, M. (2012). *Education and power*. New York: Routledge.

Auerbach, E. and Burgess, D. (1985). The Hidden Curriculum of Survival ESL. *TESOL Quarterly*, 19(3), 475-495

Babaii, E. and Sheikhi, M. (2017). Traces of neoliberalism in English teaching materials: a critical discourse analysis. *Critical Discourse Studies*, 15(3), 247-264.

Bori, P. (2018). *Language textbooks in the era of neoliberalism*. Abingdon: Routledge.

Boxer, D. and Pickering, L. (1995). Problems in the presentation of speech acts in ELT materials: the case of complaints. *ELT Journal*, 49(1), 44-58.

Breen, M.P. (1987). Learner contributions to task design. In Candlin, C. and Murphy, D.. *Language learning tasks*. Englewood Cliffs: Prentice-Hall International.

Breen, M.P. (1989). The evaluation cycle for language learning tasks. In Johnson, R. K. *The Second Language Curriculum*. Cambridge: Cambridge University Press.

Breen, M., and Candlin, C. (1987). Which materials? A consumer's and designer's guide. In L. E. Sheldon (ed.). *ELT textbooks and materials: problems in evaluation and development.* London: Modern English Publications and the British Council.

Breen, M.P. and Littlejohn, A. (Eds.). (2000). *Classroom Decision-Making*. Cambridge: Cambridge University Press.

Brown, H., Iyobe, B., and Riley, P. (2013). An evaluation of the use of student-generated materials. *The Language Teacher online,* 37.3, https://jalt-publications.org/files/pdf-article/37.3tlt_art1.pdf

Canagarajah, A. S. (1999). *Resisting Linguistic Imperialism in English Teaching*. Oxford: Oxford University Press.

Copley, K. (2017). Neoliberalism and ELT Coursebook. *Content, Critical Inquiry in Language Studies, 15:1, 43-62*. DOI: 10.1080/15427587.2017.1318664

Cunningsworth, A. (1995). *Choosing your coursebook*. Oxford: Macmillan Education

Dendrinos, B. (1992). *The EFL Textbook and Ideology*. Athens: Grivas Publications.

Ellis, R. (2005). Principles of instructed language learning. *System* 33(2), 209–224.

Gearing, K. (1999). Helping less-experienced teachers of English to evaluate teachers' guides. *ELT Journal*, 53(2), 122–7.

Gilmore, A. (2004). A comparison of textbooks and authentic interactions. *ELT Journal* 58, 362–374

Gray, J. (2012). Neoliberalism, celebrity and 'aspirational content' in English language teaching textbooks for the global market. In Block, D., Gray J. and Holborow, M. (eds) *Neoliberalism and applied linguistics*. London: Routledge, 86–113

Gray, J. (2016). ELT materials. In Hall, Graham (ed). *The Routledge Handbook of English Language Teaching*. Abingdon: Routledge. DOI: 10.4324/9781315676203.ch7

Gray, J. and Block, D. (2014). All middle class now? Evolving representations of the working class in the neoliberal era: The case of ELT textbooks. In Harwood, N. (ed.) *English language teaching textbooks: Content, consumption, production*. Basingstoke: Palgrave Macmillan.

Guilloteaux, M. J. (2013). Language textbook selection: Using materials analysis from the perspective of SLA principles. *Asia-Pacific Education Researcher,* 22(3), 231–239. DOI: 10.1007/s40299-012-0015-3

Howatt , A. P. R. (2004). *A History of English language teaching.* 2nd edition. Oxford: Oxford University Press.

Humphries, S., Miyakoshi, M., and Miyoshi, K. (2014). Analysing textbooks using the Littlejohn framework: Viewpoints from Japanese pre-service student teachers. *MATSDA Folio 16/1*

Lantolf, J., Poehner, M. P., with Swain, M. (2018). *The Routledge Handbook of Sociocultural Theory and Second Language Development.* New York: Routledge

Littlejohn, A. (1992). *Why are English language teaching materials the way they are?* Unpublished PhD thesis. Lancaster: Lancaster University. (Also available at www.AndrewLittlejohn.net)

Littlejohn, A. (1997). Self-access work and curriculum ideologies. In Benson, P., and Voller. P. (eds.). *Autonomy and independence in language learning*. Harlow: Longman.

Littlejohn, A. (1998). The analysis of language teaching materials: Inside the Trojan horse. In Tomlinson, B. (ed.) *Materials Development in Language Teaching*. Cambridge: Cambridge University Press.

Littlejohn, A. (2011). The analysis of language teaching materials: Inside the Trojan horse. In Tomlinson, B. (ed.) *Materials Development in Language Teaching*. 2nd edition. Cambridge: Cambridge University Press.

Littlejohn, A. 2012. Language teaching materials and the (very) big picture. *Electronic Journal of Foreign Language Teaching*. Vol. 9, Suppl. 1, 283–297.

Littlejohn, A. and S.Windeatt. (1989). Beyond language learning: perspectives on materials design In Johnson, R.K. (ed.) *The Second Language Curriculum*. Cambridge: Cambridge University Press.

McGrath, I. (2002) *Materials evaluation and design for language teaching*. Edinburgh: Edinburgh University Press

Miekley, J. (2005). ESL textbook evaluation checklist. *The Reading Matrix*, 5 (2)

Nimehchisalem, V. and Mukundan, J. (2015) Refinement of the English language teaching textbook evaluation checklist. *Pertanika Journal of Social Sciences & Humanities*, 23 (4). 761-780.

Nishiyama, M. (2011). Task analysis for Japanese EFL textbooks: An application of the Littlejohn model. *Studies in English Language Teaching* 34, 21-30.

Porreca, K. L. (1984). Sexism in Current ESL Textbooks. *TESOL Quarterly,* 18(4), 705-724

Richards, J. C. (2001), Curriculum Development in Language Teaching. Cambridge: Cambridge University Press.

Rubdy, R. (2003). Selection of materials, in Tomlinson, B. (ed.) *Developing Materials for Language Teaching*. London: Continuum, 37–57

Salazar Campillo, P. (2007). Examining Mitigation in Requests: A Focus on Transcripts in ELT Coursebooks. In Alcón Soler, E. and Safont Jorda, M. P. *Intercultural language use and language learning.* Dordrecht: Springer.

Sercu, Lies (2000). *Acquiring intercultural communicative competence from textbooks: The case of Flemish adolescent pupils learning German.* Leuven, Belgium: Leuven University Press

Skierso, A. (1991), 'Textbook selection and evaluation'. In Celce-Murcia, M. (ed.) *Teaching English as aSecond or Foreign Language.* 2nd edition. Boston: Heinle & Heinle Publishers, 432–53

Tomlinson, B. (2003). Materials evaluation. In Tomlinson, B. (ed.). *Developing materials for language teaching,* 15-36. London: Bloomsbury.

Tomlinson, B. (2012). Materials development for language learning and teaching. *Language Teaching* 45, 143-179 doi:10.1017/S0261444811000528

Tomlinson, B., and Masuhara, H. (2004). *Developing language course materials.* Singapore: RELC.

Tomlinson, B. and Masuhara, H. (2018). *The complete guide to the theory and practice of materials development for language learning*. Oxford: Wiley Blackwell

Usó-Juan, E. (2007). The presentation and practice of the communicative act of requesting in textbooks: focusing on modifiers. In Alcón Soler, E. and Safont Jorda, M. P. *Intercultural language use and language learning.* Dordrecht: Springer.

Usó-Juan E. and Salazar P. (2002). Developing pragmatic competente in the EFL setting: The case of requests in tourism texts. *Estudios de Lingüística Inglesa Aplicada* 3, 103–122

Wallace, C. (2006.) The text, dead or alive: expanding textual repertoires in the adult ESOL classroom. *Linguistics and Education*, 17, 74-90.

Williams, D. (1983). Developing criteria for textbook evaluation. *ELT Journal* 37/3, 251-255.