Seeking consensus in coursebook evaluation

Fred Chambers

Selecting a suitable coursebook is not a simple task, since many people may be involved, and resistance can be strong. Thus, it is preferable for this decision to be made jointly by the whole teaching team. This article shows how a technique taken from the world of management can be used to select coursebooks in such a way that everyone involved can contribute. It shows how selection criteria can be established and weighted, and how coursebooks can be selected using these criteria.

Introduction

As McDonough and Shaw (1993: 63) note, though in different circumstances, the options to choose teaching materials may vary from totally free to extremely circumscribed: 'the ability to evaluate [them] effectively is a very important professional activity for all ELT teachers'. When teaching materials are to be used by a group of teachers, it seems sensible for them to be selected by all those who will be involved in their use. Various advantages can accrue from this, including joint ownership of the eventual decision (very important in minimizing resistance to all innovations), and increased experience of materials in use. Such decisions are usually made on the basis of models of what constitute good materials that we all, possibly unconsciously, carry in our heads. These models represent some kind of Platonic ideal, vaguely and fuzzily 'shadowed' on the back of our minds, that 'good' coursebooks should have certain features and avoid others.

In a seminal paper on macro-economic modelling, Forrester (1970) noted that there are essentially two types of models: fuzzy, or 'mind' models (the unstated models we all carry in our heads), and 'explicit' models that are stated and overtly produced in one form or another. Forrester argues that the latter are superior because, unlike mind models, they are open to inspection by others—we only guess at unstated pictures in our colleagues' minds—and because mind models are constantly changing, most of all when we are trying to explicate them during discussion. Thus, we can never be sure what the other person is carrying in their mind—even when we feel confident that we know. Like a sub-atomic experiment, the observation changes the experimental condition!

Criteria for 'good materials'

Evaluating materials is a complex process. First, it demands that we establish their relative merits from among a wide range of features (Cunningsworth 1984, Rea-Dickins and Germaine 1992, McDonough and Shaw 1993). Pedagogical factors to be considered include suitability for the age group, cultural appropriateness, methodology, level quality,

number and type of exercises, skills, teacher's book, variety, pace, personal involvement, and problem solving. Second, we have to bear in mind not only construct validity or 'the extent to which a reviewer *thinks* that a book will or will not be useful to a specified audience' (Rea-Dickins 1994: 82), but also the materials already in use. Are the texts good—not just theoretically sound, but proved in the classroom (McDonough and Shaw 1993: 79)? In addition, whenever possible we will want to consider outcomes: the achievement of students who have used the texts. Third, we need to consider whose views we wish to consider in the exercise. Low (1987) identifies up to ten parties who may have an interest. We may not want to have them all involved in our evaluation. On the other hand, we will want it to be as comprehensive as possible within the given time and resources.

Attempts to establish common preferences through discussion often lead to problems. For instance, we might have the idea that 'good' materials should have a variety of exercise types. But what constitutes 'a variety of exercise types'? Moreover, given that variety is only likely to be gained at the expense of some other feature, will those involved agree about what can be sacrificed to achieve that variety?

Different ways of reaching decisions

Vickers (1961) identifies three types of decision at play in the decision-making process:

reality decisions: when we seek to make decisions concerning how the world really is (e.g. 'Book X has 420 pages');
action decisions: when we seek solutions to questions concerning what we should do (e.g. 'We will buy book X in preference to book Y');
value decisions: when we decide that the decision we have undertaken is the best course of action (e.g. 'The methodology in book X is the best').

In selecting a coursebook we need to concede that there are no set truths, and that we are more concerned with what Vickers would consider to be action decisions ('What should we do?') and value decisions ('Is it a wise choice?'). Simons (1976) makes a related distinction when he considers:

maximizing decisions: attempting to achieve the maximum possible return, even at the risk of endangering long-term returns;
optimizing decisions: attempting to achieve the maximum return over the long term, even at the cost of not maximizing shorter-term results;
'satisficing' decisions: not attempting to maximize returns in either the short, medium, or long term, since it involves too much effort. The return expected is determined by the effort we are able or willing to make.

There are numerous ways of reaching a decision about a new coursebook. Algie (1976) suggests at least six. At one extreme it is possible to rely on instinct, or, as we should probably call it, our 'professional judgement and expertise'. Using professional judgement is important and necessary, and, in the end, all methods, no matter how constructed, are based on judgement. However, to work entirely

intuitively has its drawbacks. Intuition is not explicit. Often it is difficult to explain to others, and therefore difficult to defend. Because of its unstructured nature it can be wrong—it may be hurried, or a major factor may have been omitted from deliberation. More importantly, it tends to be an individualized approach which omits consideration of the points above concerning clarity, explicitness, and joint ownership of the decision regarding coursebook selection.

At the other extreme, there are highly precise, mathematical systems for supporting decision-making, but these are frequently beyond the competence of all but the most highly trained. Nor do they entirely remove the element of personal judgement, though judgements, for instance, on what value needs to be given to a particular variable, tend to be pushed back to lower levels in the calculus.

Making reasoning explicit

Any decision-making process must obviously ensure that a decision is reached, but we have to try to ensure that the decision is as wise as possible. For example, we could reach a decision about which text to use simply by voting and accepting the majority decision. That, however, would not ensure a consensus. Those out-voted could resent the decision. A good decision-making process will help to lead to consensus, rather than just ensuring a decision is made.

How can all these factors be weighed, and a useful conclusion drawn?

Whilst it is true, as Sheldon (1988: 245) notes, that 'materials evaluation is fundamentally a rule-of-thumb activity and that no formula, grid or system will ever provide a definite yardstick', it does seem useful to provide 'some model for hard-pressed teachers/course planners that will be brief, practical to use and yet comprehensive in its coverage of criteria' (McDonough and Shaw 1993: 53).

Complicated matters can be most clearly considered if they are written down. In this way, the 'thinking' becomes open to inspection, and it is possible, literally, to review decisions, either as an individual or as a group, and the basis on which they have been reached.

The process described here helps the joint evaluation of coursebooks by as large a constituency as possible. It is essentially a decision-making technique from the world of business (Walker 1988), adapted to help in the evaluation and selection of ELT materials. The process is simple, transparent, leads to clear decisions, and can be used by individuals or groups. It maintains the explicitness of the mathematical decision-making models, whilst retaining the centrality of professional judgement.

An example

The example presented here demonstrates how a group of teachers chose a coursebook for a new programme by working through a pro forma. The pro forma enables the explicit identification and statement of required and desired criteria, and, recognizing that all criteria are not equally important, permits possible alternatives to be weighted for relative importance (see Figure 1).

Figure 1

Features		Option A		Y/N	Option B		Y/N	Option C		Y/N
Essential		Information			Information			Information		
Intermediate level		Lower and upper intermediate		Y	Includes intermediate level exams		Y	✓		Y
Communicative approach		✓		Y	✓		Y	✓		Y
Less than £8.00		£7.50		Y	£7.99		Y	£8.50		N
Desirable	WT (1–10)	Information	SC (1–10)	WT × SC	Information	SC (1–10)	WT × SC	Information	SC (1–10)	WT × SC
Teacher's book	10	Detailed guide	8	80	Rather 'thin'	6	60			
Cassettes	8	None	0	0	Good coverage; *not very exciting*	6	48			
						3	15			
Wide variety of exercise	5	✓	9	45	Yes, but not extensive					
Additional features					Attractive layout	9	9			
Initial choice				125			132			
Risk										
Probability Seriousness (high/medium/low)		Dated: may go *out of print?*	Low	Medium	Flimsy: likely to *need replacement*	High	High			
Decision		Accept			Reject					

Step 1 *Identify possible alternatives*
The first step simply involves entering the name of the coursebook in the box marked 'Option'. In this exercise there were three options, but the form could be adapted for further options.

Step 2 *Identify essential features*
The second step is to determine the essential features expected of the coursebook, that is, those without which the book would not be considered as acceptable. It may be decided by the group, for instance, that the following factors are essential: *level* (e.g. intermediate level); *type of approach* (e.g. communicative); and *cost* (e.g. less than £8.00). These details are entered in the box marked 'Essential features'. The number of criteria has been kept small simply to demonstrate the principle, but the number of essential criteria could be much greater.

Step 3a *Identify desirable features*
The next step is to identify features that are desirable, but not essential. For instance, the group might conclude that they would prefer the coursebook to have a teacher's book and cassettes, and contain a wide variety of exercises. These features are then entered in the box marked 'Desirable features'.

Step 3b *Weight desirable features*
All the essential features are, by definition, equally important. Desirable features, however, can be absent or present in a coursebook, and if present, may be present to a different degree. The next step, therefore, is to give each desirable feature a weighting value, reflecting its desirability, on a scale of 0 to 10. In this example, the group decided to weight the desirable features they had selected thus: teacher's book (×10); cassettes (×8); variety of exercises (×5), suggesting the group considered a teacher's book extremely desirable (though not essential), cassettes very desirable, and a wide variety of exercises quite desirable.

These values were entered in the weighting (WT) column in the pro forma. (Note that while others may not agree with these values or

weighting, they were felt to be appropriate by the teachers concerned. Values and weighting cannot be right or wrong, they can only be agreed or disagreed by a particular group.)

Step 4 *Establish presence of essential features*
The next step involves ensuring that each coursebook considered has all the essential features. After inspection, in this case, it was established that coursebooks A and B met all the requirements, but that coursebook C was too expensive. Because it lacked an essential feature it could no longer be considered as a possible purchase, and was crossed off the list.

Step 5 *Establish presence of desirable features*
The next step is to determine which of the remaining books best fulfils the list of desirable features. It is likely that each book will have several, or even all, of the desirable features, but they may vary in quality and quantity. Each book is first awarded points on a scale of 0 to 10, according to how well it provides each desirable feature. In this case, after consideration, the score for each desirable feature for coursebook A was teacher's book (8), cassettes (0), and variety of exercises (9), representing a good teacher's book, with no cassettes, and an excellent variety of exercises. For coursebook B the score was teacher's book (6), cassettes (6), and variety of exercises (3), representing a teacher's book of average quality, reasonable cassettes, and limited exercise types. These values are entered in the score (SC) column in the pro forma.

Step 6 *Additional features*
One teacher, who had used coursebook B before, commented that the book had an attractive layout and was very user friendly. While this was not included in either the essential or desirable features, it was the sort of additional feature that, if all other things are equal, may determine selection. This was therefore entered in the additional features column and, after discussion, awarded 9 marks. Additional features have no weighting.

Step 7 *Calculate the initial choice*
The scores for each desirable feature are now multiplied by its particular weighting. For instance, in this example the weighting for the teacher's book was 10. Coursebook A was awarded 8 marks for this, therefore the score was 80 marks for that feature, and this was entered in the WT × SC column. Bonus points for additional features were added, and the score totalled. In this case, coursebook A scored 125 points and coursebook B scored 132 points. Hence, by a narrow margin, coursebook B was the initial choice.

Risk: the 'debit' side

An interim decision has been reached by the group, based on the positive aspects of each of the coursebooks, which have all been stated and considered. However, in all decisions there is some degree of risk, and the effect of any negative aspects on choice is not so easily calculated.

When considering risk in relation to a decision, two factors come into

play: probability and seriousness. Probability concerns the likelihood of an event happening. For instance, all other things being equal, most people would consider that the chances of having an accident in a car driven at 160 km/h are higher than in a car driven at 60 km/h, and that the seriousness of an accident at 160 km/h is likely to be higher than an accident at 60 km/h. Therefore, when driving a car fast, both the probability and seriousness factors related to an accident are higher.

The common pattern in risk-taking is that most individuals are willing to take small, non-serious risks, even when the probability of occurrence is relatively high, but will reject events involving serious consequences, even when the probability of occurrence is relatively low. Bearing this in mind, the risks attached to the purchase of each coursebook, considered in terms of seriousness and probability, have to be weighed and set off against the advantages. For reasons that will become clear, risk factors will not be given a number value, but will be rated as high, medium, or low.

Step 8 *Judge the risk associated with book purchase*
When coursebook A is investigated we may find that there is a particular risk associated with it. It may have been published some time ago, for instance, and may not be available for much longer. It would cause some difficulty if the book were selected and then went out of print. However, the consequences of this are not too serious. If the preferred text is not available we can always fall back on the current second choice—the risk is therefore rated as 'seriousness low'. In any event, the likelihood of the book not being available is not very high as the local bookshop has a good supply—the risk is therefore rated as 'probability low'.

Coursebook B also causes some concern. The book appears to be very flimsy, with poor binding. The particular teenage group for whom the book is intended can be very rough with books, and there is concern that it might become necessary to purchase replacement copies; this might not be possible, because of restricted budgets. This risk, therefore, is rated as probability 'high', seriousness 'high'.

The final decision is that the option selected, despite a slightly lower positive rating (125 against 130) is coursebook A, on the grounds that although coursebook B may be a better book for this purpose, it has too great a risk factor associated with it.

Conclusion

The selection of material to be used jointly in an ELT class should be selected by as wide a range of users as possible. This is likely to increase the wisdom associated with the selection and increase the sense of 'ownership' of the decision. The process described here has several advantages over other less structured ways of selecting the material and over other evaluation sheets. First, it is explicit. It makes the decision-makers reflect upon their criteria and then state (and possibly defend) them. Second, having identified criteria, it enables them to be prioritized. Third, it can help in the decision-making process by reducing the 'fuzziness' of the logic. It may be appealing to think that decisions

can be made 'intuitively', that intuitive decisions could be reached more quickly, and that techniques such as the one described here are too much trouble. Honest practitioners, however, must ask themselves questions concerning the quality of such quick decisions.

A further advantage of the process demonstrated here is that it can be used for other purposes in the language school, such as selection of computer hardware or software, video equipment, or even for staff recruitment. In effect, it can be used whenever a judgement in action is being made between identified alternatives. The disadvantage is that it is slower than the intuitive process, and so should only be used for more important decisions. Determining which those are brings us back to questions of judgement: the fact that a judgement is explicit and algorithm-based does not mean that the judgement aspect is reduced or removed, since it is not an automatic process. Professional judgement remains a central responsibility in deciding which features are essential; which desirable; what weight and score to give; what are the risks; and how we score their seriousness and probability weightings. The final 'judgement' is on whether or not to accept the decision provided by the process; if it feels 'intuitively' wrong, and this feeling is strong enough, the option is to disregard the decision. However, the judgement will have been reached openly and hence more democratically, with all the benefits of accrued wisdom brought into the decision-making process. The fact that so many judgements are still required is evidence that these matters should not be treated lightly and need to be spelled out fully and publicly.

Received August 1995

References

Algie, J. 1976. *Six Ways of Deciding.* Birmingham: British Association of Social Workers.

Cunningsworth, A. 1984. *Evaluating and Selecting EFL Teaching Materials.* London: Heinemann.

Forrester, J.W. 1970. 'Understanding the counter-intuitive behaviour of social systems' in The Open Systems Group. *Systems Behaviour* (3rd edn.). 1981. London: Paul Chapman.

Low, G. 1987. The need for a multi-perspective approach to the evaluation of foreign language teaching materials'. *Evaluation and Research in Education* 1/1: 19–29.

McDonough, J. and **C. Shaw.** 1993. *Materials and Methods in ELT.* Oxford: Blackwell.

Rea-Dickins, P. 1994. 'Evaluation and ELT'. *Language Teaching* 27/2: 7–20.

Rea-Dickins, P. and **K. Germaine.** 1992. *Evaluation.* Oxford: Oxford University Press.

Sheldon, L.E. 1988. 'Evaluating ELT textbooks and materials'. *ELT Journal* 42/4: 237–46.

Simons, H. 1976. (3rd edn.) *Administrative Behaviour.* New York: Free Press.

Vickers, G. 1961. 'The Sixth Elbourne Memorial Lecture: Judgement'. *The Manager* 61/1: 31–9.

Walker, P. 1988. 'Decision making and problem solving' in D. Stewart. 1988. *Handbook of Management Skills.* Aldershot: Gower Press.

The author

Fred Chambers works at the Chichester Institute of Higher Education, where he is the co-ordinator of the MA (ELT Management) programme. He has taught overseas in Zambia, Ascension Island, Saudi Arabia, and Malaysia, and has visited many other countries in a professional capacity. After training as a teacher of Business Studies his experience overseas led him to become interested in ELT, and from there to an interest in management issues in ELT. He has a BEd in Business Studies, an MA in applied linguistics, and an MPhil in soft systems methodology.