

Reviewed work:

Assessing Reading. (2000). J. Charles Alderson. Cambridge: Cambridge University Press. Pp. 398. ISBN: 0773470107. \$30.00

Reviewed by:

Mary Jane Hogan
IELTS Chief Examiner, Australia

The real question we are asking when we look at assessing reading is: What distinguishes a good reader from a poor reader? Implicit in this question is an even more fundamental question: What are we doing when we read? Assessment is an attempt to answer the first question, but if we cannot at least try to answer the second, we do not know what we are assessing, and any measure or description of reading proficiency we suggest is meaningless.

These questions are central to J. Charles Alderson's *Assessing Reading*. This study in the Cambridge University Press Language Assessment series (edited by J. Charles Alderson and Lyle F. Bachman) is a comprehensive, up-to-date survey of the issues and research in reading assessment. It explores the nature of reading and the variables that affect reading, and discusses and exemplifies issues in real-life test design, although without offering a blueprint or "best method", not even, as the author says, a set of practical guidelines for item writing or text selection (p. 357). This is just as well, since if it did, the book might have been dismissed as a *how-to* manual that of necessity would have had to curtail its breadth and to select one among many approaches to reading test development. Instead, it presents a thorough and academic study of matters at the heart of what we are doing when we say we are assessing reading.

Unlike many studies in language testing, Alderson's book has an entry in the index under "literacy". In the practicalities of language and literacy teaching and testing in the classroom, it has often seemed that language testing and literacy assessment inhabit separate universes (perhaps indicated by the cloud over the very term "literacy tests" so that the term collocates much more easily with "assessment"). Not all writers, and few teachers, are convinced that tests (a testing event as opposed to ongoing assessment of particular achievements, or evaluation of goals and programs) deserve a place in literacy programs at all. Alderson records the warning of Barton (1994) that parents and educators should be cautious of standardised testing, especially if it is divorced from a context. He notes that teachers "often view tests with suspicion (not always rationally)" (p. 257). Alderson clearly does not share this view, and believes that the difference between language teaching techniques and testing techniques is overstated (p. 203, and see his discussion of Grellet's influential *Developing Reading Skills* (1981) in Chapter 9). Literacy teachers and developers will find much useful and interesting discussion here. Alderson explores models of literacy and their implications for assessment, as well as the importance of L1 literacy for L2 reading. He explores informal methods of literacy assessment in Chapter 7, "Techniques

for testing reading", and is courageous enough to raise the issues of reliability and validity in literacy assessment. Alderson would, I think, agree that "reading" and "literacy" are not synonymous terms in all circumstances, but that insights into assessment of the one for the assessment of the other are mutually useful (p. 269).

Assessing Reading takes up an argument first presented by Alderson in 1990, that of the difficulty of testing individual skills in isolation. Alderson discusses several of the most widely used taxonomies of supposed reading skills in Chapter 1 (including Munby, 1978, and Grabe, 1991) and notes:

Such lists or taxonomies are seductive because they offer an apparently theoretically justified means of devising test tasks or items, and of isolating reading skills to be tested. They also suggest the possibility of diagnosing a reader's problems, with a view to identifying remediation. They are potentially very powerful frameworks for test construction and will doubtless continue to be so used (p. 11).

He then goes on to cite the many criticisms of such taxonomies: the lack of empirical data to support them, their lack of definition, that they are rarely as discrete as the taxonomies would suggest, that it is almost impossible to isolate what skills are operationalised by what test items, and that analysis of test performance does not support such a separation of skills. Alderson himself leans to the view that reading involves several overlapping "skills", which are used in conjunction with each other as necessary (p. 13). He takes the issue up again in Chapter 2 as part of an examination of the nature of reading, and roundly demolishes any remaining notions that reading skills can be isolated, identified, defined, and tested in specific items:

Moreover ... analyses of test performance do not reveal separability of skills, nor implicational scales nor even a hierarchy of skill difficulty. Thus there are statistical and judgemental reasons for doubting whether skills can be measured separately, or whether sub-skills of reading can be shown to exist and be related to the ability to answer particular sorts of test questions. Indeed whether test questions can unambiguously be said to be testing particular skills is quite unclear. ... This issue is crucial to the assessment of reading: if we are not able to define what we mean by the 'ability to read', it will be difficult to devise means of assessing such abilities (p. 49).

In the face of such a resounding judgment, can reading test designers continue to try to identify which reading skills are being tested by certain item types?

Alderson places this discussion of the separability of skills in the context of asking and answering two questions about reading in Chapters 1 and 2: What is reading? What variables affect the nature of reading? He looks at reading as *product* and as *process*. In looking, for example, at whether reading involves top-down or bottom-up processing, he suggests that reading achievement and proficiency tests, by their very nature, are more focused on product than on process, and therefore much of the argument of *how* we read has contributed little to tests that look at *what* we have read. He takes up the topic again in the final chapter of the book, and examines several recent studies which, in attempting to identify and define reading strategies, have succeeded in confusing the skills and strategies issue further. Part of the problem, of course, is one of terminology, and it certainly cannot be said that applied linguists are averse to

re-inventing the meanings of terms. It appears that no one is clear about the difference between a reading skill and a reading strategy, much less about the reading construct of a test. Alderson is excited by the possibility that "if we could identify strategies we might be able to develop diagnostic tests, as well as conduct interesting research" (p. 307). He points out, however, that there is an inherent problem in testing strategies: in standardised testing, responses are either correct or incorrect, but "it is very far from clear that one can be prescriptive about strategy use" (p. 307). The reader should remember here what was noted in Chapter 1, that text does not have "meaning" of itself, but that this meaning is "created in the interaction between a reader and a text"; presumably, the reader's use of reading strategies is part of this creation of meaning. Alderson succinctly notes: "items that can allow any reasonable response are typically very difficult to mark" (p. 307). In conclusion, he makes the fairly temperate observation under the circumstances that "this underlines the need for greater clarity in deciding what are strategies and what are skills, abilities or other constructs" (p. 311).

It should be noted that the focus of the book is on reading in general, but that Alderson is careful to distinguish as necessary between reading in a first language and reading in a second. Chapter 3 explores research into testing reading, and Chapter 4 moves into defining the construct of reading ability, and in passing gives the welcome acknowledgement that by this time "the reader is likely to feel somewhat overwhelmed" (p. 116). Too true, and this is a clue to the target audience of this book. While it is a thoroughly academic study of the issues of reading assessment, it is in fact extremely useful to the testing non-specialist, opening up the complexity of the issues and always giving the reader hope that assessment is accessible, as well as valid. In Chapter 5 Alderson takes up the approach of what still remains the seminal work on language testing, Bachman and Palmer's *Language Testing in Practice* (1996) and explores issues involved in applying their principles of language test design, the Target Language Use domain approach, to developing tests of reading ability.

Chapter 6 continues to reflect the concepts of the target language use domain approach in considering test development as applied to four hypothetical real-life tests, each with different purposes. In these chapters Alderson considers in detail the relationship of text and task, and returns to this question in Chapter 7, where he suggests that in a "communicative" approach to testing, texts should be selected that the target candidate is likely to read, and tasks should be designed in response to the question: "What would a normal reader of a text like this do with it?" (p. 256). Chapter 7 describes and exemplifies most of the common reading task types, with their strengths and weaknesses. Chapter 8 looks at various ways in which reading test constructs have been operationalised in real tests, and explores the developmental scale implicit in them. Chapter 9 rounds off the book by exploring the skills and strategies confusion cited above but also looks at ways in which the thorny area of *process* in reading might be explored.

On the subject of process, one recent finding might be relevant. Schema theory, dealt with in Chapter 2, presents a new way of looking at one phenomenon in the IELTS reading modules, namely, that candidates dislike items of the True/False/Not Given or Yes/No/Not Given type (Merrylees, 2003). If readers comprehend what they read more effectively when they recognise the schema or create one that meets their existing knowledge, could it be that items of the T/F/NG type in fact cause an overload of content schemata demand? Each item has to be processed independently of each other item, although all relate to the overall schema of the text.

Could it be that it is simply exhausting to do a "search-and-find" for the schema of each item and match it to the already acknowledged or created schema of the text? Anderson and Pearson (1988) point out that in an unintegrated set of propositions, the processing of a new proposition makes this search-and-find harder, while adding a new proposition to an integrated set does not add search time. Thus, reading the text is easier than dealing with the items. This has significance if it indicates a split between interacting with the text as an authentic task compared with interacting separately with the items.

Finally, *Assessing Reading* is a highly readable book, written by one who is not only one of the leading researchers in the field of language testing but also a writer possessed of a sense of humour; the reader begins to suspect from the first chapter that the book will be an enjoyable read when we are assured that understandings of text are open to many interpretations, since "How else can we account for the existence of lawyers as a profession?"

References

- Alderson, J. C. (1990). Testing reading comprehension skills (Part One). *Reading in a Foreign Language*, 6(2), 425–438.
- Anderson, R. C. & Pearson, P. D. (1988). A schema-theoretic view of basic processes in reading comprehension. In P. L. Carrell, J. Devine & D. E. Eskey (Eds.), *Interactive approaches to second language reading* (pp. 37-55). Cambridge: Cambridge University Press.
- Bachman, L. F. & Palmer, A. S. (1996). *Language testing in practice*. Oxford: Oxford University Press.
- Barton, D. (1994). *Literacy: An introduction to the ecology of written language*. Oxford: Basil Blackwell.
- Grabe, W. (1991). Current developments in second-language reading research. *TESOL Quarterly*, 25(3), 375–406.
- Grellet, F. (1981) *Developing reading skills*. Cambridge: Cambridge University Press.
- Merrylees, B. (2003). An impact study of two IELTS user groups: Candidates who sit the test for immigration purposes and candidates who sit the test for secondary education purposes. *IELTS Research Reports*, 4, Canberra: IDP: IELTS Australia.
- Munby, J. (1978). *Communicative syllabus design*. Cambridge: Cambridge University Press.