Reviewed work:

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

Reviewed by:
Brendan Moloney
RMIT University Language Centre

J. Charles Alderson shares the trait of other leaders in educational research in his ability to provide a clear and accessible overview of a wide body of research. In *Assessing Reading*, Alderson presents a general introduction to reading assessment of first and second language learners which will be of interest to teachers, researchers, literacy specialists, curriculum designers and writers, and those involved in assessment. The strength of the work is surely the skill with which Alderson draws together major debates on reading and assessment, the lucidity of the prose, the writer's depth of knowledge about the subject, and the practicality of the ideas concerning reading assessment practices.

In Chapter 1, "The nature of reading", Alderson broaches important theoretical questions about reading assessment: whether reading assessors should focus on process or product; how test designers can best approach the testing of levels of understanding; whether the skills of reading can be tested separately; and, if so, of what these skills of reading consist. Alderson concludes the chapter by suggesting that adherence to one theoretical view of reading is unlikely to be adequate for test designers. Incidentally, the printing of most key concepts in bold in the first chapters is (unfortunately) abandoned in later chapters.

Alderson discusses "variables that affect the nature of reading" (pp. 32-84) in Chapter 2. He identifies the main variables in reading assessment as related to reader, text and task. For instance, he sees background and general knowledge, motivation and language proficiency as variables relating directly to the reader; genre and text-type, lexical density and topic variables as those associated with texts; and the language and types of questions as variables related to the task. The point Alderson makes in discussing reading variables, based on a general view of reading as both a product and process (as is discussed in Carrell, Devine and Eskey's [1988] excellent collection of articles on reading research) is that it is important to be aware of many influences on text, task and reader in order to minimize test bias and maximize reliability. In other words, test designers should seek to sample as many different texts, tasks, topics, test methods and so on as possible, in as varied a set of situations as possible (p.81)

Chapter 3, "Research into the assessment of reading", is where Alderson provides a theoretical framework for reading assessment. One of the biggest questions for researchers of reading assessment is whether the general skill of reading can be broken down into sub-skills. Alderson

outlines different research positions from, for example, Carroll (1993) and Drum, Calfee and Cook (1981), who see reading as consisting of various skills, and in doing so essentially urges assessors to be cautious about assessing various skills like critical reading, reading for gist or specific ideas. He points out that the reliability and validity of assessments aimed at evaluating a "macro-level of understanding" is reasonably well established. However, the testing of different reading skills, for instance, critical reading, is much more complex and it is difficult to demonstrate reliably a connection between the assessor's goal and the skill being tested (p.114). Alderson challenges test designers to display vigilance and caution in making reading tests reliable; he points out that there are countless ways in which researchers have claimed to measure the same skills but that test items are very different.

Chapter 4, "The reader: Defining the construct of reading ability", focuses on defining terms commonly used in reading assessment, for example, construct, constructs of reading, and constructs and test specifications. The general idea in this chapter is that assessors should "make explicit the theoretical framework underlying the test" (p. 124). Alderson critically examines the test specifications of the European Commission's DIALANG diagnostic approach towards reading "UCLES" the first certificate in English (FCE) and the IELTS test. Finally, he outlines descriptors of reading proficiency, where he comments on the tendency for assessors to promote both global and specific skills, and approaches to communicative reading tests. It is interesting that Alderson does not spend time in this chapter on analyzing the language of assessment descriptors and bands and the politics of high stakes reading tests generally. Neither does he address the issue that a student's academic background or learning purpose is often mismatched with the aims of high-scale research (cf. Alderson and Urquhart, 1985).

Chapter 5 steers towards "frameworks for test design". Here, Alderson seems excited about the possibility of combining communicative learning strategies and "real-world" target language (in comparison with the higher stakes assessment methods of the former chapter). While evidently an exponent of target language use in the testing of reading, Alderson writes clearly and even-handedly on the strengths and limitations of this framework and admits that it is difficult to discuss generally the idea of reading test design through one framework, as there are many others capable of being used. He bases much of this chapter on frameworks developed by Bachman and Palmer (1996). In order to be able to relate measurement or assessment to the real world, we need some sort of framework within which we can compare our test or assessment procedure with the real-world activity and outcome. For instance, students might keep a portfolio of reading tasks or perform small reading exercises with real-life text types. This follows a growing body of research (e.g., McNamara and Deane, 1995) on the success of assessment which encourages learner autonomy.

In Chapter 6, "Tests in the real world: test purposes", Alderson presents various stages of test development -- in essence an inventory of steps in the design process. Primarily, he examines steps in the development process by tracing a range of different educational settings including second language post-graduate students aiming to go on to further study, government based reading assessments focused on employment and training, and first and second English language learners in a secondary school setting.
One minor criticism is that Alderson could have spent more time in explaining various pitfalls in the development process as this is the stage at which most assessment tasks fall down. For

instance, re-formulating feedback so that it informs changes in a particular test, discussing sample sizes of informants, and examining how designers of small-scale tests might develop a test blueprint would have all been helpful. If the initial explanations are not as detailed as they might have been, Alderson's illustration of ways in which assessment develops in different situations compensates. He suggests that once specifications and tasks have been made, "they will need to be edited, revised, trialled, analysed, revised, retrialled, and possibly reanalyzed, before they can put into operation" (p 200).

Chapter 7 is the most "hands-on" for teachers, as it discusses "techniques for testing reading". Although a lengthy chapter at 70 pages, it includes many important examples of reading test exercises or tasks and is of value as a model for in-class assessment. Alderson is comprehensive in discussing the most popular methods of assessing reading, for instance, the use of cloze, multiple-choice, short answer, ordering, computer testing, information transfer, and recall. However, rather than just providing a simple "how-to", Alderson discusses the benefits and liabilities of using certain test items - on the premise that a clear framework of assessment has already been established. It is in his analysis of potential mismatches between the intentions of the assessor, the questions and test techniques, and the reader's performance that he highlights the difficulties of assessing reading, and the necessity for explicitness in test construction. While Alderson points out that designing language test items is more difficult than developing teaching materials, in principle he sees all reading exercises as potential test items.

While the previous chapter focuses on small-scale test design, in Chapter 8, "The development of reading ability", Alderson presents a critique of high-stakes summative tests. Alderson surveys national frameworks for reading in the UK, the ACTFL proficiency guidelines, the Framework of ALTE, the IELTS test and, finally, Cambridge approaches to English as a Foreign Language and the Certificate of Communicative Skills in English. He also outlines debates on reading progress, using terminology and labels to define reading, and the importance of provision of sample tests, framework descriptions and test specifications.

The final chapter of the book, Chapter 9, "The way forward: Assessing the interaction between reader and text: processes and strategies", centers on alternative reading assessment methodologies. Alderson treads carefully by outlining recent research on evaluating a learner's process of reading, as opposed to traditional concerns with the product (or reaction) to what is read. New movements in reading testing towards more functional, communicative modes of examination are sensibly reviewed and speculation is made on the role that computer literacy will have on test taking.

To conclude, it is likely that Alderson's *Assessing Reading* will become a standard and essential text for teachers of reading in a foreign language for some time. It is interesting, well-written and informative. The greatest strength of the book is Alderson's extensive knowledge and research into reading and his ability to provide clear and common sense views of assessment. The book is a well-balanced mix of theory and practice and I recommend it highly for those involved in reading assessment; from writing small-scale tests, to undertaking research to developing higher-stakes assessments. The long list of references to other academic works on reading assessment found in the bibliography is an additional benefit.

References

Alderson, J. C. & Urquhart, A. H. (1985). The effect of students' academic discipline on their performance on ESP reading tests. *Language Testing,* 2(2), 192-204.

Bachman, L. F. & Palmer, A. S. (1996). *Language testing in practice*. Oxford: Oxford University Press.

Carrell, P. L., Devine, J., & Eskey, D. E. (Eds.). (1988). *Interactive aproaches to second language reading*. Cambridge: Cambridge University Press.

Carroll, B. (1993). *Human cognitive abilities*. Cambridge: Cambridge University Press.

Drum, P. A., Calfee, R. C., & Cook, L. K. (1981). The effects of surface structure variables on performance in reading comprehension tests. *Reading Research Quarterly,* 16, 486-514.

McNamara, T. & Deane, D. (1995). Self-assessment activities: Toward autonomy in language learning. *TESOL Journal*, 5(1), 17-21.