# Assessing Reading

*J. Charles Alderson*

CAMBRIDGE
UNIVERSITY PRESS

PUBLISHED BY THE PRESS SYNDICATE OF THE UNIVERSITY OF CAMBRIDGE
The Pitt Building, Trumpington Street, Cambridge, United Kingdom

CAMBRIDGE UNIVERSITY PRESS
The Edinburgh Building, Cambridge CB2 2RU, UK http://www.cup.cam.ac.uk
40 West 20th Street, New York, NY 10011–4211, USA http://www.cup.org
10 Stamford Road, Oakleigh, Melbourne 3166, Australia
Ruiz de Alarcón 13, 28014 Madrid, Spain



First published 2000

Printed in the United Kingdom at the University Press, Cambridge

*Typeface* 9.5/13pt Utopia [CE]

*A catalogue record for this book is available from the British Library*

*Library of Congress Cataloguing in Publication data applied for*

ISBN 0 521 59000 0 hardback
ISBN 0 521 59999 7 paperback

# Contents

CHAPTER ONE

# The nature of reading

## Introduction

I am not the first person to say that an overview of the study of the nature of reading is impossible. The sheer volume of research on the topic belies any individual's ability to process, much less to synthesise, everything that is written. Similarly, the number of different theories of reading is simply overwhelming: what it is, how it is acquired and taught, how reading in a second language differs from reading in a first language, how reading relates to other cognitive and perceptual abilities, how it interfaces with memory. All these aspects of reading are important, but will probably never be brought together into a coherent and comprehensive account of what it is we do when we read. Added to this are the inevitable complications when we consider the complexities of analysing texts: since the nature of *what* we read must have some relation to *how* we read, then text analysis must be relevant to theories of reading and to research into reading. Yet the simple phrase 'text analysis' covers an enormous range of study within linguistics, which again no individual can hope to overview.

Any review, therefore, of 'the nature of reading' is bound to be somewhat pretentious, and this introductory chapter will inevitably be selective, rather than exhaustive. Yet consider the dilemma for anybody wishing to assess reading. In order to assess the **construct** – the ability we wish to test – we need to know what the construct is. In order to devise a test or assessment procedure for reading, we must surely appeal, if only intuitively, to some concept of what it means to

read texts and to understand them. How can we possibly test whether somebody has understood a text if we do not know what we mean by 'understand'? How can we possibly diagnose somebody's 'reading problems' if we have no idea what might constitute a problem, and what the possible 'causes' might be? How can we possibly decide on what 'level' a reader is 'at' if we have no idea what 'levels of reading' might exist, and what it means to be 'reading at a particular level'? In short, those who need to test reading clearly need to develop some idea of what reading is, and yet that is an enormous task.

The fact is, however, that if we wait until we have a perfect understanding of our constructs before we begin to devise assessment instruments, then we will never begin test construction. Some might say: 'Good. Better not to start than to design something invalid that may do harm.' And we might have sympathy with such a position, yet the plain fact is that assessment of reading is necessary – we will look at the multitude of real-world needs for this throughout this book. To refuse to get involved in designing instruments would thus be irresponsible, and risk the danger that others, with a lesser understanding of what is involved in reading, might design the instruments instead, with more calamitous results. Thus, testers have to get involved in test construction even though they know in advance that their understanding of the phenomenon – the construct – is faulty, partial and possibly never perfectible.

The consolation, however, is that by designing admittedly imperfect tests, we are then enabled to study the nature of the tests and the abilities that appear to be being measured by those tests. This will in turn hopefully lead to a better understanding of what one has assessed, which should feed back into theory, and further research. Thus by doing testing, provided that we research what we design, we can contribute to a growing understanding of the construct.

This is a fundamental tenet of this volume and other books in the series: it is only by trying to operationalise our theories and our understandings of the constructs through our assessment instruments that we can explore and develop our understanding. The corollary is that we need to look to theory in order to have some idea of what it is we are trying to test. This is what I shall do shortly. Before I begin, however, I should acknowledge that another approach to test design seems possible, and indeed, potentially more practical, and that is, rather than starting with theory, to begin with **target situation language use**. In other words, to begin by determining the situations in

which the persons to be assessed will need to 'read'; to analyse such situations; and then to devise assessment instruments which reflect reading in those target situations; and 'see' 'how well' our assessees can 'read'. Indeed, such approaches will be illustrated later in this book. Note, however, that even such an approach needs some crude notion of what we mean by the words in quotation marks: 'read', 'see' and 'how well'. 'How well' implies some sort of standard, at the very least some notion of comparison with how others read; 'see' implies that there are acceptable ways of externalising either how people are reading, or what they have understood of what they have read; 'read' implies that we know what it means to read, to process text meaning through some process of interaction with print.

Rather than continue in this vein indefinitely, we need to start somewhere, and I shall do so by considering the nature of reading.

## Process and product

It is commonplace to make a distinction between the **process** of reading, and the result of that process, the **product**. The process is what we mean by 'reading' proper: the interaction between a reader and the text. During that process, presumably, many things are happening. Not only is the reader looking at print, deciphering in some sense the marks on the page, 'deciding' what they 'mean' and how they relate to each other. The reader is presumably also 'thinking' about what he is reading: what it means to him, how it relates to other things he has read, to things he knows, to what he expects to come next in texts like this. He is presumably thinking about how useful, entertaining, boring, crazy, the text is. He may be consciously reflecting on the difficulties or ease he is experiencing when reading, and on ways of overcoming the difficulties or of continuing the pleasure. He may be completely unconscious of how he is reading, and of what is happening around him: he may be fully absorbed in 'reading'.

Evidently, many different things can be going on when a reader reads: the process is likely to be dynamic, variable, and different for the same reader on the same text at a different time or with a different purpose in reading. It is even more likely, then, that the process will be different for different readers on different texts at different times and with different purposes. Understanding the process of reading is presumably important to an understanding of the nature of reading,

but at the same time it is evidently a difficult thing to do. The process is normally silent, internal, private.

Research has focused on examining the eye movements of readers, and interesting insights have been gained from eye movement photography. Watching what the eyes are doing, however, may not tell us what the brain is doing if, in Smith's (1971) terms, 'What the Brain Tells the Eye is More Important than What the Eye Tells the Brain.'

Asking the reader to read aloud is an alternative to eye movement photography as a means of externalising the reading process, and **miscue analysis** (which analyses the mistakes readers make when reading aloud – for details see Goodman, 1969) is one method of investigating the reading-aloud process. Yet reading aloud is not the 'normal' way in which people read, and the process of reading aloud may be very different from reading silently. Externalising the private process of reading may be the only way to inspect it, yet such externalising risks distorting and changing the nature of the process.

Introspection, through think-aloud protocols or verbal retrospection in interviews, is an increasingly frequently used method of investigating the reading process, and researchers have identified different strategies that good and poor readers appear to use when reading; they have investigated the parts of text that cause problems when reading; and they have also looked at the affective issues that arise when readers are processing particular texts. Introspective methodologies have their critics and are obviously limited in how much light they can throw on the process, but, equally obviously, such methodologies have their uses.

Other research methodologies are also possible and indeed used; it is not the purpose of this chapter to review research methodologies (see Chapter 9), but simply to indicate both the importance and possibilities of examining the reading process in order to understand it, and to understand the limitations that such research must, perhaps inevitably, have.

An alternative approach to examining the process of reading is to inspect the product of reading and, often, to compare that product with the text originally read. It is sometimes said that, although different readers may engage in very different reading processes, the understandings they end up with will be similar. Thus, although there may be many different ways of reaching a given understanding, what matters is not *how* you reach that understanding, but *the fact that* you reach it, or, to put it another way, what understanding you do reach.

The problem of potentially infinite variation in processes of interpreting text is then supposedly reduced by a focus on what one has understood. Product approaches to reading have been unfashionable in recent years as research efforts have concentrated on understanding the reading process, and as teachers of reading have endeavoured to improve the way in which their students approach text. However, a great deal of research into reading earlier this century used essentially product approaches to reading, and much research into the effect of linguistic variables still concentrates on the product of reading. Both a growing realisation that processes of reading are more complex than originally assumed, and the inevitable pendulum swing in research and teaching fashions, have led to revived interest in the product of reading.

As mentioned above, earlier research into reading used a product approach. This means that researchers would typically design tests of understanding of particular texts, administer the tests to suitable informants, using particular research designs, and then inspect the relationship between the results of the tests and variables of interest.

For example, readability researchers would relate scores on reading tests to measures of the linguistic complexity of particular texts, in order to arrive at estimates of text difficulty. Researchers interested in understanding reading ability would devise text comprehension questions at various 'levels of understanding' (see below) and would then see how readers fared on these different questions. Other researchers, wishing to understand what distinguished one type of reader from another (boys versus girls, first-language readers versus second-language readers, children taught by 'whole-word approaches' versus children taught by 'phonics' methods, and so on), might compare and contrast the summaries made by their subjects after reading particular texts. What these studies have in common is that they take some measure of text understanding – test questions, summaries, even interviews – and relate that measure to other relevant variables.

There are at least two limitations to, or problems with, product approaches to reading: one is the variation in the product, the other is the method used to measure the product.

To take the matter of variation first. As we shall see in more detail in Chapter 2, it is clear that what readers understand from text varies. Obviously what people *remember* of what they have read will be affected by their ability to remember. Leaving aside variations in memory, however, and assuming that our measures of understanding

do not depend upon readers' memories, it is still the case that different readers will develop somewhat different understandings of what a text 'means'. This is at least in part because a text does not 'contain' meaning which is waiting to be discovered by an able reader. Rather, meaning is created in the interaction between a reader and a text: the text has what Halliday (1979) and Widdowson (1979) call **meaning potential**, and the potential is realised – in the product of understanding – only by readers reading. Since, as we shall see in Chapter 2, readers' knowledge and experiences influence the realisation of this meaning potential, and since readers may differ in their knowledge and experiences, then the products of reading will also necessarily differ.

Given such differences in understanding – the products – the issue is: how are we to determine (if at all) which product, which understanding is 'correct', and which is 'incorrect'? One approach popular among post-modernists is to say that all products are possible and equally 'correct', or that none are correct, and that the notion of correctness is inappropriate, or theoretically misguided. Without wishing to take sides in this somewhat philosophical argument, which clearly has some force – how else can we account for the fact that people do have legitimately different interpretations of text? How else can we account for the existence of lawyers as a profession? – there must also be some acceptance at a common-sense level that some interpretations of text are simply 'wrong': they do not represent any plausible interpretation of an author's possible intentions. The problem remains, for researchers, theorists and test constructors alike: how to decide which interpretations are acceptable and which are not? Test constructors in particular will need to be able to answer that question, since it is surely not adequate to say that somebody has only understood a text when he agrees with the test constructor's interpretation. Yet this is all too often what happens.

The second problem alluded to above is the method by which one has assessed the product of understanding. This issue will be addressed in more detail in Chapter 7, since it is central to concerns in the testing of reading. It is mentioned here to show the inevitable limitations in theories as well as tests.

If the method of assessing reading product – comprehension – involves a reader recalling what he has read without further recourse to the text (as happens, for example, in the use of recall protocols and interviews, or in some kinds of summary test), then it will be difficult

to distinguish understanding from remembering. If the method of testing is unfamiliar to readers (as happens in some cultures with multiple-choice tests, for example), then one risks a test-method effect. Similarly, if the method – as seems to happen in the case of cloze techniques and gap-filling – induces some readers to read in a particular way (paying close attention to individual words, for instance, or reading the text preceding the gap, but not the following text), then it will be difficult to generalise from a specific test performance to an ability to read, especially when assessed by other methods. It may be the case that some understandings can be assessed by some methods and not by others: can the cloze procedure, for instance, assess whether the reader has read a text critically, rather than passively? If not, obviously the view of understanding derived from the product assessed by such a method will be limited.

What is not always realised when building theories of reading upon the results of such research is that the theories do depend rather centrally on the validity of the measures of understanding used, and the 'accuracy' of the researcher's definition of 'adequate understanding'. This, incidentally, is a nice illustration both of the centrality of some means of assessing reading to the development of a theory (and the limitations therefore of such theories), and of the near circularity of using test results to build theories on which to base test construction. I shall return to this issue in later chapters.

To summarise thus far: it is possible to see reading as a process, or to examine the product of that process. Any theory of reading is likely to be affected by the emphasis that is placed on process or product. Product is easier to investigate than process, although this is not without its problems.

## Levels of understanding

It is commonplace in theories of reading as well as in everyday talk about reading to distinguish different **levels of understanding** of a text. Thus, some may distinguish between a literal understanding of text, an understanding of meanings that are not directly stated in text, or an understanding of the main implications of text. Similarly the distinction between understanding details and understanding the main idea of a text is familiar enough to teachers of reading, as is Gray's (1960) distinction between reading 'the lines', reading

'between the lines', and reading 'beyond the lines'. The first refers to the literal meaning of text, the second to inferred meanings, and the third to readers' critical evaluations of text.

Such distinctions clearly relate to the product of reading, and enable us to describe some of the observed differences in understanding among readers. They also enable the evaluation of such differences, since it is believed that inferred meanings are somehow 'deeper' than literal meanings, and that a critical understanding of a text is more highly valued by society than a 'mere' literal understanding. Such value judgements lead to an implicit (at times explicit) hierarchy of levels of understanding: the literal level being considered somehow 'lower' than critical understanding. This in turn leads to an assumption that it is more 'difficult' to reach a critical understanding of text than it is to infer meanings, and that both of these are more difficult than 'merely' understanding the literal meaning. Thus the notion of levels of understanding becomes overladen with an ordered hierarchy of increasingly valued and increasingly difficult 'meanings'. The next logical leap is from this ordered hierarchy of difficulty and value to a hierarchy of acquisition: it is very frequently assumed that readers first learn how to understand texts literally, then to infer meanings from text, and only later do they learn how to approach text critically, to evaluate text, and so on. Thus it is often asserted that the levels are ordered: i.e. one must understand the lines in order to read between them, and one had better understand both before adventuring beyond them. In fact, the empirical justification for such assumptions is very slim indeed, as we shall see in Chapter 2, but the theoretical notions are persuasive, especially to teachers of reading, and they are thus pervasive.

However, although intuitively appealing, such distinctions among 'levels of understanding' are not always easy to define. Since language is rarely completely explicit, normal language processing requires the reader to make inferences. As Bransford *et al.* (1984) show, readers of the sentence '*The floor was dirty because Sally used the mop*' will readily – some would say automatically – infer that '*the mop was dirty*', yet this statement was not made 'literally'. Similarly, writers must make assumptions about their readers' knowledge, since total explicitness would lead to enormously unwieldy use of language, and would probably make communication impossible. If readers do not possess the knowledge that writers assume, then difficulties in literal understanding will occur, even if inferences can be made.

In summary, a consideration of the nature of reading must include recognition of frequently made distinctions among levels of meaning and understanding in and from text. Test constructors, thus, must also consider the level of meaning that they believe readers ought to 'get out of' a particular text when assessing 'how well' they have understood the text in question.

## What does it mean to be able to read?

Discussions of 'levels of understanding' frequently merge into a discussion of a reader's ability to understand at certain levels. Kintsch and Yarbrough (1982), for instance, distinguish levels of comprehension: it is possible to comprehend the words but not the meaning of a sentence, and sentences but not the organisation of the text. Kintsch and van Dijk (1978) relate the former to 'microprocesses' and the latter to 'macroprocesses': microprocesses have to do with local, phrase-by-phrase understanding, macroprocesses with global understanding. In fact, as mentioned above, reading researchers have frequently attempted to identify reading skills or abilities by giving subjects a series of passages, and asking them questions intended to test different levels of understanding of the passages. Thus 'the ability to make inferences' becomes defined as 'the ability to answer a question relating to meanings not directly stated in text'. There is, of course, a degree of circularity in such definitions, but that has not stopped researchers and theorists from positing the existence of reading skills and subskills from the answers to such questions. It is common to factor-analyse the results of such answers, and then to state that questions that load on the same factor measure the same skill or subskill. In such a fashion, many different lists, taxonomies and even hierarchies of skills have been developed, as Alderson and Lukmani (1989) point out. The New York City Board of Education is cited by Lunzer and Gardner (1979) as identifying thirty-six different skills. Davis (1968) defines eight skills, as follows:

1 recalling word meanings
2 drawing inferences about the meaning of a word in context
3 finding answers to questions answered explicitly or in paraphrase
4 weaving together ideas in the content

5 drawing inferences from the content
6 recognising a writer's purpose, attitude, tone and mood
7 identifying a writer's technique
8 following the structure of a passage

As we shall see in Chapter 2, however, there is a considerable degree of controversy in the theory of reading over whether it is possible to identify and label separate skills of reading. Thus, it is unclear (a) whether separable skills exist, and (b) what such skills might consist of and how they might be classified (as well as acquired, taught and tested). Nevertheless, the notion of skills and subskills in reading is enormously pervasive and influential, despite the lack of clear empirical justification.

Bloom's 'Taxonomy of Educational Objectives in the Cognitive Domain' (Bloom *et al.* 1956) appeals to similar theorising about the components of educational achievement, and his taxonomy has been enormously influential in the devising of curricula, instructional material and tests. In second-language education, Munby's taxonomy of microskills has been influential in syllabus and materials design as well as the design of language tests. Munby (1978) distinguishes the following reading 'microskills':

- recognising the script of a language
- deducing the meaning and use of unfamiliar lexical items
- understanding explicitly stated information
- understanding information when not explicitly stated
- understanding conceptual meaning
- understanding the communicative value of sentences
- understanding relations within the sentence
- understanding relations between parts of text through lexical cohesion devices
- understanding cohesion between parts of a text through grammatical cohesion devices
- interpreting text by going outside it
- recognising indicators in discourse
- identifying the main point or important information in discourse
- distinguishing the main idea from supporting details

- extracting salient details to summarise (the text, an idea)
- extracting relevant points from a text selectively
- using basic reference skills
- skimming
- scanning to locate specifically required information
- transcoding information to diagrammatic display

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.

However, as has been suggested above, they need to be treated with care. Firstly, their origins are more frequently in the comfort of the theorist's armchair than they are the result of empirical observation. Secondly, they are frequently ill defined (or undefined) and give a misleading impression of being discrete when in fact they overlap enormously (see, for example, criticisms of Munby by Davies, 1981; Mead, 1982; Skehan, 1984; and the discussion in this book in Chapter 5; and of Bloom by Seddon, 1978). Thirdly, it is frequently difficult to get expert judges to agree on what skills are operationalised by which test item (Seddon, 1978; Alderson, 1990b), and finally, analysis of test performance does not reveal separability of skills, nor implicational scales, nor even a hierarchy of difficulty or discrimination (see Alderson and Lukmani, 1989; Alderson, 1990b; Alderson, 1990c). Despite all these problems, a skills approach to defining reading remains popular and influential and cannot be ignored in a treatment of the nature of reading (see Chapters 2, 3, 4 and 9).

Several alternative views are possible on this issue of reading skills. One, expressed by Lunzer *et al.* (1979), is that there is no evidence that distinct separate skills exist, and that, instead, reading consists of one single, global, integrated aptitude. The second view, to which Alderson (1990c) inclines, is that 'at least part of the reading process probably involves the simultaneous and variable use of different, and overlapping, ''skills''. The division of skills into ''higher'' and ''lower'' orders, however tempting, does not seem to be justified in practice' (1990c:478). A third view is represented by Matthews (1990) who states: 'the items in Munby-based taxonomies appear to be a slightly

random and overlapping collection of strategies, skills and (chiefly) knowledge, and represent an impoverished account of the reading process.' She claims that most of what Munby calls 'skills' are, in fact, aspects of knowledge. Thus, 'understanding explicitly stated ideas' is just a more general statement of 'skills' like 'knowing the meaning of the word ''tree'' '. Matthews calls for a better understanding of what skills are required in reading, citing Eskey and Grabe's view (1988) of the importance of speed and automaticity in word recognition. She suggests that, if speed and flexibility are important, then they need to be tapped in tests of reading. It has been suggested that lists like Munby's are not processes, but products: they identify what is done, not how it is done. If this is so, it might explain why it is difficult to isolate skills of the Munby type. What needs to be isolated are the processes which lead to these outcomes.

An increasingly common view in the research literature is that reading is essentially divided into two components: decoding (word recognition) and comprehension. The latter is often described (e.g. in Gough *et al.*, 1992b) as consisting of parsing sentences, understanding sentences in discourse, building a discourse structure, and then integrating this understanding with what one already knows. This comprehension process, however, is not seen as unique to reading, but also describes the process of listening. In other words these are linguistic skills, not reading skills. The difference between listening and reading is suggested to be minimal: 'comprehension is largely a centrally-determined function operating independently of the mode of presentation of the material' (Larsen and Feder, 1940:251, cited in Gough *et al.*, *op. cit.*).

A further alternative is Carver's view that a 'simple view of reading' should be reanalysed into a three-part separability of word recognition skills, reading rate or reading fluency, and problem-solving comprehension abilities. In a number of publications (Carver, 1982, 1983, 1984, 1990, 1992a, 1992b), Carver distinguishes what he calls **rauding** ('typical' reading done under conditions wherein the individual has no difficulty comprehending each sentence) from memorising, studying, skimming and scanning. He claims that these are five different processes, and only one of these – rauding – is normal reading, where the reader is comprehending all or most of the thoughts the author intended to communicate. Carver has amassed a considerable amount of evidence to show that rate fluency abilities change as

readers develop – in other words, that reading speeds increase with reading development.

Finally, in this illustrative set of alternatives, Grabe (1991) proposes the following six component elements in the fluent reading process:

- automatic recognition skills
- vocabulary and structural knowledge
- formal discourse structure knowledge
- content/world background knowledge
- synthesis and evaluation skills/strategies
- metacognitive knowledge and skills monitoring

Among the **metacognitive skills** he includes: recognising the more important information in text; adjusting reading rate; skimming; previewing; using context to resolve a misunderstanding; formulating questions about information; monitoring cognition, including recognising problems with information presented in text or an inability to understand text. Self-regulation strategies like planning ahead, testing one's own comprehension, and being aware of and revising the strategies being used are also said to be typical reading strategies of fluent readers. We will discuss the evidence for these views in more detail in Chapter 2 of this book.

## What do we do when we read?

If theorists are not (yet) agreed on what skills are involved in the reading process, is it at least possible to find some consensus on what happens when we read? What kinds of tasks characterise the activity involved in reading?

Clearly, reading involves perceiving the written form of language, either visually or kinaesthetically (using Braille). Here we already encounter the first problem: do readers then relate the printed form of language to the spoken form? If so, then once that translation has taken place, reading is the same sort of activity as listening, and the only specific aspect of reading that we need to concern ourselves with as testers is the process of transformation from print to speech. One argument, put forward by theorists like Smith (1971), is that readers proceed directly to meaning, and do not go via sound. They

claim that readers can process print much faster than sounds, and so there would be an upper limit on the speed with which we read if we had to go from print to sound. Fluent reading is frequently done at speeds up to three times as fast as many people speak in everyday conversation.

However, research has consistently shown that listening comprehension does not break down with accelerated speech. Carver (1982), for example, shows that there are optimal rates of processing prose, and they are roughly equivalent for reading and listening, at 250 to 300 words per minute. The results of such studies challenge the views of Goodman (1969, 1982) and Smith (1971) and allow the question of whether we access meaning directly or via sound to be revisited.

We have all experienced the sensation of sounding out, possibly subvocally, difficult words, or parts of text where we have to concentrate. Does such subvocalisation constitute normal activity (which we are usually unaware of), or does it only occur when we encounter difficulties, when we need the extra support of the subvocally heard sounds? There is a growing consensus in the recent cognitive psychology research literature that all reading requires what is called 'early phonological activation': in other words, that readers typically identify the sound of words as part of the process of identifying their meaning. An issue frequently discussed is whether this phonological identification proceeds independently of and in parallel to the use of semantic and other cues (the 'modular' approach), or whether it is sequential, proceeding in stages – i.e. sound is recognised first, then meaning. Research is unclear on the matter, but the view one takes on this presumably affects what one considers essential to assess when looking at reading success or abilities.

Recent accounts of the fluent reading process tend to emphasise that it is rapid, purposeful, motivated, interactive (in terms of component skills as well as in the relation between knowledge and the printed word), it is comprehending (readers expect to understand), it is flexible, and it develops gradually (it is the product of long-term effort and gradual improvement).

When we are reading, we are clearly engaged in a great deal of mental activity, some of it automatic, some of it conscious. For example, we may consciously decide to skip a page or two in a rather boring text, we may decide just to focus on the headlines in a newspaper, or to read the end of the detective story first before reading the introduction. We may scan through a telephone directory ignoring all

names except the one we are looking for; or we may read every letter and word of a memorandum we are writing to our boss, in which we want to be sure we have made no spelling mistakes, and have expressed ourselves diplomatically but clearly.

These conscious strategies involve a deliberate choice of process or task, each of which may involve different constellations of skill and knowledge (being able to spell words in English, for example, or knowing the order of the alphabet). Such strategies may be semi-conscious, or at least recoverable to consciousness, as when we try to figure out the meaning of a word we have never met before by thinking about the context in which it comes, its form, the sort of word it is (noun, verb and so on) and the sort of meaning it is likely to have. We may consciously decide to look the word up in a dictionary, or not to worry about its 'exact meaning', since we have sufficient idea of what it must mean to be able to continue reading without disruption.

Other activities are not amenable to consciousness – hence the use of the term **automaticity**. We are not normally conscious of processing the distinctive features in each letter in English text, for example, yet word recognition for the normal reader must involve some process of discriminating visual shapes. When we are absorbed in a novel we are not normally conscious that we are visualising the setting – the faces, dress, voices of the characters, the location of the action, the surrounding scenery – yet evidence suggests that we do precisely this, and that what we visualise becomes part of our meaning for what we are reading. Researchers seek to identify and characterise these processes and strategies, and useful lists have been developed in recent years (see for example Harri-Augstein and Thomas, 1984; Nevo, 1989; Storey, 1994).

There are two broad approaches available for assessment for those who feel that the view of reading as a series of strategies and activities is correct, or at least relevant to their purposes. One is the analytic approach: to seek to test whether readers successfully engage in, or master, those aspects of the process which testers consider to be important. Thus one might seek to devise test items which explore whether a reader can successfully deduce the meaning of unknown words from context. One might devise tasks that require readers to scan rapidly through a number of headlines in order to identify the one(s) that are relevant to a particular need or topic. In other words, one seeks to isolate and identify components of the reading process

relevant to the purpose for which one is testing (see Chapter 6 for more on testing purposes). Some aspects will, however, be easier to test than others: can one, for example, successfully test whether readers are visualising settings 'appropriately' when reading a short story? Can one assess whether readers are fully absorbed in a novel, with no sense of their surroundings, or are just pretending for the sake of the assessor?

The other broad approach is to recognise that the act of assessing itself risks disturbing parts of the process one is wishing to assess, and to acknowledge that individual readers may well not need to engage in a particular activity in order to read 'successfully' (they may already know the meaning of the word, they may find an irrelevant news story interesting). Such an approach would entail seeking to simulate as far as possible the conditions in which one is interested – reading newspapers in order to get an overview of the day's events, scanning TV guides in order to plan the evening's viewing – and then assess whether the reader had successfully completed the task. The assumption would be made that if the task was successfully completed, then either the reader would of necessity have engaged in the sorts of processes of interest or had not, and such processes were not necessary. We return to this difference of approach later in this volume, in Chapters 4, 5 and 6.

## Top-down and bottom-up processing

Much has been made in reading research over the last twenty years or so of an apparent dichotomy between two different approaches that may be taken by readers. One is the bottom-up approach, and the other is the top-down approach. The latter owes much to the work of Smith (1971) and Goodman (1969, 1982), who emphasise in their writings the importance of the contribution made by the reader to the reading process, and who downplay the importance traditionally ascribed to the printed word.

**Bottom-up** approaches are serial models, where the reader begins with the printed word, recognises graphic stimuli, decodes them to sound, recognises words and decodes meanings. Each component involves subprocesses which take place independently of each other, and build upon prior subprocesses. Subprocesses higher up the chain cannot, however, feed back into components lower down (identifi-

cation of meaning does not lead to letter recognition, for example). This approach was typically associated with behaviourism in the 1940s and 1950s, and with 'phonics' approaches to the teaching of reading that argue that children need to learn to recognise letters before they can read words, and so on. In this traditional view, readers are passive decoders of sequential graphic–phonemic–syntactic–semantic systems, in that order.

On the other hand, as we shall see in Chapter 2, much research has emphasised the importance in reading of the knowledge that a reader brings to text. Models of reading that stress the centrality of this knowledge are known as **schema-theoretic models**. They are based upon schema theory, which accounts for the acquisition of knowledge and the interpretation of text through the activation of schemata: networks of information stored in the brain which act as filters for incoming information (for much more detail, see Bartlett, 1932; Ausubel, 1963; Hudson, 1982; Carrell, 1983a, Carrell *et al.*, 1988). In this view, readers activate what they consider to be relevant existing schemata and map incoming information onto them. To the extent that these schemata are relevant, reading is successful. **Top-down** approaches emphasise the importance of these schemata, and the reader's contribution, over the incoming text. Goodman (1982), for example, calls reading a 'psycholinguistic guessing game', in which readers guess or predict the text's meaning on the basis of minimal textual information, and maximum use of existing, activated, knowledge. Smith (1971) claims that non-visual information transcends the text, and includes the reader's experience with the reading process, knowledge of the context of the text, familiarity with the structures and patterns of the language and of specific text types, as well as generalised knowledge of the world and specific subject matter knowledge.

A typical statement of the top-down approach can be found in Schank (1978):

> We would claim that in natural language understanding a simple rule is followed. Analysis proceeds in a top-down predictive manner. Understanding is expectation based. It is only when the expectations are useless or wrong that bottom-up processing begins. (Schank, 1978:94)

However, many psychologists and psycholinguists now question the usefulness of schema theory to account for, rather than provide a

metaphor of, comprehension processes. One issue is *how* prior knowledge is called up from memory, and how it is then used in understanding. The problem is that schema theory does not lead to explicit definitions or predictions of processes of understanding, although it has clearly provided a powerful incentive to research into the products of understanding for first- as well as second-language readers.

Partly as a result, recent research tends to emphasise the important contribution of bottom-up or data-driven processing to fluent reading. In particular, numerous studies of eye movements using sophisticated instruments have consistently shown the importance of rapid and automatic processing of most of the words on the page: one estimate is that fluent readers process some 80% of content words and 40% of function words (in English). What distinguishes good from poor readers is not the number of letters in a fixation, nor the number of words fixated per page, but the speed of the fixation – the automaticity of word recognition – and the processes that occur during fixation. It has been suggested that after initial word identification, but still during the fixation, good readers move onto higher-level prediction and monitoring, as well as planning of subsequent fixations. This is thought to be because they use less capacity to analyse the visual stimulus, and therefore have other resources available for other sorts of processing.

Not only are good readers rapid in their word recognition, they are precise as well. Readers take in letter features of short words simultaneously and appear to recognise all the letters in a word. The ability to recognise words rapidly and accurately is an important predictor of reading ability, especially with younger first-language readers, and even for college-level students.

In fact, however, neither the bottom-up nor the top-down approach is an adequate characterisation of the reading process, and more adequate models are known as **interactive models**, in which every component in the reading process can interact with any other component, be it 'higher up' or 'lower down'. Processing, in fact, is now thought to be **parallel** rather than **serial** (Grabe, 1991:384). Rumelhart's (1977) model, for example, incorporates feedback mechanisms that allow knowledge sources (linguistic as well as world knowledge) to interact with visual input. In his model, a final hypothesis about the text is synthesised from multiple knowledge sources interacting continuously and simultaneously. Stanovich (1980), on the other

hand, has developed an **interactive compensatory model** in which the degree of interaction among components depends upon knowledge deficits in individual components, where interaction occurs to compensate for deficits. Thus, readers with poor word recognition skills may use top-down knowledge to compensate. (However, the evidence that such compensation does in fact occur is controversial, as we shall see in Chapter 2.)

Although Goodman's model is often characterised as a top-down model, and Smith's popularisations acted as useful correctives to excessively bottom-up approaches in the 1970s, Goodman himself (1982) rejected the label, and claimed that his model assumed that the goal of meaning is the construction of meaning which requires interactive use of grapho-phonic, syntactic and semantic cues to construct meaning. Readers are not passive identifiers of letters and words but active constructors of their own knowledge. He saw reading as a complex process of **sampling** the text for graphic clues, **predicting** grammatical structures and meaning, **confirming** the validity of the hypotheses advanced and **correcting** the hypotheses as necessary as text sampling proceeds.

Less proficient readers often appear 'word-bound'. Traditional psycholinguistic models such as Smith's claimed that such readers need to take more risks, but more current views suggest that these readers are not yet efficient in bottom-up processing. They do not recognise the words sufficiently rapidly and accurately (and for second-language readers there might well be graphic as well as lexical problems anyway – there are too many new forms for students to attend to). Guessing will not overcome this deficiency and lead to automatic recognition – there are no short-cuts to automaticity, although some research has attempted to improve automatic recognition.

More recent approaches to reading have begun to investigate the importance of the visual input once more. It is recognised that letters are not processed serially in order to identify words (Samuels and Kamil, 1988) – there are syntactic and semantic effects on word recognition, so that related pairs of words will be recognised more quickly than unrelated pairs, and in word recognition errors, substitutions are often of the same syntactic category as the word being substituted. It has been shown that good readers do not simply sample text – they do not skip over words in normal fluent reading: 'the single immutable and non-optional fact about skilful reading is that it involves relatively complete processing of the individual letters of print'

(Adams, 1991:105). What seems to matter is the speed of recognition, the automaticity of the process.

Poor readers are distinguished from good ones by: poor phonetic decoding; insensitivity to word structures; and poor encoding of syntactic properties (Vellutino and Scanlon, 1987). They see reading difficulties as a linguistic problem, involving a failure to recognise how particular structures encode information, and not as a problem of insufficient background knowledge or insufficient top-down strategies.

Thus, the pendulum swings. It is clear that both bottom-up and top-down information is important in reading, that the two interact in complex and poorly understood ways, and that the balance between the two approaches is likely to vary with text, reader and purpose (see Chapter 2). Given the emphasis placed by researchers and teachers alike on more top-down approaches in the 1970s and 1980s, it is likely that we will soon see some change of emphasis as the importance of text recognition receives more attention from research and pedagogy.

What are the implications for assessment? Clearly there are diagnostic issues: causes of poor reading can be hypothesised to be more bottom-up or more top-down, depending upon one's model and the data available, and diagnostic testers would do well to pay attention to both possibilities, rather than to concentrate on one. One can envisage situations where poor reading may be due to poor bottom-up strategies or inappropriate application of background knowledge, and knowledge of which approach has prevailed might be useful.

However, reading achievement and proficiency tests may well be less influenced by notions of the nature of the process and the strength of the arguments in the debate, since I would argue that, by their very nature and purpose, at least as presently conceived, such tests concentrate on product rather than process (but see Chapter 9). It is, moreover, difficult to envisage a reading test that would *require* students to adopt either a bottom-up or a top-down approach, since it may well be that either can result in a given understanding.

It might be useful for testers to ask themselves, when looking at test items they have devised: is this a top-down or a bottom-up item? Would top-down reading give a better chance of getting this item right or wrong? But it is highly unlikely that any test item involving meaning would involve only one or the other approaches. It is most likely that there will be an *interaction* between textual clues and the reader's knowledge.