Comprehensions and Interpretations

A.H. Urquhart

College of St. Mark and St. John, Plymouth

This paper argues that 'reading comprehension' and its assessment can usefully be divided into two aspects: 'comprehensions' and 'interpretations'. The first term refers to different standards or levels of comprehension the reader adopts to suit different purposes of reading. It is argued that the emphasis in both teaching and assessing reading on 'deep comprehension' is limiting and harmful. The second term refers to different readings of the same text, either by different readers or the same reader at different times, resulting from either a different background knowledge or a different preoccupation at the time of reading. The paper explores some of the consequences for testing of recognising these different aspects of reading.

A few years ago it was virtually a cliché to state that reading was often tested but seldom taught. Nowadays, more attention is paid to reading, and test to many reading teachers is a rude word. Yet testing goes on, and must go on, while students either wish or have to be entered for public EFL examinations. In this paper I am partially concerned with testing, but also with the broader area of reading assessment, in which I include both formal tests, informal teacher tests, and students' own assessment of their reading.

Most people concerned with assessing reading would probably agree that the primary concern was with comprehension, as opposed to, say, the ability to read aloud, or the syntax of text. However, notions as to what constitutes comprehension differ radically, and this clearly affects our idea of what it means for a student to read a text in a satisfactory way.

In this paper, I argue that it is not helpful to think in terms of a unitary, 'total' comprehension. Instead I want to talk about comprehension and interpretation, and to use both terms in the plural, for reasons which I hope will soon become obvious.

By 'comprehensions' I mean the different products of the reading process, the results of the different standards which readers set themselves, partly because of their purpose in reading, and partly because of the nature of the text. Readers may, for example, set themselves different levels of comprehension — depending on whether they are reading a newspaper, or a textbook set for an examination. In the same newspaper, they may read general news, editorial, sports page, etc., at different levels of comprehension. The standards which comprehensions depend on are more or less under the conscious command of the reader. We can, if we wish, examine our daily tabloid with all the attention of the most devoted text analyst; on

Alexander URQUHART has worked in East Africa, Nepal, the United States and Britain, and has taught EFL, TEFL, and Applied Linguistics. He is at present Senior Lecturer at the College of St. Mark and St. John, Plymouth. Sandy Urquhart's academic interests include text analysis, pedagogic grammar, and the teaching of reading. He is co-editor (with J.C. Alderson) of Reading in a Foreign Language (1984), Longman.
the other hand, we may decide to treat a text designed as a textbook in paleobotany as a piece of light bedtime reading.

It is now widely accepted, at least in theory, that comprehension is affected crucially by the reader's relevant background knowledge. This knowledge may relate to the reader's general cultural background. For example, an Indian may understand differently from a North American (cf. Steffensen and Joag Dev 1984). Again, comprehension may be affected by academic background: an engineer may read a text in some ways differently from a non-engineer (cf. Alderson and Urquhart 1985). These different readings of the same text I propose to call 'interpretations'. Such interpretations are nowadays often discussed in terms of schema theory.

I also want to use this term to refer to the different readings of the same text made by the same reader at different times, and presumably as the result of a different mental set. I am surely not the only reader to find on reading, say, an academic paper for the second time, that my first and second readings are considerably different. Sometimes material which I 'remember' turns out not to be there; sometimes, to my embarrassment, a text turns out to include material which I have criticised the author for omitting. If your reaction to this is to say that I am a bad reader, I will have to agree (but will mutter about motes in other people's eyes). The fact is that in this, as in other areas, we see what we want to see. Interpretations in my sense are not under the control of the reader, which is a major factor separating them from comprehensions.

DEFINITIONS OF COMPREHENSION

Before proceeding, it will be useful to look at some more or less formal attempts to define reading comprehension.

Bormuth (1969) puts forward an 'operational definition' of comprehension intended to be free of woolly introspection. Comprehension is 'an increase in information as the result of reading'. Reading texts are language texts, hence comprehension is a response to the language system. Responses to different aspects of the language system can be made overt by different types of comprehension question, e.g. Rote, Transform, Semantic Substitution, etc. The comprehension 'unit' thus consists of the text, the question, and the response.

It is easy to make fun of Bormuth's system, and some of his 'rigorous' questions are frankly ludicrous, e.g. 'What caused the breaking of the diminutive boy's leg?' Bormuth's value in this discussion is his very explicit assertion that comprehension equals the ability to answer comprehension questions. Many of us would be uneasy about accepting this in its entirety, though remarks found in the literature that answering 10 questions correctly out of 10 equals 'perfect' comprehension seem to go most of the way with Bormuth. There is no doubt, however, that the assessment of reading, as opposed to comprehension itself, requires an overt response.
Carroll (1972), another tester, attempts to separate comprehension from other aspects such as memory and inferencing. Remarks such as 'a certain amount of intellectual effort that... goes beyond sheer comprehension' suggest that for Carroll, comprehension is a passive manifestation. He further distinguishes between 'adequate' and 'total' comprehension. By the first, he means recognition of the linguistic information which exists in a partial text. So, for example, given the sentence, *The Fundalan added an are to his plot*, 'adequate' comprehension would consist of recognition of Subject-Verb-Object relationships, the fact that the suffix 'an' may signify 'a person originating from', etc. 'Total' comprehension would occur when this information was related to the total context, thus letting us know who or what the Fundalan was, etc.

Carroll’s account breaks down very quickly. As he himself admits, relating 'adequate' comprehension to a wider context may involve inferencing, which he has ruled out. In addition (though he does not say so) it involves memory, which he has also ruled out. His account is psychologically unconvincing, requiring us to believe that the reader will be content with virtually grammatical information until later enlightenment, so that in normal human terms his 'adequate' comprehension is not 'adequate'. Moreover, in many cases, the wider context never stops expanding, so Carroll's 'total' comprehension is never total.

A recent collection of papers on reading comprehension (Flood 1984) contains at least three very different definitions of comprehension. According to Rumelhart (1984) “readers are said to have understood the text when they are able to find a configuration of hypotheses (schemata) which offer a coherent account for the various aspects of the text.” This is the currently most fashionable definition of comprehension, and is equivalent to my ‘interpretations’.

Calfee and Curley (1984) define comprehension in terms of recognition of text structure. The reader's task in comprehending a text is said to consist of finding the organisational structure and modifying it for his own purposes. This is interesting for its recognition of purpose. But there are comprehensions that do not seem to require much recognition of structure, e.g. those resulting from scanning.

Baker and Brown (1984) do not actually define comprehension, but rather describe nine activities crucial to good comprehension. Some of these are: establishing the purposes for reading; modifying reading rates and strategies in accordance with different purposes; identifying the important elements of the passage. The ninth activity is particularly interesting in the context of this paper, and is expressed as 'selecting appropriate standards for assessing one's level of comprehension'. This is clearly related to what I have called 'comprehensions'.
ASSUMPTIONS ABOUT COMPREHENSION

While formal definitions of comprehension are important, they are often produced by theorists or researchers, and thus may not have an immediate effect on the practice of teaching reading. Perhaps more significant for the argument in this paper are assumptions about comprehension apparently lying behind statements by writers of reading material. I now want to move on to consider some of these assumptions, which I think are widely held, even if not always explicitly stated. I would summarise them as follows:

1. There is one kind of reading which is ‘better’ than all the other kinds; or, perhaps, there is only one kind of activity which really deserves the name ‘reading’.

2. Associated with this, there is such a thing as ‘total’ or ‘perfect’ comprehension of a text.

3. This perfect comprehension exists independent of a particular purpose.

4. While at times, a less than perfect comprehension is acceptable, it is normally the case that the more total the comprehension the better.

As support for the existence of these assumptions, I cite the following evidence. The distinction is often made between ‘skimming’, ‘scanning’ and ‘reading’. For example, Sticht (1984) argues that claims for the possibility of reading much faster than listening rest on a confusion between skimming and scanning on the one hand, and reading on the other. Skimming and scanning, Sticht argues, can be accomplished faster than listening; ‘reading’ can not. It seems to me that this position implies that (a) skimming and scanning are not ‘reading’; (b) that real ‘reading’ implies a slower, probably more careful activity than do skimming and scanning. On the other hand, I would argue that skimming is one kind of reading, scanning another, and both are different from the slow, careful reading for which Sticht seems to reserve the term.

Further, Fry (1963) — discussing comprehension in the context of faster reading —states that 100% gained on his comprehension exercises equals ‘perfect comprehension’, that 50% equals ‘poor comprehension’, and that, (by simple arithmetic), 70/80% comprehension is the figure to be aimed at. This position has been restated much more recently by Whitney and Davies (1984), again talking about faster reading:
READING EFFICIENCY: a chart

Use this chart to calculate your reading efficiency for this text:

<table>
<thead>
<tr>
<th>% COMPREHENSION</th>
<th>EFFICIENCY</th>
</tr>
</thead>
<tbody>
<tr>
<td>If you think you understand</td>
<td>You are probably reading...</td>
</tr>
<tr>
<td>0—40% of the text</td>
<td>far too quickly</td>
</tr>
<tr>
<td>40—60%</td>
<td>too quickly</td>
</tr>
<tr>
<td>60—80%</td>
<td>about the right speed</td>
</tr>
<tr>
<td>80—100%</td>
<td>perhaps not quickly enough</td>
</tr>
</tbody>
</table>

Nuttall (1982) seems to accept this view, even if a trifle reluctantly: ‘... what score is to be considered adequate? The general view seems to be that about 70% is enough’.

There are two points to be made at this juncture. First, the writers above seem to be contradicting my assumption 4. I do not think they are, however, since they are talking about a special context, in which the student is trying to increase his reading speed, so that a drop from ‘perfect’ comprehension is justified. Fry actually claims that 50% of material found in newspapers and magazines is redundant (what he bases this claim on I do not know). Hence, presumably, 50% comprehension is in a sense equal to 100%.

The second point is that I am almost certainly being unfair to these writers, all of whom would consider that speed, and to some extent comprehension level, should be flexible. However, all of them are to some extent responsible for a belief that a single ‘level’ of comprehension is acceptable for all or most purposes. But there is no reason to accept this. Even if we could agree on a way of measuring ‘70%’ comprehension, I would argue that there are times when 5%, or 1% is adequate. And this is true even in the restricted context of ‘faster reading’. (One of my few successes with faster reading involved persuading a Syrian postgraduate student that he should be ruthlessly selective and read only those parts of the text relevant to questions I had pre-set. He was delighted with his new-found ability (even though by any measure his comprehension of the whole text must have been well below 50%) and actually started to move away from his habit (all too common) of grinding through each text at about 80 wpm.)
As further evidence for the existence of the four assumptions I outlined above — In a rubric in her very useful book *Authentic Reading* (1982), Walter states, ‘... Do this exercise, which needs a better understanding of the text than the Summary Skills exercises.’ It might seem trivial, but I think the use of the word ‘better’ here is significant, implying as it does that ‘more detailed’ equals ‘better’. I think that this is a very common assumption, related closely to the view that there is a perfect comprehension and that the closer we approach it the better. This is a view which I want to look at next.

The reader may well ask: ‘What’s wrong with ‘perfect’ comprehension? All other things being equal, surely 100% comprehension (if it can be measured) is better than 50% or 10%?’ One response to this is to say that other things seldom are equal, and that the comprehension has to be paid for some way, usually in terms of speed. Many EFL readers read painfully slowly, desperately anxious to ‘understand everything’. This is often explained in terms of cultural background, but the emphasis in EFL reading on slow, careful reading may well play a part.

Another response is that this emphasis on one kind of reading does not match our behaviour outside the classroom. Now obviously, the classroom is never a true model of the real world. Nevertheless, emphasis on one kind of comprehension as the ideal, when it is not the ideal in our non-classroom life, may seem to harbour dangers. This is perhaps particularly evident when the ‘ideal’ held up by different writers turns out to be not the same. For example, Williams (1986) states: ‘the concept of ‘an accurate image’ of the writer’s intentions is an ideal (of comprehension)’. This is a view which would receive wide support, and there is no doubt that in many cases it is valid. Williams’ ideal reading behaviour, in which the reader strives to re-create the author’s message, is similar to the behaviour of what Widdowson (1984) has called ‘the submissive reader’. However Fransson (1984) clearly regards as the ideal the ‘deep processer’, whose typical recall (reflecting his reading behaviour) is described as follows:

The student summarises his main conclusions from reading the text. He explains his thoughts and reflections while reading the text, and summarises the parts of information that he has found most interesting.

As opposed to deep processors, surface processors try to give ‘a neutral and complete summary of the content of the text’. It is not possible to match all these descriptions perfectly. However, it does seem that Fransson’s surface processor has something in common with Williams’ reader striving to accommodate the author, while the deep processor — intent on taking from the text what he finds interesting — has more in common with what Widdowson has called the ‘dominant reader’.

The obvious answer is that we must not insist on a single model, since different standards may be appropriate at different times.
NORMAL READING OR DISCOURSE ANALYSIS?

This very strong emphasis on 'deep' reading can result in our adopting abnormal reading strategies, making us behave more like discourse analysts than normal readers. Consider the following text and accompanying questions, taken from the Watson/ Glaser Test of Critical Thinking (Watson and Glaser 1951).

The first newspaper in America, edited by Ben Harris, appeared in Boston on September 25, 1690 and was banned the same day by Governor Bradstreet. The editor's subsequent long fight to continue his little paper and print what he wished marks an important episode in the continuing struggle to maintain a free press.

The text comes from the section of the test which deals with inferences. The testee must read the text, then decide whether the following set of statements are, with reference to the text, TRUE, FALSE, PROBABLY TRUE, PROBABLY FALSE, OR INSUFFICIENT DATA (i.e. don't know).

1. A copy of the first issue of Ben Harris' newspaper was promptly brought to Governor Bradstreet's attention.

2. The editor of this newspaper wrote articles criticising Governor Bradstreet.

In a paper I wrote several years ago (Urquhart 1979) I argued that this type of activity, devised and put into practice by educational psychologists, was a far more interesting and more valid reading exercise than much of the stuff resulting from more linguistically oriented approaches. I still, in fact, think that is the case. However, this type of question raises problems.

Let us look at the two statements, starting with 2. It might on casual reading have occurred to us that this was, in fact, the cause of the banning. More detailed thought, however, will usually lead to our coming up with other possibilities, e.g. criticism of the government in England, articles revealing state secrets. Question 1 seems less in doubt. Surely this is true? But the governor might have been informed of the contents of the paper before publication, and might have decided in advance to ban it.

I have used parts of this test with classes of students as part of a study of inferences in texts. While the texts and questions certainly generate discussion, and may deepen insights into the nature of inferences in discourse, one slightly unexpected result is that everyone becomes extremely tentative — TRUEs and FALSEs go out the window, and everyone settles for PROBABLY this and PROBABLY that, sprinkled with a few INSUFFICIENT DATAs.
I referred above to a ‘casual’ reading of the text, and some readers may take this in a derogatory sense. But what I am trying to argue here is that for a text like this, a casual reading may be very appropriate, and the comprehension which results from it a very valid comprehension. In contrast, the much more tentative comprehension which results from an investigation of the implied information in the text may tell us more about the way we construct texts, and may in some ways be deeper, but it is certainly not normal. It is very likely that in real life reading we read into texts all sorts of information which introspection would lead us to qualify.

Investigating the nature of textual evidence like this leads to one putting heavier weight on a text than it was ever designed for. Deep comprehension can be unnecessary, and from the point of view of someone wanting to elicit normal reading, positively dangerous.

The Watson/Glaser Test is a test of ‘critical thinking’, so presumably ‘normal comprehension’ is not its main target. The problem becomes more acute for tests such as the RSA Communicative Use of English Test, which set out to test ‘ordinary use of language’. Such a test, apart from the requirement that it use normal, authentic texts of the sort the testees are likely to come across in normal life, also requires that they be required to show ‘normal comprehension’. Sometimes, as I suggest below, the very test mechanism seems to mitigate against this.

To indicate some of the problems, I now wish to examine part of the reading test (advanced level) from the RSA examinations, November 1985.

The Ways that Cold can Kill

by Tim Rayment

Cold weather can kill you in ways you may never have thought of. As Britain entered winter’s chilliest phase last week, safety experts issued warnings that explosions, carbon monoxide poisoning, fires and falls could make an untimely end more likely than in any other season. The blast that killed eight people in Putney, south London, alerted the nation to the danger of gas leaks.

But carbon monoxide accidents, which are less dramatic than explosions, kill nearly three times as many. According to British Gas, there were 54 deaths from poisoned air and 20 in explosions in a year. All but six of the poisonings happened between October and March. Poor maintenance of equipment and inadequate ventilation are generally to blame.

“People try to seal themselves off in winter,” said a British Gas spokesman. “The silliest thing they do is to block up airbricks. Appliances have to breathe like you and me, and if the room is hermetically sealed you can get incomplete combustion, and carbon monoxide.”
Mains gas explosions, such as that at Putney, are seven times more likely in winter, according to Dr. Philip King of Manchester University, Institute of Science and Technology.

Perhaps the most spectacular accident was last week's collapse of a bridge over the river Loire in France, apparently because of frost.

"Cold makes all materials contract," explains Tony Rochester, senior bridge engineer at the Department of Transport. "But if the temperature is very low it can make metal very brittle, which can cause a failure." Bridges built in Britain since 1978 must be able to cope with temperatures as low as -18°C inland to -24°C in Scotland.

Snow can be a bigger problem. The Building Research Establishment, part of the Environment Department, has issued new guidance on allowing for the strain, which can be equivalent to a crowd of people on the roof.

Water is another unsuspected danger. A fire that killed 25 old people in France this month is being blamed on a short circuit caused by water leaking from a burst pipe.

The London Fire Brigade answered more than 800 flood calls during an eight-hour thaw last week. "Frozen pipes can be prevented by keeping central heating on low 24 hours a day, and proper lagging and insulation", the brigade advised.

 Particularly at risk from the extreme cold are the nation's down-and-outs. With an estimated 40,000 homeless people in the Greater London area, it had been mooted that London Regional Transport should open selected tube stations as night-time shelters, but LRT said yesterday that no suitable stations had been found.

Hypothermia, however, seems for most people an exaggerated risk. Dr. Robert Irvine, consultant adviser on geriatric medicine to the Health Department, said most people in hospital with hypothermia, defined as a core body temperature under 35°C, were also found to have some other illness.

Falls are far more frequent. The John Radcliffe Hospital, Oxford, has issued so many crutches that stocks are running out. It is urging patients who not longer need support to return their crutches.

The outlook is for snow turning to sleet or rain, and becoming less cold.
6 Look at these statements.  
Are they true?

According to the article.

a) Gas explosions are more common than carbon monoxide accidents.
b) Gas explosions are more likely to occur in winter than in summer.
c) A gas explosion caused a bridge to collapse in France last week.
d) Bridges built in Britain before 1978 are unsafe in very cold weather.
e) Water can indirectly be the cause of fire.
f) Most people who are suffering from hypothermia in hospital are likely to die.

Tick (√) TRUE, FALSE or DON’T KNOW beside each letter in the answer column.

<table>
<thead>
<tr>
<th></th>
<th>True</th>
<th>False</th>
<th>Don’t Know</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
a |     |       |            |
b |     |       |            |
c |     |       |            |
d |     |       |            |
e |     |       |            |
f |     |       |            |

I have to say that some of these questions gave me quite a lot of difficulty (readers are invited to try them themselves) and I am still not sure whether I have arrived at the answers required by the test-setters. These problems are relevant to my argument, so I shall go through the main ones.

First of all, the rubric. ‘Look at these statements. Are they true?’ This looks like a straight, unhedged question: ‘Are they true for you in the light of your knowledge?’ However, this would mean that we wouldn’t need to read the text, and is corrected by ‘According to the article....’ But a certain element of doubt has been introduced. Now we must decide what the rubric means by ‘Are these statements true according to the article?’ I think this can legitimately be taken to mean either ‘Does the writer present information which guarantees the truth/falsity of the statements?’ or, perhaps more likely, ‘On the evidence of the text, does the writer believe these statements to be true?’ It is clear that one is well and truly involved in a guessing game.

Turning now to the actual questions, three give me particular problems:

Q.a: A quick reading left me with the impression that the opposite was the case — after all, more people are killed by gas poisoning, so it seems reasonable to assume that carbon monoxide accidents are actually more common. So False. But wait a minute. The text never compares actual numbers of explosions
Comprehensions and Interpretations

versus gas leaks, and it is conceivable that numbers of small explosions occur harmlessly, or at least non-lethally. So should I tick Don't know? I'm still inclined to False but my confidence has been damaged. (If this was part of the Watson/Glaser test, I would opt for Probably False.)

Q.d: Here I am inclined to choose True, but with considerable reservations. What does the tester mean by 'very cold weather'? At about the time I write this, the London-based BBC has been twittering about Arctic conditions on the Air Ministry roof, with temperatures as low as -6°C. Clearly they believe this is very cold. But it is doubtful whether it would endanger bridges. Then what does the tester mean by 'unsafe'? What is 'unsafe' to an engineer, assessing the risk of bridge collapse is almost certainly not the same as unsafe in layman's terms. The author of the article is a journalist: he almost certainly doesn't know much about it. The same is probably true of the tester. Is the best answer then Don't know? (Again, in a Watson/Glaser test, I would choose Probably True.)

(f) This one has me truly beaten. I'm not at all clear as to what the text means, and suspect that the tester has come to different conclusions from me. One could argue that the article implies that one doesn't often die of hypothermia, since even being in hospital is normally the result of another disease. Or one might argue that since the patients have more than one disease, they are likely to die anyway. Or one can throw up one's hands and choose Don't know.

I would conclude the following from my difficulties:

1. I feel when trying to answer these questions that the tester is getting between me and the article, forcing me to try to guess what he's thinking.

2. I don't know how subtle he's trying to be; in other words, I don't know what level of comprehension he's aiming at.

3. While I think there is evidence that the tester is looking for a fairly surface kind of comprehension, based on the rough suppositions which we normally make when reading texts like this, there is no way the testee can be sure of this.

3. There would seem to be no constraints on the tester, no limits in theory to the subtlety he employs. There are, of course, practical constraints in terms of the proportion of testees who pass the paper.

4. The exercise took far more effort on my part than I would normally spend on a newspaper article like this. That is, the comprehension it seems to be forcing me to aim at is a different comprehension from what I would normally extract from such a text.

When formulating questions, test-setters tend, naturally enough, to dig deeper in the text than the normal reader would. The problem is then that the questions force
the testee into this same kind of intensive reading, which often brings difficulties absent from normal reading. Tester and testee, by the nature of the game, become involved in discourse analysis.

**TASK-BASED ASSESSMENT**

In the context of the teaching or assessment of reading the word 'task' is often used to describe an activity dependent on the reading text but different from answering comprehension questions. Thus the reader may be asked to fill in a table, label a diagram, write a letter, etc. Nuttall (1982), discussing tasks as an alternative or supplement to questions in comprehension teaching, makes three points in their favour: they add flexibility and thus interest to the reading programme; they can be used to get an indication of process rather than product; they give the student a clear purpose for reading.

I think using tasks in reading assessment may solve some of the problems I have discussed above. This will involve, however, a more precise definition of 'task', and this in turn requires a brief look at what is meant by 'communicative reading' and in particular 'communicative reading tests'.

It would be widely accepted that a communicative situation must include, among other components, a Sender, a Message, and a Receiver — in reading terms a writer, a text, and a reader. Another component is Purpose, with the complication that there are two purposes, the author's and the reader's. (This, of course, is implicit in discussions of 'dominant' and 'submissive' readers.) All this has been at least implicitly accepted in a great deal of reading instruction long before the advent of any 'communicative movement'. Hence one's feeling sometimes that there is nothing new about 'communicative reading'.

When we consider testing of reading, the situation becomes more complicated. The testees have to adopt two communicative roles. As readers, they play the role of receivers; the message is the text and the sender is, of course, the original author. When however they attempt to answer questions set them by the tester, they become senders themselves. The message is now their response to the questions, and the receiver is the tester. Thus:

| Sender (Author) —— Message (Text) —— Receiver (Reader) |
| TESTEE |
| Sender (Responder) —— Message —— Receiver (Tester) |

In this second role, testees may be placed under a severe disadvantage: they often do not know the purpose of the communication, and are thus unable to adjust their
message to fit. Related to this, they do not know who the receiver is. (This is important since, when answering a number of test items, I have found myself trying to estimate just how good a reader the item setter is.)

I don't think this is just pseudo-academic hairsplitting. It seems to describe my reaction to trying to answer the RSA questions above. I felt there that my problems were not with the author, but rather revolved around what I was supposed to say to the testers. In a real sense I didn't know who they were or why I was communicating with them. Hence my problem with the level of comprehension I was being asked to display.

Thus according to this argument, for a reading test to be 'communicative' the testees must be given a communicative task which specifies who they are communicating with and why. Otherwise a test may use authentic texts but be no more communicative than traditional tests.

The tasks discussed by Nuttall are not necessarily tasks in my sense, though many of them could be made so by supplying the student with an audience, imaginary if necessary, and a purpose.

If 'communicative tasks' were employed in testing, the following advantages would be gained:

1. By providing a clearly outlined purpose, they establish the context or limits of assessment. In other words (in my terms) they establish which comprehension is being sought. And since basically I have been arguing that comprehension can only be judged in relation to a particular purpose, it follows that it is only through tasks that we can assess comprehension.

2. They should establish constraints on the testers, limiting them to what can be considered as 'normal comprehension'.

3. If the task is explicit enough, its use should reduce the presence of the tester in the assessment situation, cutting down on the 'guess what I'm thinking' element.

4. If, as it should be, the task is reasonable and recognisable in terms of the testees' experience of real-life situations, the test will be more in line with the requirement that 'communicative' reading tests should assess the testees' ability to handle normal language in normal situations.

There are two objections to task-based assessment. The first is theoretical: by having the testees perform text-related tasks, we are assessing Application (Bloom 1956) rather than Comprehension. One blunt answer is 'So what?! A more reasoned one is to argue that it is impossible to assess 'pure' comprehension, since even to respond to comprehension checks, the testees must apply their comprehension. In other words, comprehension and response to comprehension
checks and tasks are always separate, and the movement from one to the other always involves application. It is as impossible to test ‘comprehension’ as it is to test ‘competence’.

A more practically oriented objection, as Nuttall points out, is that the tasks tend to involve activities other than reading (e.g. writing), so that the assessment would not be able to distinguish between reading and the other activity. This problem must be admitted. It is particularly the case in tests such as the RSA, which attempt to separate language skills. Part of the problem can be got round by the marking procedures adopted. However, the problem remains for integrative approaches in general.

TEXTS AND TASKS

Some texts lend themselves to tasks better or more naturally than others. For example, instructions lend themselves to being tested by having the testee obey them, as pointed out by Mohammed and Swales (1984). Various other text-types have activities associated with them in normal contexts; time-tables, for example, can be tested only by having the testee find certain times. The increase in such functional texts in course-books, e.g. Reading Resources (Forrester 1984) may be partly in response to a desire to relate reading to recognisably normal activity.

Many texts (and not just literary ones) do not, however, relate easily to such overt activities. Advertisements do not, for example. Nor do the majority of newspaper articles. Nor, in fact, do most expository texts. In such cases, the teacher/assessor must use some imagination in constructing an imaginary but reasonable task.

INTERPRETATIONS

So far I have been discussing comprehensions, and aspects of testing or assessing them. I now want to turn to my other aspect of reading, which in the main concerns differences in comprehension between different readers. In the Scottish Higher English examinations, comprehension texts with their associated questions used to be referred to as ‘interpretations’ (they may be still). As one might expect, no allowance was actually made for ‘interpretation’ — you either got the answer right or wrong. However, ‘interpretations’ seems to be a good term to describe the way different readers may arrive at different conclusions, based on the same written material, may select different aspects, may remember different parts, etc.

This, of course, is a common theoretical position nowadays; indeed, it is virtually a cliché. It is based in part on the observation that the information in text is never complete, but must be completed, made into discourse, by the reader. This applies to all texts, from children’s stories to scientific reports. Examine, for example, the following children’s text:
The Proud Princess

Long ago, the King of Ireland had a lovely daughter. But she would not choose a husband. She thought no one was good enough for her. The King was very cross. He asked the three princes in the land to tea at his castle. “You must choose one of them,” he said.

The only overt clue to intersentential relationships is the presence of but between S1 and S2. The way I read this text, I must supply material similar to the following:

Long ago, the King of Ireland had a lovely daughter. (“Now normally, in such a case, he would have no difficulty marrying her off, which would be a good thing, because that way he could maybe form a new alliance, or possibly get some money, or simply rid himself of the necessity of providing his daughter with food, accommodation, education, etc.”) But she would not choose a husband. (“Why not? Because...”) She thought no one was good enough for her. (“And that is why she was considered to be a ‘proud’ princess). The king was very cross. (“Why? Because by refusing to marry, she was denying him all the possible advantages mentioned above. So what action did he take to try and solve the problem which was making him cross?”) He asked the three princes in the land to tea at his castle. (“What was the good of doing this? Was he going to ask their advice? No. These were to be the choice of bridegroom. And, to the princess...”) “You must choose one of them,” he said.

Much of this, of course, depends on cultural presuppositions. A reader from a culture in which there is no relationship between feminine beauty and availability of bridegrooms might be puzzled. Similarly, readers who did not have access to the stereotype that kings chose grooms for their daughters would have to structure the text in some different form. (See Steffensen and Joag Dev 1984.)

The way in which comprehension is affected through the text being related to the particular background knowledge, or knowledge structures, of a reader is an important part of schema theory. The definition of comprehension earlier cited from Rumelhart belongs to this approach. However, rather than give yet another account of this theory, I should like to concentrate on how common the experience of different interpretations is to many of us. My reason is that this is a crucial aspect of reading comprehension, and yet I suspect many practising teachers are put off by being referred to what may appear to be another esoteric theory, produced by workers in some remote discipline such as Artificial Intelligence.

I should therefore like to begin with a quotation from an American historian, writing long before the advent of schema theory. Talking, appropriately enough, about the way in which ideas from Western culture were misunderstood or distorted when exported to other cultures, McNeill (1963) remarks: “It would be strange if such were not the case. Even within the western European world,
differences of language make perfect translation impossible. Indeed, anyone who has read reviews of his own writing is likely to conclude that accurate communication between men of the same profession and native tongue seldom occurs.” (My own academic production is nothing like as distinguished as McNeill’s, but this is an experience I have myself shared.)

Why do such ‘misinterpretations’ occur? It is surely not because of perversity or defective knowledge of syntax, semantics, rhetoric, and so on. It is easy enough to see that it comes from the fact that readers in certain situations come to a text with an overriding interest or purpose, and that this causes them to interpret everything relevant in such a way that it conforms to their wishes. We see what we want to see.

Thus there are at least two reasons why we sometimes do not comprehend a text as the author perhaps intended us to: our cultural background may be different; and our current preoccupations may blind us to certain aspects of the text, or at the very least, to concentrate on some aspects at the expense of others.

This is easy enough to accept in theory. However, in our activities as practising teachers of reading, it is all too easy for us to slip into the attitude that there is one correct interpretation of a text, and one way of approaching a text. If pressed, many of us would identify this one meaning with ‘author’s intended meaning’, and it is this position which I want to discuss now. It has been stated forcefully by Williams (1986), part of whose discussion is quoted below. Talking about the possibility of allowing for more than one understanding (i.e. interpretation) of a text, Williams says:

First, it does not seem helpful to discuss the question in general terms. The nature of the text concerned is crucial. The distinction between ‘closed texts’ that aim at arousing a precise response (e.g. a set of instructions) and ‘open texts’ which may have no such aim (e.g. a novel) is relevant here (cf. Eco 1981).

Second, a view that accepts any meaning attribution as correct risks devaluing the writer. The writer has information which he wishes to make ‘accessible to his supposed reader’ (Widdowson, p. 221). The extent to which a reader reconstructs and ‘follows’ the writer can be taken as a measure of ‘correct meaning attribution’. The reader’s reconstruction may never be an accurate image of the writer’s intentions (cf. Williams 1983: 178). This does not invalidate the approach, since the concept of an ‘accurate image’ of the writer’s intention is an ideal. It aims to safeguard the writer’s position as an equal but absent partner in the reading process, rather than condone the misrepresentation of the writer by the reader; in short, it aims to safeguard communication. If the idea that a reader can misunderstand a text is denied, then the writer is relegated to the status of a provider of cues for which any
response is ‘correct’. Such a denial would also make it very difficult to answer the question of what the ultimate purpose of improving reading strategies is.

In response to this, I should like to make the following comments:

a) Williams’ first point — that with certain texts some looseness of interpretation is permissible, but not with others — is, I am sure, widespread. However, I do not believe it is tenable. If one accepts that comprehension is a process whereby text information is related to reader’s knowledge and preoccupations, then this applies to the comprehension of *all* texts, not just some. Hence the distinction between ‘closed texts’ and ‘open texts’ just cannot hold.

It is true that, in the case of some texts, the consequences of our interpretation may be more immediately striking than is the case with other texts. If I interpret a foreign detective novel in a way very different from the way a native would interpret it, then the consequences are not likely to be very severe. In fact, if nothing jars in my interpretation, i.e. there is nothing which my internal monitoring tells me must be wrong, then I simply will not notice anything amiss. (In fact, I would argue there was nothing amiss). If I ‘misinterpret’ a study text for an examination, I may receive a reduced mark, or even be accused of not having done the necessary reading. If I interpret instructions for the electrical wiring of my house in a way radically different from that intended by the writers of the instructions, I may go out with a bang. But there is no point in telling me that such instructions are ‘closed texts’, not permitting individual interpretations. You might as well tell me not to read them.

Actually, instructional texts which result in the reader performing physical tasks are often self-correcting. If you fit together two bits of wood as, you believe, you are being instructed, and holes which should align do not do so, then this is good evidence that your reading is also out of line. But this is not a feature of the text itself; rather, it is evidence that such texts elicit responses which themselves are ‘closed’, to the extent that your Do-It-Yourself bed, for example, may end up with two legs pointing upwards.

b) Secondly, it is not the case, as Williams asserts, that we always have to pay attention to the views of the author. Whether we do or not depends on whether, for the purpose of the moment, we are adopting what Widdowson has called a dominant or a submissive attitude as a reader. If we go to a text to find out the name of the capital of Rumania, which it incidentally contains, then we are simply using the text as a carrier of a particular bit of information which we want, regardless of what message the author wants to convey. Hence, whether or not we wish to acknowledge the writer’s original intention depends on what type of comprehension, in my use of the term, we are seeking to attain.

However, it is certainly true that in many reading situations, we attempt to discover what the author intended. What is very important to realise is that this does not
conflict in any way with schema theory, or what I am calling ‘interpretations’. Readers striving to make sense of the text, to understand the author’s message, have no choice but to relate text to their existing knowledge. They are not being perverse; they simply have no alternative.

With the best will in the world, we are bound to ‘misunderstand’ texts; it happens all the time, is a normal part of communication, and will not change because of classroom teachers of reading saying, ‘No, you’re wrong. The author means...’ by which, of course, we mean, ‘The teacher thinks the author means...’.

Thus perhaps my deepest disagreement with Williams rests with his apparent belief or assumption that in some way the teacher is in control of students’ comprehension. I certainly think that teachers should try to aid students’ recognition of the common meaning of the code; I think they should try to expand their students’ awareness of the possibilities of extracting meaning from text; I think it is perfectly correct to discuss the efficiency of certain reading approaches, or the consequences of some applications. But I feel very strongly that ultimately the interpretation is up to individual readers. And I think this not because of adherence to some libertarian philosophy, but because in reality, that is all that is possible.

To illustrate this, I shall go anecdotal. Some years ago when discussing reading comprehension with a group of foreign teachers of English, I described an occasion when my class of African secondary school students was reading a text from Steinbeck’s *Cannery Row*. The passage describes how Danny got drunk, wandered down the main street breaking shop windows, and was arrested. He went along quietly because ‘he had a great respect for the law’. In jail, Danny whiled away the tedium by catching bedbugs, squashing them against the cell wall, adding legs, arms, etc., and naming them after various civic dignitaries, but never after the police or the judiciary because Danny ‘had a great respect for the law’. Eventually Danny and his jailer got drunk together and both ‘escaped’ from the prison.

Most of my students read this text with solemn faces and then proceeded to the multiple-choice questions. One student, however, smiled broadly at parts of what he was reading. In recounting this to the group of teachers, I said that it could be argued that his smiling was better evidence of comprehension than a score on the test items. (It’s a very debatable argument, but I would still claim there is something in it.) One teacher disagreed, not with my main argument, but with my interpretation of the text. It was not an amusing text, she claimed. It was a serious passage, expressing Steinbeck’s sympathy with the common man who, in his struggle against oppressive authority, never loses his noble faith in the rule of law.

Now I believe she was wrong in her interpretation; but then, she believed that I was wrong. Perhaps if I had been a student in her class, I might have kept quiet, but I doubt if I would have allowed her to change my mind. (Of course, if I had been one
of her students, the chances might be higher that we would have shared the same interpretation."

It is true that in the somewhat special situation of the reading classroom, it is always possible to produce more data which may succeed in altering a reader's interpretation. But we must always remember that the vast bulk of reading goes on outside the classroom, and there the teacher's authority is zero.

**DISTORTIONS AND ELABORATIONS**

I have in the discussion above sometimes referred to 'misinterpretations', and this is likely to leave the impression that the effects are generally harmful. A useful corrective to this view is supplied by McNeill:

Discrepancies between the reader's familiar knowledge and an author's presuppositions will always introduce such disproportions, and a book that attempts to deal with so large a subject as the history of the world invites misunderstanding on an unusually massive scale. Yet it is only when others take aspects of what an author has thought and said in order to develop, twist, and reinterpret his ideas to fit their own predilections and answer their own problems that the cold type of a printed page leaps to life; and, if this happens often enough, a single book such as this may become a real force in the cultural history of mankind. Without irony, therefore, I hope my book may be richly and repeatedly misunderstood.

In the experiment they conducted to measure the effect of different cultural background knowledge on comprehension, Steffenson and Joag Dev label as 'elaborations' information which a reader adds to make sense of a text produced within his own culture. Information from his own culture supplied to make sense of a culturally alien text — that is, information which is different from that possessed by the writer of the text — they label 'distortions'. Other writers, e.g. Bransford et al (1984) refer to both types as 'elaborations'. I think the latter position (which holds basically that whatever information a reader uses to make sense of a text is valid, and is certainly better than not making sense of the text) is better, and in fact is the more practical. Again I would point out that on the whole we are not in control of readers, and as teachers perhaps the most we can do is encourage them to make sense of texts.

**GENERAL IMPLICATIONS**

What effect are views such as these likely to have on the teaching of reading? Not a lot, if the teacher accepts that his job consists of, among other things, demonstrating what kinds of information texts can contain, how this information is realised and signalled, and how best the student can extract it. The approach does tend to mitigate against the reading teacher seeing himself as an objective arbiter
judging what the student has got out of a text.

As far as the testing of reading is concerned, I think that one implication of the argument is that the testing of inferential information should be viewed with suspicion, since different inferences may be equally successful in making the text meaningful for particular readers. This does not, of course, rule out the discussion of inferential information. Nor, in fact, does it apply to all inferences, since some seem to be inherent in certain forms of expression (cf. Urquhart 1981). But I think I have come round to something like Carroll's position — that if one wishes to conduct conventional reading tests, then one should limit them to tests of information linguistically 'committed' to the text.

THE LSP CLASSROOM

What I said above about the classroom teacher goes for the general teacher of L2 reading, whose class may consist of students from several cultures, each differing from each other, so that in terms of interpretation, the teacher's personal view of the text is one among a number of possibly differing interpretations. Things are rather different for the teacher of LSP (Languages for Specific Purposes) reading. As I said above, culture can be defined in more than one way: it can apply to general social background, or in a specific sense to, say, the academic background of an individual. It is in this latter sense that Widdowson has claimed that scientists belong to one culture regardless of nationality. It is a claim interestingly anticipated by McNeill, who I quote for the last time: "...certain branches of thought (particularly the natural sciences) and of the practical arts (particularly engineering) were sometimes genuinely shared by men of very different cultural backgrounds and personal predilections."

Thus an LSP group may consist of students from many different cultural backgrounds in the sense of social or national culture, who nevertheless all belong to the same academic culture — such as engineering. In such a case, the teacher is in more than one way the alien, and the one whose interpretation of the text is likely to differ most from any other member of the class. In theory, this should not be an impossible situation to live with. However, complaints from many teachers of LSP suggest that it is an uncomfortable situation. It is, of course, one of the strongest arguments in favour of team teaching.

LSP TESTING

When we come to assessing LSP students in any formal or semi-formal way, we must assume that the test-setter who does not belong to the academic culture of the students will interpret the text differently from the students and from professionals in that field. The theory tells us that this will be the case; so does common sense. As long as 'comprehension' could be defined as 'information linguistically committed to the text', this was not a problem. Once it is accepted that part of an interpretation
is dependent on the student's background knowledge, then it becomes a major problem.

Part of the problem is shown by the British Council's ELTS Study Skills Modules, one of the few examples of LSP tests. While the texts are specialised (their authenticity is in some cases doubtful) the questions remain very much the questions which might be asked by someone with a degree in Liberal Arts, in other words the typical EFL test setter. But we cannot be sure that these are the sort of question which a professional in the field would ask. In the series of reading experiments conducted by Alderson and Urquhart (1985) we sometimes used gap-filling tasks. The gaps were designed to test the students' ability to gather information from as wide an area in the text as possible, i.e. not just from the immediate environment. Even with such an apparently simple test instrument, we were struck by the doubt that possibly a professional in the field related to a particular text might choose to delete other items than the ones we did.

Zuck and Zuck (1984) have produced evidence that content specialists rate texts in their field in ways different from EFL teachers, and produce recognisably different types of questions. Thus, for example, content experts rated a text in their own specialist field as very difficult while EFL teachers rated it as comparatively simple. When constructing questions aimed at the text, the EFL teachers tended to produce items aimed at specific bits of information, while the content experts produced broader, more general questions. The overall moral is that one cannot guarantee the validity of an LSP test simply by making use of authentic texts; the tasks, too, must be authentic, and this means involving content experts.

The overall moral would seem to be 'Preserve flexibility'. We must be flexible in allowing many varieties of reading goal, and hence many kinds of comprehension. And we must be flexible enough to realise that our own interpretation of a text is only one among many, and may not be the most appropriate one for a particular purpose.

REFERENCES


Calfee, P.C. and R. Curley (1984) Structures of prose in content areas. In Flood (Ed.)


Sticht, T.G. (1984) Rate of comprehending by listening or reading. In J. Flood (Ed.)


