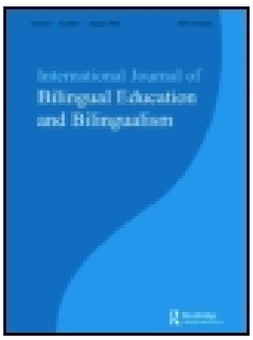

International Journal of Bilingual Education and Bilingualism

ISSN: 1367-0050 (Print) 1747-7522 (Online) Journal homepage: https://www.tandfonline.com/loi/rbeb20

Testing of Inferencing Behaviour in a Second Language

Jane A. Kembo

To cite this article: Jane A. Kembo (2001) Testing of Inferencing Behaviour in a Second Language, International Journal of Bilingual Education and Bilingualism, 4:2, 77-96, DOI: 10.1080/13670050108667720

To link to this article: https://doi.org/10.1080/13670050108667720

Testing of Inferencing Behaviour in a Second Language

Jane A. Kembo
Moi University, Faculty of Education, PO Box 3900, Eldoret, Kenya

The term 'inferencing' has been used in many texts and teaching books to mean a process or a discrete skill in reading and implies the process of gap-filling. Other texts call this 'pragmatic inferencing', meaning the incorporation of world knowledge into the meanings reconstructed during the processing of a text. This paper utilises the term after Winne *et al.* (1993) to mean everything a reader does in the process of reconstructing the meaning of a text. Our definition is synonymous with reading. Inferencing is a complex process and testing its products may never be accurate or even simple. The problems of testing SL inferencing may result from assumptions made by testers on the nature of reading, or test types to the presumptions and problems that readers bring into the testing situation. The study used 300 final year secondary school students who are SL speakers of English and administered two reading tests, one culturally familiar and the other culturally unfamiliar, based on three narrative texts per test. Four categories of inferences were tested in four different sections 'A' to 'D'. The results showed that certain inference types were more difficult to make. Even Short-Answer Questions presented peculiar problems. Readers did significantly better on culturally familiar texts than culturally unfamiliar texts. The ability to identify the locus of an answer was not an adequate requisite for arriving at an acceptable answer.

The main aim of the study was to examine the inferencing behaviour of final-year secondary school students who are second language learners of English on culturally familiar and unfamiliar texts. The sub-aims included: (1) determining how successfully readers could identify the locii of their responses and judge question difficulty, and (2) establishing whether the readers' ability to identify their response locii has any significant relationship with their overall score or scores in the two tests.

Introduction

This paper argues that there are many problems attendant to the testing of SL inferencing. It suggests that the results of a test are not all a consequence of the reader's ability or inability to understand the text, but may be a result of inadequate language facility, or writing ability (writing as opposed to talking where there are many things that are taken for granted; the topic is immediate; the context helps the listener to fill in many of the gaps) and/or the pressure of the testing situation itself as well as the presumptions made by the test-maker and the mismatch between the reconstructions of the test-maker and the test-taker (cf. Shohamy, 1985; Williams,1981). These factors which are only some of the elements may not allow the reader to translate his spontaneous reactions to a given text onto paper.

Inferencing in a second language presents peculiar problems because a plethora of factors influence the outcomes (meanings) that readers reconstruct from a given text. The meanings they are able to express are often not the whole mean-

1367-0050/01/02 0077-20 $16.00/0

ing that they take out of a reading, and the processes through which they arrive at these outcomes are themselves largely unobservable and inaccessible both to the reader and the tester of reading. Moreover, the processes are many and may vary from individual to individual. But first, what is inferencing?

Inferencing

Inferencing involves the mental process of recognising text, or integrating ideas from various parts of a text in order to reason out (Thorndyke, 1973–4) and construct an appropriate message from the propositional content of the text. To do this, the reader will sometimes use background knowledge. The term inference can also be used to designate the end product of the process, where it signifies a reader's reaction to a text or question item. In the latter case, inference will refer to an answer.

According to Oakhill and Garnham (1988) and Chikalanga (1991) a text is no more than a collection of words (lexical items) and paragraphs until a reader draws inferences. Inferencing is essential to textual comprehension (Farr *et al*., 1986; McIntosch, 1985), and Winne *et al*. (1993: 53) claim that 'it is the cornerstone of reading comprehension', with 'even the simplest type of literal comprehension' demanding 'that we engage in inferencing'. Furthermore, Weaver and Kintsch (1991: 235) claim that during the reading of narrative and expository texts comprehension is inference-driven as the reader may make 'as many as 12 to 15 implicit inferences for every expressly mentioned statement in the passage'. In this paper, therefore, the term inferencing is synonymous with reading.

As most of what is in a text is implicit, understanding implied meanings presupposes an understanding of explicit meanings and the ability to distinguish main ideas crucially depends on skills such as understanding relations between parts of a text (Matthews, 1990; Munby, 1978; Oakhill & Garnham, 1988). For example, if we consider the text below:

(1) Ashen-faced he clutched his chest.
(2) Thirty minutes later the stethoscope did not register a heartbeat.

Several inferences can be made from proposition (1):

(a) that 'he' is in shock – signalled by (ashen-faced);
(b) that 'he' is in pain – (clutched);
(c) that 'he' is probably having a heart attack;
(d) that 'he' refers to the same 'he', i.e. that he clutched 'his own chest'.

Sentence (2) suggests the following:

(a) that a nurse or doctor has been called in; or
(b) that 'he' has been rushed to a health institution signalled by (stethoscope);
(c) that either of the two actions above has been done by someone else;
(d) that some kind of examination has taken place – signalled by the presence of stethoscope;
(e) the time lapse between proposition (1) and (2) is thirty minutes (thirty minutes later)
(f) that the man is already dead – (the stethoscope did not register a heartbeat).

To arrive at some of these inferences we either try to put together what is explicitly stated in the text or use what we know of the world, diseases and processes to arrive at interpretations that are only signalled/implied by the text. For example, we are not told that 'he' is dead (f), but we arrive at this conclusion by doing the latter. However, to arrive at (e) we integrate the information in both (1) and (2). The reader need not go out of the text at all to get to that meaning.

Thus the ideal reader makes use of synctatic, semantic and discoursal clues in the texts, in addition to identifying and activating appropriate background knowledge which is brought to bear on the propositions of the text, which are structurally linked and divided into semantic units (Norman & Rummelhart, 1975; Rummelhart, 1977).

How the reader processes these 'interpretational' or 'meaning' propositions cannot be overtly observed because it all takes place in the brain. In a second language situation we may not be sure how this information is processed exactly; whether some readers, depending on their linguistic proficiency, which is key in reading comprehension (Bossers, 1991; Carrell, 1989; Eskey, 1988; Hacqueboard, 1989), do this processing in their L1 and somehow translate it to the L2, or whether they process text straight in the L2.

Inferencing and Testing

In the testing of inferencing (reading) the test-maker is, therefore, attempting to have the reader do two things: understand the propositions contained in the question, as well as use parts of his reconstructions of the text to suit the demands of the tasks. In any case, in the testing of reading comprehension attempts are made to make overt that which has taken place 'behind the eyeball', and what is elicited is always a small amount of what a reader understands from the whole. The testing of reading in a second language takes many forms (Multiple-Choice, Cloze, Gap Filling, etc.) and these forms make a difference (Shohamy, 1985). Below we discuss SAQs.

Short answer questions

Short-Answer Questions (SAQs) have been thought to be the best way of testing in-depth processing especially the testing of understanding that involves the integration of different propositions, and/or requiring the integration of appropriate world knowledge. Apart from limiting the chance of guessing (Weir, 1993, Weir & Roberts, 1984), SAQs 'force' students to interact with a text as they puzzle out the responses for themselves (Carrell, 1989). I think this is because SAQs do not give readers alternatives to choose from but an opportunity to find their own reactions to the text and task. They are further considered to limit problems that arise from language proficiency which prevent readers from expressing what they have understood. How successfully they do this is the question. Although Carrell *et al*. (1989 in Chikalanga, 1991; Hughes & Porter, 1983) point out that Short Answer questions, because of their open-endedness, directly reflect students' mental processing of textual information, better than, for example, Multiple Choice, the language of the questions is often a critical issue. In addition, whether a reader's response is limited to one word, a phrase or a short sentence, his language ability still plays a pivotal role in the success of that

response. Weir and Porter (1993), however, argue that SAQs avoid 'extended writing involved in task completion' which 'interfere with any inferences' that are likely to be made about a reader's processing ability. The example below will help in the subsequent discussion.

Text:

Johanna came to look for Maina yesterday. He said it was important that Maina rings his family. His wife is in hospital.

SAQs:

(1) Why did Johanna go to look for Maina?
(2) Why, do you think, Maina's wife was in hospital? Or
(3) What do you think was wrong with Maina's wife?

The response to question (1) demands that the reader uses what is in the text to come up with a reasoned-out answer such as:

Response: Maina needs to know about his wife's illness and hospitalisation.

If we consider questions (2) and (3), we are asking the reader to understand the following propositions from the utterance or sentences above:

(1) Maina is married.
(2) His wife has something wrong with her.
(3) She is undergoing/ receiving medical attention.

Meanings (2) and (3), are not directly in the passage and the reader has to make use of his world knowledge to reconstruct them from the propositions of the text.

To answer question (2) above the reader has to realise that there is no direct proposition in the text that responds to this task. He or she has also to put more than one proposition together in addition to using their world knowledge to come to a response. The reader's task is to choose the most appropriate response that will provide a reasonable response within the constraints of the textual information available. In answering this question the reader has to:

(1) Understand the individual lexical items (Ridgway, 1996).
(2) Understand the syntactical, semantic and discoursal clues which lend coherence to the text.

Beyond this the reader has to activate appropriate schemata against which to map the propositions of the text in order to come up with a reasonable inference. The reader has to realise that the question **WHY** is demanding a causal response. In other words, in this SAQ, the reader must not only fully comprehend the text, but must understand the task adequately to meet its demand. To further complicate this scenario, the reader must have adequate productive linguistic ability to express succinctly what is understood from the text.

It will be observed that in this situation three things obtain. The language ability of the reader plays a critical role in processing of text. Secondly, language ability dictates whether what is demanded by the comprehension task in understood. Thirdly, success in adequately answering the question depends on language proficiency, but may also depend on the reader's 'test-taking' skills (Amer, 1993; Rogers & Bateson, 1991) such as recognising where the answer to a

particular question is located and how to fit his understanding to the question task. In the following paragraphs we describe the study from which the thesis of this paper arises.

The Study

The subjects

Three hundred students were selected from the 4th (final year) streams of four National and Provincial secondary schools. In Kenya, National schools, based on perfomance at the end of the primary school examinations, admit students from all the eight provinces of the country. In other words, the National schools admit only the best performers from each province. There are about 25 National schools in total. The Provincial schools admit those who have not been absorbed by the National schools and they take 85% of their students from the same province, and 15% from others; Kenya has eight provinces in total. As a result of the feedback from the pilot study, it was decided that a motivated sample of only Provincial and National schools be used in the study. Both draw their pupils from the common pool; both are government funded and maintained. Students from the district schools were not sampled in the final study. The schools used in the study are summarised in Tables 1 and 2.

Table 1 School types used

Sex	*National*	*Provincial*	*Total*
Mixed	0	1	1
Girls	1	0	1
Boys	1	1	2
Total	2	2	4

Table 2 Sample of students used according to school type

School type	*Initial number*	*Wastage*	*Final number*
Provincial Day/mixed	84	08	76
Provincial Board/boys	72	12	60
National Boys	72	22	50
National Girls	137	23	114
Total	365	65	300

Each of the schools in Table 1 had four streams in the final year and two classes were randomly chosen for the study. All the students in the study varied in age from 16.9 to 17.5 years, and all had spent at least 11 years using English in school. All of the students had used English for at least 8 years as a medium of instruction and learning. In all cases, the students had an L1 other than English and a working knowledge of Kiswahili, which most of them had studied in school. From the initial 365 students in the study, 65 subjects were lost through lack of complete

tasks, absence from some of the tests, or simply not identifying themselves as required by the research.

Materials

Two inference tests, based on narrative texts, were administered. The first test was based on a set of three narrative texts which had been deemed to be culturally familiar, while the second test was based on passages judged unfamiliar. Cultural familiarity, following Kroeber and Cluckhorn (1963) and Pritchard (1990) was defined as a way of perceiving the world based on common patterns that are acceptable to a particular group (Brooks, 1975) to which an individual belongs. In this way, one can loosely talk about a Kenyan, British or even African perspective or culture. The familiar texts were what could be termed African familiar, while the second set of texts would have been more familiar to British students.

Familiar and unfamiliar text types

In the culturally familiar test, three texts were selected from seven passages extracted from the African Writers' Series by a panel of four people all qualified in teaching English as a second language (two of these were teachers of English as a second language in different parts of Africa). The panel further determined that all the chosen texts were adequately African in their orientation for the average African to be able to identify with since the passages treated their content from what could generally be called an African standpoint and the cultural contexts were African.

The second group of texts were judged to be culturally unfamiliar. They were texts that would have been culturally familiar to British students and had been selected from Form 3 class readers. From a group of six texts, three were chosen by the same panel. Both texts types were ranked on a scale of (1) to (5) for and selected on the basis of:

(1) Completeness – where this refers to an episode or episodes that seemed to form a complete story, having an identifiable beginning, a middle, and a resolution;
(2) Interest-appeal as a story to the target group;
(3) Potential to yield the number of questions required;
(4) Comprehensibility of the text to the target group.

In addition to choosing texts the panel also generated (based on descriptions of the inference types and sample questions given) at least five questions for each inference category. From discussions of these, a corpus of questions for each inference category was built. The questions were further tested through piloting with two Kenyan Form 4 classes from the target population. Adjustments were made following this in terms of clarifying questions and adding options to the marking scheme for questions that had more than one acceptable response. The pilot, using the Equivalence (parallel forms reliability) (Bachman, 1990) showed that the tests had a reliability of 0.82 and 0.68 respectively for the familiar and unfamiliar texts. Once the tests were revised on the information obtained in the pilot, they were pretested again using the same method. The reliability was then 0.92 and 0.82.

Narrative texts were chosen for their inherent ability to sustain interest and motivation (Benton, 1991). Narrative texts were also a genre that the Kenyan Form 4 student would be familiar with; thus their choice, for us, reduced the chance of a possible extraneous factor, that of unfamiliarity with the genre.

The difficulty of the texts in each test ranged from 'low' to 'high' with ranges in readability ratings from 121 to 163. This was thought ideal for catering for all reading abilities among the subjects (Aukerman, 1972; Klare, 1963). In addition, the study utilised teacher validation which gives as valid a measure as many test formulae. In fact, Davidson and Kantor (1982: 182) argue that the difficulty of texts is not just a function of their measured properties such as sentence length and vocabulary.

The tests

In each test there were two booklets, an answer booklet and a question booklet. The tests were prepared according to the taxonomy of infererences (Chikalanga, 1991).There were four sections 'A', 'B', 'C' and 'D', each testing a different type of inference category. 'A' in both tests examined pronominal inferences, 'B' – pragmatic inferences (or elaborative and explanatory inferences), 'C' textually explicit (Logical Informational) inferences and 'D' textually implicit (Logical Explanatory) inferences. Table 3 gives a summary of the inference categories tested.

Table 3 Inference types tested

Section of test	*Inference type*
A	Pronominal-personal pronouns-text based
B	Elaborative-Informational and Elaborative/Explanatory inferences Text and Knowledge based
C	Logical Informational-(textually explicit) text based
D	Logical Explanatory-(Textually implicit)-Text based

Inference types tested

Sections 'A' and 'C' tested pronominal and logical informational inferences respectively and were both textually explicit, while 'D', which tested textually implicit inferences demanded the integration of various propositions in the texts for the reader to come up with an answer. In other words, the answer was there, but the reader had to engage in some reasoning before coming up with an acceptable inference. Section 'B' tested elaborative informational and explanatory inferences which involved pragmatic inferencing; that is, the incorporation of world knowledge or a reader's schemata into what was read in order to arrive at a response. We give an example below to illustrate the differences between 'B' and 'D' in the tests that were administered.

Example text:

Maria was driving without lights when the police stopped her.

(1) Question: What does her refer to?
Response: Maria (Pronominal inference)
(2) Question: What kind of driver is Maria?
Response: A careless one. (Elaborative Informational)
(3) Question: What do you think the policeman said to her?
Response: The response will vary from reader to reader but the reader must draw from knowledge other than what is in the text. (Elaborative Explanatory)
(4) Question: What should Maria have done?
Response: Switched on her lights. (Logical Explanatory – Textually implicit)

Response (3) is not textually explicit; but it is in the text. The reader needs to incorporate world knowledge into the text to come up with the response. The last example is a response that requires the reader to integrate what is in the text. The reader need not go out of the text at all. It should be mentioned that the difference is not always this clear-cut between Logical/explanatory inferences and Elaborative/informational and elaborative/explanatory inferences, but in testing these inference types, one of the things that was done by the panel of four was to discuss the questions set and to modify or eliminate those that did not clearly belong in one category.

Each inference section of each test consisted of 20 questions per test. Thus each test had a total of 80 inference questions from the four sections. The readers were given as much time as they needed to attempt each item on the two tests. The two tests were administered within two weeks of each other in each school, with each test administered simultaneously to all testees in the same school.

Short-Answer Questions were used, with the questions themselves phrased in what was considered simple and accessible structure and language. Students' responses were expected to involve single words, short phrases or very short simple sentences at the most. We considered that in this way, any mis-inference on the part of the reader would be a result of poor reading rather than a consequence of language deficiency. The SAQ format was deemed the best method for 'forcing' the reader to give their own reconstructions of the text.

The third part of the task involved students in making a personal jugdement of question difficulty by section or inference category. Each reader was asked to place a tick in a box provided at the end of each section of the test indicating the difficulty of that section. The difficulty index ranged from: 'very easy' to 'very difficult'.

Very easy	Easy	Between easy and difficult	Difficult	Very difficult

The second part of the task involved identifying the sources of their inferences. As has been indicated in a previous paragraph, there were basically three sources of responses or reconstructions: the text, the text and general knowledge, and general knowledge alone designated '1', '2' and '3' respectively. At the end of each response brackets were provided for the reader to put in the response source number ('1', or '2' or '3') e.g.

Example: Who was the culprit in the story?

_____ [ ].

So the student would respond to the question and try to determine the source of the answer which he would then write in the brackets provided as shown above.

Section 'A', the pronominal inference responses, could be lifted from the text. The reader was required to identify the antecedent that corresponded to the pronominal in question. Section 'C' which was also textually explicit, demanded that the reader identify the proposition that corresponded to the question cue and lift it in response. Section 'D' also tested textual inferences, but these were implicit and required some reasoning with the text and understanding more than one proposition. Thus it involved text integration. Section 'B' contained what would generally be termed pragmatic questions which demanded the comprehension of textual information as well as the incorporation of world knowledge to reach an acceptable response.

The language test involved 50 vocabulary and 50 grammar items using ELBA (1974) to gauge reader language ability. Language, especially SL2 vocabulary (Carrell, 1989; Curtis, 1987; Eskey, 1988; Swaffar, 1988) has been found to be a major determinant of SL reading (Alderson & Urquhart, 1984; Devine, 1987; Grabe, 1991; Laufer & Sim, 1982).

Results

(1) We began this discussion by stating that culture is a way of viewing the world that is collective and is conditioned. To test the hypothesis of no difference in performance between the culturally familiar and unfamiliar inference texts the study run a *t* test on the means. The test returned a *t* May 19:1 statistic of 13.31 (df = 299; $p < 0.000$) against a critical value of 1.980 at *p* May 19:1 0.05 level. There was a significant difference in the performance of the two inference tests in favour of the familiar. The fact that students did significantly better in the culturally familiar texts strengthens the position of other researchers (Halazs, 1991; Malik, 1990; Steffenson & Joag-Dev, 1984, among others) that culturally familiar information is easier to relate to and comprehend than culturally unfamiliar information. The correlations, for example, of answer location scores and inference scores in the pragmatic and the textually implicit responses in 'B' and 'D' illustrate this point adequately; they were higher for the familiar texts. They underscore the fact that L2 reading and the reconstructions that a reader makes of the propositions of a text, among other factors, are culturally dependent.

(2) What was also evident was that not all cultural aspects were equally familiar to a whole group of people. Although one can talk about an African culture, the specific details of tradition, practices, beliefs differ from place to place. Thus, it is possible that a reader from the same culture may still possess no relevant knowledge with respect to particular aspects of the culture. Similarly, within mainstream cultures, for instance in Kenya, there are particular sub-cultures created by ethnicity, class, education or even employment. Hence, when it comes to questions touching on specifics, readers may not have appropriate schemata to bring to bear on the text.

(3) The results of the inference tests showed that there were close correlations

between both vocabulary and inferencing, and grammar and inferencing. The Pearson Product Moment returned a r of 0.538 and 0.562 for vocabulary and inferencing in the familiar and unfamiliar tests respectively, and an r value of 0.460 and 0.500 for grammar and the two tests. Both values were tested at the 95% confidence Interval (N = 300) and were significant. Total language (vocabulary and grammar) correlations stood at 0.617 and 0.581 for the familiar and unfamiliar tests respectively. The study fitted Anova GLM with all-two-way using backward and forward selection and examined the relationships between inferencing and vocabulary, grammar, sex and school type. By checking the assumptions associated with GLM (e.g. normalcy of distribution, randomness of distribution along the line of perfect fit), the regression tests showed that vocabulary on its own contributed 52.5% ($p = 0.000, p < 0.05$)) of the reading results. Grammar on its own contributed 35% ($p = 0.000$, $p < 0.05$)) of the reading results in these tests. Vocabulary and sex together contributed only 5.90% ($p = 0.0016$) of the results and sex only less than 1% ($p = 0.352$) Students who had a strong lexical base in English did generally better on the inference tests than those who had a weaker lexical base.

(4) Although the language of the questions was judged to be accessible to the students, and the test format (SAQs) familiar, it was evident that in the testing of reading the language ability of the students plays a pivotal role both in the comprehension of texts and in the ability to adjust textual information appropriately to meet the demands of inference tasks. Students who had poor language ability were hard pressed to write their responses succinctly enough to gain the full marks even though the grammar of the language was not part of what was being examined and no mark was awarded for language.

(5) More than 60% of the readers were able to identify their response sources for all categories of the inference tests. The category whose response locus was most successfully identified was the pronominal (section 'A') and the textually explicit (section 'C') inferences. It is interesting to note that even for these textually explicit questions, although some 36% of the readers successfully identified the response locii, they were not always able to identify the appropriate responses in 42% of the cases, and therefore, the correlations between answer locii and inference scores were non-significant.

There were significant relationships between answer location and inference for the textually explicit inferences ($r = 0.503$ and 0.528) in 'C' sections for the familiar and unfamiliar tests respectively. Of interest is the lack of correlations between the pronominal answers and the pupils' abilities to identify the locii of the responses given, because most of those who successfully identified the answer locations did not always get the right responses. The null hypotheses were also accepted for relationships between answers and pragmatic (Section 'B') and logical explanatory (textually explicit or section 'D') response locii. The correlation between answers and location scores was 1 point above the critical level (0.365 against 0.364) for the Unfamiliar texts.

The total Answer-Location scores and correct responses were also significant at 0.491 and 0.477 for the familiar and unfamiliar tests respectively. The lowest

Table 4 Relationship between answer location and inferencing scores (Familiar test)

Test section	*Correlation (r) values*
A	0.263
B	0.286
C	0.503
D	0.328

Table 5 Relationship between answer location and inferencing scores (Unfamiliar test)

Test section	*Correlation (r) values*
A	0.365
B	0.245
C	0.528
D	0.249
Totals	0.477

responses were returned for the pragmatic sections ('B') of the tests as shown in Tables 4 and 5.

Although overall it is true that many of those who were able to identify the sources of their responses made acceptable inferences, it was evident that the ability to identify or guess the locus of the response was not a sufficient ground for arriving at an appropriate response to meet the demands of the inference task in question.

(6) Readers found it more difficult to make pragmatic inferences than other types of inferences mentioned above. Surprisingly, though, there were readers (31%) who proved very efficient at making pragmatic inferences, but were poor at arriving at appropriate textual inferences (cf. Chikalanga, 1992).

(7) Question: 1. What is the religion of the people in the story? Give a reason. (See text 3 – Test 1)
Responses:
(a) Some are traditionalists, others are non-religious e.g. gossipers.
(b) Hindu – 'they speak the language'.

Question 2: Why did Blackie put his arms around Emma's shoulders at the end of the passage?
Response:
(a) He had negative attitudes.

(8) Responses such as the examples above often made it difficult for the tester to decide whether comprehension of both (i) the narrative texts and (ii) the inference questions had occurred, or whether the reader had been able to understand the one, and not comprehend the other. Alternatively it was not always possible to decide whether, if the student had been able to under-

stand both the texts and the questions, had then been unable to express that understanding because of deficient language. Thus, in the testing of reading in a second language, proficiency in the language plays a major role in how successfully readers express their understanding of both texts and tasks. In our case the readers were not asked to respond in their L1s because of many reasons: one, there were 40 languages represented in the sample and using L1 would have brought about enormous marking problems; two, the readers had used English at least for eight years as a medium for learning and were soon going to sit for their final examinations in the language. It was presumed that they were adequately proficient in the target language to be able to operate in the tests given. Three reading in one language and responding in another is likely to introduce confounding factors in the study and this we avoided.

(9) Reader jugdement of question difficulty returned results that demonstrated on the whole, that the pronominal inferences were the easiest, the logical/informational inferences in section 'C' followed; with the elaborative informational and explanatory inferences in section 'B'. The most difficult inference type for the readers was the section 'D' of both tests which involved text integration. It can be observed below that there was considerable variability in reader judgement of question difficulty. Furthermore, it can be seen from Table 6 that the categories with more variability were the textually based inferences, 'A', 'C' and 'D'.

Table 6 Reader judgement of difficulty

Section	*Mean (Fam.)*	*Unfamiliar*	*Sd (Fam.)*	*Sd (Unfam.)*
A	2.587	2.551	0.689	0.781
B	3.322	3.281	0.608	0.629
C	3.149	3.234	0.614	0.720
D	3.410	3.590	0.662	0.738

Discussion

The testing of reading in a second language, therefore, seems to be influenced by many factors, as the results above show. These problems range from assumptions of testers about testees and reading in general, the restrictions testing formats place on both testees and what can be successfully tested, the language of both the cues, the texts and the readers, and the conclusions that are made when the testees' responses have been scored.

The testing of reading, especially in a second language, assumes, firstly, that what is tested represents what the reader has understood from the text. As a reader processes a text he understands individual words (Ridgway, 1994) and these are rapidly formed into sentences and meaning units. From these, discoursal constraints are used to 'knit' the individual meanings of the text into a whole and so form a global picture of the message. It may be, however, that the sum of the individual understandings is not always equal to the whole, nor that

the whole is easily or successfully divisible into discrete parts. Statman (1998) argues that readings of a text by the test-maker, test-taker and test-marker (scorer) tend to result in different meanings and distortions of the original text by the questions asked leading to multiple interpretations.

Secondly, reading tests seem to presume that what is understood globally can be isolated in bits and pieces and expressed. These will vary from one reader to the next, depending on such factors as pre-set reading patterns, sensitivity to rhetorical structure, to connectives, and other rhetorical devices (Statman, 1998: 196). The details that test-takers end up remembering are at times a product of their interest (Wade *et al*., 1993). This can result in readers paying attention to irrelevant detail which to them become important. Besides, students are not always balanced between efferent and efficiency reading. While the test-maker may read with an eye on information and ideas, the test-taker may be influenced by aesthetic reading (Statman, 1998: 197).

Thirdly, what is understood can be moulded to fit the demands of question cues formed from the tester's own understanding and reconstruction of the text. It should be borne in mind that this understanding, in the first instance, arises among other things, from the background knowledge of the test-maker. While the test-maker may have one interpretation, test-takers may have other interpretations some of which may only be partially what the test-maker was aiming at eliciting. Some of the responses therefore may only be partially acceptable, therefore because, according to Fransson (1984), they reveal different levels of understanding. In the Unfamiliar Test, an account is given of a house whose brooding presence and outward appearance frightens people at night. Readers were asked the following question in section 'D' 12:

What feeling does the house create in people at night?

The following responses illustrate the point above.

(1) It creates a feeling of being frightened [*sic*].
(2) The house creates a frightening feeling inside people.
*(3) That is was a cemetery.
*(4) It creates panic and fear since most people fear the dead people in the graveyard.
*(5) The mansion seemed hostile and strangely brooding.

The responses with the asterisk do not meet the demands of the question, but they all reveal partial understanding of the task in the question and a failure to meet that demand. It can be argued that responses (3) and (5) have not met the final step of fitting information to the question. Response (4) is one in which the test-taker both meets the demand of the question up to 'fear', but goes on to explain; this is something not asked for in the question. It seems that a point that is very clear to a test-maker may be missed by testees or only partially understood or recognised.

If one argues following Urquhart (1996) that there are as many interpretations to a text as there are readers, then one would perhaps be able to go further and say that in a second language situation these interpretations are further complicated by factors that we discuss below. From the foregoing results, it is possible to see that reading in a second language depends a great deal on the language profi-

ciency of the reader. A reader, thus, may understand a text, understand the question and be able to express both that which he has understood from the text and the question; or he may not be able to make 'a perfect fit' between what he understands of the text and what the question demands. Another student may be able to 'understand a text' and not understand the questions set on it, or not know where to look for the propositions that meet the demands of a particular question or questions. The fourth reader type may not be able to express what he has understood because he lacks language ability that would enable this.

Secondly, it can be argued that in reading each reader brings to the text a unique set of circumstances and these will mediate the meanings that he takes out (Benhardt, 1995; Benson, 1988; Devine, 1988; Oded & Stavans, 1996; Steffenson *et al.*, 1979). Testing what is read apparently 'forces' processors to think of meanings that they themselves may not have got from the text or focused on. Spolsky (1994) argues that between the test-maker, and test-scorer there are likely to be three more texts created:

Original writer	Text 1
Test-maker	Text 2
Test-taker	Text 3
Test-scorer	Text 4

with each student creating a different text from the original. Open-ended questions lend themselves to a plethora of answers (Statman, 1998: 201) all of which can be justified. Even in MC questions, the mental representations that readers have at the initial reading may be different from the second. This is because the MC questions act as information source which may not have been in the first representation (Gordon & Hanauer, 1994) especially in cases where it has been observed that readers survey questions before reading texts given for comprehension texts. When this happens, selective processing results.

In a test situation the result is divergent responses to a text cued by the test-maker's interpretations. This begs the question of fair representation of what can be called reading ability. This is not to argue that the parameters of an acceptable response are not set by textual propositions. They are; but between what is appropriate and what is not, there are a lot of possibilities.

From the foregoing results, it is possible to see that reading in a second language depends a great deal on the language proficiency of the reader (Alderson & Urquhart, 1984). On the one hand, a reader may understand a text, understand the question and be able to express both that which he has understood from the text and the latter. On the other hand, he may not be able to make 'a perfect fit' between what he understands of the text and what the question demands. Another student may be able to 'understand a text' and not understand the questions set on it, or not know where to look for the propositions that meet the demands of a particular question or questions. The fourth reader type may not be able to express what he has understood because he lacks language ability that should enable this. It seems that two arguments follow from this. One, that normal reading is hardly ever done for the sake of answering questions which themselves are a reflection of another's reconstruction of a given text. People read selectively for what they want to get out of a text. It seems that the testing of reading introduces constraints making readers do so for the sake of the tester. I want to propose that this psychological pressure forces, especially the

'medium to poor' readers, into perhaps poorer comprehension of what is placed before them. Two, from the foregoing results, it is apparent that reconstruction of textual meaning, and the ability to fit what is understood to the demands of a question, are different skills that require training. Understanding of a text will be varied because readers notice different things about any given text, and bring differential schemata to the reading tasks (Arner & Khousam, 1991; Attariba & Forsythe, 1993). This alone, apart from their language proficiency, will mean there are perhaps as many subtle reconstructions of a text as there are readers. Are all of these valid within the constraints of the text? (cf. Anderson, 1991).

In a second language it seems that expressing that which forms our understanding is also a skill that learners must be trained in. Understanding is covert, but there are many times in a second language when speakers and learners say, 'I know what I mean but I cannot quite express it'. It seems that to be able to express what they understand readers must be trained to bring to the surface what is 'behind their eyeballs'. This skill calls not only for a wide vocabulary, which most second language learners lack at this level, but also the ability to make appropriate selections of lexical items which enable the making of the most appropriate responses, especially in short answer questions such as the ones that were used in this study. SAQs have been thought to solve many of the problems inherent in testing reading using other means such as long and open-ended answers. This may not be true. Whereas a reader might meander in open-ended questions and perhaps get to what they wanted to say, SAQs do not allow for this; they call for a succinctness that many students of English as a second language find difficult. Key in this lack is the limited vocabulary for the majority of the readers. This makes adequate language proficiency not just crucial for comprehension (Anderson & Freebody, 1979; Ayodele, 1984; Bossers, 1991; Brisbois, 1992; Clarke, 1980) but also pivotal in the production of responses that meet the demands of set tasks (Benito *et al*., 1993; Bryant & Bradley, 1985; Clarke, 1979).

The alternative may be to test readers in their L1. But this is not always practical. In this case, readers in this exercise were not asked to express understanding in their L1 because: one, they were a mixed linguistic group roughly representing 40 Kenyan languages. This would have presented enormous problems of scoring. Secondly, the students in the study had all been studying and learning in English for at least 11 and 8 years of their school lives respectively. They were expected to sit for their 'O' Level examinations in another 6 months and to do this in English. In a number of cases the readers have been noticed to have undergone subtractive bilingualism. Thus they are not fluent in either of their two languages. In a number of cases, they are more fluent in informal Kiswahili, which is the language of wider communication in Kenya.

From the results of this study, there is ample evidence that the ability to understand questions and what they demand from the reader is a peculiar skill; Amer (1993) calls this 'testwiseness'. The need for this skill and training in it are further supported by Raphael and Pearson (1982), Raphael (1982) and Benito *et al*. (1993). Training in Question-Answer Relationships is important in improving both comprehension and response skills. Benito *et al*.'s observation that in many classrooms teachers ask questions but rarely do anything with the students' responses except acknowledge their 'correctness' or otherwise (1993: 21) seems to be further supported by Durkin (1978) who posits that although teachers

expect comprehension it is hardly ever taught, while Bolagun (cited in Onukaogu, 1987) states that in many classrooms reading is either poorly handled or ignored because teachers themselves are poorly trained in the handling of this central skill. A group of 52 teachers with whom the researcher was conducting an inservice course in the teaching of reading admitted that they were often not sure how to teach reading. More than 46 of them assumed that because reading is an activity that each learner does for himself it grows naturally once students learn to decode, and hence it need not be taught.

The differential performance that was observed in the readers indicates that the different types of inferences demand different mental processes and effort (cf. Lumley, 1993). It is possible that there are different reading skills required for meeting the demands of the different inference tasks within the reading continuum. This was shown, for instance, by the fact that the pragmatic inferences and the questions which required readers to integrate more than one proposition from the text were the most difficult. This may also have been a result of language which could facilitate the 'seeing' of the connections in the propositions in the latter case, or the false belief on the part of some of the readers that all meaning resides in the text (Carrell, 1988). The differential cognitive demands may explain why some students could do well in pragmatic questions and not as well on textually based tasks, or vice versa.

Readers were able to locate sources of their responses but not able to identify acceptable answers even when these were textually explicit. This shows in the low correlation between the location scores and the inference scores in section 'A', the pronominal responses. It is possible that although being able to keep a focus on the references in a text is crucial to understanding it, the two skills are simply separate. Alternatively, it may be that because many of the answers to each section of the question paper were the same, readers responded by guesswork without giving their responses too much thought. Furthermore, it could even be that this is a poorly developed skill among the readers in this study, or that identification of an answer, even though when trained should improve reader ability to identify appropriate responses, is not in itself an adequate condition for making acceptable inferences. The relationship between logical explanatory (implicit) inferences and answer location scores seem to point to the fact that when texts are unfamiliar to readers they may need to rely more on the text hence be more text-bound.

Language ability has been found in several studies to be a major determinant of successful reading (Alderson & Urquhart, 1984; Benhardt & Kamil, 1995; Bossers, 1991; Brisbois, 1992; Hacquebord, 1989). Below a certain threshold of language proficiency (Clarke, 1980; Cummins, 1979) comprehension processes are not used effectively. When there is a language deficiency, readers are not able to make use of syntactic, contextual, semantic and discoursal clues (Cziko, 1978; Devine, 1984; Haynes, 1993) to arrive at appropriate responses. Besides these language disadvantaged readers are forced to rely more on bottom-up processes which prevent them from identifying and accessing appropriate schemata to map the text against.

Finally, one factor that is often taken for granted is that of time. In testing reading test-makers often envisage model or near-model answers on texts that they themselves have spent a lot of time analysing and even discussing (as was the

case in these inference tests) and it is often the pressure of time which forces readers into 'looking' for what they consider to be appropriate responses instead of aiming at comprehension (in its widest sense) thus engaging in selective processing of a given text. When this happens, is it not possible that part(s) of the text may not be processed at all?

Conclusion

In conclusion, the testing of reading in a second language is complicated because there is no one test type that is able to get at exactly what a reader understands. Secondly, what a reader understands may not always be expressed succinctly enough to meet the requirements of the tasks set. This, however, does not mean that readers have not 'understood' the text. It may even be that what a reader understood from a text was not elicited by the questions set. Each test type seems to call for familiarity with it, lack of which may impact on reader performance. Reading test scores may reveal perhaps only reading ability in a particular genre or rhetorical organisation and the particular questions set, but may not always reveal with certainty the global reading capability of a particular reader. The very idea of testing seems to place psychological pressures on readers that may prevent them from accessing meaning. To add to this complex scenario is the fact that to date, although more is now understood about reading, it is not enough for comprehending the process because most of what goes on in reading occurs unobserved and what we try to assess as the ability is perhaps only a very small part of that process. We assess what we call the product, which is mediated by, among other factors: language, questions, and question types, personal reading styles, reading concepts and testwiseness. Is it even possible to test all that one understands from what they read through processes that are not entirely accessible to observation or conscious on the part of the readers? While these questions and others remain, the option seems to be for test-makers to resort to multi-pronged tests such as Mcs, with SAQs as well as protocols where these can be applied together. However, this is not always possible and the problems remain: that of finding the best way(s) to gauge the complex 'skill' that is reading ability, an ability whose workings we have not understood fully (Manguel, 1996).

Correspondence

Any correspondence should be directed to Dr Jane A. Kembo, C/O Aga Khan Education Service, Tanzania, P O Box 125, Dar Es Salaam, Tanzania (pdcea@cats-nets.com).

References

Alderson, J.C. and Urquhart, A.H. (1984) *Reading in a Foreign Language*. London: Longman.

Altariba, J. and Forsythe, W.J. (1993) The role of cultural schemata in reading. In *Cognition and Culture: A Cross-cultural Approach to Psychology* (pp. 145–155). New Jersey: Erlbaum.

Amer, A.A. and Khouzam, N. (1991) The effect of EFL students' reading styles on their reading comprehension performance. *Reading in a Foreign Language* 10 (1), 967–979.

Amer, A.A. (1993) Teaching EFL students to use test-taking strategy. *Language Testing* 10 (1), 71–77.

Anderson, N. (1991) Individual differences in second language reading strategies. *Modern Language Journal* 75, 460–472.
Anderson, R.C. and Freebody, P. (1979) Vocabulary knowledge. *Technical Report No. 136.* Urbana: University of Illinois, Center for the Study of Reading.
Ayodele, S.O. (1984) Reading characteristics of learners from two widely differing backgrounds. *Reading* 18 (2), 115–127.
Aukerman, R.C. (1972) *Reading in the Secondary School Classroom.* London: McGraw Hill.
Bachman, L.F. (1990) *Fundamental Considerations in Language Testing.* Oxford: Oxford University Press.
Bernhardt, E.B. and Kamil, M.L. (1995) Interpreting relationships between L1 and L2 reading: Consolidating the linguistic threshold and the linguistic interdependence hypothesis. *Applied Linguistics* 16 (1), 15–34.
Benito, Y.M., Foley, C.L., Lewis, C.D. and Prescott, P. (1993) The effect of instruction in question-answer relationships and metacognition on social studies comprehension. *Journal of Research in Reading* 16 (1), 20–29.
Benson, J. (1988) The psychometric and cognitive aspects of test-wiseness: A review of the literature. In M.H. Kean (eds) *Test-Wiseness.* Bloomington, IN: Phi Delta Kapan.
Benton, M. (1991) Reader response: Theory and practice. *Proceedings of The Role of Language and Literature in the School Curriculum* Seminar, 25–27 February, Moi University. Nairobi, Kenya: British Council.
Bolagun, I. (1987) Reading in English as a second language. *Journal of English as a Second Language* 1 (4), 7–12.
Bossers, B. (1991) On thresholds, ceilings and short-circuits: The relation between L1 reading, L2 reading and L2 knowledge. *AILA Review 8, Reading in Two Languages.*
Brisbois, J. (1992) Do first language writing and second language reading equal second language reading comprehension? An assessment dilemma. Unpublished PhD Dissertation, The Ohio State University.
Brooks, N. (1975) The analysis of language and familiar cultures. In Lafayette (ed.) *The Cultural Revolution in Foreign Language Teaching* (pp. 19–31). Skokie, IL: National Textbook Company.
Bryant, P. and Bradley, L. (1985) *Children's Reading Problems – Psychology & Education.* Basil Blackwell.
Carrell, P. (1988) Interactive text processing: Implications for ESL/second language classrooms. In P. Carrell, J. Devine and D.E. Eskey (eds) *Interactive Approaches to Second Language Reading.* Cambridge: Cambridge University Press.
Carrell, P.L. (1989) SLA and classroom instruction: Reading. *Annual Review of Applied Linguistics* 1988 9, 233–242.
Chikalanga, I. (1991) Inferencing in reading: Across-cultural study. Unpublished PhD thesis, University of Reading.
Chikalanga, I. (1992) A suggested taxonomy of inferences for the reading teacher. *Reading in a Foreign Language* 8 (2), 697–709.
Chikalanga, I.W. (1993) Exploring the reading ability of ESL readers. *Reading in a Foreign Language* 10 (1), 931–966.
Clarke, M.A. (1980) The short-circuit hypothesis in ESL reading – or when language interferes with reading performance. *Modern Language Journal* 64, 203–209.
Clarke, M.A. (1979) Reading in Spanish and English: Evidence from adult ESL students. *Language Learning* 29, 121–150.
Cummins, J. (1979) Linguistic interdependence and the educational development of bilingual children. *Review of Educational Research* 49, 222–251.
Curtis, M.E. (1987) Vocabulary testing and vocabulary instruction. In M.G. McKeown and M.E. Curtis (eds) *The Nature of Vocabulary Acquisition* (pp. 37–51). Hillsdale, NJ: Erlbaum.
Davidson, A. and Kantor, R.N. (1982) On the failure of readability formulas to define readable texts: A case study from adaptations. *Reading Research Quarterly* 17 (2), 187–209.
Devine, J. (1987) General language competence and adult second language reading. In J.

Devine, P.L. Carrell and D.E. Eskey (eds) *Research in Reading in English as a Second Language*. Washington, DC: TESOL.
Devine, J. (1988) A case study of two readers: Models of reading and reading performance. In P.L. Carrell, J. Devine and D.E. Eskey (eds) *Research in Reading in English as a Second Language*. Washington, DC: TESOL.
Durkin, D. (1978–79) What classroom observations reveal about reading comprehension instruction. *Reading Research Quarterly* 14 (4), 482–528.
ELBA (1974) *English Language Battery*. London: Oxford University Press.
Eskey, D.E. (1988) Holding in the bottom: An interactive approach to the language problems of second language readers. In P.L. Carrell, J. Devine and D.E. Eskey (eds) *Research in Reading in English as a Second Language*. Washington, DC: TESOL.
Farr, R., Carrey, R. and Tone, B. (1986) *Recent Theory and Research into the Reading Process: Implications for Reading Assessment*.
Fransson, A. (1984) Cramming or understanding? Effects of intrinsic and extrinsic motivation on approach to learning and test performance. In C.J. Alderson and A.H. Urquhart (eds) *Reading in a Foreign Language*. London: Longman.
Garner, R. (1987) *Metacognition and Reading Comprehension*. Norwood, NJ: Ablex.
Goodman, K.S. (1967) *Reading: A psycholinguistic guessing game. Journal of the Reading Specialist* 6 (1), 126–133.
Gordon, C. and Hanauer, D. (1994) The effect of multiple choice questions on mental model construction in EFL reading tests. Paper presented at the Academic Committee for Research into Language Testing, Kiryat Anavim, Israel.
Grabe, W. (1991) Current developments in second language reading research. *TESOL Quarterly* 25 (3), 375–406.
Hacquebord, H. (1989) Reading comprehension of Turkish and Dutch students attending secondary schools. Doctoral Dissertation, University of Groningen.
Halazs, L (1991) Understanding short stories: An American-Hungarian cross cultural study. *Empirical Studies of the Arts* 9, 143–163.
Haynes, M. (1993) Patterns and perils of guessing in second language reading. In Huckin *et al*. (eds) *The Nature of Vocabulary Acquisition* (pp. 89–106). Hillsdale, NJ: Erlbaum.
Hughes, A. (1989) *Testing for Language Teachers*. Cambridge: Cambridge University Press.
Hughes, A. (1993) Testing the ability to infer when reading a second or foreign language. *Journal of English as a Foreign Language* 10 & 11, 13–20.
Hughes, A. and Porter, D. (1983) *Current Developments in Language Testing*. London: Academic Press.
Klare, G.R. (1963) *The Measurement of Readability*. Iowa: Iowa State University Press, Ames.
Kroeber, A.L. and Kluckhorn, C. (1963) *Culture: A Critical Review of Concepts and Definitions*. New York: Vintage.
Laufer, B. and Sim, D.D. (1982) Does the EFL reader need reading strategies more than language? Some experimental evidence. Paper presented at the Annual Meeting of the American Council on the Teaching of Foreign Languages, New York. Alexandria, VA: ERIC No. ED 228 848.
Lumley, T. (1993) The notion of subskills in reading comprehension tests: An EAP example. *Language Testing* 10 (3), 212–232.
Manguel, A. (1996) *A History of Reading*. London: Harper Collins.
Malik, A.A. (1990) A psycholinguistic analysis of the reading behaviour of EFL proficient readers using culturally familiar and culturally non-familiar expository texts. *American Educational Research Journal* 27, 205 –223.
Matthews, M. (1990) Skill taxonomies and the problems for the testing of reading. *Reading in a Foreign Language* 7 (1), 511–517.
McIntosch, E. (1985) What do practitioners need to know about current inference research. *The Reading Teacher* 38 (8), 755–760.
Munby, J. (1978) A problem-solving approach to the development of reading comprehension skills. Presentation at the University Teachers of English in Israel (UTELI) Regional Meeting, Jerusalem
Norman, D.A. and Rummelhart, D.E (1975) *Explorations in Cognition*. San Francisco: Freeman.

Oakhill, J. and Garnham, A. (1988) *Becoming a Skilled Reader*. Oxford: Basil Blackwell.
Oded, B. and Stavans, A. (1996) The effect of 'false' schema activation on the construction of meaning. *System* 22 (4), 497– 507.
Onukaogu, X. (1987) Reading in secondary school. *Jessel. Journal of English as a Second Language* 27, 23–42. Lagos.
Pritchard, R. (1990) The effects of cultural schemata on reading processing strategies. *Reading Research Quarterly* 15, 273–295.
Raphael, T.E. (1982) Question-answering strategies for children. *The Reading Teacher* 36, 186–190.
Raphael, T. and Pearson, D.P. (1982) The effect of metacognitive awareness training on children's question answering behaviour. *Technical Report, 238*. Urbana, IL: Centre for the Study of Reading, University of Illinois.
Rawl, E.H. (1984) Test taking strategies can be the key to improving test scores. *NASSP Bulletin* 68 (468), 8–112.
Ridgway, T. (1994) Theory and foreign language reading comprehension. *Reading in a Foreign Language* 10 (2), 55–76.
Ridgway, T. (1996) Thresholds of the background knowledge effect in foreign language reading. (Unpublished article).
Rogers, W.T. and Bateson, D.J. (1991) Verification of a model of test-taking behaviour in high school seniors. *Journal of Experimental Education* 59, 331–49.
Rummelhart, D.E. (1977) Toward an interactive model of reading. In S. Dornic (ed.) *Attention and Performance* V1 . Hillsdale, NJ: Erlbaum.
Shohamy, E. (1985) *A Practical Handbook in Language Testing for the SL Teacher*. Shaked, Ramat, Aviv, Israel.
Spolsky, B. (1994) Comprehension testing or can understanding be measured? In G. Brown, A. Politt and J. Williams (eds) *Language and Understanding* (pp. 139–153). Oxford: Oxford University Press.
Statman, S. (1992) Convergent tests and the divergent student. *System* 20 (2), 211–221.
Statman, S. (1998) Tester and testee: Two sides of different coins. *System* 26 (2), 195–204.
Steffenson, M.S. and Joag-Dev, C. (1984) Cultural knowledge and reading. In J.C. Alderson and A.H. Urquhart (eds) *Reading in A Foreign Language*. Singapore: Longman.
Steffenson, M.S., Joag-Dev, C. and Anderson, R.C. (1979) A cross-cultural perspective on reading comprehension. *Reading Research Quarterly* 15 (1), 10–29.
Swaffar, J. (1988) Readers, texts and second languages: The interactive processes. *Modern Language Journal* 72, 123–149.
Thorndike, R.I. (1973–4) Reading as reasoning. *Reading Research Quarterly* 3, 499–545.
Urquhart, A.H. (1996) Defining the good reader. In *To Read in a Foreign Language. Triangle 14*. Paris: Goethe-Institute/The British Council/Ens-Credif, Diffusion Didier-Erudition.
Wade, S., Schraw, E., Gregory, B., William, M. and Hayes, M.T. (1993) Seduction of the strategic reader: Effects of interest on strategies and recall. *Reading Research Quarterly* 28 (2), 125–145.
Weaver C.A. and Kintsch, W. (1991) Expository text. In R. Barr, M.L. Kamil and Mosenthal and P.D. Pearson (eds). *Handbook of Reading Research* (pp. 230–245) (Vol. 2). New York: Longman.
Winne, P.H., Graham, L. and Prock, L. (1993) A model of poor readers' text-based inferencing: Effects of explanatory feedback. *Reading Research Quarterly* 28 (1), 53–75.
Williams, D. (1981) Factors related to performance in reading English as a second language. *Language Learning* 31 (1), 31–50.
Weir, C. (1993) *Developing and Understanding Language Tests*. Hemel Hemstead: Prentice Hall.
Weir, J.C. and Porter, D. (1993) The multidivisible or unitary nature of reading: The language tester between Scylla and Chabrydis. *Reading in a Foreign Language* 10 (2), 1–21.
Weir J.C. and Roberts, J. (1994) *Evaluation in ELT*. Oxford: Blackwell.