How using dictionaries affects performance in multiple-choice EFL tests

Hilary Nesi University of Warwick and Paul Meara University College Swansea

The studies reported in this paper replicate and extend the paper by Bensoussan et al (RFL 2,2), which examined the relationship between students' performance in reading tests and their use of dictionaries during the tests. Bensoussan et al found that availability of dictionaries did not significantly affect students' scores on the tests or the time they took to complete. The present study confirms the first finding, but found that students who had access to dictionaries took significantly longer to complete the test. The authors attempt to account for their findings by reference to aspects of the experimental conditions, test item types, usability of particular dictionaries, and the users.

INTRODUCTION

The studies reported here are based on a paper by Bensoussan, Sim and Weiss (hereafter BS&W) (1984). BS&W had noted with concern the apparent conflict between examination administrators and language teachers. Examination administrators were reportedly against the use of dictionaries in the examination room because they thought that dictionaries might help the candidates too much, whilst also wasting precious working time. In contrast many EFL teachers wanted learners to be able to use dictionaries in examinations. This was seen as a test of the candidate's ability to function in a natural reading situation: under normal non-examination circumstances dictionaries would be accessible to the learner.

Like the examination administrators, BS&W assumed that learners would be both helped and hindered by dictionary use in examinations. On the positive side they expected that "the permitted use of monolingual and/or bilingual dictionaries would significantly raise examination scores"; on the other hand they also foresaw that "the use of dictionaries would significantly increase the time taken to complete a test". When they conducted a number of experiments to test their hypotheses, however, these assumptions were seriously called into question.

In their pilot study, 900 first year students answered multiple choice questions on ten reading passages. Half used monolingual dictionaries, half used no dictionary at all. In their Study 1 (N=91), Study 2 (N=670) and Study 3 (N=740) use of a monolingual or bilingual dictionary was optional during the reading tests. The scores in these tests were compared with the type of dictionary consulted and in Studies 2

Hilary NESI is a lecturer in the Centre for English Language Teacher Education at the University of Warwick. She holds an MSc in ESP from the University of Aston, and is currently involved in research on the productive use of learners' dictionaries.

Dr Paul MEARA was head of the department of Applied Linguistics at Birkbeck College London; he is now Director of research in the Centre for Applied Language Studies, University College of Swansea.

The authors may be contacted via Hilary Nesi, Centre for English Language Teacher Education, University of Warwick, Coventry CV4 7AL.
and 3 the length of time taken to complete the test was also noted. In Study 1 the type of dictionary consulted was also compared with the number of words students had originally stated that they intended to look up and the number of words they in fact looked up.

In all four studies no significant difference was found between the test scores of dictionary users and those who did not use dictionaries. There was also little correspondence between test scores and the time taken to finish the test, although students who did not use a dictionary tended to finish fastest, students who used bilingual dictionaries tended to be slowest and there was a slight tendency for slower students to obtain lower marks. This last finding led the authors to conclude that “students who choose to use bilingual dictionaries tend to be slower and weaker in taking reading comprehension tests in English and, by extension, in reading English texts”. The authors admitted, however, that this was speculative, and in the tests as a whole bilingual dictionary use was not found to have any bearing on results.

BS&W’s research has been presented as providing evidence that dictionary use makes little difference in reading tests: students were just as fast but no more proficient when they consulted a dictionary in the examination room. These results did not confirm the fears of examination administrators, but neither were they particularly helpful to the English teachers who advocated dictionary use in examinations. There does not seem to be much point in advising students to use dictionaries if using dictionaries does nothing to improve test scores. Moreover these puzzled teachers would have every reason to enquire why the test scores did not improve, if, as BS&W maintain, dictionary use is normally beneficial to readers working with the right level of text.

The data collected by BS&W provided convincing evidence against the two original hypotheses, but, perhaps because of the immense scale of the project, the researchers did not focus on those details necessary to answer the questions raised by the rather disturbing findings.

Dictionaries are designed to help readers read more efficiently, and reading tests are designed to test reading efficiency; in order to account for the surprising failure of dictionaries to improve reading test scores in this experiment, we need to know more about the dictionaries used, the skills the candidates employed, and the reading comprehension test itself. None of these variables is discussed in any detail in BS&W’s paper, and for this reason we decided to recreate the conditions of BS&W’s project, working on a smaller scale but paying particular attention to the interface between candidates, questions and dictionaries.
EXPERIMENTS

Two studies are reported here. In both cases subjects took a reading test on the final day of a four week professional EAP course at Warwick University. All the subjects were non-native speakers of English. Most were postgraduate students, and all intended to continue their studies at British universities. The test used in both studies consisted of two texts (812 words in total) and 15 multiple choice questions. The texts were both taken, unadapted, from New Scientist, and had a readability level of 9th grade (text 1) and 11th grade (text 2), according to the Fry readability formula (Stauffer, Abrams and Pikulski 1978).

EXPERIMENT ONE

Test administration

83 overseas students participated in this study. 40 took the test without access to dictionaries. The remaining 43 were allowed to use either their own monolingual dictionaries – the Oxford Advanced Learner’s Dictionary (OALD), Longman Dictionary of Contemporary English (LDOCE), Longman Active Study Dictionary (LOSD) – or bilingual dictionaries (Japanese, French, Turkish, Chinese, Polish, Korean or Thai). All subjects took the test in the same examination room. Students were allotted a maximum of 60 minutes for the test. Those with access to a dictionary were asked to indicate the words which they looked up in the dictionary while working on the test.

Data analysis

As in the study by BS&W, monolingual and bilingual dictionary use was related to:

   i) test score
   ii) the amount of time subjects took to complete the test.

We also compared test score with quantity of dictionary use (i.e. the number of words looked up), and noted which words subjects had chosen to look up.

Results

Although dictionaries were available for half the subjects, not all of this group actually used them. This means that the subject population can be divided into four groups:

\[ \begin{array}{lll}
\text{DICMo} & \text{= monolingual dictionary users} & n = 19 \\
\text{DICBi} & \text{= bilingual dictionary users} & n = 9 \\
\text{DICNo} & \text{= dictionary available but not used} & n = 15 \\
\text{NOTAv} & \text{= no dictionary available} & n = 40 \\
\end{array} \]

Data from these groups is summarised in Table 1.
Table 1: mean correct comprehension items (max = 15)

<table>
<thead>
<tr>
<th>Group:</th>
<th>DICMo</th>
<th>DICBi</th>
<th>DICNo</th>
<th>NOTAv</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean Score:</td>
<td>10.7</td>
<td>10.8</td>
<td>10.8</td>
<td>10.9</td>
</tr>
<tr>
<td>s.d.</td>
<td>1.5</td>
<td>1.9</td>
<td>1.7</td>
<td>2.1</td>
</tr>
</tbody>
</table>

As in the studies by BS&W, the difference in scores between those who used dictionaries and those who did not was non-significant.

BS&W reported that there was no difference between high scorers and low scorers in the number of words they looked up. We therefore divided our dictionary users into high scorers (13-15 n=3), medium scorers (10-12 n=12) and low scorers (6-9 n=8). Table 2 shows the mean number of words looked up by each group.

Table 2: mean number of words looked up

<table>
<thead>
<tr>
<th>Group:</th>
<th>high</th>
<th>medium</th>
<th>low</th>
</tr>
</thead>
<tbody>
<tr>
<td>mean no of look ups</td>
<td>1.6</td>
<td>6.3</td>
<td>2.3</td>
</tr>
<tr>
<td>s.d.</td>
<td>0.8</td>
<td>6.0</td>
<td>1.7</td>
</tr>
</tbody>
</table>

The data suggests there might be a tendency for high scorers and for low scorers to use their dictionaries less than intermediate scorers. However, there is considerable variation between subjects, and generalisations seem unreliable: among the medium scorers, for instance, look-ups ranged from 0 to 23 words.

As BS&W's Study 1, bilingual dictionary users seemed to have used their dictionaries slightly more than average; they looked up a mean of 6 words (sd=7.7, range=23).

BS&W also reported a correlation between speed of completion and score achieved, with faster candidates gaining a higher average score than their slower companions. We therefore divided our subjects into three groups according to the time span within which they submitted the test: fast (submitted within 40 minutes, n=36), medium (submitted within 60 minutes, n=33) and slow (submitted only when required to do so at end of test, n=14). Table 3 shows the relationship between completion speed and score.

Table 3: mean correct comprehension items

<table>
<thead>
<tr>
<th>Group:</th>
<th>fast</th>
<th>medium</th>
<th>slow</th>
</tr>
</thead>
<tbody>
<tr>
<td>mean score</td>
<td>11.5</td>
<td>10.4</td>
<td>9.8</td>
</tr>
<tr>
<td>s.d.</td>
<td>2.1</td>
<td>1.3</td>
<td>1.5</td>
</tr>
</tbody>
</table>
There is a significant difference here: \(F(2,82) = 5.43, p < .01\). Further analysis showed that this effect is due to a significant difference between the fast group and the slow group. The medium group is not different from the other two.

This finding agrees with BS&W. In other ways, however, our data does not agree with theirs. In particular, BS&W report that there was no relationship between dictionary use and the time taken to complete the test. In our data there appears to be a rather closer correspondence. In the fast group, only 19% of the subjects used a dictionary; in the middle group, 36% used a dictionary; while in the slowest group 64% of the subjects consulted the dictionary. This discrepancy clearly deserved further investigation.

**EXPERIMENT TWO**

**Background**

In order to resolve this apparent contradiction between our findings and those of BS&W, we decided to repeat the experiment, gathering more precise information regarding subjects' speed of completion. We were also concerned that BS&W's random allocation of subjects to the experimental groups may have introduced some uncontrolled variables into the design of the experiment. In BS&W's Study One the three conditions – without a dictionary, with a monolingual dictionary and with a bilingual dictionary – were selected “randomly”, and in the later studies there was a certain amount of free choice in dictionary use: “Of those students not using dictionaries at all, some students decided they did not need a dictionary because it was too time-consuming, while others simply forgot to bring them and would really have preferred to use a dictionary” (p267). In our first study, following BS&W, the group to be allowed access to dictionaries was also selected randomly, and within that group those who opted to use bilingual dictionaries were those who happened to have their own dictionaries with them. It is likely that these students made a habit of using their bilingual dictionaries, carrying them with them wherever they went, and these subjects may have been less confident and weaker than those who did not carry dictionaries. We found that there was also within the dictionary-access group a number of students who considered that they should not have been allotted dictionaries, because they felt themselves to be too advanced. They very ostentatiously pushed the dictionaries aside and indicated that they wanted to join the group who had not been allotted dictionaries. In our second study, we therefore decided:

a) to match the two groups according to language ability;

b) to conduct the test in two separate rooms, so that both groups were unaware that the other group was taking the test under different conditions;

c) to allow access to one type of dictionary only (OALD), rather than a range of monolingual and bilingual dictionaries; and

d) to record accurately each subject's completion time.
Test administration

65 overseas students participated in this study. On the basis of scores on a previous test they were divided into two groups of matching ability. The two sets of subjects took the test in separate examination rooms. One group (31 subjects) was not given access to dictionaries during the test. The other group (34 subjects) was given access to OALD during the test.

Prior to the test all 65 subjects were asked to underline on a wordlist those words which they were not familiar with. The wordlist contained all lexical words in the text and question paper, with the exception of common words (those in Bands 1 and 2 of Hindmarsh’s Lexicon (Hindmarsh 1980)).

Subjects were then given a maximum of sixty minutes to complete the test. They were required to indicate completion time on the test paper.

On completion of the test subjects in the group with access to dictionaries were required to indicate on the wordlist those words which they had in fact looked up.

Data analysis

As in the studies conducted by BS&W, test scores were analysed according to:

i) whether the subject had access to a dictionary;
ii) the amount of time the subject took to complete the test.

An analysis was also made of the words subjects indicated that they were not familiar with, and which words they actually looked up.

Results

Subjects in both groups were evenly matched on the basis of the presessional course entry-test scores. Results from the pretest wordlist activity confirmed that there was no significant difference in the make-up of the two groups. The first group, who were to be given access to dictionaries, indicated that a mean of 16 words were unfamiliar to them (sd 7.6), the second group, who were not to be given access to dictionaries, were unfamiliar with a mean of 17.9 words (sd 8.2).

The subject population can be divided into three groups:

\[
\begin{align*}
\text{DICU} &= \text{dictionary users} \\
\text{DICNo} &= \text{dictionary available but not used} \\
\text{NOTAv} &= \text{no dictionary available}
\end{align*}
\]

\[
\begin{align*}
n=29 & \quad n=5 \\
n=31 &
\end{align*}
\]

The test results of these groups are summarised in Table 4.
Table 4: mean correct comprehension questions

<table>
<thead>
<tr>
<th>Group:</th>
<th>DICU</th>
<th>DICNo</th>
<th>NOTA\text{av}</th>
</tr>
</thead>
<tbody>
<tr>
<td>mean score</td>
<td>11.0</td>
<td>12.6</td>
<td>10.7</td>
</tr>
<tr>
<td>s.d.</td>
<td>2.3</td>
<td>2.2</td>
<td>2.4</td>
</tr>
</tbody>
</table>

As in our first study and in those conducted by BS&W, there was no significant difference in comprehension scores between those who had access to a dictionary and those who did not. However, our results regarding dictionary use and speed of completion, which are summarised in Table 5 below, contradicted the findings of BS&W.

Table 5: time taken to complete the reading task (mins)

<table>
<thead>
<tr>
<th>Group:</th>
<th>DICU</th>
<th>DICNo</th>
<th>NOTA\text{av}</th>
</tr>
</thead>
<tbody>
<tr>
<td>completion time</td>
<td>37.3</td>
<td>38.4</td>
<td>25.8</td>
</tr>
<tr>
<td>s.d.</td>
<td>7.8</td>
<td>8.0</td>
<td>4.3</td>
</tr>
</tbody>
</table>

Whereas BS&W reported that dictionary use made no significant difference to the speed of test completion, our subjects took significantly longer to finish the test if a dictionary was available, irrespective of whether they used it or not [F(2,62)=24.4, p<.001]. Some possible explanations for this puzzling result are put forward in the next section.

DISCUSSION

BS&W made certain assumptions about the effects of dictionary use. They believed, among other things, that:

1. All students would prefer to avail themselves of the opportunity to use a dictionary during a test when permitted.
2. The permitted use of monolingual and/or bilingual dictionaries would significantly raise examination scores.
3. The use of dictionaries would significantly increase the time taken to complete a test.”
   (BS&W 1984:270)

However, the results of the studies conducted in Israel suggested that these assumptions were false:
"According to the three studies, the use of a dictionary has no significant effect on reading comprehension test scores based on multiple-choice questions. Neither does its use affect the time students need to complete the test. Moreover, even when permitted to use a dictionary, many students (mostly those with relatively high English proficiency) did not wish to do so." (BS&W 1984:270)

Our studies have confirmed two of these findings. Test scores were not significantly affected by dictionary use, and dictionaries were not popular with all subjects; 15 of the 43 subjects in Study 1 and 5 of the 34 subjects in Study 2 were permitted to use dictionaries but chose not to use them.

However, BS&W's other finding was not confirmed in these studies. In both tests, dictionary users took considerably longer to complete the test. We can only speculate about the reasons for this difference. It is possible that the BS&W's subjects were more efficient dictionary users than our subjects, and therefore wasted less time when using their dictionaries. It is also possible that BS&W's subjects were under more pressure to work at speed, and dictionary users, when intermingled with non-users, may have been sensitive to the pace of the examination room. In our Expt 2, where subjects with access to dictionaries worked in a different room from those without access to them, it was found that the whole group worked more slowly, including those subjects who did not in fact look up words. One interpretation of this behaviour is that, as the majority of subjects were dictionary users and thus worked at a slower pace, the non-dictionary users were not spurred to work more quickly by the sight of their colleagues submitting their completed papers.

Another possible explanation for the differences is that our subjects were making greater use of their dictionaries than their Israeli counterparts. However this does not appear to be the case as far as BS&W's Study 1 is concerned, as monolingual dictionary users were reported to have looked up a mean of five words, and bilingual dictionary users a mean of 13 words in the three hour tests, a figure that does not differ greatly from our results: in a one hour test, subjects looked up an average of four words in Expt 1, and 3.2 words in Expt 2. (In the other three studies by BS&W, the number of words subjects looked up was not reported.)

BS&W were surprised that dictionary use did not help their subjects in the test, and speculated that the students' lack of dictionary skills, their lack of knowledge of syntactic rules, or the difficulty of the test itself might have caused dictionary users to fare no better than those without access to a dictionary. The results certainly suggest a breakdown at some point in the process. If reading tests are designed to measure the learner's ability to comprehend text, and if dictionaries are designed to aid reading comprehension, it is not unreasonable to assume that the test, the
dictionary, or the user is failing in its purpose when dictionary use cannot improve reading test scores. Analysis of our data suggests that the responsibility lies with all three agents: the test, the dictionary and the user.

i) The Test

BS&W suggests that some of their testees failed to cope with the text, and could not use their dictionaries effectively, because the text contained too high a proportion of unknown words. This explanation can be ruled out here. BS&W based their suggestion on Johns’ claim (1980) that “when more than approximately 50 per 1000 words are unknown, perception of overall structure may be effectively blocked, which in turn means that there is not enough in the way of context to allow successful guessing”. Although Johns’ “threshold effect” may have operated in the experiments recorded by BS&W where some subjects identified as many as 68 unknown words per 500 – 700 word text, dictionary users in our study identified a mean of only 6.7 unknown words in Expt 1, and a mean of 16 words in Expt 2: the texts totalled 812 words. This suggests that the subjects recognised enough vocabulary to permit the successful application of guessing techniques – techniques which also help in the identification of meaning during dictionary consultation.

Most communicative reading tests are primarily concerned with testing reading skills rather than language knowledge. On the whole the student is tested on his understanding of the meaning of the text rather than on what he knows about the language. Questions which require the reader to recognise the function of a grammatical structure may be included, but questions depending on the understanding of individual lexical items are generally avoided because they do not enable the tester to generalise about the learner’s overall reading ability. The specific lexical item may be one of a tiny number that one learner knows, yet it might be excluded from another learner’s vast mental lexicon.

BS&W’s paper does not give details of the types of text subjects had to read, or the types of test question they had to answer. Analysis of our own test suggested, however, that the choice of test might be a very significant factor in determining whether dictionary users derived benefit from their dictionaries. Only five of the fifteen questions in our test depended to any degree on the comprehension of individual words. The other questions in the test required the reader to analyse language functions, to process anaphora, or to extract factual information. For such questions comprehension of overall context seemed to be more important that the ability to define a key word or expression. The five questions which did seem to require detailed understanding of specific lexical items are reproduced here in full:
4) Lines 18-20 state that “laminated plywood components are slotted together like a giant jigsaw puzzle”. This is so that:
   a) it can be manufactured anywhere
   b) any damaged part can be replaced (CORRECT ANSWER)
   c) people can choose different designs
   d) the car is corrosion and dent proof

6) The word “configurations” in line 25 is most likely to mean:
   a) strengths and weaknesses
   b) components
   c) shapes and sizes (CORRECT ANSWER)
   d) colours and forms

7) The word “just” in line 46 could be replaced by:
   a) alone
   b) right
   c) only (CORRECT ANSWER)

13) The expression “in conjunction with” in lines 84-85 means:
   a) in co-operation with (CORRECT ANSWER)
   b) in competition with
   c) in co-ordination with
   d) in cahoots with

14) The word “assessed” (line 89) is most likely to mean:
   a) marked
   b) criticised
   c) measured
   d) judged (CORRECT ANSWER)

Note that some of the words in these questions are more crucial than others. Question 4 is difficult to answer correctly without some understanding of slotted, and a knowledge of the word jigsaw helps to contribute to that understanding, but dent-proof can only help in the elimination of alternative d), and laminated is not useful at all (although a subject who did not know the meaning of the word would not be able to guess that it was not important). In question 13 knowing the meaning of in conjunction with is not enough; the subject is required to make fine distinctions between co-operation, co-ordination and cahoots.

Surprisingly, there was not a great deal of correspondence between the words that dictionary users indicated that they had looked up and the words needed to answer these five questions correctly. Table 6 lists “keywords” (i.e. words which needed to be understood before the question could be answered correctly) and shows the
number of subjects who indicated that they had looked up these words. Table 6 also indicates that in most cases, subjects who looked a word up answered the question correctly. Tests which contain a large number of items where knowledge of individual words is crucial may be more affected by the availability of a dictionary than tests with few items of this type.

**Table 6: look-up rates and correct answers for “keywords”**

<table>
<thead>
<tr>
<th>key word</th>
<th>looked up</th>
<th>correct</th>
</tr>
</thead>
<tbody>
<tr>
<td>jigsaw (puzzle)</td>
<td>12</td>
<td>8</td>
</tr>
<tr>
<td>slot together</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>corrosion proof</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>dent proof</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>laminated</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>configurations</td>
<td>9</td>
<td>8</td>
</tr>
<tr>
<td>in conjunction with</td>
<td>6</td>
<td>4</td>
</tr>
<tr>
<td>in cahoots with</td>
<td>9</td>
<td>4</td>
</tr>
<tr>
<td>in coordination with</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>assessed</td>
<td>8</td>
<td>3</td>
</tr>
<tr>
<td>Total</td>
<td>57</td>
<td>37</td>
</tr>
</tbody>
</table>

**ii) The Dictionaries**

Leaving aside the question of whether the dictionary definitions were comprehensible to the users, it would appear that the dictionaries did not always supply the information necessary to answer the test questions correctly. We have no data on the entries in the bilingual dictionaries, but the monolingual dictionary definitions did not always give the candidate the lexical information he needed to answer the test question correctly. This was particularly the case with OALD entries, where the examples were limited in number and not always applicable to the technological topics of the *New Scientist* texts. For example, OALD ignored the (non-figurative) sense of *slot* which was relevant to question 4 (whereas the LDOCE examples at *slot* managed to express the idea that it is easy to remove something that has only been slotted into place). OALD also provided no examples for *in conjunction with*, and only provided examples with a financial theme for *assess*. (LDOCE again guided the reader to the correct answer for question 14 by including the synonym “judge” in definition 2). LASD too lacked necessary information; there were no entries for two of the five keywords that subjects indicated that they had looked up. (Both LDOCE and OALD, however, provided guidance for answering question 6, by mentioning “shape” in their definitions of
configuration. By their provisos “disreputable” and “dishonest” they also steer the reader away from the in cahoots with alternative in question 13).

The question of whether the correct answers to the test questions were themselves accurate definitions of the words in the text remains open to debate.

iii) The Users

Table 6 indicated that not all instances of dictionary use result in correct answers to the questions. Our data does not provide a means of determining whether lack of dictionary skills led students to make mistakes, but we can identify two other possible reasons for this result: the inadequacy of the dictionary definitions themselves (discussed above), and the users’ failure to look up other important keywords. Of the twelve subjects who looked up jigsaw puzzle, four answered the question incorrectly, and this may well have been because the question depended more heavily on an understanding of the word slotted – which none of the twelve went on to look up.

Some dictionary users did not look up keywords at all. Although it is not usually taught as a dictionary skill, the ability to identify relevant words in a text is just as important as the ability to find their meanings in the dictionary. Subjects choose to look up a variety of words apart from keywords. Some of these words were relevant to the test questions, others apparently only of interest to the subject. The data does not provide us with information regarding subjects’ motives for selecting words to look up, but it would appear that some were taking the opportunity to use the dictionary to learn new vocabulary, or looked up words to answer their own questions about the meaning of the text, despite the fact that they were working under test conditions. According to BS&W, the subjects in their experiments looked up words to answer test questions, and were not motivated by “the desire purely to understand the text”. In this study, the subjects were mature and self-motivated students, used to taking responsibility for their own learning. Perhaps they used the dictionary in the test situation in the same way as they would use it in “real life”, without too much regard for the demands of the test.

The majority of the words looked up occurred in text 1 and most of these words fell in the opening paragraphs. This suggests that subjects were more enthusiastic about looking up words at the beginning of the test, and lost interest later on. We can only speculate as to whether subjects became bored, disillusioned, or perhaps more confident as they worked through the test. It should be noted that although proportionately more of the slower subjects in Expt 1 used dictionaries, they did not in fact look up more words. There is therefore no evidence that dictionary use was in itself responsible for slowing down the performance of these subjects. It is possible that they took longer to look up words than their faster colleagues, but this would be symptomatic of their general slowness in tackling the test.
CONCLUSION

BS&W's analysis of their data did not relate the words their subjects looked up to the demands of the test questions, they did not check to see whether the dictionaries their subjects used dealt adequately with the words their subjects look up, and they did not check whether look up of keywords in the test resulted in correct answering. All these considerations are important if we want to find out why dictionary use did not affect reading comprehension test scores. In our studies, it appears that dictionary use did not affect test scores primarily because the test itself was made up of items which were not likely to be affected by the availability of a dictionary. However, in some cases where dictionary use might have aided the subjects, either the dictionaries themselves did not provide the necessary information, or the users failed to identify the words in the text which were most crucial for correct answering of the test questions.

REFERENCES


