Readability of Science Textbooks in Forms 1 and 2

Abstract

Alexandra Sollars

The readability of science textbooks, used by pupils attending Forms 1 and 2 in state schools, was investigated. According to the Flesch Reading Ease formula and the Fry readability graph, This is Science 1 had a mean reading age of 12.7 years ($\sigma = 1.47$) and 11.5 years ($\sigma = 1.54$) respectively, whereas This is Science 2 had a mean reading age of 12.6 years ($\sigma = 1.2$) and 11.6 years ($\sigma = 1.70$).

Seven versions of two close tests were administered to a sample of 316 Form 1 girls from four area schools and 397 Form 2 girls from the same area schools and a Junior Lyceum. Two methods of scoring were used. It was found that the pupils may understand This is Science 1 if teacher's help is available. This is Science 2 was too difficult for pupils of average and below-average ability but within the capabilities of high-ability pupils.

PURPOSE OF THE STUDY

Many research studies suggest that science texts provide serious problems of comprehension and readability for pupils (Hart, 1980; Walker, 1980; Reid *et al*, 1983). Readability — the quality of the text which determines understanding — depends on different factors. Vocabulary is very important in determining the difficulty of a science textbook. Not only is the usual vocabulary used difficult, but the introduction of unfamiliar scientific and technical words together with the use of complex sentence structures creates further problems. A pupil may also find difficulties because of his mood, interest and intelligence, the type of print and the page layout (Carré, 1981). These findings have instigated the examination in some detail of the readability of the pupils' science textbooks used in Forms 1 and 2 in Maltese state schools.

The Science Textbooks

Pupils in the first two years of their secondary education follow an Integrated Science course. All state schools make use of the same textbooks — *This is Science 1* (Dobson, 1981a) and *This is Science 2* (Dobson, 1981b). According to the author,

> "the pupils' books are quite short and written in a vocabulary and with a syntax that should make them useful to all but the remedial reader and it is hoped that even these pupils can be helped in their science lessons by the diagrams and experiment instructions".
>
> *(Dobson, 1980, p. 1)*

Further on, he states that "*This is Science* caters for the pupils of average and below-average reading ability" (p. 11) who, however, can understand science. The author also asserts that the aim of the science course has been to produce "attractive book(s) that encourage pupils to read — for pleasure, for background information, [and] for help in carrying out experiments" (p. 2).

It should be noted that one of the reasons for the selection of these books for Maltese pupils was that the language level was easier compared to the previous texts, *Science for the 70s* Books 1 and 2, as the syntax and vocabulary used were simpler. Hence it was expected that the majority of pupils would understand the books (Ventura, 1985).

Maltese Pupils' Reading Ability

Since the science course is in English and examinations are set in this language, and as readability involves a measure of the match between the text and its readers' ability, determining the reading ability of the Maltese pupils in the second language becomes very important.

Unfortunately, no recent reliable measures of the pupils' reading ability are available. According to Ventura (1985), the only recorded standardised reading ability test was administered to Maltese pupils in 1968. Officials from the Education Department constructed an English Word Recognition Test of 110 words selected from two word recognition tests by Burt and Schonell for which British pupils' norms were known. The test was set to a representative sample of 753 Maltese pupils, between the ages of seven and eleven years, in both private and state schools (Falzon, 1968). The results of the test showed that the reading age of seven-year-old Maltese pupils was about a year-and-a-half lower than that of the British counterparts. This gap increased with age and, on entering a secondary school, an 11-year-old Maltese had a reading age equivalent to an 8-year-old British pupil. When the same word recognition test was used in 1979 in the lower streams of twelve secondary schools, the results showed that boys of median age 12 years 2 months had a reading age corresponding to that of an English 7-year-old.

Girls of a median age 12 years 1 month had a reading age which corresponded to a British pupil of seven-and-a-half years (Ventura, 1985).

The only data available on the pupils's level of understanding of the science textbooks is from an unpublished study by Ventura (1985) who administered a cloze test to 152 Form 2 pupils in the second best streams of five area schools (3 boys' and 2 girls' schools). In the cloze test, a passage from the pupils' books was used. The results showed that the sample of pupils chosen was finding *This is Science 2* too difficult. But as the sample was not a representative one, the results obtained may not have been a true picture of the actual situation.

Due to the meagre data on Maltese pupils's reading ability in general, and on the readability of the textbooks of the *This is Science* course in particular, it was decided to investigate further the readability of the science textbooks in Forms 1 and 2 and indirectly to obtain an indication of the female pupils' ability to read and understand simple texts in English.

Main Purposes of this Study

This study is designed:

(a) to check the reading age of *This is Science 2* by applying readability formulae to passages chosen at random from all the possible chapters in the textbooks. The readability formulae applied to the texts were the Flesch Reading Ease formula and the Fry Readability graph;

(b) to check whether Maltese pupils are able to understand these textbooks. For this purpose it was decided to administer cloze tests because several research studies discussed below favour this method, they are easy to prepare and straightforward to score. Also, cloze test scores are comparable to results from readability formulae.

Readability of Textbooks

What is Readability?

Readability is not a simple measure of some property of the written material but it is a function of characteristics of the text, the readers and the interactions between readers and text. According to Harrison (1980), those aspects of a text which make it easy for a reader to understand what is being conveyed by the author include: legibility of print, illustration and colour, vocabulary, conceptual difficulty, syntax and organisation.

The person's understanding of a book is not only determined by the factors related to the text itself, but also by the individual's own knowledge and abilities. His level of familiarity with written and oral language structures and vocabulary influence the degree of comprehension. This factor applies very much to Maltese students since English is their second language. Interest and motivation are also vital in reading, especially with low ability students.

Measuring Readability

Because science lessons involve a lot of practical work, demonstrations, blackboard work and clearing up before the end of the lesson, and a packed syllabus has to be covered for the exam, teachers tend to use the textbook as an aspect of homework. The findings of the *Effective Use of Reading Project* (Lunzer and Gardner, 1979) suggest that in spite of the potential value of reading for the individual, science teachers do not rate it highly as a method of teaching and in fact use it very little. A book, which is to be used without the help of the teacher, can only be effective if it is appropriate in its level of difficulty and satisfying to the reader. Not only textbooks, but even books for wider reading, reference and research need to be suitable for a given age group and a range of pupil characteristics (Hart, 1980).

Even if the teacher has a knowledge of the reading ability of each pupil, deciding whether a text is suitable or not is difficult. Objective methods of predicting text difficulty are therefore required to help the teacher. Foremost among these methods one finds readability formulae, charts and graphs or cloze tests.

Readability Formulae

Readability formulae, charts and graphs are based on counts of language variables in a piece of writing in order to obtain a measure of the probable difficulty of the text for readers. They are predictive measures because no participation by readers is required (Klare, 1974-75). The most commonly used formulae assume that the two most important determinants of text difficulty are vocabulary and sentence structure. Measures of word frequency, word length (such as average number of syllables per word) and sentence length can predict the difficulty of a text. However a correlation does not imply a causal relationship and formulae can be "fooled" by specific texts such as poems (Harrison, 1980). Irrespective of some possible drawbacks, formulae are often made use of. Five of the more popular ones are the Dale-Chall Formula, the Flesch Formula, the Lorge Formula, the Mugford Readability Chart and the Fry Readability Graph. The Flesch formula and the Fry graph have been applied to the science textbooks *This is Science 1* and *This is Science 2* and will be discussed later on.

Cloze Tests

The term "cloze procedure" was first used by Wilson Taylor in 1953 to describe a new way of testing the readers' comprehension. He derived the word "cloze" from "closure" in Gestalt

psychology, the tendency by persons to close or complete an incomplete circle or any other familiar object (Oller, 1979; Wainman, 1979). In the cloze technique blanks are placed in prose where words in the text have been omitted and examinees are asked to insert the word which seems most suitable in each blank space. The proportion of correctly-guessed words gives an indication of the extent to which the reader has understood the passage.

Unlike any formula, cloze scores are sensitive to the variations within a specific group. Variations can arise because of the group's background knowledge of the subject-matter of the test, or the group's overall level of linguistic comptence if it happens to differ from the norm. Unlike most readability formulae, cloze tests are not "fooled" by simple long sentences or by short but unfamiliar words. The cloze procedure is an important method of finding out how well a specific group can understand a book. Oller (1979) quotes Klare, Sinaiko and Storulow (1972) who point out that the average score of a group of subjects on a cloze test is an *actual* measure of readability whereas the formulae are means of *estimating* difficulty levels. The cloze procedure, which will be considered in more detail below, has been applied to two separate passages taken from the science textbooks in order to find out the pupils' level of comprehension.

Flesch Reading Ease Formula (1948)

The Flesch formula is one of the best-known and most widely used readability measures because of its high validity and reliability. Flesch chose a difficulty index which did not relate to age, but to an imaginary comprehension score out of 100: the lower the score, the more difficult is the passage. Flesch found that the two variables which correlated most highly with difficulty are the average number of syllables per word and sentence length. He incorporated these two variables in an equation which gives a reading ease score (Harrison, 1980, p. 72).[1] The reading ease score can be related to age levels by an arithmetical transformation.

Fry's Readability Graph (1977)

According to Fry, the purpose of his readability graph is "to aid teachers and editors to help children or adults read better by giving them material on the proper difficulty level" (Fry, 1977, p. 243). It is helpful because it is easy to work out and the readability index is represented visually on a graph. The graph can be applied from grade one (age 6) up through the college years (18 to 22 years). The two variables entered in the graph are the number of syllables in a continuous passage of 100 words, and the number of sentences in that passage. The information is plotted on the graph and the area where the two lines intersect gives the approximate American grade level. The UK reading level is found by adding five to grade level.

It can be assumed that readability formulae are quite reliable because they correlate fairly well with each other. Validity can be found by correlations between formulae, comprehension scores, cloze scores, oral reading errors, observer judgement and with written passages of known difficulty. Fry's readability graph gives a high correlation coefficient, 0.81 with cloze rankings (Bormuth levels). It also correlated 0.95 with the Flesch and 0.85 with the Dale-Chall formulae (Fry, 1977).

Cloze Procedure

Many researchers have shown that the cloze procedure gives more reliable information about the difficulty levels of samples of text than any other method yet devised, and the technique works much better than the readability formulae (Oller, 1979). This is because the cloze procedure not only measures the difficulty of the passage itself but also measures what a reader brings to the passage in terms of content knowledge. Construction of cloze tests will now be briefly discussed.

Selecting the Passage

One or two passages about 250 words long are selected at random, however this may not be possible if the text is too brief or continuous prose is kept short.

Deletion Procedure

Deletion usually takes place at equal intervals, say every fifth, seventh or tenth word. This procedure is called the fixed-ratio method. By this method, the type of words which are deleted are not pre-calculated. Wainman (1979) showed that the cloze method of mechanical deletion produces results for second-language learners as good as the results obtained by deleting content words. Such tests are easier to prepare and a wider variety of language categories can be studied.

A sentence of two at the beginning of the text can be left intact to serve as a brief lead in before the deletions begin. Some researchers prefer leaving a few undeleted sentences as lead out at the end of the passage. The minimum number of deletions for a quite reliable cloze test is 35 (Harrison, 1980).

Fifth-word deletion has been the most widely used rate in research work, but it may not be the best for the classroom. Klare *et al.* (1972) suggested that the seventh-word rather than the fifh-word deletion is better for readers who are quite weak or

if the passage is likely to be found difficult (Harrison, 1980).

The best estimates of readability may be obtained by "clozing" every word in sample texts (Oller, 1979). That is, if the text is 250 words long, with every seventh word deleted, it is possible to construct seven versions of the text with 35 blanks each, by deleting the first, eighth, fifteenth . . . and so on for the first version; the second, ninth, sixteenth . . . for the second version; and so on until the seventh, fourteenth, twenty-first . . . are deleted for the seventh and last version. In this way, the cloze test measures the difficulty of every single word, phrase and sentence in the passage.

Each time a word is omitted, a standard length of blank is used. The space commonly left is twelve or fifteen letter-spaces, depending on the size of the type.

Scoring Procedures

Two types of scoring techniques are the exact word method and the contextual appropriate method. By the *exact word method*, a word is scored correct only if it is the exact one which appears in the original passage. This may be considered too rigid, especially when alternatives can be just as good or better than the word in the text. However, verbatim scoring is simpler, faster and more reliable and valid than any other method (Harrison, 1980).

In the *contextual appropriate method*, any word which fully fits the total surrounding context is considered correct. This method relies on a certain degree of subjective judgement by the person correcting the texts. However high correlations are observed between scores generated on this basis and scores based on the exact word technique (Oller, 1979). Researchers who have used the cloze method as a measure of second-language proficiency often preferred scoring by the contextually appropriate method (Oller, 1972; Wainman, 1979).

Criteria to Assess Understanding

Once the cloze tests have been corrected, it is important to interpret the results obtained and to know what the scores mean. Different researchers offer different suggestions (Bormuth, 1968; Oller, 1979; Harrison, 1980).

Anderson (1971), quoted by Oller (1979), suggests the following interpretations for the cloze scores obtained by non-native students and these criteria are used in this study. A cloze score of 53% or above (by the exact word method) corresponds to the *independent level* of reading. Passages corresponding to this level are said to be suitable for the student's study in an unsupported context. A score between 44% and 53% is, according to

Anderson, appropriate at the *instructional level*. Passages or textbooks at this level can be effective if they are read in the presence of a teacher who can define, explain, simplify and illustrate with reference to the pupils' experience. If the score is less than 44%, it falls in the *frustration level* and this implies that the texts are too difficult and the reader cannot learn much from the book without a great deal of specific help.

Research Methodology

Selection of Samples from the Pupils' Texts

A minimum of three samples of 100 words each were chosen at random from every chapter. Thus any change in the level of difficulty in the same chapter was catered for. Each passage chosen had to deal with one specific idea. However this was not always feasible as certain chapters did not have passages which were long enough. In all forty-seven 100-word samples, twenty-four from *This is Science 1* and twenty-three from *This is Science 2*, were subjected to the analysis required by the Flesch formula and the Fry Readability graph.

Using the seventh-word deletion for the cloze tests and having a minimum of thirty-five deleted words in every version, meant that the sections had to be at least 245 words long. It was thought best to use 7th-word deletion rather than 5th-word deletion since the cloze tests were to be given to second language speakers. Two or three sentences at the beginning of the cloze test were to serve as a brief lead in before deletions commenced. Every deleted word was replaced by a blank 15-typed spaces wide.

The section entitled "Looking at Animals" was selected from the second chapter of *This is Science 1*. The author of the books says that this section "is for reading by the pupils. It could be set as 'quiet work' or as homework . . ." (Dobson, 1980, p. 57). This section was one of the very few passages of continuous prose of suitable length (280 words).

The second cloze test was prepared from a section of the sixth chapter of *This is Science 2*. The section is entitled "Energy from the Sun". The cloze test was 282 words long which included the first two sentences (37 words) as a brief lead in. As with the first cloze test seven versions were set and every deleted word was replaced by a blank 15-spaces wide.

The Sample of Students

Four girls' area schools with Forms 1 and 2 were chosen out of the nine available ones. The second cloze test was also carried out at a girls' Junior Lyceum.

The first cloze test was distributed to 316 Form 1 girls of average age 12 years 1 month. The top three or four streams were chosen, depending on the school population and the Headmistresses' opinion on the students' level of understanding of English. Table I shows the total number of Form 1 girls in the different streams.

Research Findings

Results of Readability Formulae

The mean reading age for *This is Science 1* is 11-and-a-half years ($\sigma = 1.54$) according to the Flesch Reading ease formula and 12 years 8

Test	School	Numbers of pupils				Total
		Stream A	Stream B	Stream C	Stream D	
Cloze I	Area	102	96	85	33	316
Cloze II	Area	89	84	80	39	219
Cloze II	JL	25	25	27	26	105

Table I - Sample of girls in Forms 1 and 2 according to school and stream.

In order to keep certain extraneous variables - such as intelligence, motivation and attitude towards science - constant, the girls selected for the first cloze test also participated in the second cloze test a year later when they were in Form 2. Due to circumstances beyond control, it was impossible to have the same number in Forms 1 and 2, because of transfers to the Junior Lyceums or other schools, promotions or demotions to a different stream or absenteeism. The girls' average age was 13 years. The second cloze test was also applied to 105 Form 2 girls in four streams of a Junior Lyceum. Their average age was also 13 years. Table I shows the total number of Form 2 girls in the different schools and streams.[2]

Application of the Cloze Tests

Copies of the seven versions were distributed as evenly as possible to the pupils in the four chosen area schools and Junior Lyceum. Every pair of girls was given the same version in order to prevent any copying. The first cloze test was held between June and July 1984 whereas the second cloze test was held the following year in June 1985. The pupils were given about 30 minutes to complete the cloze test. It was more than ample time, and most girls - especially the weaker ones - handed in the cloze test before the bell rang.

Since the cloze test was given to second-language speakers, it was decided to score the versions both by the exact word method and the method for contextual appropriateness. In the second method of scoring, any insertion that fully fit the surrounding context was considered correct.

months ($\sigma = 1.47$) according to the Fry readability graph. According to the Fry graph the easiest passage has a reading age of 8 years and the most difficult of 13 years. The Flesch formula gives a wider range with the easiest passage having a reading age of 10 years 5 months and the most difficult of 16 years 4 months. Slight discrepancies between the two readability formulae are to be expected, for although they show very high correlation, no two formulae show perfect correlation. The most frequent reading age is 13 years, which is higher than the average age of girls in Form 1.

The mean reading age for *This is Science 2* is 11 years 7 months ($\sigma = 1.70$) according to the Flesch reading ease formula and 12 years 7 months ($\sigma = 1.16$) according to the Fry readability graph. The easiest passage has a reading age of 8 years and 10 years 6 months according to the Fry graph and the Flesch formula respectively. The most difficult sample has a reading of 14 years and 14 years 11 months, depending on the readability formula applied. The most frequent reading age of the passages is 13 years, which agrees with the girls' mean age in Form 2.

Wide divergences from the average exist even within the same chapter. The average variation in reading age in the chapters of *This is Science 1* is 2.5 years with the maximum being a 5-year discrepancy in reading age level (by Fry graph) in Chapter 3. The average difference in reading age of *This is Science 2* is 2 years, although passages taken from the first chapter show a variation of 5 years (by Fry graph).

This implies that whereas the author of the science course (Dobson, 1980) says that the books cater for the pupil of average and below-average reading ability and that they have been written in a vocabulary and with a syntax suitable for all, the formulae applied in this study indicate otherwise. Unless teacher support and guidance are available, an average 12-year-old would find it extremely difficult to understand passages suitable for 16-year-olds.

Readability formulae give the aproximate reading age level for a person whose first language is English. Since formulae are predictive measures of readability of texts, they do not take into account the readers' background and therefore they do not distinguish between a first-language and a second-language speaker. According to both the Fry graph and the Flesch formula, *This is Science* has a suitable average reading age level for UK readers. However, considering the results quoted previously about the Maltese pupils' reading ability (Ventura, 1985), both Book I and Book 2 are too difficult for Maltese pupils.

Results of Cloze Test I (Area Schools)

According to the readability formulae, the passage chosen for cloze test I has a reading age level between 12.3 and 13.2 years (by the Flesch formula) and 11 and 12 years (by the Fry graph). The global results, using two methods of scoring, are shown in Table II.[3]

When the passages were corrected by the exact word method, the global score obtained in cloze test I was 12.5 (maximum score 35). The easiest version was the fifth, with an average score of 15.0, and the most difficult proved to be the firsdt of 15.0, and the most difficult proved to be the first ,version with a very low mean score of 9.2.

The scores of cloze test I range from 27.4% to 42.9%, with an average equivalent ot 36.1%. The percentage marks of four out of seven versions fall with the 35% to 40% criterion, which - according to Bormuth - might be too low and readers show signs of frustration and inattention when faced with such difficult texts. Children scoring below 40% will not learn much from a book unless a lot of help is available. The marks obtained in the cloze test are indicating that girls attending the area schools are facing serious problems which can only cause low interest and motivation in the subject.

Verbatim scoring may be considered too strict a method for second language speakers and therefore the cloze tests were scored by a second method. Words which fit in the context written were also marked as correct. The global score obtained in cloze test I by the appropriate word method was 14.2. The easiest version remained version 5, with an average score of 16.1. The most difficult versions were the first and seventh, both giving an average score of 11.7.

The percentage marks of cloze test I range from 34.4% to 46%, with an average equivalent to 40.7%. The percentage marks of four versions fall in the 40% to 45% criterion. Bormuth (1968) suggests that scores equivalent to 44% represent the instructional level of reading. Harrison (1980) considers the 40% - 45% criterion as the most important in a classroom situation because very high percentages in cloze tests on school texts are not very common. Marks falling within this range therefore indicate that the text is suitable for studying only if help and advice are readily available.

The results of cloze test I indicate that girls in the area schools are encountering severe problems which can lead to a low level of understanding if the teacher in class does nothing to help them internalise the scientific concepts and comprehend the language used in the textbook. Certainly, if pupils are asked to use the textbooks on their own, for example by setting them reading tasks for homework, the text is too difficult for them.

Cloze Test/ School	Scoring by the Exact Word Method		Scoring for Contextual Appropriateness	
	×	% mark	×	% mark
Cloze Test I (Area Schools)	12.5	36.1	14.2	40.7
Cloze Test II (Area Schools)	8.5	24.3	11.3	32.3
Cloze Test II (Junior Lyceum)	17.1	48.9	22.5	64.3

Table II - Global scores obtained in the Cloze Tests using two methods of scoring.

Results of Cloze Test II (Area Schools)

According to the Fry graph and the Flesch formula, the passage chosen for cloze test II has a reading age between 12 and 14 years and 12.8 and 14.9 years respectively. Therefore the readability formulae predict a passage beyond the reading age of an average 13-year-old pupil. As can be seen from Table II, the global score obtained in this cloze test was 8.5 (maximum score 35) when the passages were corrected by the exact word method. The easiest version was the first, with an average score of 10.8, whereas the most difficult was the fourth, with an average score of 6.4.

The percentage mark of all the seven versions of this cloze test is below 31%, with the lowest being equivalent to 18.3%. These percentage marks fall well below 35%. Bormuth considers such a percentage too low for any learning to take place. The passage used in the cloze test proved to be too difficult for the girls in the area schools. The results confirm the estimations given by the readability formulae and they indicate that the language is causing a formidable barrier for science to be comprehended by the girls in the study.

When cloze test II was scored for contextual appropriateness, the global score increased slightly to 11.3. The easiest version was the third, with a mean score of 14.0, while the most difficult version remained the fourth, with an average score of 8.9. Five out of the seven versions remained below 35% and only the third version increased to 40.1%. Although there was an 8 percent increase in the mark, the passage remained at the "frustrational" level and was therefore beyond the pupils' ability.

Result of Cloze Test II (Junior Lyceum)

The scores achieved in the cloze test by Junior Lyceum girls were higher than those obtained by the girls attending area schools. The global score was 17.1 when verbatim scoring was used. The girls found the fourth version, with an average score of 15.6, as the least facile, and the sixth version, with a mean score of 18.7 as the easiest.

All the versions obtained a percentage mark greater than 45%, with the highest being equal to 53.5%. These marks are beyond the 40%-45% criterion but slightly belwo the 55%-60% criterion. The 40% to 45% score range indicates that the text is suitable for study at the instructional level. As expected girls of a higher ability have a better knowledge of English and are capable of understanding the passage, even though the readability formulae predicted a higher reading age level.

The girls in the study have a wider vocabulary range than those in area schools and this is evident from the scores obtained when the cloze tests were marked using the appropriate word technique. The average score increased to 22.5. The easiest version was the sixth, with an average of 25.1, whereas the first version obtained a mean of 19.8. The percentage marks obtained in the cloze test by the second method of scoring, showed a 10% increase in all the versions, with the sixth version obtaining a very high 71.8%. The marks of all seven versions were highr than 55%. Such percentages indicate that even though the pupils did not write the author's words, they were able to supply an alternative which fitted the context. This result shows that these girls can comprehend such and similar texts with little or no help from the teacher.

Reliability and Validity

The reliability of the cloze procedure was estimated by the Kuder-Richardson formula 20 (KR-20). The results were considered separately as cloze test I (area schools), cloze test II (area schools) and cloze test II (Junior Lyceum). Scores obtained by the exact word method were applied such that each item was either correct or wrong. Reliability of cloze test II (Junior Lyceum) was also calculated on the scores obtained by the method of contextual appropriateness. As each cloze test had seven versions, reliability of every version was determined because of possible fluctuations within the same cloze test.

The cloze test on the selected passage from *This is Science 1* shows a very high reliability, r = 0.851 (range r = 0.811 to r = 0.919). The high reliability coefficient indicates that the cloze test is consistently measuring the readability of the text and that the results obtained by the girls in Form I are a true picture of the situation in the area schools.

The reliability of the cloze test on the passage chosen from *This is Science 2*, for girls attending both the area schools and the Junior Lyceum, is lower than that of cloze test I. The scores of cloze test II (area schools) gave an average reliability of 0.733 (r ranging from a low 0.599 to a very high 0.913). The low reliability could have resulted from factors such as lack of interest, low motivation and boredom because the cloze test proved to be very difficult to comprehend.

Assuming lack of interest is not a problem in the case of girls attending the Junior Lyceum - especially the ones in the top stream - the low reliability coefficient (r = 0-689) obtained in cloze test II could have been due to the homogeneity of the group. Ebel (1965) says that a group having a wide range of ability will give a higher reliability coefficient than one which is more homogeneous in ability. Sample number could also have influenced reliability.

Reliability coefficient increased to 0.812 when scores obtained by the contextually appropriate method were used in item analysis. With the exception of one version, all the others gave a higher and more realistic reliability. It seems that with second language speakers, the contextually appropriate method is a better way of testing the pupils' level of comprehension of a text.

Concurrent validity of the cloze tests was determined as a correlation coefficient by finding the extent of relation between performance in the cloze test and annual examination results in English and Science. Cloze tests I and II showed a higher correlation with English ($r = 0.717$ and $r = 0.665$) than with Science ($r = 0.656$ and $r = 0.622$). The decrease in validity of cloze test II may have been due to the sample itself, who found this cloze test beyond their ability and hence showed low motivation and interest.

This study has shown that the cloze procedure can be a reliable and valid method to investigate the readability of a text and the pupils' level of comprehension of English. The results were corroborated by a different study carried out on a small scale in which Form 2 pupils in an area school and a Junior Lyceum had to explain certain scientific and non-scientific words from their science textbook.

Conclusions and Recommendations

The potential value of a book in the classroom is determined by at least three groups of factors: factors associated with the text itself such as the extent to which it is written in a simple and clear way, and legibly and attractively produced; the ability of the children to read and understand; and how the teacher presents the text within a lesson (Harrison, 1979). This study has shown that, although the author of *This is Science* says that the books are aimed for pupils of average and below-average ability, it appears that Maltese female pupils in area secondary schools are facing serious problems because of the language used. The texts are presenting difficulties which are preventing the pupils from understanding and learning science from them. As the sample in the study are all girls, it would be interesting to find out if boys too are having these problems and whether sex differences exist.

Science teachers may be well aware of the limitations language puts on teaching and learning of science. However, they need to be more alert to the problems caused by these and other books, worksheets and handouts. When the pupils have to tackle work on their own, they will not show any progress unless they can fully comprehend what they are asked to do. Also, if the pupils' intrinsic

motivation is low, providing books which have a high level of prose difficulty is more likely to lead to non-comprehension and frustration. If teachers wish to encourage a greater use of printed sources in science, easier texts must be used to contribute to independent learning.

A suggestion which could be followed is to reduce the emphasis made on English. At present, English is used in textbooks, worksheets and national examinations. However, pupils of below-average ability find this too demanding. Perhaps, it would be better if the science curriculum for the less able pupils - including textbooks, handouts and setting of exams - is planned and taught in Maltese. Having a textbook written in Maltese by local authors could cater for pupils in the secondary schools because local examples and situations could be included together with examples related to or present in our island. Vocabulary, sentence structure and other factors which make a book readable should be taken into consideration. This may mean a lot of preparation and a change in approach. However, it may also be the solution to the problems that are being faced by the less intelligent pupils. After all science teachers should primarily be concerned with the development of scientific concepts rather than with pupils' language difficulties.

Notes and References

1. Reading Ease Score = 206.835 - (0.846 × SYLLS/100W) - (1.015 + WDS/SEN) where SYLLS/100W = average number of syllables per 100 words
   WDS/SEN = average number of words per sentence.
2. For further details on the sample number according to school and the number of respondents for each version, refer to: Sollars, A., - *Readability of Science Textbooks in Forms I and 2*, pp. 39-41.
3. For a detailed picture of the results obtained in the separate versions of each cloze test, refer to: Sollars, A. (*op. cit.*), pp. 58-68.

Bormuth, J.R. (1968) "Cloze Test Readability: Criterion Reference Scores", *Journal of Educational Measurement*, 5, 3, pp. 189-196.
Carré, C. (1981) *Science*, London: Wardlock Education.
Dobson, K. (1980) *This is Science: Teachers' Guide*, London: Macmillan.
Dobson, K. (1981a) *This is Science 1*, London: Macmillan.
Dobson, K. (1981b) *This is Science 2*, London: Macmillan.
Ebel, R.L. (1965) *Measuring Educational Achievement*, Englewood Cliffs, N.J.: Prentice Hall.
Fry, E. (1977) "Fry's Readability Graph: Clarifications, Validity and Extension to Level 17", *Journal of Reading* 21, 3, pp. 242-252.

Harrison, C. (1979) "Assessing the Readability of School Texts" in Lunzer, E. and Gardner, K. (eds) *The Effective Use of Reading*, London: Heinemann Educational for the Schools Council.

Harrison, C. (1980) *Readability in the Classroom*, Cambridge: Cambridge University Press.

Hart, W. (1980) "Reading in Science" in Prestt, B. (ed) *Language in Science*. UK: The Association for Science Education.

Klare, G.R. (1974-75) "Assessing Readability", *Reading Research Quarterly*, **1**, 10, pp. 62-102.

Lunzer, E. and Gardner, K. (eds) (1979) *The Effective Use of Reading*, London: Heinemann Educational for the Schools Council.

Oller, J.W. (1972) "Scoring Methods and Difficulty Levels for Cloze Tests of Proficiency in English as a Second Language", *The Modern Language Journal,* **56**, 3, pp. 151-158.

Oller, J.W. (1979) *Language Tests as School*, London: Longman.

Reid, D.J.; Briggs, N. and Beveridge, M. (1983) "The Effect of Picture upon the Readability of a School Science Topic", *British Journal of Education Psychology*, **53**, pp. 327-335.

Sollars, A. (1987) *Readability of Science Textbooks in Forms 1 and 2*, University of Malta.

Ventura, F. (1985) *Curriculum development: A case study from Science*, Mimeo.

Wainman, H. (1979) "Cloze Testing of Second Language Learners", *English Language Teaching Journal*, **33**, 2, pp. 126-132.

Walker, N. (1980) "Readability of College General Biology Textbooks: Revisited", *Science Education*, **64**, 1, pp. 29-34.