Further Progress Towards a Standardised Curriculum-based Measure of Reading: calibrating a new passage reading test against the New South Wales Basic Skills Test

ALISON MADELAINE & KEVIN WHELDALL*, Macquarie University Special Education Centre, Macquarie University, Sydney, Australia

ABSTRACT Results are presented from a pilot study and a larger scale study aligning a curriculum-based passage reading test, the Wheldall Assessment of Reading Passages (WARP), with the literacy components of the New South Wales Basic Skills Test (BST). A strong relationship between the WARP and BST Literacy is demonstrated, particularly for Year Three students. Representative (approximate) norms for students in Years One to Five on the WARP are reported, based on the results from a school with a BST literacy profile that is very similar to that for the state of New South Wales as a whole.

Introduction

Previous research in the area of curriculum-based measurement (CBM) has identified reading aloud measures as the measures of reading progress that correlate most highly with standardised reading tests including tests of reading comprehension (Deno, Mirkin, & Chiang, 1982). These results have been replicated in other studies, strengthening claims for the validity of curriculum-based passage reading tests (PRTs) as a means both of measuring reading competence and of indexing progress in reading (Deno et al., 1982; Fuchs, Fuchs & Maxwell, 1988; Parker, Hasbrouk & Tindal, 1992; Shinn, Knutson, Good, Tilly & Collins, 1992).

A curriculum-based passage reading test typically requires students to read a grade level passage from a basal reader for one minute. The number of words read correctly

* We would like to thank the principals, staff and students of the two schools involved in the research reported in this paper. We would also like to thank Dr Terry Burke, formerly Deputy Director General of the New South Wales Department of Education and Training, for his permission to use the Basic Skills Test data in our research.
in that minute is the index of student reading performance. For a detailed review of the literature on curriculum-based measurement of reading, see Madelaine and Wheldall (1999).

Traditionally, grade level passages from basal readers have been used to generate PRTs. Several authors have discussed the disadvantages of this method, however, which include variations in readability (between basal reading schemes, books at the same level in the same scheme, and within books), possible student familiarity with passages and the fact that passages taken from books represent only part of a story (which can affect comprehension) (Fuchs & Deno, 1994; Hasbrouk & Tindall, 1992; Mehrens & Clarizio, 1993; Wheldall & Madelaine, 1997). It should also be noted that using passages strictly from grade level text necessitates a changing metric from year to year since the number of words typically read correctly per minute will obviously be different at different levels of grade level text.

More recent research has suggested that CBM need not necessarily be literally based in the actual curriculum per se and that the use of a set of generic passages might overcome some of the disadvantages identified above (Fuchs & Deno, 1994; Wheldall & Madelaine, 1997). Moreover, such a set of passages, carefully written so as to be at the same (or very similar) level of difficulty, would also allow reading progress to be tracked over a number of years on the same scale or metric. This is particularly important when working with low-progress readers.

The development of the Wheldall Assessment of Reading Passages (WARP) (Madelaine & Wheldall, 1998; Wheldall, 1996; Wheldall & Beaman, 2000; Wheldall & Madelaine, 1997; Wheldall & Madelaine, 2000) constitutes an attempt to meet the need for such a test. The WARP is a relatively new reading assessment instrument comprising a series of specially written passages, each passage being exactly 200 words in length. Although the WARP passages are not curriculum-based measures in the strong sense of CBM, they do sample the broader curricular domain of reading. As Arthaud, Vasa and Steckelberg (2000) put it, CBM uses “materials drawn from the students’ curriculum, or materials of comparable difficulty that are similar to the curricular materials used in daily instruction” (p. 206–7).

The main aim of the WARP is to provide a means of tracking and monitoring the performance of older low-progress readers toward functional literacy, typically achieved at around age ten and a half years (Wheldall & Madelaine, 2000). Previous research on the WARP has established the reliability and validity of the passages (Madelaine & Wheldall, 1998; Wheldall & Madelaine, 2000). Madelaine and Wheldall (1998) reported parallel forms reliability of 0.94–0.96 and criterion validity coefficients of 0.83–0.87 with Neale Analysis—reading accuracy and 0.67–0.72 with Neale Analysis—reading comprehension. Wheldall and Madelaine (2000) report both high parallel forms reliability (0.94–0.96) and internal consistency (0.97–0.99), and good criterion validity (0.78–0.80) with Neale Analysis—reading accuracy. In order to use the WARP to measure reading achievement and to track reading progress, some tentative norms, preferably grade-based, need to be established.

Previous Attempts at Establishing Norms for Oral Reading Fluency

There have been several attempts to establish norms for oral reading fluency in order to provide educators with an indication of how far above or below ‘the average’ an individual student’s oral reading fluency may be said to be (e.g., Hasbrouk & Tindal, 1992; Marston & Magnusson, 1988; Shinn, 1988). These studies have been problem-
atic in that the curriculum content (text passages) accessed was not constant across all schools and, hence, are open to the same criticisms of using passages from basal readers as discussed earlier. It could be argued that differences in reading curriculum would be controlled for over so many schools, that any such differences in text level would effectively even out. Even if this were to be the case, (and it is more optimistic hope than verifiable established fact) employing the norms established from these global data, based on a multiplicity of passages, as representative norms for the specific set of passages used in a particular school is a risky and unwarranted extrapolation. There may also be problems associated with the representativeness of the samples of students employed in these studies, however large they were.

An alternative approach, however, is to use the same set of generic passages for all students. Of course, the norms resulting from this process will only reliably relate to this specific set of passages of equal difficulty level, as is the case with the WARP. Preliminary norms have been reported for primary school students’ oral reading rate on the WARP (Madelaine & Wheldall, 2001). Results were based on the testing of over 1000 children in New South Wales schools. Data indicated that children’s reading scores (number of words read correctly per minute) increased as a function of age, from a mean of 84 words per minute in Year Two to a mean of 139 words per minute in Year Six, where a ceiling effect began to become apparent.

The New South Wales Basic Skills Test

Another approach to reading assessment is included within the New South Wales Basic Skills Test (BST) (New South Wales Department of Education and Training, 2000) which comprises tests of literacy (reading and language) and numeracy (number, space and measurement). The BST is administered to all primary school students in state schools and to many students in Catholic sector and independent sector schools in Years Three and Five in August of each year. The literacy component tests language (including spelling, punctuation and grammar), and reading. The reading component tests student’s understanding of a range of written texts used in the primary key learning areas (New South Wales Department of School Education, 1997).

The BST aims to provide feedback to parents, teachers and schools in a number of areas, including: individual children’s achievements in the areas of literacy and numeracy, their performance in comparison to state norms, identification of those students in need of additional assistance, as well as providing a means of monitoring literacy and numeracy standards over time (New South Wales Department of School Education, 1997). BST results are reported using the same scale for Year Three and Year Five students, with Year Three students scoring in bands one to five and Year Five students scoring in bands one to six. In 1997, the then New South Wales Minister for Education and Training, the Hon. J. J. Aquilena, commented, “There is no parallel in Australia or elsewhere, to date, for the quality of the tests, the amount of information provided and the size of the test population.” (New South Wales Department of School Education, 1997, p. 3).

If performance on the WARP were to be shown to be related to performance on the BST, a potential means of calibrating the WARP becomes apparent. Moreover, if a sample of students closely representing the state parameters on BST were to be identified, then it is not unreasonable to assume that such a sample would be likely to be representative of the state as a whole for WARP performance.
As previously noted, Madelaine and Wheldall (2001) provide tentative norms for the WARP. Although the students in the sample of over 1000 came from many different schools and could be considered to be representative, there is no way of knowing, however, just how representative they were of the state as a whole.

This paper reports the results of a pilot study involving two schools and a subsequent major study involving one of the two pilot study schools. The purpose of these studies was to align the WARP with the New South Wales Basic Skills Test and to provide more representative approximate norms for the WARP.

Pilot Study

Method

A pilot study was conducted in 1999 involving two primary schools, one in an inner city working class area with a relatively high proportion of students from non-English speaking backgrounds, and one in a middle class suburban area. The aim of this simple study was to provide data to establish the relationship between performance on the WARP and performance on the BST, prior to the completion of the major study. The sample comprised all students who were present on the days the tests were administered.

Sample and Procedure

The sample of students from School One, a Sydney Catholic Primary school, comprised 283 students in Years One to Five. The sample from School Two, an independent school in Sydney, comprised 273 students in Years One to Five.

Students in both schools were administered two WARP passages by trained data collectors. As some of the students at School Two had recently been assessed on the set of five WARP passages employed in previous studies (Madelaine & Wheldall, 1998; Madelaine & Wheldall, 2001; Wheldall & Madelaine, 2000), two other passages of approximately the same difficulty level were chosen for this pilot study. At approximately the same time (within two weeks), students in Years Three and Five at both schools were assessed on the BST in their schools as part of the state-wide initiative.

Instruments

Wheldall Assessment of Reading Passages (WARP) (Wheldall, 1996). The WARP is a new curriculum-based measure of reading comprising (currently) 21 passage reading tests of 200 words each (although only two passages were administered in the pilot study). A sample passage is reproduced as Appendix A. The instructions for administration are reproduced as Appendix B. Scores indicate the number of words read correctly in one minute (WPM) averaged, in this case, over the two WARP passages administered.

The New South Wales Basic Skills Test (BST). The BST is completed in August (Term Three) each year by all Year Three and Year Five students (with a few exceptions such as students with severe learning difficulties) in all State schools and additionally by
students in many Catholic and independent schools. This study is only concerned with the BST measures of literacy (language and reading). The BST results for literacy are reported using several measures. An overall score is reported for literacy, with the reading and language raw scores representing the number of correct responses on each part of the test. In addition, skill bands reflect:

... abilities, knowledge and understanding grouped according to increasing complexity, difficulty and utility. The higher skill bands, five and six, represent the more difficult skills. Each skill band therefore represents, at a standard of achievement, a group of skills which the student has been able to demonstrate. (New South Wales Department of School Education, 1997, p. 11)

Results and Discussion

The two WARP passages were correlated at 0.98 in School One and at 0.97 in School Two, indicating excellent alternate forms reliability.

Of the total of 113 Year Three students, from the two schools combined, there were complete data on 109. Only the results for the students for whom complete data sets exist are reported. WARP score, averaged over the two passages, correlated with BST literacy overall at 0.69 and with the BST Literacy Bands at 0.64. The WARP mean score correlated slightly more highly with the BST Language raw score (0.69) than with the BST reading raw score (0.59).

Similarly for Year Five, of the 105 students from both schools, there were complete data sets on 102, on whom these analyses are based. The WARP, averaged over the two passages, correlated with BST literacy overall at 0.73 and with the BST literacy bands at 0.77 and again correlated slightly more highly with the BST language raw score (0.66) than with the BST reading raw score (0.61). These results confirm, as expected, a strong relationship between performance on the WARP and performance on the BST (Literacy).

The performance of School One was noticeably lower than that for School Two on both the WARP and the BST measures for the Year Three students (WARP: $t = 2.49$, $p < 0.02$; BST: $t = 3.51$, $p < 0.001$). Significant differences were not found on either measure at Year Five level, however ($t = 0.66$, $p > 0.5$ for the WARP and $t = 0.26$, $p > 0.5$ for the BST). More importantly, on examination of the performance of School One on the BST compared with the results for the state as a whole, it was apparent that the performance of School One, at both the Year Three and Year Five levels, was very similar to that for the state as a whole.

The mean (and standard deviation) for literacy for School One in Year Three ($N = 55$) was 49.9 (SD = 6.4) compared with 50.3 (SD = 6.4) for the State as a whole ($N = 60,922$). For Year Five ($N = 53$) the mean overall literacy score was 58.4 (SD = 6.7) compared with 56.6 (SD = 6.4) for the State ($N = 58,537$). An examination of the comparable data for the previous year (1998) confirmed the similarity of this school’s performance with that of the state as a whole at both the Year Three and Year Five levels.

As predicted, the WARP was shown to be strongly related to the BST. Moreover, the performance of School One students on the BST was shown to be highly representative of the performance of students in the state as a whole. The results from this pilot study prompted a more extensive study involving students in Years One to Five from School One, the following year (2000).
Main Study

Method

This study was completed as part of a larger data collection exercise which will form the substance of a subsequent paper.

Participants. The participants comprised 261 students in Years One to Five from the same Catholic Primary school in Sydney as in the pilot study who were present on the days that testing took place. The numbers of students in each grade were as follows: 37 in Year One; 47 in Year Two; 65 in Year Three; 56 in Year Four; and 56 in Year Five. The sample had a mean chronological age of 106 months (SD = 16.13) and comprised 52% girls and 48% boys. The performance of this school on the BST had already been shown to be very similar to that of the state of New South Wales as a whole, for Year Three and Year Five, in both 1998 and 1999, as demonstrated above.

Given the similarity of the school’s performance on the BST to the state over a number of years, it is reasonable to assume that the students in each year at School One are likely to be representative of the state at each grade level. The results for Year One should be treated with a degree of caution, however. The sample is smaller (N = 37) than that at other grades (N = 47–56) due to Year One in this year (2000) having a greater number of students experiencing difficulties with English who were, as a result, excluded by their teachers from the WARP testing.

Procedures. The same instruments were used in this study as were used in the pilot study. Data collectors were trained in procedures for administering the WARP (see Appendix B for procedures). In total, 21 WARP passages were administered in random order over two sessions on separate days, 11 in the first session and ten in the second. The WARP was administered on an individual basis according to the procedures for administering passage reading tests (See Appendix B). At approximately the same time (within two weeks), the BST was undertaken by students in Years Three and Five during August 2000. Note: Three Year Three students and one Year Five student were not present on the day the BST was given.

Although 21 WARP passages were administered, data for three passages (only) will be reported in this paper. These three passages have been identified as the basal WARP measure and are a subset of the original five passages upon which much of the previous research on the WARP is based (Madelaine & Wheldall, 1998; Madelaine & Wheldall, 2001; Wheldall & Madelaine, 2000). The administration of 21 passages was carried out to identify a further 10 passages of similar difficulty level to the three basal passages to allow weekly progress monitoring over one term. These findings are outside the scope and remit of the present paper and will be reported separately in a subsequent paper.

Results and Discussion

The three WARP passages all intercorrelated with each other at 0.97, again indicating excellent alternate forms reliability.

Relationship between WARP and BST. For Year Three students, the WARP correlated with BST literacy overall at 0.85 and with the BST Literacy Bands at 0.82. The correlations between the WARP and the BST component literacy scores were 0.78 for the language raw score and 0.82 for the reading raw score.
TABLE I. Mean WARP Scores for each BST Band–Literacy (n)

<table>
<thead>
<tr>
<th>Band</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Year 3</td>
<td>32 (2)</td>
<td>65 (17)</td>
<td>87 (20)</td>
<td>111 (10)</td>
<td>138 (13)</td>
<td>N/A</td>
</tr>
<tr>
<td>Year 5</td>
<td>110 (16)</td>
<td>130 (20)</td>
<td>139 (12)</td>
<td>176 (7)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Years 3 and 5</td>
<td>32 (2)</td>
<td>65 (17)</td>
<td>97 (36)</td>
<td>124 (30)</td>
<td>139 (25)</td>
<td>176 (7)</td>
</tr>
</tbody>
</table>

The data for Year Five also indicated a good relationship between the WARP and BST, with correlations of 0.66 between the WARP and the BST literacy overall, 0.62 between the WARP and the BST literacy bands, 0.61 between the WARP and the BST language raw score, and 0.54 between the WARP and the BST reading raw score. These lower correlations at Year Five level may be indicative of ceiling effects becoming evident on either or both variables (WARP and/or BST) for some higher achieving students. Whatever the case, the WARP does not predict BST performance at Year Five as well as at Year Three.

Analysis of WARP means by BST band gives an indication of typical WARP performance for each band. As expected, WARP scores increase progressively as band level increases, both for Year Three and Year Five (See Table I). Data such as these could be used to predict how students in Years Three and Five are likely to perform on the BST. The mean WARP scores for each BST band are not always similar when the Year Three and Year Five data are compared. This is particularly apparent in the data for bands three and four, and is not consonant with the claim that the BST uses the same scale for Years Three and Five. These data, however, should be treated with a degree of caution as the WARP means are based on small group sizes, as can be seen in Table I, and (as already noted) WARP predicts BST band less well at Year Five level (0.62) than at Year Three level (0.82). Table I also provides WARP mean scores for each literacy band averaged over the total sample of Year 3 and Year 5 students, given the claim that the same scale is employed at both ages, yielding estimates based on larger band sample sizes for half of the cells. (The correlation between WARP and BST literacy overall for the combined sample was 0.81, and 0.79 for BST literacy band.)

Approximate WARP norms by grade/year. The sample of students at School One was again shown to be representative of the State population as a whole, in some important respects. While the school is a Catholic sector school as against a state school, with a larger population of students who may be categorised as NESB, only one student from Years Three and Five, tested in 2000, had “lived in Australia for four years or less and never or only sometimes speak English at home” (New South Wales Department of Education and Training, 2000). More importantly, the average scores for the school on BST literacy (and numeracy for that matter) were very similar to State averages at both Year Three and Year Five levels in 2000 and have consistently been so in recent years, as previously shown.

The mean (and standard deviation) for BST literacy for this school in Year Three (N = 65) was 50.2 (SD = 6.1) compared with 50.0 (SD = 6.5) for the State as a whole (N = 60,499). Similarly, for Year Five (N = 58) the mean overall literacy score was 55.3 (SD = 5.0) compared with 56.2 (SD = 6.4) for the State (N = 59,783). For Year Five (only), separate figures are provided for Reading and Language as sub-components of literacy; again the school and State means are similar for both measures: Reading—school 54.5 (SD = 5.2) compared with 56.4 (SD = 6.9) for the State; Language—56.6
TABLE II. Mean age and mean WARP scores by grade

<table>
<thead>
<tr>
<th>Grade</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>n</td>
<td>37</td>
<td>47</td>
<td>65</td>
<td>56</td>
<td>56</td>
</tr>
<tr>
<td>C.A (months)</td>
<td>M  82</td>
<td>92</td>
<td>104</td>
<td>116</td>
<td>128</td>
</tr>
<tr>
<td></td>
<td>SD  (4.65)</td>
<td>(3.82)</td>
<td>(3.24)</td>
<td>(3.63)</td>
<td>(3.75)</td>
</tr>
<tr>
<td>WARP mean M</td>
<td>37</td>
<td>78</td>
<td>93</td>
<td>121</td>
<td>130</td>
</tr>
<tr>
<td></td>
<td>SD  (24.90)</td>
<td>(38.98)</td>
<td>(35.10)</td>
<td>(42.40)</td>
<td>(33.72)</td>
</tr>
<tr>
<td>50% range limits (WPM)</td>
<td>16–56</td>
<td>45–104</td>
<td>66–111</td>
<td>94–155</td>
<td>107–154</td>
</tr>
</tbody>
</table>

(SD = 5.7) for the school compared with 56.4 (SD = 7.2) for the State. A similar picture obtained in previous years, as we have already noted.

These latest results confirm that the literacy performance of students at this school is very similar to that of students in the state as a whole. This being the case, a representative set of norms for the WARP may be approximated, since a strong relationship between the WARP and the BST has been established in both the pilot study and the main study at both Year Three and Year Five levels. Table II shows the mean WARP scores for each grade. The data show a steady progression from Year One to Year Five, with a possible ceiling effect becoming evident in the results for Year Five (see Fig. 1). The mean WARP score for Year One may not necessarily be representative due to some students being excluded from the testing (as explained above). Table III also provides 50% range limits for each grade. These have been determined by calculating the scores that lie at the first and third quartile for each grade and may be particularly useful for identifying the bottom 25% of students in each grade.

Concluding Discussion

Taken together, the results from both the pilot and the main study provide convincing evidence of a strong relationship between WARP performance and performance on the BST at both Year Three and Year Five levels. The relationship is particularly strong between the WARP basal measure and overall BST literacy performance in Year Three ($r = 0.85$). This is not unexpected since BST literacy measures reading performance.

![Fig. 1. Performance on the WARP by grade.](image)
and the WARP has been shown consistently to correlate highly with other measures of reading in validity studies (Madelaine & Wheldall, 1998; Wheldall & Madelaine, 2000). These latest findings constitute, in effect, further evidence for the validity of the WARP as a measure of literacy, specifically reading, performance. Consequently, the WARP lends itself for use by teachers as an approximate measure of likely performance on the BST Literacy.

School One has been shown to demonstrate, repeatedly over at least three years, a BST literacy profile that is very similar to that for the state of New South Wales as a whole. Given the strong relationship between the WARP and the BST, this means that the results for this school may serve as a reasonable proxy for the State for performance on the WARP and thereby providing approximate representative norms. This logic has been employed to derive the approximate norms for each grade detailed in Table II and shown graphically in Fig. 1.

Particular confidence may perhaps be attached to the results for Year Three since the BST and the WARP were administered at School One at about the same time (August 2000) and the relationship between WARP score and BST overall literacy performance was particularly strong \( r = 0.85 \). Thus we may conclude that the average WARP performance of students who are over half way through Year Three is likely to be about 93 words read correctly per minute. Year Three students who produce a score of 65 or lower are likely to be in the bottom 25% of students for reading and are, hence, perhaps likely to be of particular concern to teachers. Putting it another way, Year Three students who score around 90 words read correctly per minute are likely to achieve in band three on BST Literacy whereas students who score 65 or lower are likely to be placed in band two or band one.

Similarly, the average WARP performance of students who are over half way through Year Five is likely to be about 130 words read correctly per minute. Students who produce a score of 106 or lower are likely to be in the bottom 25% of students in terms of reading performance. Moreover, if functional literacy is typically achieved at about ten and a half years old, or about half way through Year Five (Wheldall & Madelaine, 2000), then a WARP score of about 130 words read correctly per minute is indicative of functional literacy, an appropriate long term goal for low-progress readers, for example.

In this paper, we have sought to calibrate the WARP against the New South Wales Basic Skills Test and to use the data from a school performing typically for the State as a whole (at both Year Three and Year Five levels) to generate approximate norms for use by teachers and others when evaluating the performance of individual students. Data collected contemporaneously with the data informing the results presented here may serve to extend further the utility of this model of curriculum-based assessment by providing a series of ten additional WARP passages. These ten passages are of roughly equal difficulty level with the basal WARP measure and are also very highly correlated with the basal measure. By administering one of these ten progress passages at weekly intervals over a term, and graphing the results, performance toward the goal of functional literacy may be effectively tracked.

Correspondence: Professor Kevin Wheldall, Director, Macquarie University Special Education Centre, Macquarie University, Sydney, NSW 2109, Australia (e-mail: kevin.wheldall@mq.edu.au).
REFERENCES


Appendix A
A Sample Passage from the WARP
Last week Dad decided that buying things cost too much money. “We could make all sorts of things if I had a workshop,” he said. “Just a little one with a workbench and a few tools.”

Mum looked worried. She did not think it was a very good idea. “Great idea, darls. The only problem is that good tools cost a lot of money.” And Mum smiled. She thought that would be the end of Dad’s idea but she was wrong.

“No problem, love,” said Dad. “My friend has a shed and some tools that he wants to sell. He’ll let me have the lot for just three hundred dollars.”

Mum gulped. What could she try next?

“Why not go and get the shed and the tools right now,” she said. “Then you could fix the toaster and the dryer. And then why not look at the back gate. It keeps sticking.”

Mum went on and on. Dad looked a bit sick and then he gave in. “Perhaps not after all, Darls. There’s a really good show on the television tonight. Maybe next week.”

Mum smiled. She knew that she would not hear about the shed or the tools again.
© Kevin Wheldall, 1996.
Copying prohibited without permission.

Appendix B
Directions for Administering and Scoring the WARP Passages
Give the student the copies of the appropriate WARP passages that do NOT have the numbers on the right hand side of the sheet. The student sheet should not have a passage title. Do not tell the student the title of the passage.

Ask the student to read the passage:

“I’d like you to read this passage as quickly and as carefully as you can. I’ll start recording when you read the first word.”

Begin timing when the student says the first word in the passage. Mark the last word read at the end of one minute and ask the student to stop reading. If a student stops or struggles with a word, wait three seconds, supply the word and mark one error.

Scoring Procedures
Use the number column to assist in calculating the number of words read in one minute. To calculate the number of words read correctly in the first minute (WPM), subtract the total number of errors made in the first minute from the total number of words read in the first minute.

If a word is omitted, score 1 (one) error. If several consecutive words are omitted (e.g. a line is skipped), subtract the total number of words omitted, and score 1 (one) error. If a word is inserted, do not count the inserted word, and score 1 (one) error. If the word order is reversed, score 1 (one) error.