

4. Final test

4.1 Test concept

An assessment of the reading skills of Improved Reading course participants several months after they attended the course serves four purposes.

1. To ascertain whether they can still achieve the performance that was measured at the end of the course, the participants are given an English text that is similar in degree of difficulty to the texts read during the course. (The results are shown in Section 4.2)
2. To ascertain whether the skills learned in the course can be applied to German texts, a text is selected which corresponds in degree of difficulty to the texts read in the course. (Please refer to Section 4.5.)
3. To ascertain whether the skills learned in the course can be applied to texts with a higher degree of difficulty, another comparative German text with a higher level of grammatical complexity is selected. A comparison is then made of the WPM and ERR rates achieved for both German texts. (The results are shown in Section 4.3)
4. To more precisely establish the performance of Improved Reading course participants compared with non-course participants, the two German texts that were used to ascertain WPM and ERR rates are given to two control groups consisting of test subjects with no knowledge of techniques to improve reading efficiency. The control groups are also asked to read the English text so that they are familiar with the test procedure. These results were not taken into account in the evaluation. (Results are shown in Section 4.4)

Calculation of the degree of difficulty of the selected texts

The degree of difficulty of the texts was calculated in two ways:

1. The average length of the sentences was calculated by adding together the number of words and dividing them by the number of sentences.
2. The average grammatical complexity of the two German texts was established by allocating numbers to the structural elements of the sentences according to their grammatical status. The total of these numbers represents the complexity of the sentence.

The numbers were allocated according to the following scheme:

MC (main clause) = 1 (example: *the boy laughed.*)

MC linked to another MC = 1 (The boy laughed, *and his mother looked at him in amazement.*)

SC (subordinate clause) right recursive = 2 (example: The boy laughed, *because he was happy.*)

SC left recursive = 3 (example: *Because he was happy,* the boy laughed.)

SC embedded = 4 (example: The boy, *unobserved,* continued to play.)

If the SC is a matrix clause for other dependent and subordinate units, the structural elements of the sentence are allocated numerical values to reflect their position in the matrix clause.

subordinate clause = 2

right recursive clause = 5

left recursive clause = 6

embedded clause = 7

This process then continues at the next level:

subordinate = 3, 4

right recursive = 8, 11

left recursive = 9, 12

embedded = 12, 13 and so on.

Note: This numbering system makes it possible to highlight differences in the degree of difficulty of texts. However, it is impossible to draw the conclusion from a comparison of the figures that one text is a certain number of times more complex than another text, i.e. 'Text A is four times more complex than Text B'. This is not the kind of comparison that we are attempting to make. In addition to sentence length and grammatical complexity, it is also necessary to define content.

1. This means that the subjects dealt with in the texts must be of a general nature so that the test subjects can have no prior knowledge of the reading matter. (For example, a text on the subject of the influence of the press would be unsuitable because the majority of the two control groups are students of journalism and communication history, and are therefore more familiar with the subject matter of the texts than other people would be.)
2. In order to make an objective assessment of text content, the selected texts were in denotative language. Non-literary texts enable the identification of core statements and the testing of comprehension on the basis of relevant questions.
3. In contrast to the questions in the comprehension tests, the authors attempted to avoid repeating text elements in the answers. In addition to the elements of the text that the test subjects were able to recall, the tests provided information about the subjects' assimilation of knowledge from the text *and* their processing of this knowledge into applicable knowledge.

The texts (test documents are included in the Appendix)

Text 1: "Stress"

The text comprises 903 words and consists of 43 sentences. The average length of each sentence is therefore 21 words. This and several other texts were provided by Improved Reading Centres. It corresponds to the degree of difficulty of the texts used in the comprehension tests. When the texts are selected, it is necessary to ensure that no prior knowledge of the subject matter required. The text is of a general nature and deals with the phenomenon of 'stress'.

The method of questioning is the same method used in the comprehension tests. The questions related to details contained in the texts that are repeated word for word in the answers. This method was selected to test reading comprehension because it enables a comparison with the Improved Reading course participants' performance during the course. The test subjects who had not attended an Improved Reading course were also requested to read these texts. However, their results were not evaluated. The non-participants were asked to read the texts to familiarise themselves with this type of test.

Text 2: "Wie hart arbeiten die Japaner wirklich?"

This text was taken from Heidemarie Colmsan-Freyberger, Haruko Kishimoto, Peter Krebs, Susanne Krebs, Manfred Pohl: Japan. Daten - Bilder - Perspektiven. Mit einem Landesportrait by S. and P. Krebs, Munich, Lucerne 1982, p. 105f. It was slightly modified for the purpose of the test.

The text comprises 753 words and consists of 61 sentences. The average length of each sentence is therefore 12.3 words. The average sentence complexity is 3. The most complex sentence has a complexity rating of 10.

The text provides an explanation and assessment by a Japanese citizen of working hours in Japan and the attitudes to work of his fellow countrymen, based on the widespread belief in the Western world that the Japanese are very hard-working people. All test subjects are likely to be familiar with this belief. However, few are likely to be aware of what it is really like to work in Japan. The information about the working habits of the Japanese, and the anecdote at the end of the text, do not require any previous knowledge and are therefore comprehensible to all.

The reading comprehension questions relate to the core statements of the text. The multiple choice answers deviate from the wording of the text.

Text 3: "Wir und die Anderen"

The text was published under the title 'Welt in Bewegung' in Marion Gräfin Dönhoff: Welt in Bewegung. Berichte aus vier Erdteilen, Düsseldorf, Cologne 1965, p. 5f. (It was necessary to

change the title to prevent misunderstandings in questions due to the statement about the world having become a dynamic place. (Question 3. Test documents are included in the Appendix.)).

The text comprises 468 words and consists of 18 sentences. The average length of each sentence is therefore 26 words. The average sentence complexity is 12.5. The most complex sentence has a complexity rating of 61.

The text contains the comments and opinions of the author about the travel habits and lifestyles of people in the western hemisphere compared with people in Africa, Asia and the communist states. It is a very general text and therefore does not require any special prior knowledge.

The reading comprehension questions relate to the core statements made in the text. Most multiple choice answers diverge from the wording of the text. It was necessary to repeat the wording in some cases in order to clarify which of the answers was correct.

4.2 Comparison of the reading efficiency of course participants in Text 1 with the values that were achieved at the beginning and end of the course

The test papers were numbered in the order that they were handed in so that they could be evaluated anonymously.

The values stated for the test subjects' WPM and ERR rates at the beginning of the course relate to the first two comprehension tests taken during the course. Because the subjects were unfamiliar with this type of test, both the first and second tests were taken as the basis of assessment. However, the results of Test 2 are not purely baseline values because learning progress had already been made.

The WPM and ERR rates at the end of the course correspond to the Improved Reading Centres policy of determining overall course results on the basis of the results of the last comprehension test and, in some cases, the average values of the last two, three or four tests. (See section 2.1).

The WPM and ERR rates in the final test were calculated on the basis of the test subject's notes on 'time required' and 'number of correct answers' on the test paper. The test subjects were therefore not aware of whether their test results corresponded to the results that they achieved in the course.

Tab. 15: WPM and ERR rates at the beginning of the course (comprehension tests 1 and 2), at the end of the course and in Text 1 of the final test.

No. 1	Beginning of course		End of course	Final test
WPM	213	308	469	248
ERR	170	246	375	198

No. 2	Beginning of course		End of course	Final test
WPM	169	260	669	260
ERR	101	182	468	104

No. 3	Beginning of course		End of course	Final test
WPM	199	276	546	304
ERR	159	193	400	213

No. 4	Beginning of course		End of course	Final test
WPM	199	299	422	280
ERR	139	239	295	112

No. 5	Beginning of course		End of course	Final test
WPM	164	176	300	243
ERR	115	106	240	146

No. 6	Beginning of course		End of course	Final test
WPM	270	429	688	352
ERR	189	276	550	106

No. 7	Beginning of course		End of course	Final test
WPM	166	343	542	188
ERR	100	172	379	113

No. 8	Beginning of course		End of course	Final test
WPM	171	235	378	195
ERR	154	188	340	59

No. 9	Beginning of course		End of course	Final test
WPM	152	218	336	312
ERR	92	174	235	156

No. 10	Beginning of course		End of course	Final test
WPM	184	429	877	304
ERR	129	343	702	182

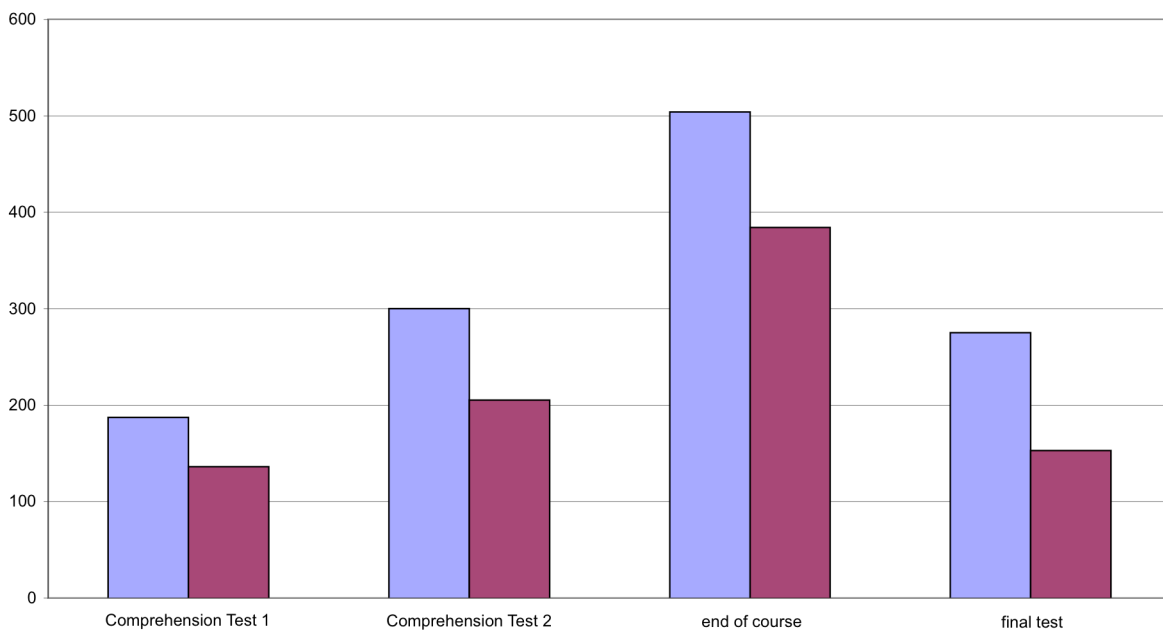
No.11	Beginning of course		End of course	Final test
WPM	209	324	456	342
ERR	167	130	410	308

No. 12	Beginning of course		End of course	Final test
WPM	164	230	403	304
ERR	98	115	281	213

No. 13	Beginning of course		End of course	Final test
WPM	169	376	472	248
ERR	152	301	312	74

The following diagram shows the average WPM and ERR rates of Improved Reading course participants who took part in the final test at the beginning of the course (Comprehension Tests 1 and 2) at the end of the course and in the final test. (Blue column = WPM; red column = ERR).

Diagram 34: Average WPM/ERR - Beginning of course, end of course, final test



It shows that the results achieved by the final test subjects were not even close to the results that they attained at the end of the course. The average WPM rates in the final test were slightly lower than the average WPM rate in comprehension test 2. The average ERR rate was slightly higher than the average rate attained in comprehension test 1.

WPM and ERR deviate to a greater extent in the final test than they do in the tests at the beginning and end of the course. This shows a poorer performance in the comprehension questions.

When interpreting this decrease in performance, it is necessary to take the comparative tests involving non-course participants that are detailed in the following sections into account. Since the test subjects who had attended the course achieved far better results than non-course participants, on the whole, in the two German texts (Section 4.4), the lower values achieved in the English texts compared to the results at the end of the course can be primarily attributed to the fact that the course was in a foreign language and the degree of difficulty of the 'Stress' text.

The fact that the two-day course, comprising 12 hours in total, was held in a foreign language led to a short-term positive effect on the participants' foreign language proficiency. Verbal communication in English and the focus on English language texts enhanced the course participants' language skills. However, if this immersion in a foreign language is not continued, the course participants not only have to accustom themselves to using the reading techniques that they learned again, but also to the foreign language.

The degree of difficulty of the texts used in the comprehension tests - as described in Section 2 - was ascertained on the basis of sentence length (words per sentence) and the number of words of more than two syllables. Unfamiliar words are a factor of influence for non-native speakers, whose vocabulary is generally less extensive than that of native speakers, and they are an obstacle to comprehension of a foreign language text. The same applies for grammatical complexity, which is not taken into account when establishing the degree of difficulty. Unfamiliar vocabulary and grammatical structures may therefore explain to some extent the decline in performance of the course participants in the final test.

4.3 Comparison of participants' Effective Reading Rate in Texts 2 and 3

A comparison of WPM and ERR rates attained by course participants in Text 2 ("Wie hart arbeiten die Japaner wirklich?") and Text 3 ("Wir und die anderen") is shown below. This comparison shows the extent to which the learned techniques can be applied to texts with a higher degree of difficulty.

Tab. 16: WPM and ERR attained by Improved Reading course participants in Texts 2 and 3

No. 1	Text 2	Text 3
WPM	430	344
ERR	351	246

No. 2	Text 2	Text 3
WPM	452	330
ERR	407	165

No. 3	Text 2	Text 3
WPM	335	330
ERR	335	264

No. 4	Text 2	Text 3
WPM	301	296
ERR	271	266

No. 5	Text 2	Text 3
WPM	258	255
ERR	206	230

No. 6	Text 2	Text 3
WPM	532	468
ERR	426	374
No. 7	Text 2	Text 3
WPM	196	244
ERR	176	122

Nr. 8	Text 2	Text 3
WPM	301	176
ERR	271	158

No. 9	Text 2	Text 3
WPM	377	330
ERR	302	264

No. 10	Text 2	Text 3
WPM	312	330
ERR	281	297

No. 11	Text 2	Text 3
WPM	393	468
ERR	314	374

No. 12	Text 2	Text 3
WPM	532	296
ERR	479	178

No. 13	Text 2	Text 3
WPM	425	351
ERR	425	176

Diagrams 35 and 36 show the WPM and ERR rates attained by course participants in Text 2 (blue rectangle) and Text 3 (pink square) in the final test.

Diagram 35: WPM of course participants in Text 2 and Text 3

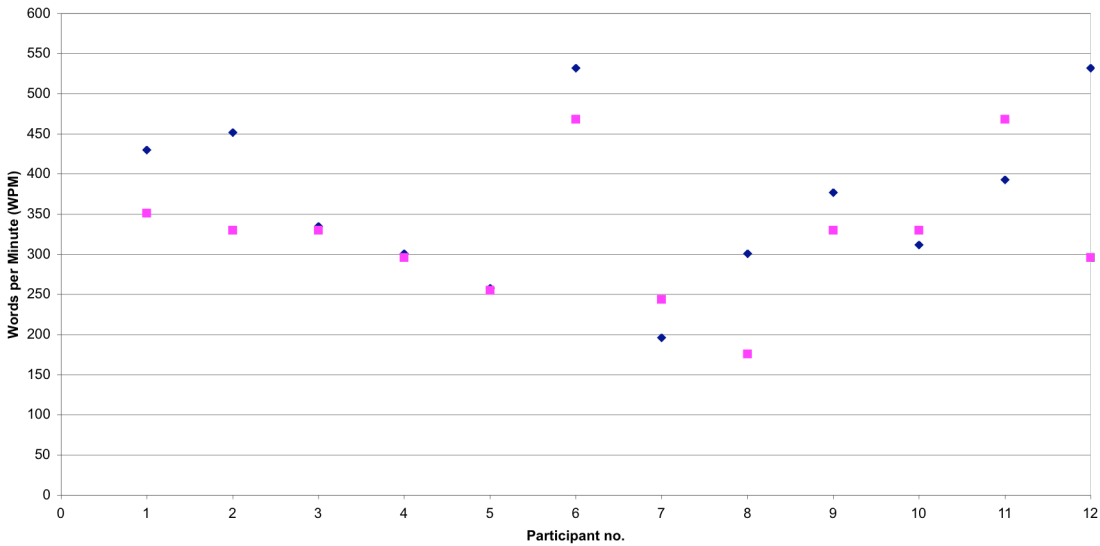
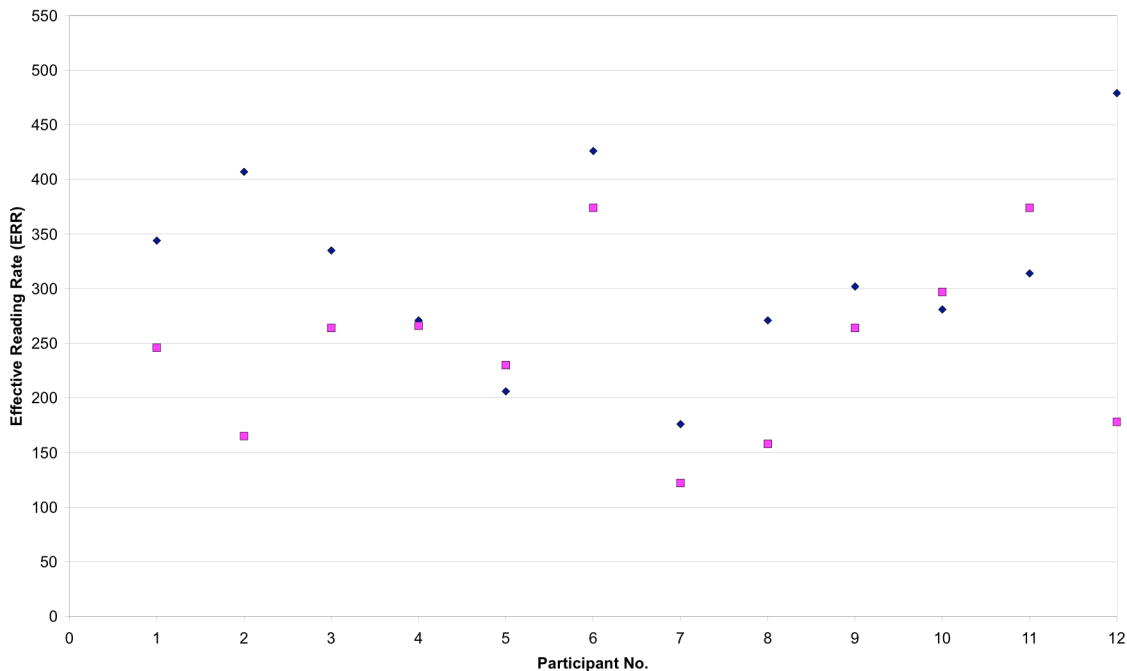


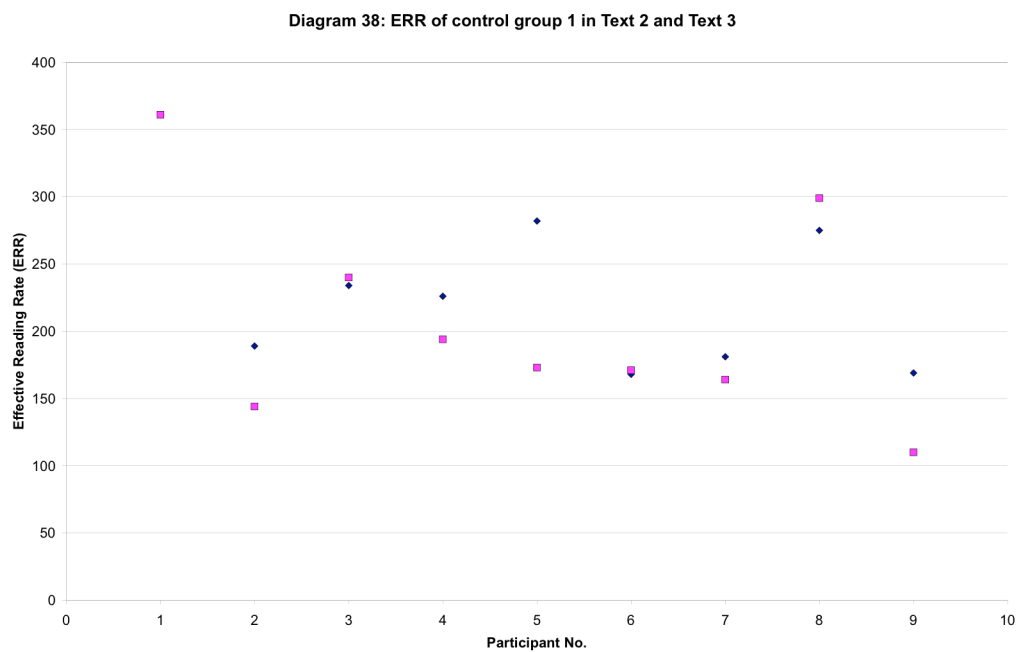
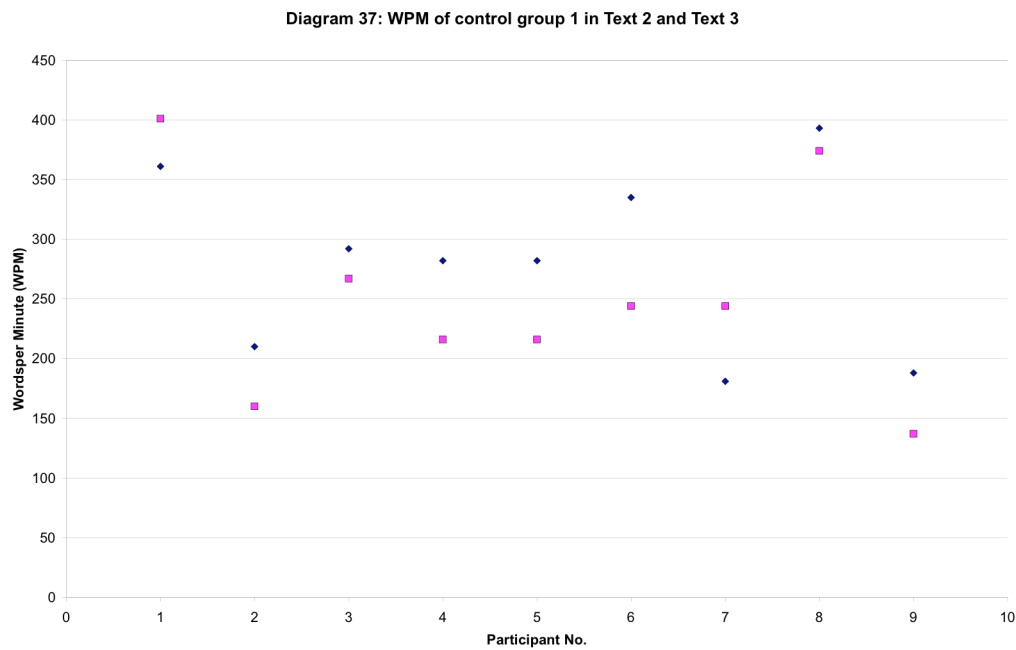
Diagram 36: Effective Reading Rate (ERR) of course participants in Text 2 and 3



A discrepancy exists in most cases between WPM and ERR in Texts 2 and 3. In the majority of cases, it is somewhere in the region of ERR 50 (or 50 WPM). In extreme cases, it is higher than 200. The results of most test subjects in Text 2 - the text with less complex grammatical structures - are higher than those for Text 3. In order to measure whether this discrepancy is due

to limitations on the use of the reading techniques for more complex texts, the performance of non-course participants in Text 2 and 3 must be taken into consideration.

Diagrams 37 and 38 show the WPM and ERR rates of control group 1 in Text 2 and Text 3 (see 4.4.1 for a description of control group 1).



Diagrams 39 and 40 show the WPM and ERR rates of control group 2 (see 4.4.2 for a description of control group 2).

Diagram 39: WPM of control group 2 in Text 2 and Text 3

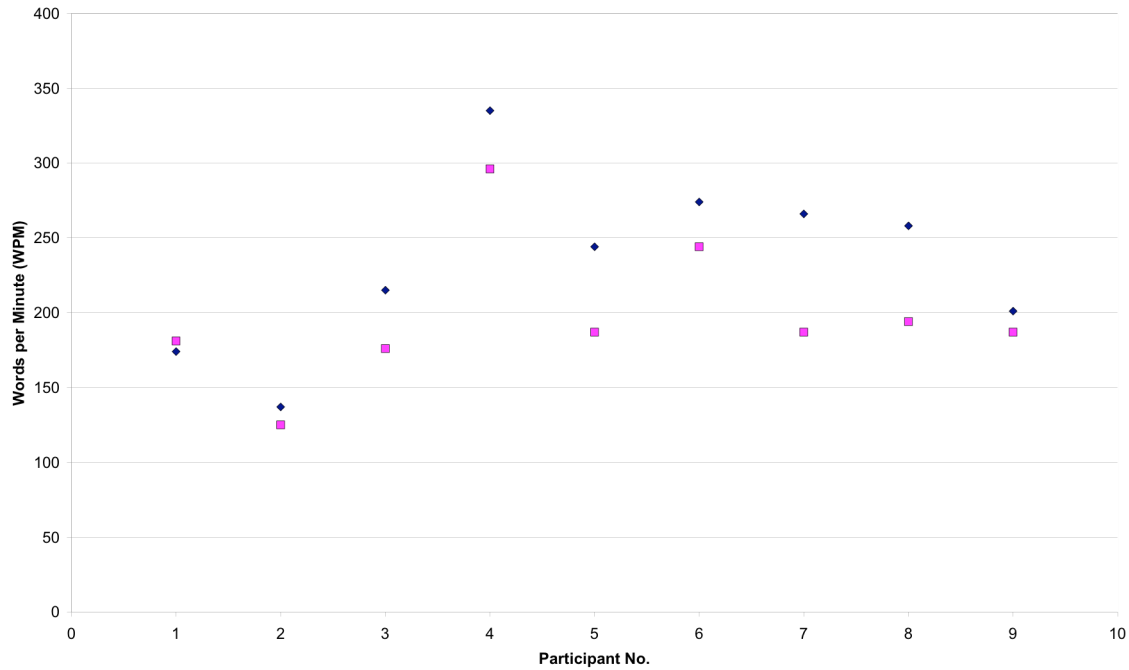
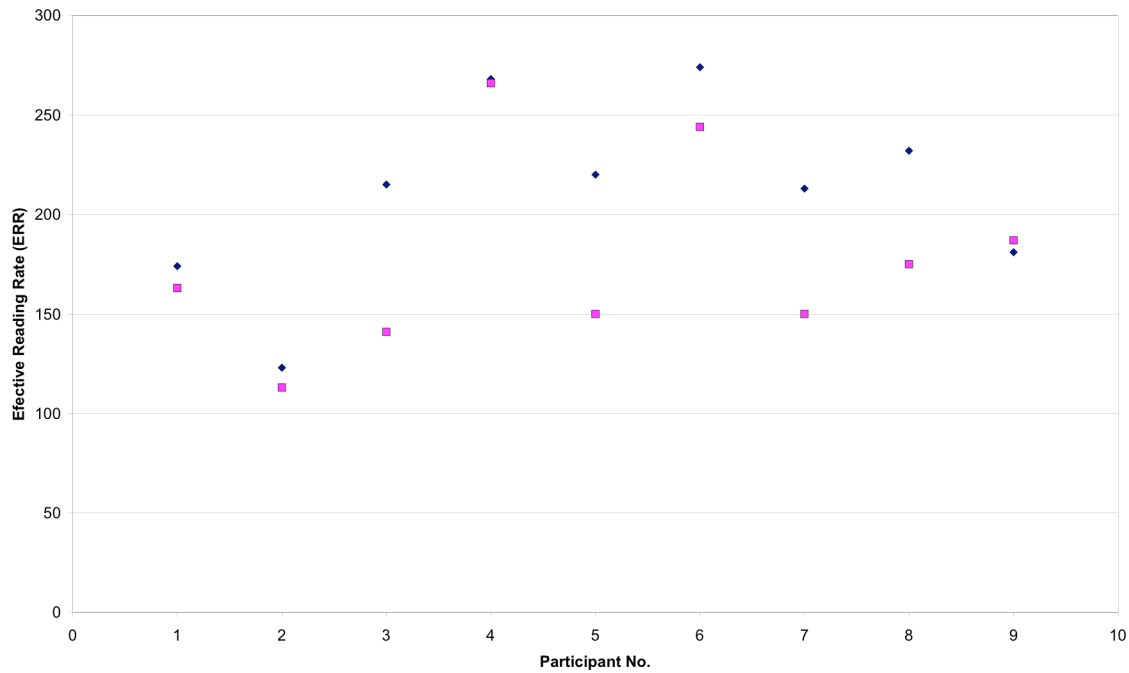


Diagram 40: WPM of control group 2 in Text 2 and Text 3



In control groups 1 and 2, there is also a discrepancy between WPM and ERR in Texts 2 and 3. However, it only exceeds 70 points in one case (Diagram 38). This indicates that *some course participants were not able to apply the reading techniques that they learned in the course when reading grammatically more complex texts. In the majority of cases, the discrepancy between the performance of course participants in Text 2 and 3 is similar to the discrepancy in the results of non-course participants.*

4.4 Results of the comparison between Improved Reading course participants and non-course participants

This section compares the performance of course participants in Texts 2 and 3 with that of non-course participants. The non-course participants can be divided into two groups. The members of the first group are university graduates, due to graduate in the near future or have a higher education qualification. The members of control group 1 therefore have more advanced academic qualifications than the Improved Reading course participants. The members of control group 2 are undergraduates in the first to fourth semesters. The segmentation of the survey into two groups enables a more precise assessment of the reading efficiency attained on the basis of the Improved Reading course.

4.4.1 Result of the comparison between course participants and control group 1 (advanced academic qualifications)

Description of the control groups

1. There are thirteen members in the group of Improved Reading course participants who took part in the final test. Eleven are students, one is a graduate with a degree in geography and one is an employee of the university. This person stated on the test paper that he/she did not do any reading between May and August due to an eye problem.

The eleven students are in the following semesters:

Sem. 1-4 : 2 students

Sem. 5-8: 3 students

Sem 9-12 : 3 students

Sem. 13+: 3 students

These eleven people are studying humanities and social sciences, law, biology and medicine, as well as IT and mathematics.

The group of course participants were mainly students in late semesters studying for various degrees.

2. Control group 1 consisted of three students in the final exam phase, five PhD students and a Dr.phil. With the exception of one person who was writing a doctoral thesis in the field of mechanical engineering, all were on humanities and social science courses.
Control group 1 therefore had more advanced academic qualifications than the group of course

participants and, as humanities and social science undergraduate and post-graduate students, did an extensive amount of reading.

Tab. 17: WPM and ERR rates achieved by control group 1 (advanced academic qualifications) in Texts 2 and 3.

No. 1	Text 2	Text 3
WPM	361	401
ERR	361	361

No. 2	Text 2	Text 3
WPM	210	160
ERR	189	144

No. 3	Text 2	Text 3
WPM	292	267
ERR	234	240

No. 4	Text 2	Text 3
WPM	282	216
ERR	226	194

No. 5	Text 2	Text 3
WPM	282	216
ERR	282	173

No. 6	Text 2	Text 3
WPM	335	244
ERR	168	171
No. 7	Text 2	Text 3
WPM	181	244
ERR	181	164

No. 8	Text 2	Text 3
WPM	393	374
ERR	275	299

No. 9	Text 2	Text 3
WPM	188	137
ERR	169	110

Diagram 41 compares the WPM rates of course participants and control group 1 in Text 2 (The blue rectangles represent the results of the course participants, the pink squares denote the results of control group 1).

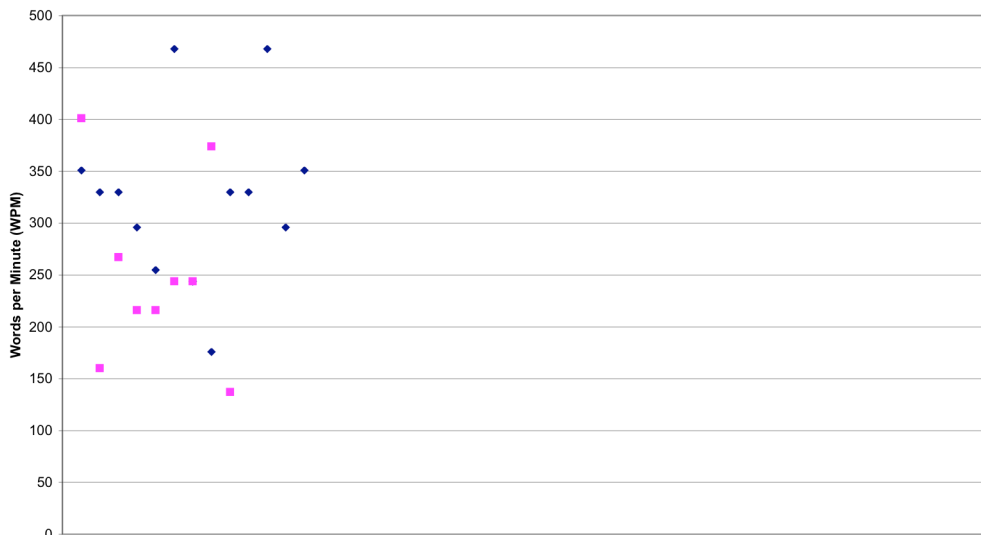
Diagram 41: WPM of course participants and control group 1 in Text 2



The five best results were achieved by course participants. Most course participants achieved results in a range of 300 to 450 WPM (average: 370 WPM). Control group 1 achieved results in a range of 200 to 360 WPM (average: 308 WPM). *The WPM rate of course participants was considerably higher than that of control group 1.*

Diagram 42 compares the WPM rates of course participants and control group 1 in Text 3

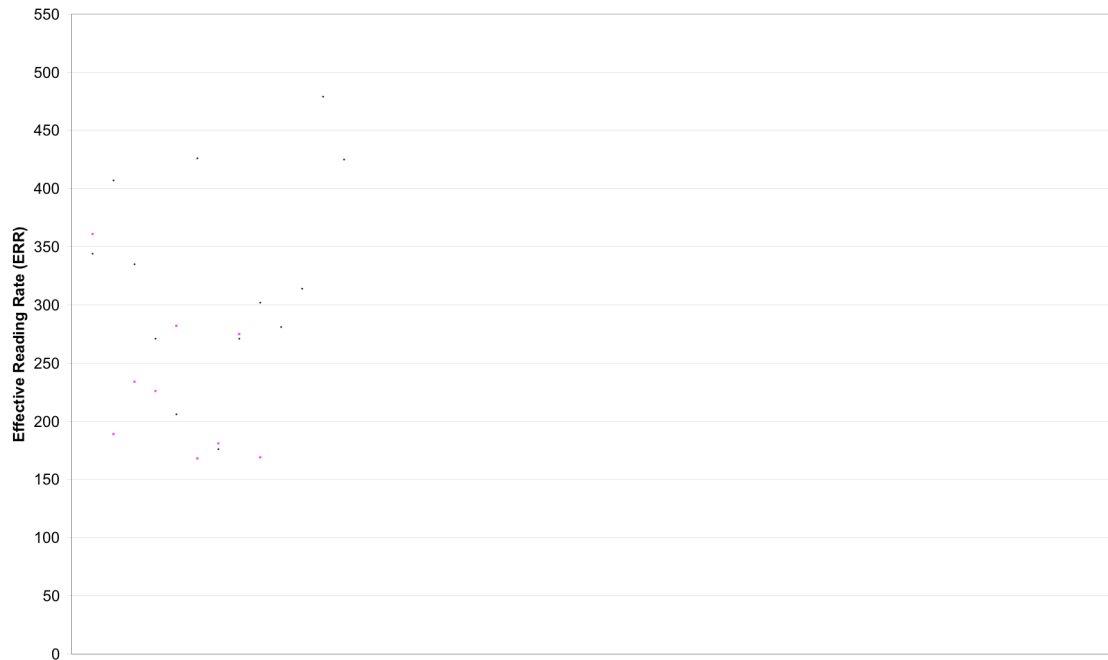
Diagram 42: WPM of course participants and control group 1 in Text 3



The two best results were attained by course participants, the third and fourth best results by members of control group 1. The course participants' results were mainly in a range of 250 and 350 WPM (average: 138 WPM). Control group 1 achieved results in a range of 160 to 270 WPM (average 224.5 WPM). As was the case in Text 2 *the WPM rate of course participants was higher than that of control group 1.*

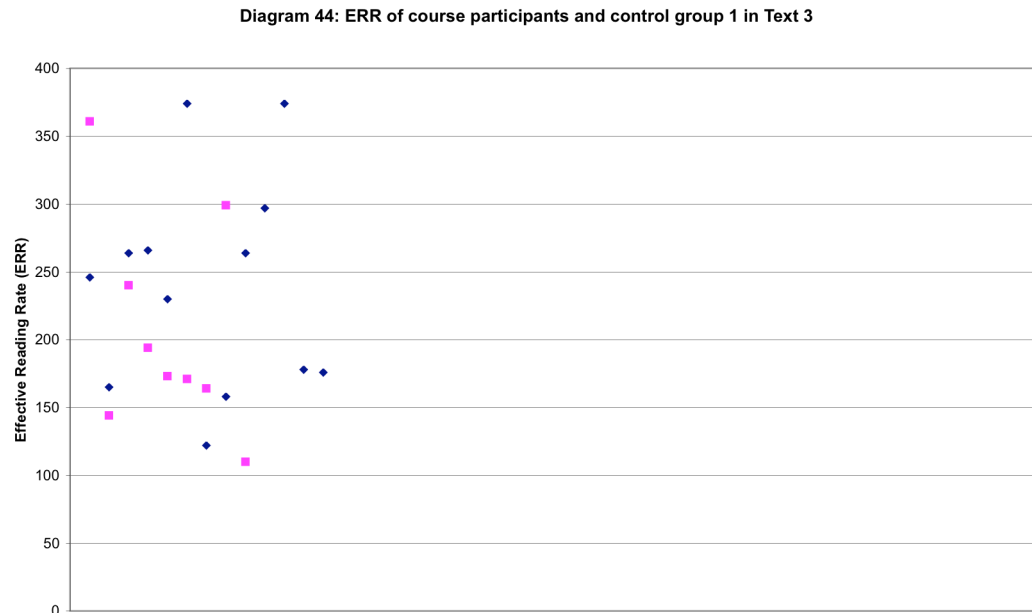
Diagram 43 compares the ERR of course participants and control group 1 in Text 2.

Diagram 43: ERR of course participants and control group 1 in Text 2



The four best results were achieved by course participants. With one exception, control group 1 achieved ERR rates of between 170 and 275 (average ERR: 215.5). The results of course participants can be divided into two groups. One group achieved ERR of between 275 and 350 (average ERR: 302.6) and the second group between 400 and 480 (average ERR: 434.3). *These groups therefore achieved results that were 2 and 1.5 times better than those of control group 1.*

Diagram 44 compares the ERR of course participants and control group 1 (Text 3).



The best results were achieved by two course participants and one member of control group 1. Control group 1 mainly achieved results ranging from 140 to 240 (average ERR: 172.9). The ERR results achieved by course participants varied considerably. They ranged from 160 to 300. This range can be sub-divided into two parts.

The smaller of the two comprises four results ranging between ERR 160 and 175, and the larger comprises six results ranging between ERR 230 and 300 (average ERR: 261.2). *The course participants therefore achieved results that were 1.5 times better than the average results of control group 1. However, just under one-third of course participants achieved results similar to the average results of control group 1.*

4.4.2 Result of the comparison between course participants and control group 2 (undergraduates in early semesters)

Description of the control groups

1. Control group 2 comprises a total of nine persons. These people are all humanities or social science students in semesters 1 to 4.
2. Please refer to 4.4.1 for a description of course participants.

Tab. 18: Results achieved by control group 2 (undergraduates in early semesters) in Texts 2 and 3.

No. 1	Text 2	Text 3
WPM	174	181
ERR	174	163

No. 2	Text 2	Text 3
WPM	137	125
ERR	123	113

No. 3	Text 2	Text 3
WPM	215	176
ERR	215	141

No. 4	Text 2	Text 3
WPM	335	296
ERR	268	266

No. 5	Text 2	Text 3
WPM	244	187
ERR	220	150

No. 6	Text 2	Text 3
WPM	274	244
ERR	274	244

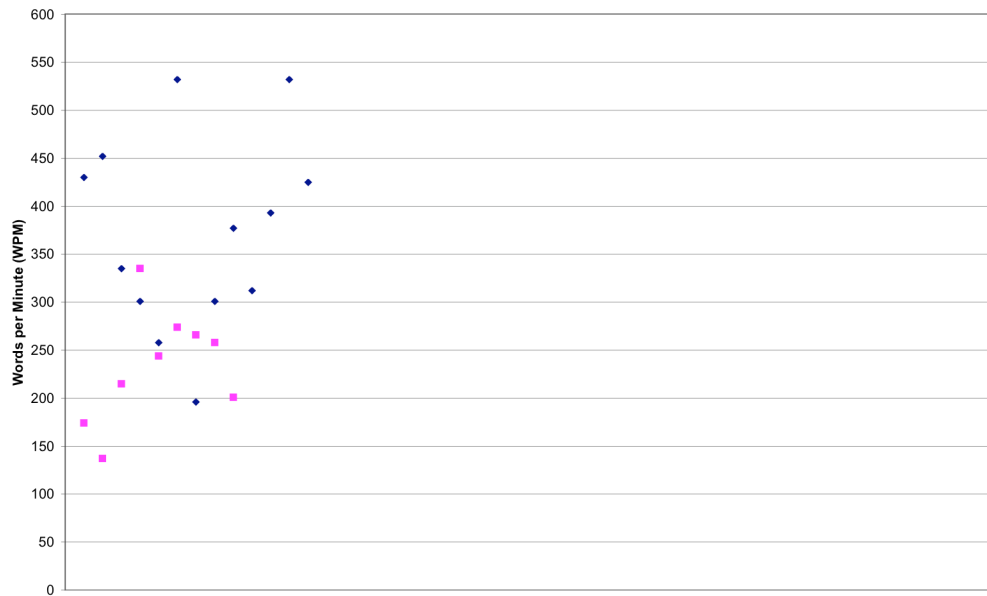
No. 7	Text 2	Text 3
WPM	266	187
ERR	213	150

No. 8	Text 2	Text 3
WPM	258	194
ERR	232	175

No. 9	Text 2	Text 3
WPM	201	187
ERR	181	187

Diagram 45 compares the WPM achieved by course participants and control group 2 in Text 2.

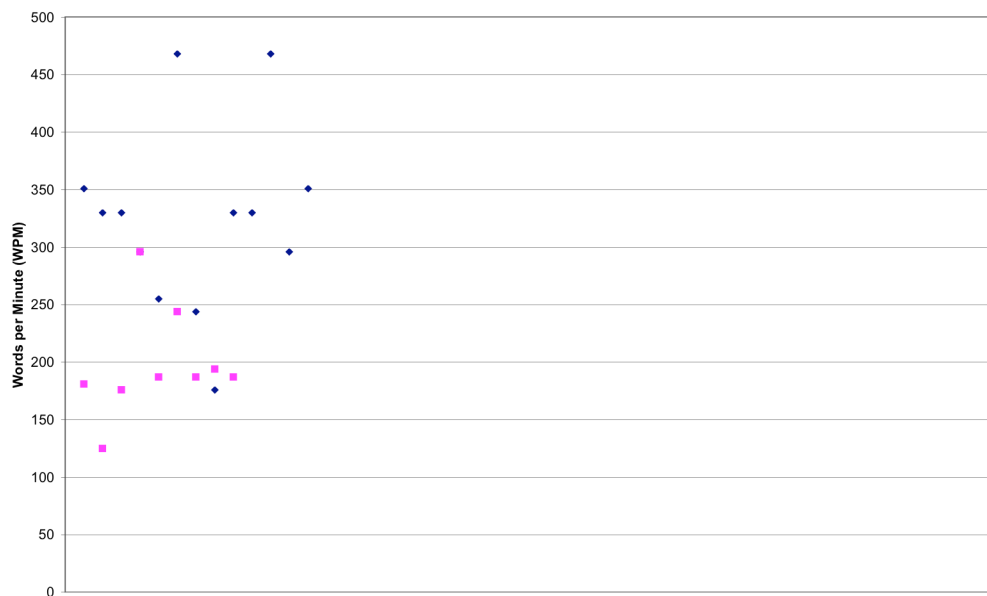
Diagram 45: WPM of course participants and control group 2 in Text 2



The seven best results were achieved by course participants. The course participants achieved results in a range of 300 to 450 WPM (average: 369.6 WPM). Most members of control group 2 achieved results in the range of 200 to 275 WPM (average: 243 WPM). *The results achieved by course participants are therefore considerably better than those achieved by control group 2. There is no overlapping in the centroid range of the two groups.*

Diagram 46 compares the WPM rates of course participants and control group 2 in Text 3

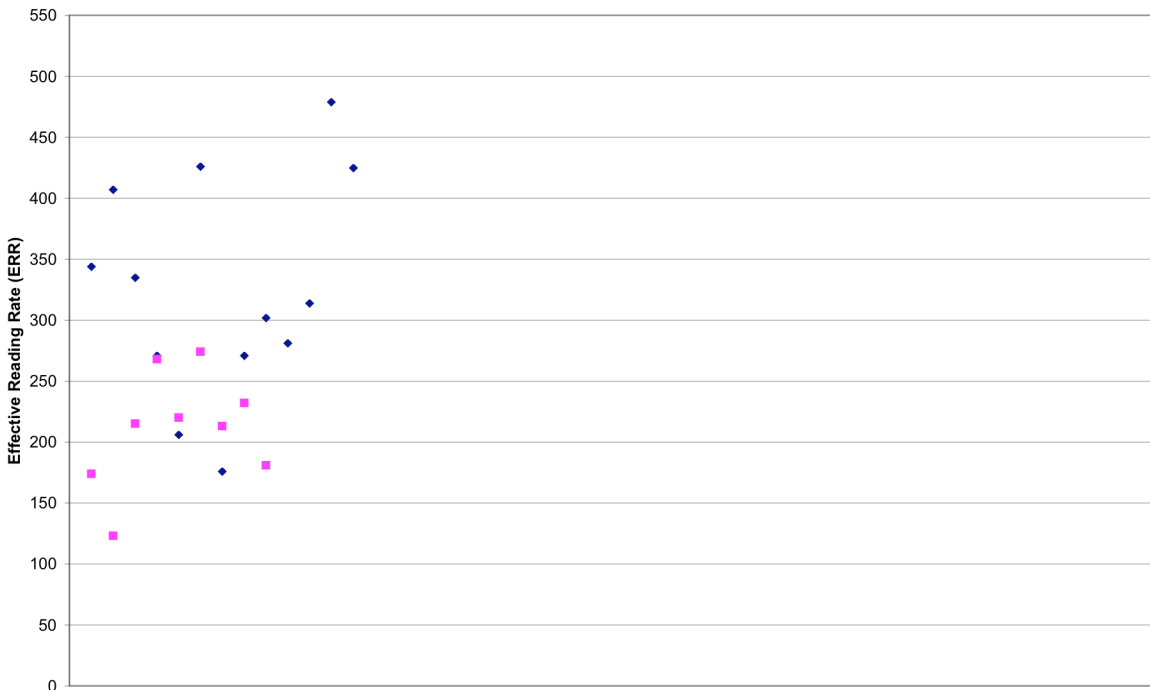
Diagram 46: WPM of course participants and control group 2 in Text 2



The eight best results were achieved by course participants. Control group 2 mainly achieved results ranging from 175 to 250 WPM (average: 193.7 WPM), and course participants achieved results ranging from 250 to 350 WPM (average: 318.7 WPM). As in Text 2, the results of course participants are considerably higher than those of the control group. The previously separate centroid ranges have aligned slightly and meet at 250 WPM.

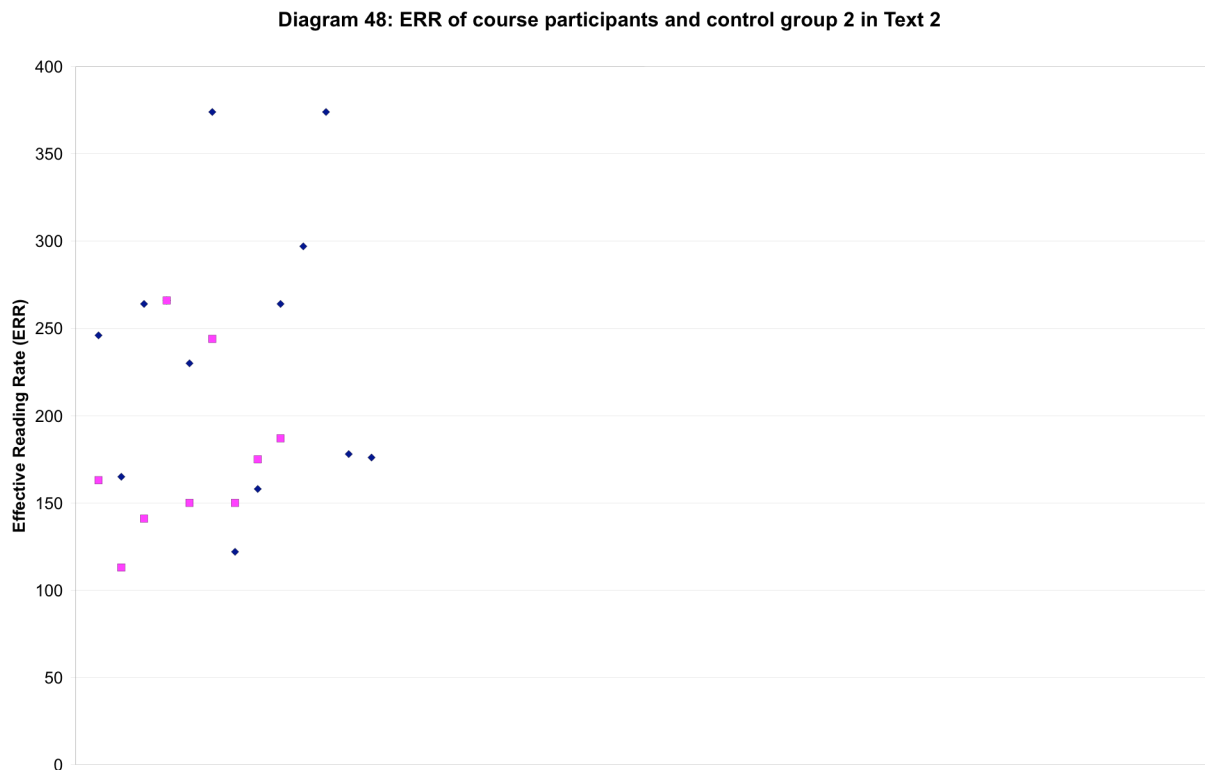
Diagram 47 compares the ERR of course participants and control group 2 in Text 2.

Diagram 47: ERR of course participants and control group 2 in Text 2



Course participants achieved the nine best results. The ERR of control group 2 mainly ranged from 175 to 230 (average ERR: 205.8). The results of course participants can be divided into two groups. The larger group of seven results ranges from 275 to 350 (average ERR: 302.6), and the smaller group of four results ranges from 400 to 480 (average ERR: 434.3). *Both groups achieved results that were considerably better than the control group. The ERR of the smaller group was more than twice as high, and that of the larger group almost 1.5 times higher.*

Diagram 48 compares the ERR of course participants and control group 2 in Text 2.



The three best results were achieved by course participants. The course participants' results can be divided into two almost equally sized groups. The group of better ERR results comprises six results ranging from 230 to 300 (average ERR 261.2), which is more than 1.5 times higher than the results of the control group. The group of lower results comprises four results ranging from 150 to 175 and is therefore in the centroid range of control group 2, where ERR was between 140 and 190 (average ERR 161).

4.5 Summarised interpretation of the final test results

The comparisons of the test results of Improved Reading course participants and non-course participants are decisive for the ascertainment of the effectiveness of the *reading techniques* taught in the course (Section 4.4). They show *that course participants can read the texts under the same test conditions faster and more effectively than undergraduates in early semesters. Most course participants were also able to read the texts faster and more effectively than the control group with advanced academic qualifications.* In Text 2 - the text with the less complex grammatical structure - the Effective Reading Rate (the product of reading speed and comprehension) of course participants was 1.5 to 2 times higher than that of both control groups. In the grammatically more complex Text 3, most course participants achieved results that were 1.5 times better than those of both control groups. The fact that course participants assessed

their pre-course reading skills as 'requiring improvement' (Section 1.3, Diagram 2) emphasises the effectiveness of the course.

Since the final test was carried out several months after the course, the positive results achieved by course participants in the German tests show *that participation in the course permanently improves reading efficiency*. The results also lead to the conclusion that *the techniques learned on the basis of English teaching materials can be effectively used when reading German texts*.

The comparison of the results achieved by course participants in Text 2 and Text 3 (Section 4.3) support the proposition *that the learned reading techniques can in many cases be effectively applied when reading texts of a higher grammatical complexity*. *The comparison also shows that around one-third of the course participants were not able to apply these techniques to the grammatically more complex texts*. The comparison between course participants, undergraduates in early semesters and students with higher academic qualifications confirms this finding. It shows that a small proportion of course participants read complex texts with the same efficiency as the majority of members of both control groups.

In the English part of the test (Section 4.2) the *course participants did not attain the results that they achieved at the end of the course*. Since the German part of the test verifies the learning effect of the course, it can be assumed that *the decline in performance is primarily due to the fact that the course is conducted in a foreign language*. Immersion in verbal and written English during the twelve-hour course has the short-term effect of enhancing foreign language proficiency, which has a positive impact on performance development during the course. A few months after the course, this performance-enhancing effect no longer exists, and the course participants have to get used to the English language again in the final test.

Finally, it is necessary to point out the limitations of the final test, once again. Firstly, it does not enable *the precise establishment of the course participants' rate of improvement*. In order to precisely establish rate of improvement, the results of the German part of the test would have to be compared with results attained when reading German texts at the beginning of the course. No such results exist because the course participants attended an English language course. To ascertain the degree of improvement, we recommend that these tests are performed on the participants of the German Improved Reading courses which have been offered since the beginning of winter semester 2002/03.

Secondly, the final test only measures the effectiveness of the reading techniques used. The *influence of reading strategies* (see Section 2.1), which are also taught in the course *cannot be ascertained in this test*. Reading strategies are only effective if the individual reads large volumes of texts and they cannot be examined in a 15-minute test. An evaluation of all aspects of the Improved Reading course must give consideration to reading strategies as additional course content to which considerable time and weighting are accorded.