The Effect Of Summarization On Intermediate EFL Learners' Reading Comprehension And Their Performance On Display, Referential And Inferential Questions

Zargham Ghabanchi, Ferdowsi University of Mashhad, Iran
Fateme Haji Mirza, Etrat Pre-University Center, Iran

ABSTRACT

This study examined the effect of summarization as a generative learning strategy of the readers' performance on reading comprehension, in general, and reading comprehension display, referential and inferential questions in particular. The subjects in this study were 61 high school students. They were assigned to two groups - control and experimental – each given the same texts taught by one of the researchers during ten sessions. In the control group, learners automatically used their own self-preferred strategies; but the experimental group was taught how to summarize the paragraphs. Then all were post-tested on their achievement of the instructed texts. The results revealed that the use of summarization did not have a significant effect on the readers' performance on display and inferential questions. As for the referential questions, however, the results demonstrated a significant effect for the use of summarization.

Keywords: Effect, Summarization, Comprehension, EFL learner, Display Questions, Referential Questions, Inferential Questions, Intermediate Subjects

INTRODUCTION

The conditions of meaningful learning require an instructional method that must elicit the cognitive processes in the learner. This research aims at exploring that area of research, focusing on the effect of summarization of reading comprehension and EFL learners' performance on reading comprehension questions. Since different comprehension questions require different levels of cognitive processing, the aim of this study can be rephrased into exploring the effect of summarization as a generative strategy on Iranian readers' level of text processing. Such a depth of text processing is defined in terms of comprehension at the level of sentences (through display questions), inter-sentential relationships (through referential questions), and inference in relation to world knowledge (through inferential questions). For a complete discussion, refer to the section entitled "Factual, Referential and Inferential Questions".

The present study tries to investigate the extent to which summarization, as a generative strategy, may facilitate L2 reading comprehension in the Iranian context. Of course, applying the procedure is possible after the learners have been instructed how to do the job. The study will then try to determine the degree of relative effectiveness of the strategy in question on the basis of the obtained results on three types of reading comprehension tests, including display, referential and inferential questions (Cf. part display, referential and inferential questions).

Spiro (1979, cited in Mahmoudi 2002) states that skilled readers constantly change their way of processing to accommodate the demands of a particular text; less skilled readers tend to over-rely on either bottom-up or top-

down procedures in one direction, which produces ineffective ways of interacting with a text. So pedagogically, this study can be of help toward the first steps in clarifying some non-linguistic abilities that make a distinction of good readers and good reading strategies. This study will also reveal the readers' conceptions about reading and their ability to make use of different ways of dealing with a text while reading that text for comprehension.

RESEARCH HYPOTHESES

No (1): EFL learners' use of summarization as a generative study strategy has no effect on their performance on reading comprehension display questions.
No (2): EFL learners' use of summarization as a generative study strategy has no effect on their performance on reading comprehension referential questions.
No (3): EFL learners' use of summarization as a generative study strategy has no effect on their performance on reading comprehension inferential questions.
No (4): EFL learners' use of summarization as a generative study strategy has no effect on their global reading comprehension proficiency.

SUMMARIZATION

The process of summarization involves the extraction of the gist and main themes of what is read while integrating the details into a coherent whole. Summarization depends on basic language skills, inferential abilities, and knowledge and engagement with texts (Brown and Palinscar, 1985, cited in developing reading comprehension).

Otero (2008) believed the global tasks, like summary writing, which would make the students actively connect and integrate units of information present in each document, induce long-term learning. This was because a deep processing of the texts had taken place when answering the global task; i.e., relevant information was read more slowly rather than constructing a very isolated understanding of the main units of information in the documents.

The importance of rhetorical structures in summarization has been observed by Brown, Day and Jones 1983; Slater and Graves 1989; Dole et al. 1991; and Sharp 1999 cited in Sharp, 2004). McGee and Richgels (1985) comment '…research…has shown that the structure of text and how adeptly a reader recognizes that structure affects the amount of information the student remembers'. Leon and Carretero (1995 cited in Sharp, 2004) note the importance of summarization in 'helping readers to differentiate between important and unimportant information, as well as in the organization and recall of information'.

Mani (n. d.), in explaining the methods for evaluating the text summarization task, suggested assessing mainly the coherence and informativeness of summaries. He explains that summaries that are extracts may be extracted out of context, in which case coherence problems may occur, such as dangling anaphors and gaps in the rhetorical structure of the summary. He also explains that the measure of informativeness of a summary is to assess how much information from the source is preserved in the summary.

GENERATIVE LEARNING THEORY

Wittrock (1974) introduced and elaborated this model of learning. He claims that there are many happenings in a learner's brain to transfer concepts learned previously; i.e., the learner's existing schemata to something comprehended completely as an integrated new idea (cited in Grabowski, n.d.). Within this framework, teaching becomes the process of leading learners to use their generative processes to construct meanings and plans of action (Wittrock, 1992). Wittrock (1991) states "the generative model is a model of the teaching of comprehension and the learning of the types of relations that learners must construct between stored knowledge, memories of experience, and new information for comprehension to occur" (p. 170).

Generative learning involves students in higher level thinking processes and helps learners to integrate new knowledge within the structure of old knowledge (Schott, n.d.). Generative learning strategies can be broken down into four elements: recall, integration, organization, and elaboration. Summarizing is mentioned as an integration-

type strategy (Generative learning, n. d.).

FACTUAL, REFERENTIAL AND INFERENTIAL QUESTIONS

Farhady (1998) classifies the comprehension questions into factual, referential and inferential based on the kind of information processed in the text. Farhady defines a factual question as an instrument to check the testees' understanding of the factual information provided in the text. In fact, such questions are intended to check whether the testees understand who is doing what to whom, when, and where (Farhady, 1998). The answer for such questions can be directly extracted from the text. Such questions are often used at the elementary levels of language education (ibid.).

Referential questions, according to Farhady (1998), require the reader to move beyond the level of sentence comprehension to understand the relationship among the sentences. That is, the students are required to make references from one sentence to the preceding or following sentences.

Inferential questions, however, are designed to check the information provided in the text (ibid.).The answer to this type of question cannot be directly extracted from the text.

Since summarization makes the readers re-read the text to find the gist, resulting in deeper textual understanding (Special connections, n.d.), and the three chosen types of questions are classified by Farhady based on the kind of information processing in the text (1998), the researcher guessed there might be a convincing relation to check the efficiency of the strategy on just the chosen types of questions in the study rather than other kinds.

PARTICIPANTS

All participants were EFL high school students educated at level 8 at Jihad Institution in Neishabour, studying the mid-lessons of "*Interchange 2*". The subjects were divided into control and experimental groups. The control group consisted of 30 subjects, whereas the experimental group included 31 students.

THE PROCEDURE

To conduct the study, the experimental group - the *summarization* group - was first taught how to apply the strategy of developing summaries of the paragraphs of a given text. Then the subjects were asked to generate a summary of each paragraph of the text containing the most important ideas, concepts, events and key facts in one or two complete sentences. It has been stressed that a good summary should include the key concepts or events and their relationships.

The second group - the *control* group - received the same reading materials, but without the required instruction for using and implementing summarization. In other words, none of the subjects in the control group used a generative summary writing strategy.

TESTING MATERIALS

The proficiency test used as a pre-test had been taken by the institution itself in which the experiment was done, as the level seven final achievement test taken on the series books of "*Interchange*". To be sure of the homogeneity of subjects, of all 76 students, the researcher selected 61 who got marks from one standard deviation below the mean to one standard deviation above the mean.

The post-test was used as an achievement test using the readings taught during the treatment. It had 51 items consisting of 17 displays, 17 referential and 17 inferential questions.

Ten texts from among those taught during the experiment were chosen to develop the final test. The first four texts were obtained from *Interchange 2;* i.e., taught in the institution. The first 21 questions of the post-test were written based on these four texts and were then validated in a pilot study administered to a group of 14 students

in level nine in order to evaluate the effectiveness of items and to decide which items or choices needed to be revised or deleted.

The last 30 questions of the post-test (questions 22 through 51) were obtained from the reading sections of two Michigan tests. These texts were used in classes, not with their own real questions, but with new questions written by the researchers.

In order to calculate the reliability coefficient, SPSS 15 for Windows was used. The reliability coefficient, using Cronbach Alpha as a meticulous measure of internal consistency, was estimated to be 0.78, which is an acceptable Alpha-Cron, knowing that in experimental research, having an Alpha-Cron higher than 0.65 is at an acceptable level (Waineer, 2006).

To estimate the validity of this test, the researcher validated it against the Michigan test (the last 30 questions of the reading comprehension final test that was administered in the previous session). These two parts of the test were administered to our subjects one after another in two sessions and the degree of the correlation coefficient between the two sets of scores, using Pearson formula, turned out to be 0.7231 (P-value = 0.003), which is an acceptable estimate of criterion-related validity indicated in the concurrent validity of the first part (Interchange reading questions).

The post-test was administered in two sessions at the end of the experiment due to time limitation. First the Michigan questions and then in the next session the Interchange questions were administered.

RESULTS OF THE PRE-TEST

The extreme cases of the subjects having taken their final test of level seven to enter level eight were removed of the study and those subjects who scored between one standard deviation above and below the mean were selected as the subjects of the study. The descriptive data appear in Table 1.

Table 1: Descriptive Statistics of the Pre-test Subsections and the T-test

Descriptive Statistics of the Pre-test Subsections					t-test for Equality of means		
Grouping	**N**	**Mean**	**S D.**	**Std. Error Mean**	**t-observed**	**df**	**Sig (2-tailed)**
Experimental Reading	30	14.35	3.08	.21	-0.293	60	0.723
control	31	14.04	4.36	.34			
Experimental overall	30	69.10	6.32	.41	1.052	60	0.155
control	31	69.04	6.72	.31			

RESULTS OF THE POST-TEST

The candidates in the two groups took the same reading comprehension test after the instruction. To obtain statistically based evidence to test the hypotheses, the technique of T-test was utilized. The results appear in Table 2.

Table.2: Descriptive Statistics of the Post-Test Subsections and the Related T-Test

	t-test for Equality of means		
	t-observed	**df**	**Sig (2-tailed)**
Reading	-0.293	60	0.723
overall	1.052	60	0.155

The table shows that the value of significant level for display questions (0.273) and for inferential questions (0.681) is more than the probability value of 0.05.

P-value for display questions = $0.273 > \alpha = 0.05$
P-value for inferential questions = $0.681 > \alpha = 0.05$

Therefore the first and fourth null hypotheses are not rejected. However, as is clear from the table, the value of significant level for referential questions (.002) and for all three types of the questions overall is highly less than the probability value (0.05).

P-value for referential questions = $0.002 < \alpha = 0.05$
P-value for overall questions = $0.009 < \alpha = 0.05$

Therefore, the second and fourth null hypotheses are rejected.

DISCUSSION

As the results of the study revealed, the use of summarization as a generative study strategy for reading comprehension had no significant effect on display questions, meaning both the control and experimental groups showed almost the same performance on display questions. Answering these types of questions may only require focusing on linguistic forms at the level of words and sentences, checking only vocabulary and grammar. That is why in normal reading comprehension texts, the number of display questions in comparison with the other two types of questions is relatively low (Mahmoudi, 2002). Another supporting point is that display questions are easier than the other two.

As for the referential questions, the investigation showed that the use of summarization as a generative study strategy had a significant effect on the experimental group. It can be argued that since referential questions require an understanding on the part of the readers, as Farhady (1998) states, to make references to the preceding and following sentences in a text, as mentioned previously**,** the use of summarization makes the reader better prepared to see this relationship. In fact, a higher level of processing is needed to answer referential questions in comparison to the other types of questions. Consequently, summarization may encourage deeper engagement with a text and encourage students to re-read as they construct a summary and seems reasonable to help readers go beyond the more common and more readily available unidirectional way of processing a text. In this study, the use of summarization as a generative study strategy for reading comprehension proved to be capable of enabling the readers to process the text more deeply, while extracting the gist and main themes of what is read (as putting aside the irrelevant details) and integrating the details into a coherent whole. Perhaps it helped the readers to overcome the initial over-reliance on bottom-up processes and helped them to step back and look at the organization of the paragraph as well.

The investigation of the third null hypothesis detected no significant role for the use of summarization in the readers' performance on inferential questions. Although the mean performance of the experimental group was a little higher than the control group, the difference was not big enough to reach the significant level. This outcome can be viewed in different ways:

First, as the functional definition of the term also indicates, and as mentioned previously defining them, such questions are very demanding on the readers. They not only have to understand the relationship between sentences and the organization of the text, but also to combine this understanding with their world knowledge, their information about the world, and how it works. Therefore, it can be argued that the use of this particular complementary activity has had no bearing on the readers' world knowledge and much less on its activation.

Second, Bachman (1990) states that some aspects of processing a text may go beyond the linguistic ability of a reader. She says that to answer reading comprehension questions involving inferences is not only a matter of linguistic knowledge, but utilizing other sources of knowledge. Bachman relates performance on such questions to "strategic competence"- the capacity of making the most efficient use of available abilities in carrying out a

linguistic task. She considers "strategic competence" as ability within the realm of general cognitive abilities. Therefore, to answer inferential questions is very demanding on the readers. It presupposes a lot of knowledge of the world to be imparted by the readers, in addition to the information provided by the text. Because this knowledge has very little to do with the text, it seems understandable for the summarization not to have a significant effect on such questions (ibid.). The highly demanding nature of such questions is also revealed by descriptive statistics; i.e., both groups performed very poorly on inferential questions after the treatment as seen on Table 2.

Third, answering inferential questions requires a lot of world knowledge (Farhady, 1998) and, more importantly, a lot of reading practices in the target language. Widdowson (1979) regards inferencing as a high order process that goes well beyond the text and is closely related to the target language. It can be claimed that, as the results also demonstrate, the use of summarization as a generative complementary strategy has no effect on this complex process; i.e., inferencing.

As for the fourth hypothesis, the results showed a significant relationship between the use of summarization as a generative complementary strategy and readers' performance on all reading comprehension questions as it was expected that the summarization group would score higher on the post-test since that generative learning strategy required the most comprehensive cognitive effort to assimilate the material with prior experience, beliefs, and knowledge (Ritchie and Volkl, 2000 cited in Mcgriff, n. d.).

This research hypothesis finds its importance in the fact that a reading comprehension text is not usually followed by only one type of question, but by a combination of the three, with more and more emphasis on referential and inferential questions as the proficiency of the readers grow. Therefore, reading comprehension instructors can include the summarization as a post-reading strategy in their lesson plan and benefit the possible results, such as increases in learning and comprehension.

The results indicate that summarization as a generative strategy had no significance on learners' performance of display questions. In fact, given the highly sentence-bound nature of display questions, readers do not need much background knowledge to combine with the knowledge of the text to answer such questions.

The referential questions, on the other hand, require the reader to go beyond the sentence and make reference to the preceding and following sentences. The rejection of the second null hypothesis showed that summarization could help readers establish a better relationship among the sentences of a text. It helped them to see the organization of the paragraph in a better way and benefit from summarization consistently on the text in order to improve their holistic understanding of the interrelationships of the parts of the text.

The third null hypothesis dealt with inferential questions. The confirmation of the hypothesis proved that summarization had no significant effect on readers' performance of such questions. It was argued that this finding could be interpreted in line with the arguments of many experts regarding inferencing (Bachman, 1990; Grabe, 1997). To them, answering inferential questions would need a great amount of world knowledge to help readers go well beyond the text. Summarization proved to be of no significant use in that endeavor. Although the descriptive statistics showed a better gain for the experimental group, that gain was not statistically significant.

Through the rejection of the fourth hypothesis, the study implicated a significant role for the use of summarization on readers' performance on all three types of questions combined. It proved that although the gain of the experimental group in display and inferential questions was not significant, it was so much so that it established a significant relationship between the use of summarization and readers' performance on all three types of questions.

Overall, although this study needs to be replicated to make a strong statement, it approves the use of summarization as a generative post-reading activity, but once again emphasized that this usage must be very cautious and at the discretion of the teachers. On the one hand, the use of summarization as a generative post-reading activity is supported in the case of referential questions, which constitutes a large portion of the questions following a reading comprehension text, as well as in the case of the whole text. On the other hand, that strategy didn't work quite well in the case of inferential and display questions. This outcome emphasizes the difficult task of the teachers to decide where and when to use this particular kind of generative activity.

CONCLUSION

This study could, in fact, talk about the nature of questions and the level of information processing they need. The study quite indirectly showed that the level processing increases as the readers move from display questions to referential and inferential questions. It was explained that readers process the text at the sentence level while answering display questions, but they process the text at the inter-sentential level while answering referential questions. As reflected in the descriptive statistics regarding both control and experimental groups; both groups performed outstandingly right while answering display questions compared to their performance on the other two types of questions. Therefore, display questions are easier than the other two. This point proves a good idea to testers who are dealing with students. There is no doubt that it would be unfair to test an elementary student with referential questions, and even more so, to test them with inferential questions. By the same token, it is meaningless to test advanced students with display questions.

The only clear inference confidently drawn from the research literature reviewed and the results of this research is that the specificity of the strategy summarized seemed to be extremely helpful to the students, especially while answering referential questions.

AUTHOR INFORMATION

Zargham Ghabanchi received his BA in English language and literature from Mashed University, Iran. Then he received his MA in TEFL from Tehran Tarbiayyet Moderres University, Iran. He started his Ph.D. at the University of Liverpool, the UK in 1993 and Ph.D. was conferred to him in 1998 in Applied Linguistics. Now he has a chair at Sabzevar Teacher Training University, and he is the vice chancellor of Sabzevar Payam Noor University. He has published several books and articles.

Fateme Haji Mirza was born in Neishabour on September 1969. She got her BA degree of TEFL from Mashad university. She is an MA student of TEFL at Sabzevar Teacher Training University. Around seventeen years that she is teaching English in a pre-university level at an institution in Neishabour-Iran.

REFERENCES

1. Bachman, L. F. (1990). Fundamental considerations in language testing. Oxford: OUP.
2. Developing reading comprehension. (2006). Primary Framework for literacy and summarization. Retrieved February 3, 2008, from http://www.york.ac.uk/res/crl/research.html
3. Farhady, H. (1998). Constructing reading comprehension Tests. *Roshd Foreign Language Teaching Journal, 13*(49), 37-48.
4. Generative learning. (n. d.). Retrieved from http://www.ed.psu.edu/NASA/genetxt.html
5. Grabe, W. (1997). Reading research and its implications for reading assessment. Localization**:** Fairness and validation in language assessment : selected papers from the 19th language testing research colloquium, Orlando, Florida, 226-262. Retrieved from http://dialnet.unirioja.es/servlet/articulo?codigo=310551
6. Grabowski, B.L. (2oo1). Generative learning: Past, present, and future. Retrieved from http://www.aect.org/edtech/ed1/31/index.html
7. Mahmoudi, M. (2002). On the relationship between the use of L1 version of a reading comprehension text and Iranian intermediate learners' performance on display, referential and inferential questions. Unpublished master's thesis. School of Humanities, Tarbiat Modarres University, Tehran, Iran.
8. Mani, I. (n. d.) Summarization evaluation: An overview. . Retrieved February 3, 2008, from http://research.nii.ac.jp/ntcir/workshop/OnlineProceedings2/sum-mani.pdf
9. McGee, L.M., & Richgels, D. J. (1985). Teaching expository text structure to elementary students. *Reading Teacher 38* (8), 739-48.
10. McGriff, S. J. (n. d.) Using written summaries as a generative learning strategy to increase comprehension of science text. Retrieved March 6, 2008, from http://www.sjsu.edu/depts/it/mcgriff/research/QuantitativeResearch.pdf
11. Otero, R. C. (2008). Summary of integration information processes form multiple documents**.** Retrieved December, 2008, from http://dialnet.unirioja.es/servlet/tesis?codigo=7205&info=resumen&modo=popu

12. Sharp, A. (2004). Strategies and Predilections in Reading Expository Text: The Importance of Text Patterns. *RELC Journal*, 35(3), 329-349. Retrieved on November 23, 2008 from: http://rel.sagepub.com/cgi/content/abstract/35/3/329
13. Schott, M. (n. d.). Rich environments for active learning. Retrieved from: http://coe.sdsu.edu/eet/articles/reals/start.htm
14. Special connections: Teaching Paragraph summarization strategies. (2005). Retrieved February 3, 2008, from: http://www.specialconnections.ku.edu/cgibin/cgiwrap/specconn/main.php?cat=instruction&subsection=rc/paragraph
15. Waineer, G. (2006). Principles in experimental researches. Tehran: SAMT.
16. Widdowson, H. G. (1979). The process and purpose of reading. In H. Widdowson (Ed.), *Exploration in applied linguistics*. New York: CUP.
17. Wittrock, M. C. (1974). Learning as a generative process. *Educational Psychologist, 11*(2), 87-95. Retrieved March 6, 2008, from http://applij.oxfordjournals.org/cgi/content/abstract/19/3/357
18. Wittrock, M.C., & Alesandrini, K. (1990). Generation of summaries and analogies and analytic and holistic abilities. *American Educational Research Journal, 27*(30), 489-502.
19. Wittrock, M.C. (1991). Generative teaching of comprehension. *Elementary School Journal, 92* (2), 169-180. Retrieved March 6, 2008, from http://www.readingcenter.buffalo.edu/center/research/gencom.html
20. Wittrock, M.C. (1992). Generative learning processes of the brain. *Educational Psychologist, 27*(4), 531-541. Retrieved March 6, 2008, from http://www.sjsu.edu/depts/it/mcgriff/research/QuantitativeResearch.pdf