
RESEARCH ARTICLE

An Alternative Way of Teaching Reading to High School Students

El Hamydy El Mehdi¹ ✉ and Brigui Hind²

^{1,2}Literature, Arts, and Pedagogic Engineering Laboratory, Faculty of Languages, Letters and Arts, Ibn Tofail University, Kenitra 14000, Morocco

Corresponding Author: El Mehdi El Hamydy, **E-mail:** elmehdi.elhamydy@uit.ac.ma

ABSTRACT

The aim of the present study was to investigate the implementation and effect of an alternative way of teaching reading—Reciprocal Teaching (RT)—on the performance of Moroccan EFL common core students in four comprehension-monitoring skills, namely questioning, summarizing, clarifying, and predicting. The intervention, which is ten sessions long, used an instructional reading approach that combines guided practice, scaffolding, and active involvement. Seventy-two students participated in this study. These were tested on two different occasions: pre-test and post-test. The experimental group received instruction in RT, whereas the control group did not have any treatment. The scores of all participants in the pre-test were equal. However, the findings revealed that the experimental group gained comprehension-monitoring skills and reading performance, whereas the control group did not show any change from the pre-test to the post-test.

KEYWORDS

Clarifying, comprehension-monitoring, predicting, reciprocal teaching, summarizing, and questioning.

ARTICLE INFORMATION

ACCEPTED: 01 November 2022

PUBLISHED: 09 November 2022

DOI: 10.32996/jeltal.2022.4.4.6

1. Introduction

The present study is conducted within the area of metacognition and reading. It examines the effect of RT on comprehension-monitoring skills—question generation, summarizing, predicting, and clarifying—of Moroccan EFL common core students. A number of research studies have been conducted to replicate or extend the original study of RT (e.g., Palincsar & Brown, 1984; Alfassi, 1998; Spivey & Cuthbert, 2006). These researchers have asserted that RT can lead to successful learning, helping readers to regulate, monitor, and check their own understanding, which suggests that reading comprehension can be subject to remediation and improvement through the RT approach. Previous studies have been conducted in an ESL context, and their results support, to a large extent, the positive impact RT has on comprehension-monitoring skills.

Research has shown that RT improves comprehension monitoring skills (Brown & Palincsar, 1982; Palincsar & Brown, 1984; Brown, 1985; Alfassi, 1998), but most of these studies were conducted in ESL contexts. Therefore, this study examines the degree to which those claims may apply to the Moroccan EFL context. Besides that, similar studies dealing with Moroccan participants are hardly found in this aspect of reading research. This study thus seeks to find out the extent to which the RT procedure would enhance students' comprehension monitoring skills. To meet the established objectives, the following research questions have been formulated:

- 1- Is there a significant difference between the control group and the experimental group in the pre-test?
- 2- Is there a significant difference between the control group and the experimental group in the post-test?

2. Literature Review

This study scrutinises the effect of RT on comprehension monitoring skills in light of metacognition and reading. Researchers within the paradigm of reading have been working on co-operative forms of learning and teaching for at least the last six decades.

Copyright: © 2022 the Author(s). This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC-BY) 4.0 license (<https://creativecommons.org/licenses/by/4.0/>). Published by Al-Kindi Centre for Research and Development, London, United Kingdom.

RT is, indeed, one particular form of co-operative teaching. This teaching procedure was first addressed in the pioneering work of Brown & Palincsar (1982). It involves the reciprocal interaction of novice readers and experts in an overtly explicit manifestation of strategy use. RT procedure (Brown & Palincsar, 1982) has been acclaimed by a number of researchers (e.g., Glaser, 1990; Brown, 1985) as one of the most effective instructional classroom-strategies. RT is an instructional procedure developed by Brown & Palincsar (1982), which aims at helping students to foster their reading comprehension. This technique involves the use of four comprehension-monitoring skills: summarizing, questioning, predicting, and clarifying in ongoing dialogues between the teacher and the students in small groups. Both the teacher and students can take turns modelling the four skills as follows: ask questions about the text, generate a brief summary, clarify whether certain questions or statements are true or false, justify their responses from the text, and predict the meaning of the underlined words in the text according to the context in which they occur. The adult teacher's role decreases gradually, and students become more capable of assuming responsibility for their learning. However, the adult teacher still provides guidance and feedback to improve students' performance, encouraging them to shift turns while practising the four skills. The overall aim is to help students to monitor, regulate, and check their own understanding. Guided practice and active involvement are at the crux of RT (ibid). Besides, scaffolding, which emanates from Vygotskian developmental theory (Vygotsky, 1978), is perhaps the most salient theoretical element involved in RT (Palincsar & Brown, 1984). RT also provides learners with the opportunity to negotiate meaning, share responsibility for thinking and instruction, and enhance their thinking and problem-solving skills (Glaser, 1990).

Four comprehension-monitoring skills—predicting, questioning, clarifying, and summarizing— are at the essence of the RT approach (Brown & Palincsar, 1982). These skills enhance reading comprehension and allow students to monitor, regulate and check their own understanding. In this study, the four skills are operationalized as follows:

First, *summarizing* involves students concentrating on the gist of the text. While requested to provide a summary of the text, students are simultaneously asked to attend to the main idea of the text. Students would be asked to provide a brief, coherent summary of the text based on five macro-rules, such as deletion of trivia, deletion of redundancy, super-ordination, selection, or invention of a topic sentence (see Brown & Day, 1983). Second, *clarifying* relates to justifying whether certain statements are true or false. Students would be asked to check whether some questions or statements are true or false and clarify (justify) their responses using particular information from the text. Third, *questioning* refers to generating five factual questions whose answers are explicitly stated in the text based on the taxonomy of Pearson & Johnson (1978). When asking students to formulate questions on the content of the text, the instructor also aims to check the students' overall comprehension. Fourth, *predicting* implies inferring the meaning of the underlined words in the text. Students would be given a list of words and would be asked to predict (infer) their meanings according to the context in which they occur.

A number of research studies have been conducted to replicate or extend the original study of RT (Palincsar & Brown, 1984; Alfassi, 1998; Spivey & Cuthbert, 2006; Hart & Speece, 1998). For example, Palincsar and Brown (1984) conducted the first study to elaborate on the RT procedure. The findings of their study revealed momentous effects of the RT procedure on reading comprehension, namely on summarizing, answering questions, and detecting incongruities (ibid). Alfassi (1998) conducted a pivotal study in which she investigated the effect of RT on fostering high school students' reading comprehension. She found that strategy instruction was better than traditional methods in promoting high school students' reading comprehension (ibid). In addition, Spivey and Cuthbert (2006) explored the differential effect of a RT intervention on the lecture comprehension skills of college students. They reported that low-verbal ability subjects who were taught the RT method significantly improved their lecture comprehension (ibid). Hart and Speece (1998) investigated the effects of RT on post-secondary students at risk for academic failure. They found that students receiving instruction in the RT procedure outperformed students in the comparison group on reading comprehension (ibid).

3. Methodology

This part describes the methodology adopted to meet the research objectives, and thus information relative to the design, the participants, the instruments, and the procedure are presented in detail.

3.1 Research Design

This study has a quasi-experimental design in that the subjects were neither randomly assigned nor selected. This being the case, two intact classes from Wellada High School in Sale took part in the experiment. One intact class consisting of 38 students served as an experimental group, and the other intact class involving 34 students functioned as a control group. The statistical test is an independent t-test since it involves two different groups (Figure1).

| Groups | Pre-test | Treatment | Post-test |
|--------------|----------|---------------------|-----------|
| Experimental | T1 | Reciprocal Teaching | T2 |
| Control | T2 | Traditional way | T2 |

Figure1: The design of the study

3.2 Participants

3.2.1 Subjects

Seventy-two subjects took part in this experiment, and all of them were native speakers of Moroccan Arabic and had studied English for two years. Note that we made sure all the subjects were not taking extra courses in English during the intervention and that they had not taken part in similar experiments. These factors proposed that they have the same degree of congruity and homogeneity, and thus they were all equal at the beginning of the intervention. These participants were two intact classes meeting in two weekly 55-min sessions for 10 weeks. Participants in the experimental group received instruction using the RT procedure, whereas subjects in the control group did not receive instruction in the RT. However, participants in both groups received the same training sessions and the same amount of material.

3.2.2 Instructors

Two instructors took part in this experiment. Both of them are 28 years old and have approximately six-year teaching experience in high school. Every instructor taught his own class, which facilitated communication and interaction and motivated subjects to be more responsive to the intervention. The instructor teaching the experimental group received one training session, in which he was introduced to the rationale, design, and implementation. He also received a lesson plan for each of the reading passages that were taught during the intervention. The lesson plans were used to account for the teacher variable. Unlike the instructor in the experimental group, the instructor teaching the control group received no training session about the RT procedure. However, he was informed of the design, rationale, and significance of the study. He was also apprised that the intervention would last twenty sessions. He was further given the reading materials to judge their relevance to students' levels and interests.

3.3 The instruments

The instruments used to achieve the requirements of the current study included a reading selection, lesson plans, a pre-test, and an immediate post-test.

3.3.1 The Reading Selection

After agreeing on the general guidelines, the instructors were given 10 expository reading passages of approximately 240-250 words (each) that were utilized as readings throughout the intervention, in addition to the reading passages used in the pre-test, post-test, and delayed test (Appendix I). These passages incorporate different themes and topics, including citizenship, technology, sustainable development, civic education, etc. The passages were selected on the basis of their relevance to the aim of the study, as well as to the subjects' level and interest. Note that the chosen passages conformed to a tenth-grade reading level according to the Fry Readability Formula (Fry, 1977). Reciprocal students were asked to generate five factual questions about the text, using the classification of Pearson & Johnson (1978); write a three or four sentence summary about the passage on the basis of five macro-rules (Brown & Day, 1983); predict the meanings of the underlined words in the text according to the context in which they occur; and finally clarify whether certain questions or statements are true or false, using specific information from the text to justify their responses.

3.3.2 Texts used during the pre-test and post-test

The texts used in the pre-test and posttest test were not dissimilar in terms of their level and nature to those being employed throughout the intervention. The pre-test text entitled 'Radio vs. Television' dealt with two different means of communication that have made the world a small village (Appendix I-A). The post-test passage relates to the vital role the Internet plays in social life and the variety of services it provides (Appendix I-B).

Factors such as linguistic difficulty, complex grammatical structures, lexical density, and length were all taken into account. Based on the piloting, the texts were adapted in terms of complex grammatical structures and linguistic difficulty. Lexical density, a measure of the difficulty of a passage or text, was achieved by dividing the number of separate words (content words) by the sum total of words in the text and multiplying the result by 100 (Faerch et al. 1984). Lexical variation, the sum total of new words used in a passage or text, was attained by dividing the sum total of word types (words that did not occur once) by the sum total of word tokens (words) (ibid). The propositional content of the texts was estimated by the number of idea units they comprised (Table3). It should be noted that only the texts used in the pre-test and post-test test were subject to such an analysis so that we could ensure that the three texts were truly parallel in order to account for the test-effect.

| | Title | Number of words | Lexical density | Lexical variation | N° of idea units |
|-----------|-----------------------|-----------------|-----------------|-------------------|------------------|
| Pre-test | Radio vs. Television | 240 | 0,84 | 0.40 | 17 |
| Post-test | The Internet is Great | 243 | 0,83 | 0.38 | 18 |

Table3: Main features of the texts used in the pre-test and post-test.

3.3.3 Lesson Plans

The instructor in the experimental group was allotted a lesson plan for each expository reading passage. The lesson plans basically provided the instructor with useful insights on how to teach the intervention. They also served to avoid any potential impact of the teacher variable (Appendix III).

3.3.4 The pre-test and post-test

A pre-test and post-test were made use of in the present study. As this study did not involve random selection, a pre-test was administered to explore the level and the homogeneity of the students before the outset of the intervention. The pre-test was also necessary for comparing students' scores in order to see whether or not the subjects had made progress in the post-test. We paid close attention to the two tests so that we could choose the most proper testing methods that would meet the objectives of this study.

3.3.5 Testing Methods

Each test comprised four tasks, including questioning, summarizing, clarifying (true/false questions justified), and predicting (infer the meaning of words based on context). Such triangulation is used to reduce the single test method effect (Bachman, 1990) and thus would contribute to increasing the reliability and validity of the tests.

3.3.6 Piloting of the tests

The reason behind piloting was to anticipate the issues we would encounter on the day of running the tests, such as the timing, construction of the questions, length of the texts, linguistic difficulty, etc., so that the tests would fit the aim of this study and the target population. The pre-test, post-test and delayed tests were tested on a small group of common core students studying at Wellada High School in Sale. The piloting was conducted in the circumstances similar to those in the real intervention. Modifications involved the four tasks.

3.4 Procedures

This study is constituted of five different stages: the preparatory stage, the pre-test, the intervention, the post-test, and the delayed test.

3.4.1 Preparatory stage

Before the intervention, we contacted the instructors to participate in this experiment. After agreeing to participate, the teachers were informed of the aim, objectives, implementations, and procedures. This was so helpful that it opened room for improvement, modification, and remediation that were suggested in advance.

3.4.2 The pre-test administration

In the absence of random selection, a pre-test was administered to explore the level and the homogeneity of the subjects. They were told by their respective instructor (in the presence of the researcher) that the pre-test was part of a study in the form of an experiment in which they would participate. The instructor further informed the participants of the study requirements, along with the amount of time it would last. They were also assured that their scores would be kept confidential and would be utilized for the aim of the intervention. The subjects were administered the test and asked to read the instructions very carefully before writing down their answers.

3.4.3 The Intervention

3.4.3.1 The Experimental Group

This study involved one experimental group of 38 subjects. This intact group received twenty training sessions (twice a week, and one hour was allotted for each session), in which they were exposed to the treatment: RT (see section 2 above).

3.4.3.2 The Training Sessions

During the first and second sessions, the students were given a reading passage about 'Citizenship'. Students were asked to read the passage silently. After that, the teacher explained how the RT procedure operates. The instructor modelled questioning and predicting. The students only observed the instructor modelling these two reciprocal reading skills. For instance, in teaching

students to ask questions, the instructor said that self-questioning oneself while reading helps to determine if the material has been understood and that the best questions to generate should be similar to those a teacher would ask in the exam. After that, the instructor modelled how to generate five factual questions. Generating questions while reading helps students to check their own understanding. The instructor then modelled predicting. The reading passage that was distributed among students included some underlined words. The instructor modelled how to infer the meaning of the words given in the question. Every word has a synonym in the text. Thus, the instructor explained to students how to predict the meaning of any given word based on the text. He said that one should pay heed to the category of the word; if that word was a verb, then its synonym would be a verb in the text, and if that word was an adjective, then its synonym would be an adjective, and so on.

During the second session, the instructor modelled clarifying and summarizing, using the same reading passage about 'Citizenship'. As for clarifying, the instructor engaged students in a dialogue in which he uttered some statements, which students were supposed to determine their falsity or truth (Appendix IV-B). The final skill was summarizing. The instructor tried to summarize the main idea of each segment in the reading passage based on the five-macro-rules (Brown & Day, 1983). The same procedure was adopted throughout the remaining ninth reading passages. At the end of the training, most students have reached the stage at which they can model the four skills, assuming their control over strategy use, and ultimately using the strategies with little or no teacher support.

3.4.4 Control group

Like the experimental group, the control group was chosen from Wellada High School in Sale. This group includes 34 participants. The control group resumed their ordinary reading classes after the administration of the pre-test. The first expository reading passage students in the control group were given was about 'Citizenship'. During the pre-reading stage, the instructor actively engaged students in a great discussion about the topic. He asked them different questions, the aim of which was to activate their schemata, draw their attention to the topic, and anticipate what the text would be talking about. Based on their answers, the teacher highlighted relevant keywords on the board. Afterward, students were asked to read the text silently. Then, the students were requested to answer the reading comprehension questions in pairs. During pair work, the teacher made sure to monitor, correcting the work of those who had finished, answering questions as well as making sure everybody was on task. When the time was up, the teacher corrected the questions orally with the contribution of the whole class, always praising when given correct answers. As a post-reading activity, students were asked to write a role-play related to the theme of the reading text. They were equally expected to play interesting role plays (about election day, assistance, collaboration, etc.); they were immensely motivated and felt a sense of achievement when the teacher applauded them for their work. The same procedure was adopted with almost the same tasks throughout the remaining reading passages.

During the experiment, the researcher attended most sessions to see how students reacted to the instruction. Throughout the experiment, discussions were constantly held with the instructor to provide the necessary feedback for the improvement of the training. The researcher sometimes had informal conversations with the participants so as to know their reflections and perspectives on the intervention.

3.4.5 Post-test administration

At the completion of the intervention, reciprocal participants and control participants were administered an immediate post-test. The latter was roughly similar to the pre-test they took at the outset of the intervention (Appendix I-B).

3.5 Scoring and Data analysis

This section describes in detail the test components and how they were scored. It also presents how the data were submitted to SPSS.

3.5.1 Scoring

Each test had a sum total score of twenty. The test consists of four components, each of which was allocated a score based on its weighting. This being the case, question formation was given five marks. Summarizing was also given a sum total score of five points. Note that subjects who did not provide complete answers were allocated a point that corresponded to the quality and quantity of their answers. In addition, grammar and misspellings were discounted. Predicting was given four points, $\frac{1}{2}$ for each word. Clarifying constituted of three questions, each of which was assigned two points: one point for saying whether the statement is true or false, and one for clarifying one's response. The reason why the four test components received distinct weightings can be justified on the ground of the validity of each testing method and the time allocation that each testing method would take. In addition to the researcher, two other independent raters scored the tests. They were given the scoring criteria and were requested to read the text before scoring (Appendix I-D). The researcher also checked whether the two raters observed the established criteria of scoring identified in the marking schemes.

3.5.2 Data analysis

The data were submitted for analysis using SPSS. Since the data were continuous, descriptive statistics were used to measure the central tendency. As the means obtained from the data are insufficient to make cause-effect claims, we resort to inferential statistics. We ran an independent sample t-test to analyse the data. It was used to find out if there was a significant difference between the experimental and the control group.

4. Results and Discussion

4.1 Presentation and analysis of the results

Before embarking on the analysis of the results obtained from this study, it is worth noting that the data serving as the basis for the present study was quantitative. The statistical data consisted of pre-test and immediate post-test scores. Both the control and experimental group took a pre-test prior to instruction and an immediate post-test after the intervention.

4.1.1 Means and gain scores

In order to answer the research questions, analyses of both descriptive and inferential statistics were run. Starting with descriptive statistics, the findings demonstrated that there were numerical differences between the experimental group and the control group.

4.1.2 Descriptive Statistics

The analysis of descriptive statistics shows that both groups almost scored the same at the onset of the training (pre-test), which entails that all the participants had the same level of reading comprehension. The means of the immediate post-test, however, indicate that the experimental group outperformed the control group, indicating that the performance of reciprocal participants on the four comprehension-monitoring skills improved after the intervention. However, the results of descriptive statistics just indicate a numerical difference in the groups' mean scores (see Table4 and Figure2). It remains, then, to run inferential statistics to confirm if the numerical differences are statistically significant.

The descriptive statistics for the pre-test and the post-test, including group means and standard deviations for each group, appear in Table 4, and the group means are plotted on the graph in Figure 1. Accordingly, the analysis of descriptive statistics shows that both groups scored the same at the beginning of the training (pre-test). The pre-test means scores (M= 12.60 for the experimental group and M= 12.89 for the control group) suggest that all the participants had the same level of understanding and reading comprehension. The means of the post-test, however, reveal that the experimental group outperformed the control group (with a - 03.67 gain score). In other words, the participants who received the training in RT improved after the training. Nevertheless, it is worth noting that descriptive statistics only indicate a numerical difference in the groups' mean scores (see Table 4 and Figure 1). It remains, then, to run inferential statistics to confirm if these numerical differences are statistically significant.

| Group | N | Pre-test | | Immediate post-test | |
|-----------|----|----------|------|---------------------|------|
| | | M | SD | M | SD |
| Ex group | 38 | 10.78 | 1.38 | 13.44 | 1.88 |
| Con group | 34 | 10.83 | 1.07 | 10.52 | 1.13 |

Table4: Numerical summary of the data set.

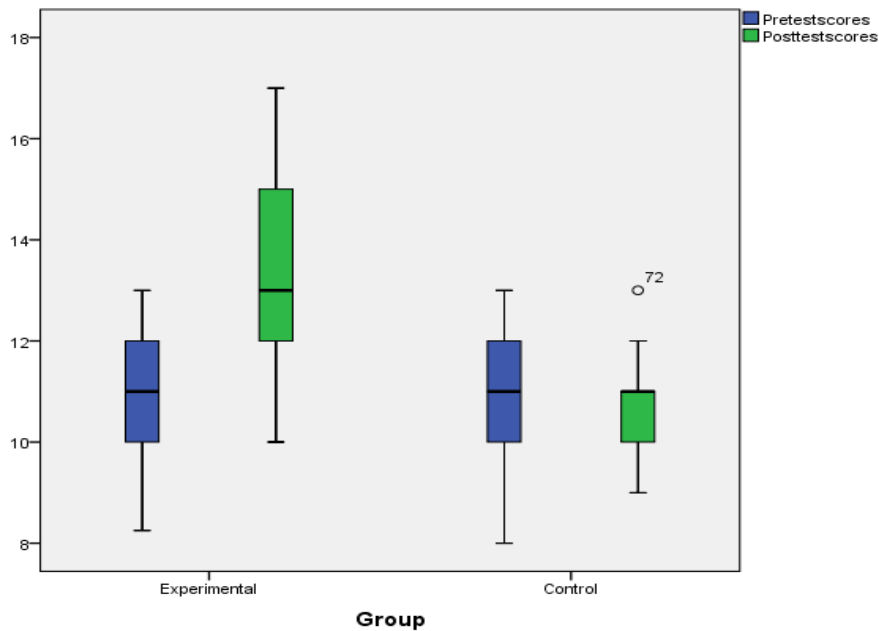


Figure1: Boxplots of the pre-test, immediate post-test, and delayed post-test of both groups.

4.2 Inferential Statistics

4.2.1 The Independent T-test

Before embarking on the statistical analysis, let us examine the data sets to see if they adhere to the assumptions of the independent t-test, namely the normality of the data and equal variances.

4.2.1.1 Normality of the data

Statistically speaking, the null hypothesis states that the population is normally distributed, whereas the alternative hypothesis states that it is not normally distributed. If the test p-value is less than the pre-defined significance level, we can reject the null hypothesis and conclude that the data from which the Ss were drawn are not normally distributed. If the p-value is greater than the predefined significance level, we cannot reject the null hypothesis. Looking at the data numerically in Table 5, there is normal distribution in the pre-test of both groups since the p-value is greater than the pre-defined significance level.

To proceed with the analysis of the results, we need to explore the normal distribution of the data sets. In this respect, two different statistical tests for the normal distribution— the Kolmogorov–Smirnov and the Shapiro–Wilk— were used. We ran these two statistical tests for both the control group and the experimental group to test the normality of the data. Table X demonstrates that the data was normally distributed for each group. The effects were significant in the pre-test of both groups, and the post-tests for both groups were normally distributed.

Tests of Normality

| Group | | Kolmogorov-Smirnov ^a | | | Shapiro-Wilk | | |
|----------------|--------------|---------------------------------|----|------|--------------|----|------|
| | | Statistic | df | Sig. | Statistic | df | Sig. |
| Pretestscores | Experimental | ,127 | 38 | ,129 | ,951 | 38 | ,099 |
| | Control | ,180 | 34 | ,007 | ,936 | 34 | ,046 |
| Posttestscores | Experimental | ,147 | 38 | ,037 | ,954 | 38 | ,119 |
| | Control | ,190 | 34 | ,003 | ,900 | 34 | ,004 |

a. Lilliefors Significance Correction

Table5: Tests of Normality

Wilcox (2003) advanced that formal tests, such as the Kolmogorov–Smirnov and the Shapiro–Wilk, often suffer from low power to detect violations, and thus we cannot be certain that the data is normally distributed by only looking at the numerical results of these tests. We need to explore graphics as well.

The histograms in Figure3 indicate that the data are normally distributed (the median is centred in the boxes, and there are equal-length tails or whiskers on both ends of the boxes). For the pre-test data, all participants are symmetrical around the median, meaning most participants scored the same. However, the control group displays some non-normality because outliers are shown, or boxes are asymmetrical around their medians. We also see some clear skewness in the X.

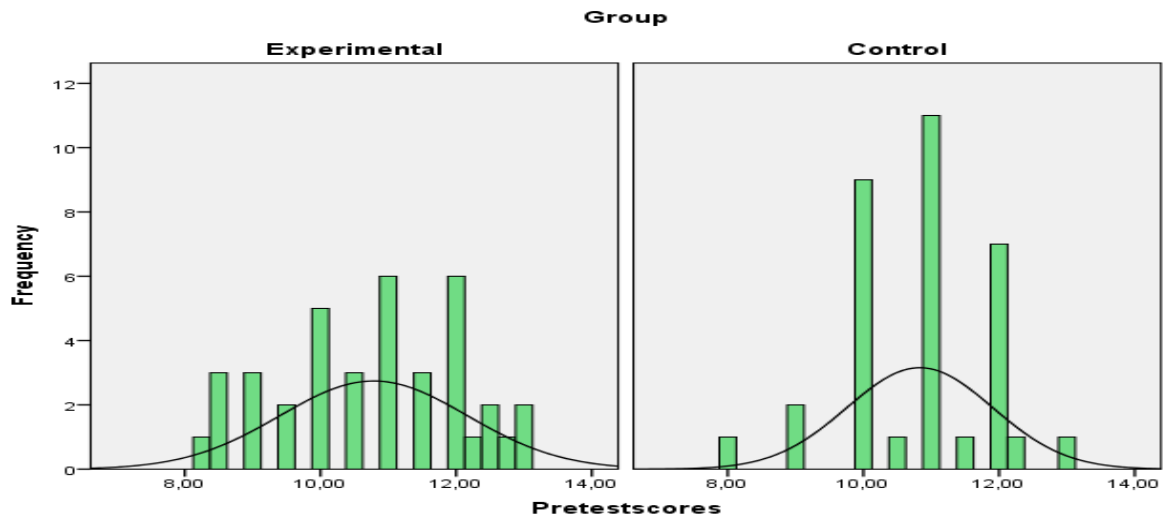


Figure3: Group pre-test histograms

4.2.1.2 Homogeneity of variances

Homogeneity of variance is assessed using Leven’s Test for Equality of Variances. In order to meet the assumption of homogeneity of variances, the p-value for Leven’s Test should be above .05. If Leven’s Test yields a p-value below .05, then the assumption has been violated.

In the Test of homogeneity of variances table, we should look under the Sig. column. If the p-value is more than .05, then we have met the assumption of Homogeneity of Variances. If the p-value is less than .05, then we have violated the assumption of Homogeneity of Variances, and we should use a non-parametric test to conduct the analysis. According to Table 6, the assumption of homogeneity of variances is met since the p-value is greater than the pre-defined significance level (see Table 6).

| | | Levene's Test for Equality of Variances | |
|----------------|-----------------------------|-----------------------------------------|------|
| | | F | Sig. |
| Pretestscores | Equal variances assumed | 3,839 | ,054 |
| | Equal variances not assumed | | |
| Posttestscores | Equal variances assumed | 9,744 | ,003 |
| | Equal variances not assumed | | |

Table 6: Leven’s Test for Equality of Variances

4.2.1.3 Independent T-test

The preliminary description of the data revealed that the experimental group outperformed the control group without being subject to any inferences (the control group, M= 10.52, Sd=1.13, N= 34); the experimental group, M=13.44, Sd=1.88, N= 38) (see table 4). However, the Independent Sample T-test was run at the inferential level, which uncovered the degree of significance between both groups.

The analysis of the Independent T-test confirmed that there were statistically significant differences between the performance of both groups in the post-test. With a t=7.843, df= 70, P=.000, and 95% CI=2.175, 3.659, we can confidently reject the null hypothesis, which says that there is no difference between the experimental group and the control group on the ground of RT.

In principle, the observed patterns in Table 7 below appear in a consistent and meaningful way. Based on the table below, one can be 95% confident that the differences are, by no means, the effect of the treatment given to the experimental group. Most

significant is the positive response in favor of the research hypothesis the current study attempts to test. In this respect, we reject the null hypothesis and conclude that the experimental group outperformed the control group in the post-test.

Table 7: The Independent Sample T-test

| | | Levene's Test for Equality of Variances | | t-test for Equality of Means | | | | | | |
|----------------|-----------------------------|-----------------------------------------|------|------------------------------|--------|-----------------|-----------------|-----------------------|-------------------------------------------|---------|
| | | F | Sig. | t | df | Sig. (2-tailed) | Mean Difference | Std. Error Difference | 95% Confidence Interval of the Difference | |
| | | | | | | | | | Lower | Upper |
| Pretestscores | Equal variances assumed | 3,839 | ,054 | -,163 | 70 | ,871 | -,04799 | ,29422 | -,63479 | ,53882 |
| | Equal variances not assumed | | | -,165 | 68,717 | ,869 | -,04799 | ,29016 | -,62689 | ,53091 |
| Posttestscores | Equal variances assumed | 9,744 | ,003 | 7,843 | 70 | ,000 | 2,91796 | ,37204 | 2,17594 | 3,65997 |
| | Equal variances not assumed | | | 8,053 | 61,695 | ,000 | 2,91796 | ,36236 | 2,19354 | 3,64238 |

4.3 Discussion

This section is devoted to the analysis and discussion of the results obtained.

Hypothesis1: there is no significant difference between the control group and the experimental group in the pre-test.

According to the first hypothesis, all the participants scored almost the same in the pre-test, which entails that both groups had the same level at the outset of the training. Basically, it was theorised that the experimental group and the control group would perform differently on the pre-test. The results obtained from both descriptive and inferential statistics indicated that the groups' performances were nearly parallel. Inferential statistics particularly showed that there was no statistically significant difference between the pre-test means of the experimental group and the control group (see Table4). Given these findings, the first hypothesis is confirmed, and therefore the study provides evidence that all participants had the same level of congruity and homogeneity at the beginning of the study.

Hypothesis2: There is no significant difference between the control group and the experimental group in the post-test.

The second hypothesis argued that the intervention would have no effect on Moroccan EFL common core students' ability to gain comprehension-monitoring skills—summarizing, clarifying, questioning, and predicting. In essence, it was theorized that the treatment would have no effect on students' scores in the post-test. The results obtained from both descriptive and inferential statistics indicated that the groups' performances were different. Inferential statistics particularly showed that there was a statistically significant difference between the post-test means of the experimental group and the control group (see Table4). This would suggest that RT improved students' comprehension monitoring skills, enabling them to self-monitor, self-regulate, evaluate and check their own understanding. The null hypothesis is, therefore, rejected.

The findings obtained support the work of Brown & Palincsar (1982) on the effect of RT on comprehension-monitoring skills. In this regard, Moroccan EFL common core students who experienced the RT intervention enhanced their comprehension-monitoring skills from pre-test to post-test, compared to participants in the control group who received no training. It seems that the experience of the RT intervention intensified students' ability to generate factual questions, provide a brief summary, clarify whether some statements or questions are true or false, and predict words' meanings based on the context. This appreciation was reflected in their performances in the immediate post-test. The scores of the experimental group drastically increased, showing improvement in the immediate post-test at the end of the intervention. The students in the control group who were not taught the RT performed roughly similarly in the pre-test and post-test, respectively.

The study in question seems to yield interesting findings. These tend to suggest evidence that RT fostered comprehension-monitoring skills of Moroccan EFL common core students, and thus it can be concluded that this metacognitive technique helps the students to monitor, regulate, and check their own overall reading comprehension. In particular, the results indicated that reciprocal students scored higher in the immediate post-test compared to students in the control group. These findings are in line with conclusions drawn by previous research conducted to examine the same issue (Brown & Palincsar, 1982; Alfassi, 1998; Hart & Speece, 1998).

5. Conclusion

This paper has examined the effect of the RT procedure on Moroccan EFL common core students' comprehension monitoring skills—question generation, summarizing, clarifying, and predicting. The findings showed that the performance of all participants was equal at the outset of the intervention. However, the subjects who were trained using RT scored higher than those in the control group in the immediate post-test.

5.1 Pedagogical implications

Our findings suggest many pedagogical implications for instructors, students, and educators in an EFL reading context. The most significant pedagogical implication is that RT is an instructional metacognitive strategy that, through decent training on the four comprehension-monitoring skills, can help the learner monitor, regulate, and check their own comprehension. RT is by far one of the most crucial metacognitive instructional techniques that teachers can implement in their classrooms, regardless of the level of students. Given that teachers are willing to adjust this method to students' ability, interest, environment, and level of learning, it can yield interesting outcomes in the end. It remains, then, to note that RT necessitates enough guided practice, provision of support and assistance, scaffolding, encouragement, etc. If all these tenets are considered, students are most likely to assume the role of their teacher.

5.2 Limitations

The quasi-experimental design generally has got many limitations, and thus this study is not an exception. The limitations identified in this paper are no random selection, a small sample, and time constraints. Because of these limitations, our findings could not be generalized, and hence they are open to discussion. However, since the design adopted is similar to classroom reality, we believe the results obtained from this study might be safely applied to any Moroccan common core reading classroom.

Funding: This research received no external funding.

Conflicts of Interest: The authors declare no conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers.

References

- [1] Alfassi, M. (1998). Reading for Meaning: The efficacy of reciprocal teaching in fostering reading comprehension in high school students in remedial reading classes. *American Educational Research Journal*, 35, 309–332.
- [2] Bachman, L. (1990). *Fundamental considerations in language testing*. Oxford University Press, Oxford.
- [3] Brown, A., L. (1985). *Reciprocal teaching of comprehension strategies: a natural history of one program for enhancing learning* (Tech. Rep. No. 334). Urbana Champaign: The University of Illinois, Center for the study of reading.
- [4] Brown, A. L., & Day, J., D. (1983). Macro-rules for summarizing texts: the development of expertise. *Journal of Verbal Learning and Verbal Behaviour*, 22(1), 1-14.
- [5] Brown, A. L., & Palincsar, A. (1982). Inducing strategic learning from texts by means of informed, self-control training. *Topics in Learning and Learning Disabilities*, 2(1), 1-17.
- [6] Faerch, C., Haastrup, K., & Phillipson, R. (1984). *Learner language and language learning*. Clevedon: Multilingual Matters Ltd.
- [7] Fry, E. (1977). Fry's readability graph: clarification, validity, and extension to level 17. *Journal of Reading*, 21(3), 242-252.
- [8] Glaser, R. (1990). The re-emergence of learning theory within instructional research. *American Psychologist*, 45, 29-39.
- [9] Hart, E. R., & Speece, D., L. (1998). *Reciprocal teaching goes to college: Effects for post-secondary students at risk for academic failure*. *Journal of Educational Psychology*, 90 (4), 670- 681.
- [10] Howell, D. C. (2002). *Statistical methods for psychology*. Pacific Grove, CA: Duxbury/Thomson Learning.
- [11] Lysynchuk, L. M., Pressley, M., & Vye, N., J. (1990). Reciprocal teaching improves standardized reading comprehension performance in poor comprehenders. *The Elementary School Journal*, 90, 469-484.
- [12] Mandel, E., Osana, H. P., & Venkatesh, V. (2013). Addressing the effects of reciprocal teaching on the receptive and expressive vocabulary of 1st-grade students. *Journal of Research in Childhood Education*, 27, 407-426.
- [13] Palincsar, A., & Brown, A. (1984). Reciprocal teaching of comprehension-fostering and comprehension monitoring activities. *Cognition and Instruction*, 1, 117-175.
- [14] Pearson, P.D., & Johnson, D., D. (1978). *Teaching reading comprehension*. New York: Holt, Rinehart & Winston.
- [15] Spivey, N., R. & Cuthbert, A. (2006). Reciprocal teaching of lecture comprehension skills in college students. *Journal of Scholarship of Teaching and Learning*, 2, 66-83.
- [16] Vygotsky, L., S. (1978). *Mind in society: The development of higher psychological processes*. In Cole, M., John-Steiner, V., Scribner, S., and Souberman, E. (eds.). Cambridge, MA: Harvard University Press.
- [17] Wilcox, R. (2003). *Applying contemporary statistical techniques*. San Diego, CA: Elsevier Science.