

Article

Empowering Educators to Sustain Reflective Teaching Practices: The Validation of Instruments

Tun Zaw Oo^{1,2}, Anita Habók³  and Krisztián Józsa^{1,3,*} ¹ Institute of Education, Hungarian University of Agriculture and Life Sciences, 7400 Kaposvár, Hungary² MTA-MATE Early Childhood Research Group, 7400 Kaposvár, Hungary³ Institute of Education, University of Szeged, 6722 Szeged, Hungary

* Correspondence: jozsa@sol.cc.u-szeged.hu

Abstract: This study aimed to investigate the sustainability of an instructional design utilizing the novel approach of the Reflective Teaching Model for Reading Comprehension (RTMRC), and to continuously develop and validate instruments for reflective measures to ensure the sustainable teaching of reading comprehension. The RTMRC design was featured based on two main parts: the reflective teaching process (planning, acting, reflecting and evaluating) and the reading comprehension process (reader, strategy, text and task). Then, a quasi-experiment (the pre- and post-test control group design) was conducted with 168 grade-9 students to assess the performance of the RTMRC in practice. The reflective questionnaire was also applied as an instrument to gather feedback from students about the instructional context, with the aim of supporting the sustainable reflective practices of the teacher. The reading achievement in the experimental group was evaluated and fostered using this model. Through the analyses of *t*-tests and Cohen's *d* effect size, it was found that the experimental group teaching with RTMRC outperformed the control group without RTMRC teaching. Using exploratory factor analysis, confirmatory factor analysis, Rasch analysis and testing measurement invariance across groups, we report that the instruments were found to be valid for measuring the effect of the RTMRC in teaching reading comprehension in English Language Teaching (ELT) in a sustainable manner.

Keywords: reflective teaching; model; reading comprehension; English language teaching

Citation: Oo, T.Z.; Habók, A.; Józsa, K. Empowering Educators to Sustain Reflective Teaching Practices: The Validation of Instruments. *Sustainability* **2023**, *15*, 7640. <https://doi.org/10.3390/su15097640>

Academic Editor: Xuesong (Andy) Gao

Received: 6 April 2023

Revised: 30 April 2023

Accepted: 2 May 2023

Published: 6 May 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

In current educational thinking, the concept of reflection is used in a variety of contexts: reflective teaching, reflective practices, reflective feedback, reflective questions and so on. This concept can be found in the work of Dewey [1], who refers to reflective practice as “assessing the grounds (justification) of one’s beliefs, the process of rationally examining the assumptions by which we have been justifying our convictions” [2] (p. 442). Understanding the high complexity of the instructional process and teachers’ sustainable development has significantly encouraged teachers to use the constructivist approaches in teaching. Accordingly, the constructive approaches aim to ensure teachers do not stop learning, instead, upgrade them to become reflective practitioners who know what has happened and what is happening in the instructional process [3]. The concept of reflection and reflective practice in education can promote sustainability in teaching and learning by encouraging teachers to constantly evaluate and improve their methods. This can lead to the sustainable development of effective instructional processes that enhance students’ learning outcomes over time [2,3].

Many teachers in education, however, are still uncertain about the idea of reflective teaching, which they interpret to simply mean thinking about the teaching–learning process, without deeper analysis. Actually, the term reflective teaching is well known. Paterson and Chapman [4] provided a clear discussion of reflective practice. They found that a person

reflecting on his/her actions is not merely observing past activity but is consciously looking with emotions and enthusiasm and combining this information with previously existing knowledge. Some studies [5–7] suggested that teachers can use reflective teaching for various purposes: reflecting on teachers' decisions, teaching approaches, teaching practices, students' recognition of the reflective teaching process and students' attitudes and learning activities. In one study [8], it was also found that reflective teaching can enhance and maintain the motivation levels of both teachers and students. Furthermore, Gordon [9] described two types of reflections, reflection in action (during the action) and reflection on action (after the action), and stated that these types are of great importance for teachers whenever they teach their students the new type of learning skill. Therefore, there is no doubt about the importance of reflective teaching for all teachers and their students.

To reflect on the current Myanmar education, it is a highly centralized system, and teaching approaches are also teacher-centered [10]. However, Myanmar is improving its education system and upgrading teachers' skills as well as curriculum development by cooperating with some international countries, such as the United States, England (for teachers' English proficiency skill development) and Japan (for curriculum development) as well as many international organizations [11]. In the Myanmar education system, English as a foreign language (EFL) is also regarded as mandatory to learn starting from the kindergarten level, and, thus, English language teaching (ELT) is essential in all sectors of education [12]. According to some studies [13–15], most of the school texts are written in English, and, thus, the English language is the one every parent in Myanmar wants their children to learn most. Furthermore, students competitively take additional English courses outside, not only for the compulsory English subject at their state schools but for further studies and work opportunities [5].

The Ministry of Education [16] hopefully created the National Education Strategic Plan (NESP) 2016–2021 for producing highly qualified citizens in Myanmar. In NESP 2016–2021, it can be clearly seen that teachers are encouraged to research with high quality, fostering the development of innovative teaching ideas and pedagogic skills. One project, Strengthening Pre-service Teacher Education in Myanmar (STEM), which was organized by the United Nations' Educational Scientific and Cultural Organization [17], mentioned that the new curricula at the upper secondary school level are based on reflective teaching practices, and, thus, "More support is needed to embed reflection in each lesson—teacher educators have acknowledged that reflection is the first element to go if they do not have enough time for the lesson" (p. 72). Recognition of this encouraged us to conduct high-quality research to evaluate the effect of reflective teaching procedures and to develop and validate reflective measures for ELT teachers in that context. In the current study, we focused on teaching reading comprehension, since reading skill is regarded as a fundamental skill for acquiring knowledge in the Myanmar EFL context, and students have to deal with reading comprehension in almost all language learning tests [18].

2. Literature Review

2.1. *Criteria for the Sustainable Development of the Effective Teaching Model*

A teaching model is an instructional design of fundamental procedures that a teacher can use in his/her teaching to complete students' effective achievement [19]. Joyce, Weil and Calhoun [20] also averred that the term 'teaching model' for a teacher means his/her systematically prepared plan of teaching procedures based on how to teach, what to teach, what kinds of activities will be given and what kinds of teaching aids will be used to be appropriate in the instructional context. According to Awla [21], the teacher should not try matching students with the teacher; however, they should try, with great effort, to learn different kinds of teaching styles to match with students. Thus, the teacher needs the instructional knowledge, skills and attitude to try, with great effort, for effective instruction. Eggen and Kauchak [22] also clarified the teaching model as a blueprint for sustainable teaching, in which the teacher is able to apply all of his/her knowledge, skills and attitude at

the uppermost level. Consequently, the specific teaching model is essential for all teachers' effective teaching.

In Eggen and Kauchak's [22] book, entitled 'Strategies and models for teachers: Teaching content and thinking skills', they firmly stated that the teaching models have some common instructional design characteristics, such as goal (to help students think critically and behave rationally), phases (for helping students succeed with the specific learning goals) and foundations (for supporting students with strong theoretical background, good intention and motivation). Furthermore, Reiser and Dempsey [23] stated some instructional design criteria, which are essential for developing a teaching model. These criteria are that a teaching model has encouraged (1) child-centered approach; (2) goal-oriented procedures; (3) meaningful performance in the instructional context; (4) the systematic evaluation of outcomes in a valid and reliable way; (5) the empirical measurement and possible self-correction; and (6) the allowance for a team effort. Based on these criteria, we decided to develop a reflective teaching model for reading comprehension in ELT.

2.2. Reflective Teaching Process

Reflective teaching is a sustainable-directed process, in which a teacher examines what occurs in teaching and learning in the classroom [24]. Actually, the word 'reflection' is based on the teacher's belief that instructional knowledge is a temporary and never-complete state, and, thus, it is a process of continuous modification based on his/her teaching practices [25]. It is clearly seen in Aliakbari and Adibpour's [26] (2018) study that reflective teaching can help teachers overcome their teaching difficulties and weaknesses in their earlier teaching. Therefore, it is necessary for teachers regarding how to reflect or what kinds of steps they need to follow in the reflective teaching process. Different authors suggest different steps within this reflective teaching process.

Taggart and Wilson [27] indicated that the teacher can take three steps in reflective teaching: planning, reflecting and evaluating. Clarke [5] also clearly averred that there are five steps in the reflective teaching process. They are identifying a problem, planning, acting, evaluating and following-up/reflecting. Dennison [28] identified similar steps: planning, acting, reflecting and feedback. Furthermore, in one study [29], it was found that there are four steps in reflective teaching, including acting, reflecting, analyzing and evaluating. According to Pollard, Black-Hawkins and Hodges et al. [30], reflective teaching involves some steps: planning, acting, reflection, analysis and evaluation. Hulsman, Harmsen and Fabriek [31] applied five steps of reflective teaching in their study, including acting, looking back on the action (reflecting), awareness of essential aspects, creating alternative methods of action (planning) and trial/testing (acting step for the next cycle). In one reflective teaching approach, three steps of critical inquiry (reflection), analysis and self-directed evaluation were also described [3]. In several studies on reflective teaching [24,32,33], the steps of planning, acting, reflecting, evaluating and feedback were outlined. Although the conceptions found in the work of these researchers do not entirely overlap, four common and important points can be identified, namely, planning, acting, reflecting and evaluating. We consider that these four are important for the reflective teaching process.

2.3. Reading Comprehension Process

Reading is a survival skill for all language learners, although there are different interpretations of reading based on different reading purposes, different background knowledge and different views [34]. The process of reading first- and second-language texts necessarily and clearly involves interactions between the reader and the text [35]. Reading is an essential skill of taking out meaningful information from the author's written text, and, thus, it cannot be separated from the word 'comprehension' [36].

In the case of reading comprehension in ELT, the teacher helps the students use their ability to extract the meaning from the reading text and properly interpret that information [37]. Therefore, Khamis [38] said that learning a language is not an easy task due to its complex rules, structures and systems. Reading comprehension is described

by Lim, Eng, Mohamed and Ismail [39] (p. 146) as “a cognitive process that takes place when an individual interacts with the text”. Connors-Tadros [40] also pointed out that “reading is an active and complex process that involves: (a) understanding the written text, (b) developing and interpreting meaning, and (c) using that meaning as appropriate to the type of text, purpose, and situation” (p. 2). Therefore, sustainable teaching of reading comprehension is not a simple and easy task for teachers; instead, they need to consider reflective teaching as an essential characteristic of effective teaching [41]. To reflect on the reading event, it is necessary for the teachers to consider what factors (to be reflected) are influencing students’ reading comprehension achievement.

Different researchers describe different influencing factors on the instructional reading event. There are three factors influencing students’ reading comprehension achievement, such as reader, text and context [42]. Furthermore, in Yusuf and Fitriasia’s [43] study, they pointed out that four main variables of teacher, strategy, reader and text influence students’ reading comprehension achievement. However, Zhang and Zhang’s [44] study clearly described only three factors influencing students’ reading events, such as text, reader and context interaction. Common key factors among those that affect the reading event can be found, as described by a range of authors [45–50]. They note that, although it is difficult to identify all factors and variables that affect students’ reading comprehension, the most common ones that can emerge during instruction for teacher reflection are reader, strategy, text and task. Therefore, we considered these factors in reading comprehension as the main components of instructional context that can be reflected during the reading comprehension instruction process.

2.4. Reflective Teaching Model for Reading Comprehension

Taking into account the two theoretical directions of the reflective teaching process (planning, acting, reflecting and evaluating) and reading comprehension process (reader, strategy, text and task), we consulted experts and developed a novel instructional design, the Reflective Teaching Model for Reading Comprehension (RTMRC) (Figure 1), in accordance with the above instructional design criteria. By following this instructional design approach, teachers can apply the four main steps of planning, acting, reflecting and evaluating to develop sustainable reflective teaching practices that can be maintained over time. This will help to ensure the sustainability of their teaching practices and promote long-term educational sustainability.

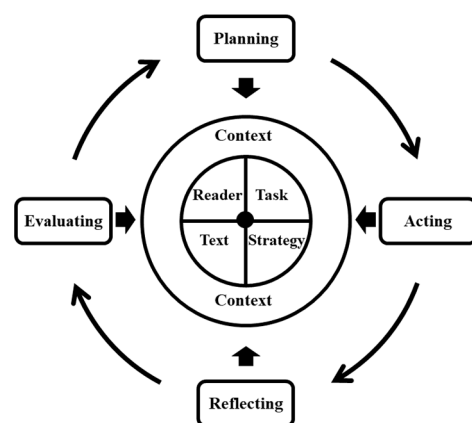


Figure 1. Reflective teaching model for reading comprehension. Source: Oo and Habók [51] (p. 133).

2.4.1. Planning

The teacher prepares a plan to teach the English reading text to the students. In this plan, the teacher can think ahead about what to teach (text), how to teach it (strategy), whom to teach (reader) and what activities the students can undertake (task). Richards and Lockhart [24] observed that during lesson planning, teachers can prepare for factors that

can be expected to arise in the lesson. Here, teachers can also reflect on and evaluate factors of reader, strategy, text and task in teaching reading. They suggest that teachers can use different types of reflective tools, such as the reflective questionnaire and teacher's diary, for the reflective step in RTMRC.

2.4.2. Acting

In the *acting* step of RTMRC instructional design, the teacher can implement the parts that were previously planned.

2.4.3. Reflecting

The teacher reflects on the state of the classroom after the actions of instruction. There are some reflective teaching tools suggested by some researchers, such as videotaping [52], diary or journal writing [53] and peer observation [54]. However, we adopted a reflective questionnaire (as the students' feedback or eyes) based on Brookfield's [55] reflective ways to obtain reflections on the *reader, strategy, text* and *task*. In this reflection, the teacher also presented reflective questions related to the reading text.

2.4.4. Evaluating

Richards and Lockhart [24] suggest that the teacher can evaluate two factors in particular, namely, the teaching–learning process (formative assessment with a reflective questionnaire) and student achievement (summative assessment with a post-test). After reviewing the students' responses to such a questionnaire, the teacher can investigate the factors (reader, strategy, text or task emphasis) that can be altered for the subsequent lesson.

2.5. Previous Findings on Reading Comprehension Achievement

Reading is an essential skill for students that plays a vital role in the sustainability of their academic and professional development [56]. Therefore, there have been studies investigating the effectiveness of instructional strategies [57–60] on students' reading comprehension. This study observed that sustainable teaching techniques, such as reciprocal teaching, interactive teaching and questioning, have proven to be effective in enhancing the achievement of reading comprehension. For example, Hamdani [61] examined the effect of reciprocal teaching on reading comprehension and found a medium effect size of Cohen's $d = 0.46$. Another study, by Decristan et al. [62], found a small effect size of $d = 0.21$ for reciprocal teaching. As for interactive teaching, a recent meta-analysis by Ceyhan and Yıldız [63] investigated the effect of interactive teaching on reading comprehension and found a small effect size of $d = 0.20$. However, the authors noted that the effect size was larger for studies that used a more intensive form of interactive teaching, such as scaffolded instruction. A recent meta-analysis by Liu [64] investigated the effect of questioning on reading comprehension and found a small effect size of $d = 0.21$. Additionally, the author suggested that the teacher should give students the right to ask questions to the teacher. Overall, these three teaching strategies could produce small and medium effect sizes (Cohen's d) on students' reading comprehension achievement. However, it is worth noting that the effectiveness of each teaching strategy may depend on several factors (instructional context), including the reader's characteristics, the implementation of the strategy, the level of difficulty of the text and the nature of the reading comprehension task [51]. Therefore, it is important for teachers to create reflective teaching practices to be able to consider these several factors (instructional context) of reading comprehension achievement.

Reflective teaching has been suggested as a way for teachers to enhance their instructional strategies and improve their students' reading comprehension achievement. For example, Oo et al. [65] examined the impact of a reflective teaching program on the reading comprehension skills of upper secondary school students. The results showed that students in the experimental group (given the reflective teaching program) outperformed those in the control group (without the reflective teaching program) on measures of reading comprehension. Another study, by Htun et al. [66], also inquired about the effects of a program

that focused on reflective teaching on the reading comprehension skills of Myanmar upper secondary school students. The results showed that students in the experimental group with reflective teaching made greater gains in reading comprehension than those in the control group. Oo and Habók [67] also explored the impact of a reflective teaching intervention on the reading comprehension skills of high school students. The findings showed that the teacher's reflective practices enhanced students' learning outcomes in reading comprehension. Hence, it is crucial to determine the combined impact of teachers' reflective practices and three method-centered teaching strategies, such as reciprocal, interactive and questioning, on students' reading comprehension achievement.

2.6. Aim and Research Questions

The purpose of this study was to examine the efficacy of the RTMRC model, which utilized three teaching strategies (reciprocal, interactive and questioning) to enhance students' sustainable development in reading comprehension skills. Additionally, the study aimed to continuously improve and validate suitable reflective assessment instruments for teaching reading comprehension in English Language Teaching (ELT). The research questions for the current research report are as follows:

- Q₁: Does the reflective questionnaire measure the factors (reader, strategy, text and task) that affect the students' sustainable reading events?
- Q₂: Do the reading tests measure the students' sustainable achievement in reading comprehension in ELT?
- Q₃: Does the RTMRC improve students' sustainable achievement in reading comprehension in ELT?

3. Materials and Methods

3.1. Instruments

A quasi-experimental design was adopted in this study, with pre- and post-tests. These test items were based on those of school textbooks prescribed by the Basic Education Curriculum, Syllabus and Textbook Committee of Myanmar [68]. These tests have 26 reading comprehension items (5 identifying referred questions, 5 sentence completion questions, 5 short-answer questions, 10 summary completion questions and 1 paragraph-writing question). The pre- and post-tests contained different question sets; however, they measured the same concepts and knowledge. During the treatment period, the teacher used a reflective questionnaire (20 items with responses on a 4-point Likert scale) to reflect on the instructional context involving reader, strategy, text and task. The reflective questionnaire was used only for the treatment group, not for the control group. This questionnaire was adapted from one developed by Richards and Lockhart [24] for investigating events while teaching reading.

3.2. Sampling, Participants and Procedures

Cluster sampling was used for this study (Table 1). In all, 168 9th-grade students (aged 13–15 years) in Myanmar were chosen as participants and were randomly assigned to an experimental group (N = 83) or a control group (N = 85). Out of 168 students, the number of male students was 78, and 90 were female students. The quasi-experimental study was conducted over the course of five weeks (25 sessions). First, we administered the pre-test to detect the initial difference between the two sub-samples. After the first measurement with pre-test, the experimental group participated in the developmental sessions using the RTMRC. The control group was not given this treatment. During the treatment period, three teaching strategies, namely, reciprocal teaching, interactive teaching and questioning, were used with the RTMRC. The students completed the reflective questionnaire (also known as the student-feedback questionnaire) three times during the treatment period, after the completion of the use of each instructional strategy (their procedures are also described as follows). The objective of the questionnaire was to help teachers reflect on the reader, strategy, text and task, thereby increasing reflection and awareness. At the very

end of the treatment period, a post-test was administered to both groups to compare their achievement.

Table 1. Cluster sampling procedures.

Cluster Sampling	Expected Sample Size
Population	About 400 9th-grade students in Sagaing, Myanmar
Groups (clusters)	Four basic education upper secondary schools in Sagaing, Myanmar
Obtaining a simple random sample	Two basic education upper secondary schools were selected
Sample	All 9th-grade students from the selected two basic education upper secondary schools

The procedures of the three teaching strategies are also described as follows.

3.2.1. Reciprocal Teaching

- Divide the class into small groups of 5–6 students each.
- Provide each group with a text to read and assign each member of the group a different role: summarizer, questioner, clarifier and predictor [69].
- Explain to the students the role of each member of the group, and model how each role works by demonstrating with a short text [70].
- Have students read the text silently, and then start the reciprocal teaching process. Each member of the group takes turns to perform their assigned role, with the group leader facilitating the discussion [71].
- The summarizer starts by giving a brief summary of the text, highlighting the main points and ideas [72].
- The questioner then asks questions about the text, based on what they have read and the summary provided by the summarizer [73].
- The clarifier provides explanations for any difficult or confusing parts of the text, using their own knowledge or by looking up information [74].
- The predictor makes predictions about what might happen next in the text based on the information they have gathered so far [73].
- After each member of the group has had a turn, the group leader summarizes the discussion and highlights any key points or ideas that emerged [72].
- Finally, have each group share their insights and conclusions with the whole class [75].

3.2.2. Interactive Teaching

- Begin by activating students' prior knowledge about the topic to be studied, using strategies, such as brainstorming, concept mapping or discussion [76].
- Provide students with a text to read and encourage them to use a variety of reading strategies, such as skimming, scanning, questioning, clarifying and summarizing, to help them understand the text [77].
- Engage students in interactive activities, such as small-group discussions or debates, to help them develop their critical thinking skills and deepen their understanding of the text [78].
- Provide opportunities for students to collaborate on reading tasks, such as peer editing or working in pairs to complete comprehension questions [76].
- Provide feedback to students on their reading comprehension skills, both individually and as a group, to help them identify areas for improvement and build on their strengths [79].

3.2.3. Questioning

- Preview the text and activate students' prior knowledge by asking questions related to the topic of the text [76].
- Introduce the questioning strategy by modeling how to ask different types of questions about the text, such as literal, inferential and evaluative questions [80].
- Provide students with a copy of the text and ask them to read it individually or in pairs [81].
- Encourage students to use the questioning strategy while reading by stopping at predetermined points and asking questions related to the text [82].
- After reading, ask students to share their questions and discuss them as a group, using evidence from the text to support their answers [83].
- Ask higher-level follow-up questions that require students to evaluate the text, make connections to their own experiences or consider alternative perspectives [80].
- Provide feedback to students on their questions and their ability to use the questioning strategy effectively, both individually and as a group, to help them improve their reading comprehension skills [80].

4. Results

4.1. Findings from the Reflective Questionnaire

Q₁: Does the reflective questionnaire measure the factors (reader, strategy, text and task) that affect the students' sustainable reading events?

To answer this research question, we performed the following analyses on the questionnaire: exploratory factor analysis (EFA), confirmatory factor analysis (CFA), some reliability and validity measures and measurement invariance across gender (male and female groups).

4.1.1. EFA

EFA aims to investigate the factors that influence students' performance [84]. In this study, we analyzed these factors and determined whether the questionnaire could measure the main factors of the RTMRC: reader, strategy, text and task. After applying the EFA, three items were eliminated from the first version of our questionnaire for failing to meet the minimum criteria of not loading above 0.3 on any factor, loading less than 0.4 on any factor and no cross-loading of 0.3 or above [85]. After the application of these criteria, we had four main components, totaling 17 items: reflection on the reader (five items), reflection on the strategy (five items), reflection on the text (four items) and reflection on the task (three items). All of these were chosen based on their factor loadings, which were all greater than 0.40. Finally, three items were eliminated because item 12, *I feel ashamed when my teacher asks me to read in English out loud alone*, had loadings of 0.461 and 0.692 for components 3 and 5. Item 13, *while the English teacher is explaining something, I understand it easily; however, it is difficult to do the reading comprehension task*, was also deleted because it also had factor loadings 0.571 and 0.421 for components 1 and 4. Finally, item 20, *the teacher gives me enough time to comprehend the reading passages*, was also deleted because its factor loading was too low. These results are shown in Table 2.

Table 2. Factor loadings from the reflective questionnaire.

Items	Components			
	Reflection on Reader	Reflection on Strategy	Reflection on Text	Reflection on Task
7. I like the English teacher to explain everything related to the reading tasks.	0.826			
3. I feel ashamed when my English teacher asks me to read the English text out loud alone.	0.765			

Table 2. Cont.

Items	Components			
	Reflection on Reader	Reflection on Strategy	Reflection on Text	Reflection on Task
16. I like the English teacher to use the blackboard/chalkboard while teaching reading comprehension.	0.712			
17. When I don't understand something while reading the English text, I like to guess the meaning by connecting with other related words.	0.693			
4. I do better at reading in English when I work with others.	0.510			
19. I like the reading techniques the English teacher uses because they help me remember the vocabulary.		0.889		
10. I like the English teacher using the relevant questions while teaching the reading text.		0.772		
15. I like the strategy the English teacher uses in teaching the reading passages.		0.729		
1. I like the English teacher's good classroom management.		0.669		
5. I can actively participate in learning reading comprehension because I hear the English teacher's voice well.		0.560		
8. I like the reading text because it is very interesting when the teacher provides us with the reflective questions.			0.843	
6. I like the reading text because it is easy to take out the questions from the reading passages to discuss.			0.785	
2. I like the reading text because it is easy to catch the main ideas to summarize it.			0.751	
14. The reading text looks difficult to understand; however, I like it because it is easy to answer reading comprehension questions after the teacher's explanation.			0.713	
9. I like learning by doing tasks (e.g., taking notes, underlining, highlighting) related to reading texts.				0.785
11. I like to participate in the collaborative activities of learning reading comprehension.				0.742
18. I like the teacher giving us various types of reading comprehension exercises.				0.576

Note: Factor loadings < 0.4 are suppressed.

4.1.2. CFA

CFA was also used through the structural equation modeling (SEM) to establish how closely the factors of reader, strategy, text and task were related to one another. In the use of CFA, Nami and Koizumi [86] suggest that non-significant chi-square (χ^2) and positive degrees of freedom (df) (one or above) should be used to indicate a good fit. In this study, the chi-square was non-significant ($\chi^2 = 234.88$, $p = 0.08$, $df = 213$). Therefore, our questionnaire could be investigated with regard to the fit indices. The following goodness-

of-fit indices were employed to evaluate model fit: goodness-of-fit index (GFI), comparative fit index (CFI) and root mean square error of approximation (RMSEA). The values for CFI and GFI range from 0 to 1, and larger values confirm a better fit. Values larger than 0.90 show an acceptable model fit. The RMSEA also indicates a model fit and also ranges from 0 to 1, but values of 0.08 or less show a good model fit [87]. In this study, the values for these goodness-of-fit indices (GFI = 1.00, CFI = 0.94 and RMSEA = 0.06) were acceptable, indicating that our CFA model was well-fitted. In this model, X^2 describes the distance between the model and the data, but it depends on the sample size: df describes the difference between the number of data points and parameter numbers to be estimated; GFI describes the minimum discrepancy function for the perfect fit; CFI describes the power of the model compared to a situation without the model; and RMSEA describes how much error or unexplained variance remains after applying the model.

In this CFA model, the item–factor correlation coefficients, which ranged from 0.46 to 0.87, are shown in Figure 2. According to Kline [88], these values are adequate if they are >0.30 . Therefore, the items and factors are closely related to one another for developing a good construct for the reflective questionnaire.

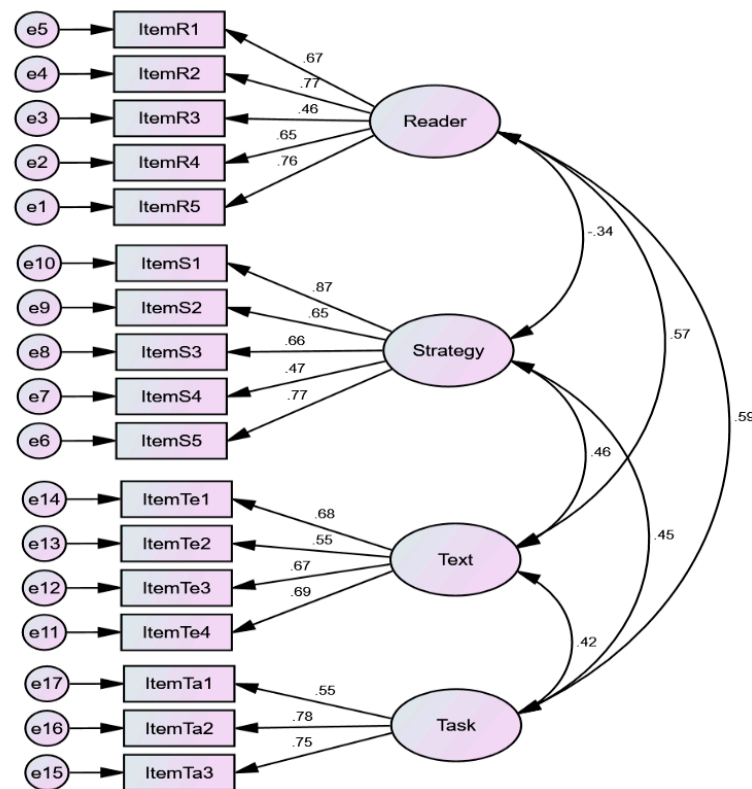


Figure 2. CFA model for the questionnaire (N = 249, for three reflections).

4.1.3. Reliability and Validity

We used IBM SPSS Statistics 23.0 to measure the reliabilities, means, standard deviations and correlations for the validity investigation of this instrument. Internal consistency reliability (measured with Cronbach's alpha) and composite reliability (CR) were estimated to evaluate reliability. The internal consistency reliability of the first three factors (reflection on reader, strategy, text) was greater than 0.70; however, for the last factor, reflection on the task, internal consistency reliability was 0.62. According to Gleaner, Morgan and Leech [89], if the value of Cronbach's alpha is greater than 0.60, this factor also has acceptable reliability. In addition, the CR values for all of the factors or constructs were greater than 0.70 (Table 3). The value found for the Kaiser–Meyer–Olkin (KMO) test showed the suitability of the data for factor analysis. We found a very good value (KMO = 0.716) for the questionnaire

(according to Gleaner et al., $KMO > 0.5$ is acceptable, $KMO > 0.7$ is good) [89]. Therefore, the influential factors for the questionnaire were considered suitable for carrying out further analysis to measure the reliability of students' reading performance.

Table 3. Convergent validity measures of the reflective questionnaire.

Factors	N of Items	Cronbach's Alpha (>0.70) *	Average Variance Extracted (>0.50) *	Composite Reliability (>0.70) *
Reader	5	0.79	0.50	0.83
Strategy	5	0.77	0.51	0.84
Text	4	0.84	0.59	0.85
Task	3	0.62	0.46	0.71
Total	17	0.74	0.52	0.94

Note: * Shows an acceptable level of reliability or validity.

Construct validity was also examined to ascertain whether the construct of the questionnaire behaved in the way predicted by the theories noted above. The convergent validities and discriminant validities were tested to establish the construct validity of the factors. According to Fornell and Larcker [90] and Habók and Magyar [91], factors in the same construct are confirmed if the AVE value is larger than 0.50 and the CR value is larger than 0.70. In the new influencing factors, reader, strategy, text and task, all AVE values were higher than 0.50 (except for 'text' factor), and their CR values were also higher than 0.70 (Table 3). Therefore, convergent validity was confirmed.

According to Fornell and Larcker [90], discriminant validity is established if the values of the square root for AVE are greater than the inter-construct correlations of the component correlation matrix. For this questionnaire, all of the values of the square root of AVE were higher than the inter-construct correlations of the component correlation matrix. Thus, the questionnaire also had good discriminant validity (Table 4).

Table 4. Discriminant validity measures of the reflective questionnaire.

Component	Component Correlation Matrix				AVE	Square Root of AVE
	Reader	Strategy	Text	Task		
Reader	1.00				0.50	0.70
Strategy	0.15	1.00			0.51	0.71
Text	0.21	0.10	1.00		0.59	0.76
Task	0.39	0.11	0.231	1.00	0.46	0.67

Based on the convergent and discriminant validities described above, this questionnaire can be considered to be a valid construct for measuring the factors (reader, strategy, text and task) that affect students' reading events. The EFA and CFA analyses also confirmed that the reflective questionnaire could measure the factors (reader, strategy, text and task) that affect the students' reading events.

Taking into account the principles of sustainability, by using the two considerations above (findings from the test and the reflective questionnaire), we found a correlation between the responses to students' achievement and reflective questionnaire and a significant difference between the results of the experimental group (treatment with the RTMRC) and the control group (without RTMRC treatment). It can also be clearly seen that the students' achievement was related to teacher sustainment of reflective practices. In terms of sustainability, the reflective questionnaire serves as a supplementary tool to aid the teacher's approach. However, the teacher's sustainment of the use of the RTMRC is of the utmost importance. Thus, we can conclude that the RTMRC, along with the use of a

reflective questionnaire, is a sustainable and appropriate approach for measuring students' achievement in reading comprehension in ELT.

4.1.4. Measurement Invariance of the Questionnaire across Gender

To analyze the measurement invariance of the student questionnaire across gender, we used a method called measurement invariance of composite models (MICOM). The main objective of the test was to ensure that both groups of males and females understood the measurements in the same manner for helping the teacher's reflection by supporting their feedback on the instructional context (student questionnaire). The MICOM process depends on latent variable scores, which are represented as composites in PLS-SEM. These composites are linear combinations of indicators, and the algorithm estimates their weights [92]. The MICOM process involves three steps: (1) establishing configural invariance assessment (ensuring that both groups have the same basic factor structure); (2) examining compositional invariance (confirming that composite scores are not significantly different between groups); and (3) verifying that composite mean values and variances are equal. When the configural (step 1) and compositional invariances (step 2) are established, partial measurement invariance is confirmed, which is sufficient for measuring its impacts on the desired outcomes [93]. If the composite has the same mean values and variance across all groups after partial measurement invariance is established, the composite is considered valid. This ultimately leads to the creation of full measurement invariance [92].

For the configural invariance assessment of this study, the same procedures were used to treat missing values and outliers and to code the data for both groups of males and females. Additionally, the algorithm settings in the software were also identical for both groups. As a result, this ensured that the configural invariance was established (recommended by previous authors, such as Ngah et al. and Keung et al.) [94,95]. For the compositional invariance of the questionnaire, a permutation test with a minimum of 1000 permutations was used. We then compared the correlation (c) and 5% of the quantile of the empirical distribution and found that the quantile was smaller than the correlation in each composite, indicating the establishment of compositional invariance [92]. Furthermore, the results were supported by permutation p -values, where all composite values of p were greater than 0.5, establishing the compositional invariance of the questionnaire across male and female groups. Based on these configural and compositional invariance results, we firmly assumed that the establishment of our student questionnaire implies partial measurement invariance (Table 5).

Table 5. Partial measurement invariance results.

Construct	Configural Invariance	Correlation	Correlation Permutation Mean	5% Quantile of the Empirical Distribution	p -Values	Partial Measurement Invariance
Reader	Yes	−0.341	−0.134	−0.93	0.605	Yes
Strategy	Yes	0.532	0.343	−0.236	0.208	Yes
Text	Yes	0.576	0.483	−0.313	0.361	Yes
Task	Yes	0.403	0.193	0.090	0.719	Yes

For verifying whether composites' mean values and variances were equal, Table 6 shows that there were no significant differences in the mean values of the latent variables between the two groups. Additionally, the p -values for all the latent variables were greater than 0.05, further supporting the lack of statistical significance between the two groups of the questionnaire. For composite invariance, the findings showed that there was no significant difference in the latent variables between both male and female groups, except for the 'Reader' construct of the questionnaire. Therefore, the full measurement invariance of the questionnaire could not be constructed despite its confirmed partial invariance construct above. Additionally, we could assume that this questionnaire would measure students' feedback without any bias between groups of gender.

Table 6. Full measurement invariance results.

Construct	Differences of the Composite Mean Value (=0)	95% Confidence Interval	p-Values	Equality of Means	Logarithms of the Composite's Variance Ratio (=0)	95% Confidence Interval	p-Values	Equality of Variance	Full Measurement Invariance
Reader	0.01	[−0.113; 0.114]	0.439	Yes	0.161	[−0.135; 0.151]	0.038 *	No	No
Strategy	0.088	[−0.115; 0.114]	0.098	Yes	0.085	[−0.166; 0.174]	0.196	Yes	Yes
Text	−0.064	[−0.116; 0.106]	0.177	Yes	0.075	[−0.143; 0.158]	0.25	Yes	Yes
Task	−0.043	[−0.113; 0.112]	0.294	Yes	0.091	[−0.15; 0.158]	0.15	Yes	Yes

Note: * $p < 0.05$.

4.2. Reliability and Validity of the Test

Q₂: Do the reading tests measure the students’ sustainable achievement in reading comprehension in ELT?

To investigate the second research question, it is important to discriminate the items that are suitable for testing student sustainable achievement and determine which items are the most difficult or the easiest for the students. We used item-response theory (Rasch analysis) and ran the Quest program to calculate estimates for both the learner ability parameters and the item difficulty levels. The distribution between the students’ ability parameters and the item difficulty levels is presented in Figure 3.

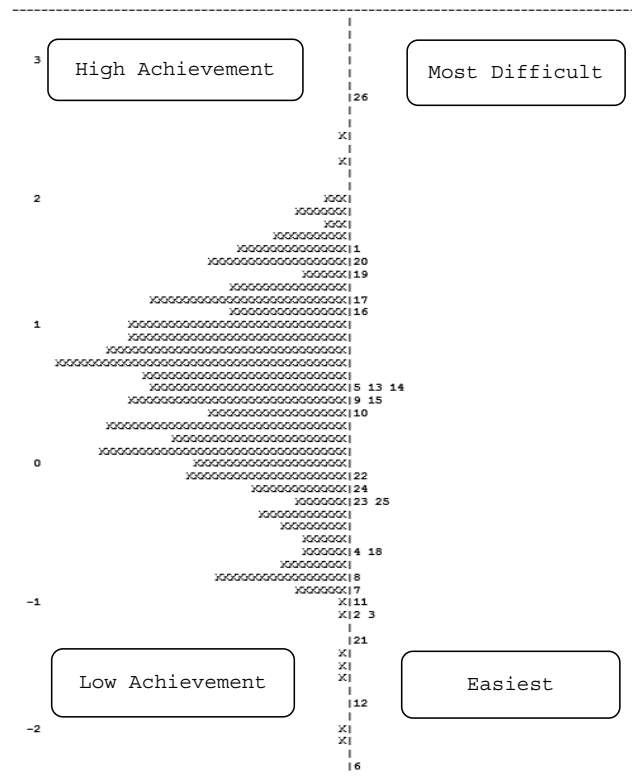


Figure 3. Person–item map indicating the person ability levels and item difficulties.

The left-hand side of the graph shows the ability points of the students, and the right-hand side indicates item difficulty. In Figure 3, it can be seen that the students had higher achievement on items of moderate difficulty, i.e., those which are neither too difficult nor too easy. Further, item 26, requiring a paragraph understanding, was the most difficult item, and items 6 and 12, of the objective type, were the easiest ones. These three outliers were eliminated in pursuit of construct validity, as the students’ achievement levels

were a little higher than zero, the logical number. On the whole, the test items showed a normal distribution. Therefore, the test items showed adjustment to the level of students' knowledge, and this type of test can be used to measure student achievement.

With the help of IBM SPSS Statistics 23.0, the convergent and discriminant validities for the test items were also measured. For convergent validity, Fornell and Larcker [90] suggested three ways of evaluating this: (1) the item reliability for each measure, (2) composite reliability (CR; measured with McDonald's coefficient omega) and (3) the average variance extracted (AVE). The factor loading for an item can assess its reliability on the underlying construct. According to Hair, Black, Babin and Anderson [96], if the factor loading of an item is above 0.50, the item is significant. In this test, the factor loadings for all of the items were between 0.615 and 0.977. For the CR of the items, Nunnally and Bernstein [97] recommend that their value should be higher than 0.70. The CR values for all of the components were higher than 0.70. All of the AVE values were also higher than 0.50. These results (factor loadings, AVE and CR values) are presented in Table 7. Therefore, convergent validity was achieved in this study.

Table 7. Convergent validity measures of the test.

Component	Items	Factor Loadings	AVE	CR
Identifying referred item questions	Item A1	0.977	0.690	0.917
	Item A2	0.734		
	Item A3	0.864		
	Item A4	0.767		
	Item A5	0.79		
Sentence completion questions	Item B2	0.754	0.580	0.768
	Item B3	0.64		
	Item B4	0.615		
	Item B5	0.977		
Short answer questions	Item C1	0.835	0.661	0.817
	Item C3	0.731		
	Item C4	0.835		
	Item C5	0.847		
Summary completion questions	Item D1	0.73	0.718	0.961
	Item D2	0.977		
	Item D3	0.834		
	Item D4	0.866		
	Item D5	0.621		
	Item D6	0.695		
	Item D7	0.977		
	Item D8	0.957		
	Item D9	0.944		
	Item D10	0.792		

Note: AVE (average variance extracted); CR (composite reliability).

Discriminant validity was measured using the heterotrait–monotrait (HTMT) ratio for the correlations. According to Kline [87], a test has significant validity if the HTMT ratios of the components are less than 0.85. In this study, the HTMT ratios for the correlations of the four main components, identifying referred questions, sentence completion questions, short-answer questions and summary completion questions, are shown in Table 8. All of the construct correlation values were less than 0.85. Therefore, discriminant validity was confirmed.

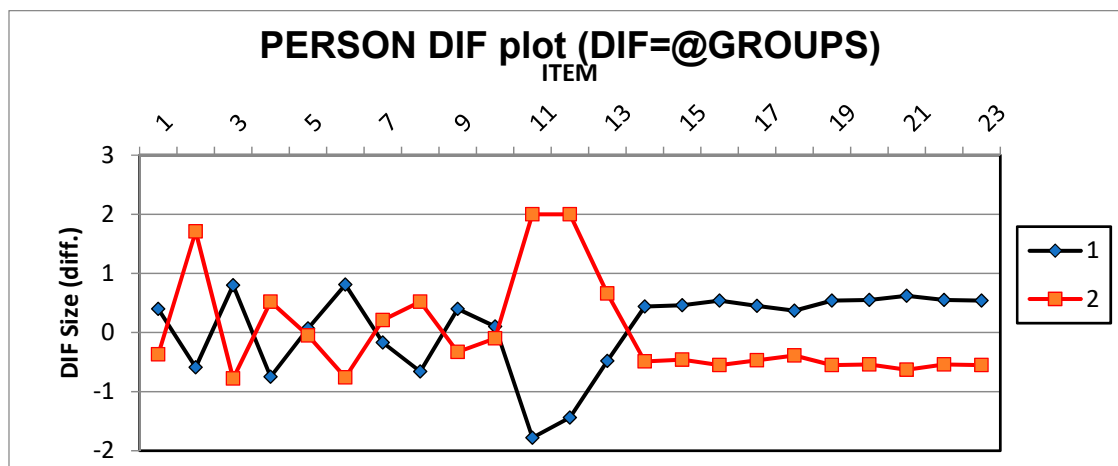
Table 8. HTMT ratios of the correlations of the constructs (discriminant validity of the test).

Construct	1	2	3	4
1. Identifying referred questions	1.00	0.76	0.63	0.65
2. Sentence completion questions		1.00	0.65	0.21
3. Short answer questions			1.00	0.52
4. Summary completion questions				1.00

Note: HTMT (heterotrait–monotrait) ratio = average heterotrait–heteromethod correlations/square root of (average monotrait–heteromethod correlation of (first construct) × (second construct)).

Measurement Invariance of the Test across Groups (Experimental and Control) and Gender

Apart from the above reliability and validity measures, the differential item functioning (DIF) of the reading test was also investigated to establish its measurement invariance across groups of experimental and control. With the help of Rasch Winsteps software, we analyzed the DIF to show the measure of each item (23 items) in the reading test responded by students from both experimental and control groups [98]. All item measures from the reading test fall between 0.20 and -0.20 logits, indicating no biased measure was found for both experimental and control groups (Figure 4). Moreover, for assessing the measurement invariance across students/gender (male and female), Figure 5 also shows that the DIF logits fell between 1 and -1.5 logits, pointing to no significant measurement variance between the male and female groups.

**Figure 4.** Differential item functioning across groups. Note: 1 (experimental group), 2 (control group).

These findings are consistent with the recommended values: (1) negligible if $DIF \leq 0.43$ logits, (2) slight to moderate if $DIF \geq 0.43$ logits, and (3) moderate to large if $DIF \geq 0.64$ logits [99]. Therefore, the results of the DIF analysis suggest that the reading test used in this study exhibits strong measurement invariance across both experimental and control groups, as well as male and female participants. The DIF logits for all items were found to be less than 0.43, indicating that the magnitude of the differences in item performance between the groups was negligible. Therefore, it can be concluded that the reading test used in this study provides a fair and unbiased measure of reading ability across all groups and genders, and it can be used confidently in future research and evaluation efforts.

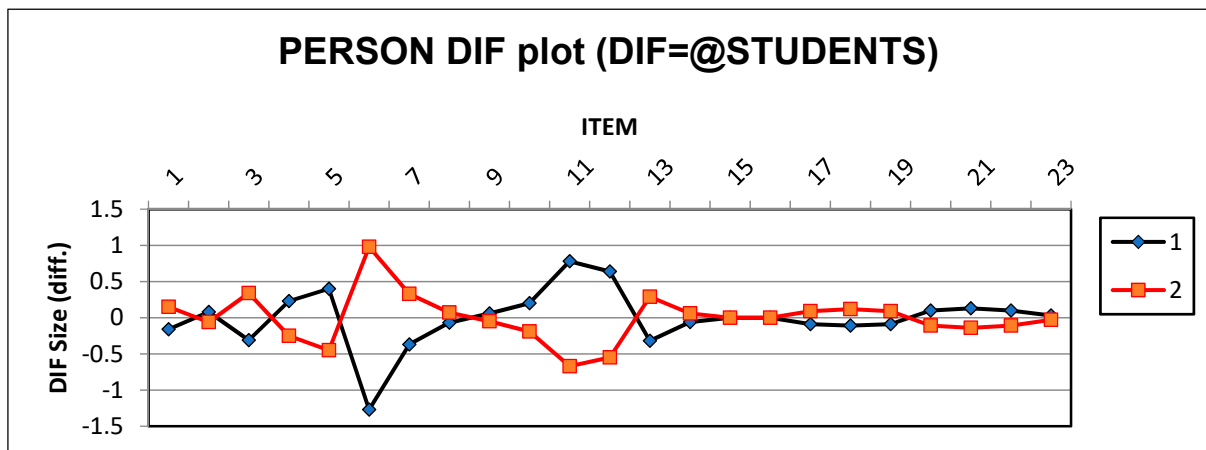


Figure 5. Differential item functioning across gender. Note: 1 (male), 2 (female).

4.3. Findings from the Tests

Q₃: Does the RTMRC Improve Students’ Sustainable Achievement in Reading Comprehension in ELT?

To answer this question, we investigated whether there was a significant difference between the results achieved by the students in the experimental group (taught with the RTMRC model) and the control group (not taught with the RTMRC model). We also measured the effect of the RTMRC approach on students’ reading comprehension achievement by comparing the students’ results before and after the treatment with that approach. Furthermore, we inquired about the impacts of teacher’s reflection (via reflective questionnaire) on students’ reading comprehension achievement.

Before the instruction with the RTMRC, we employed pre-tests to establish initial differences between the experimental and control groups. The data from the pre-tests (Table 9) were investigated using the independent samples *t*-test to establish the differences between the experimental and control groups. No significant differences could be registered between the two groups ($p = 0.386$). The mean scores (11.24) for the experimental group and (10.78) for the control group were not significantly different. Thus, the levels of reading knowledge among the students in these two groups were similar before treatment with the RTMRC.

Table 9. Findings from the pre-tests in the experimental and control groups.

Groups	N	M	SD	MD	df	Effect Size Cohen’s <i>d</i>	Sig
Experimental	83	11.24	2.46	0.46	167	0.19 (very low)	0.386 (n.s.)
Control	85	10.78	2.32				

Note: n.s. = not significant.

At the end of the study, we used post-tests to establish the effectiveness of the RTMRC and the level of significance between the two groups after treatment with the RTMRC. These findings are shown in Table 10. The data were analyzed using the independent samples *t*-test to compare the differences between the control and experimental groups. Because $p = 0.000$, there is a significant difference between students who were taught the reading texts with the RTMRC and those who were not. According to the mean difference, the experimental group ($M = 35.12, Sd = 4.05$) performed significantly better than the control group ($M = 30.44, Sd = 4.65$). We can interpret these results to indicate that teaching with RTMRC was more effective than the traditional teaching methods used in the control group. The effect size (Cohen’s $d = 1.07$) also confirmed that the RTMRC had a much greater

effect on student achievement (Cohen's *d*-values; 0.2 = small, 0.5 = medium, and 0.8 = large, suggested by Cohen, and Goulet-Pelletier and Cousineau) [100,101].

Table 10. Findings from the post-tests in the experimental and control groups.

Groups	N	M	SD	MD	df	Effect Size (Cohen's <i>d</i>)	Sig
Experimental	83	35.12	4.05	4.68	167	1.07 (very large)	$p < 0.001$
Control	85	30.44	4.65				

We also compared the results from the pre-test and post-test of the experimental group to investigate the effectiveness of the RTMRC approach. The findings are shown in Table 11. The data obtained from the pre-test and post-test of the experimental group were examined by applying a paired sample *t*-test to compare the differences between the students' achievement before and after the treatment with the RTMRC approach. Because the *p*-value is 0.000 *** ($* p < 0.05$), this indicates a very significant difference between the students' achievement. Based on the mean difference, the post-test mean value ($M = 35.12$, $Sd = 4.05$) is much higher than that of the pre-test ($M = 11.24$, $Sd = 2.46$). Therefore, it can be said that the students achieved more as a result of the treatment with the RTMRC approach. Regarding effect size, Cohen's *d* value is 7.12. This means that teaching with the RTMRC approach has a significant effect on the students' reading comprehension achievement in ELT.

Table 11. Results from the pre-test and post-test of the experimental group.

Experimental Group	N	M	SD	MD	Effect Size (Cohen's <i>d</i>)	df	Sig
Pretest	83	11.24	2.46	−23.88	7.12 (very large)	82	$p < 0.001$
Posttest	83	35.12	4.05				

Based on the reflective questionnaire to gather students' feedback, we conducted a regression analysis using SmartPLS to examine the impact of different types of reflections on students' reading comprehension achievement. The results showed that reader reflection ($\beta = 0.16$, $p < 0.05$) and strategy reflection ($\beta = 0.11$, $p < 0.05$) significantly impacted students' reading comprehension achievement. However, text reflection and task reflection did not have a significant impact on students' reading comprehension achievement (Figure 6). These findings suggest that reader reflection and strategy reflection may be particularly effective methods for enhancing comprehension among students. The lack of a significant impact of text reflection and task reflection on reading comprehension achievement may indicate that these types of reflection may not be as effective in improving students' reading comprehension skills.

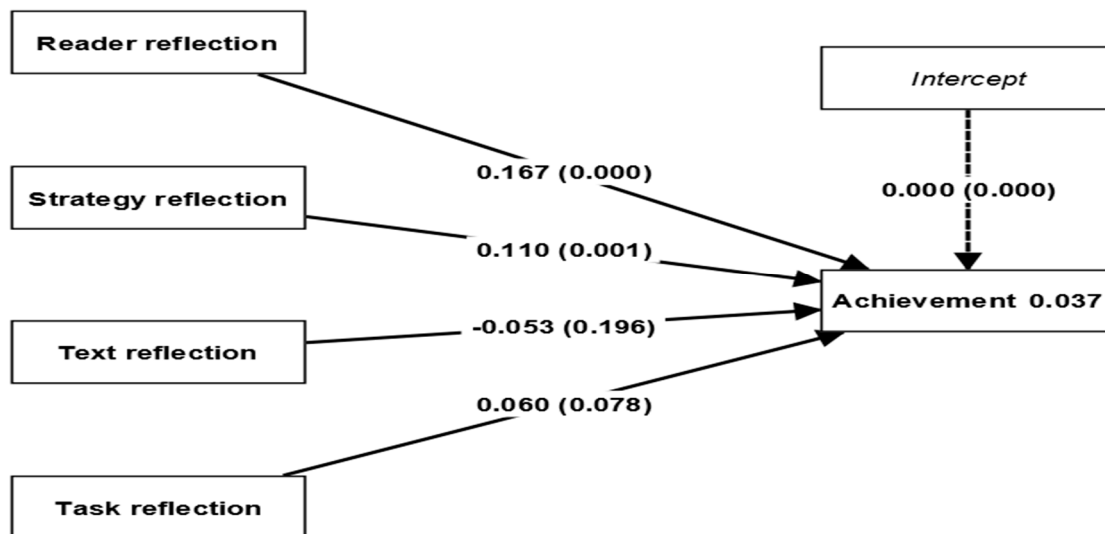


Figure 6. Regression model showing the impacts of the reflective questionnaire on reading comprehension achievement (N = 249).

5. Discussion

In this study, we investigated the sustainable effectiveness of the RTMRC approach and continuously developed valid instruments for the reflective measures in RTMRC instruction. According to Badia [102], one or more assessment tools are required for the use of a model (e.g., Brookfield's model) in research implementation to improve teaching practices. Therefore, to investigate the effectiveness of the RTMRC teaching model, we adopted a quasi-experimental approach of pre- and post-test design, employing an instrument (a reflective questionnaire), and we validated them in different ways.

For the pre-test and post-test, the same content was used with different question sets. Each test had 26 items. In the analysis, we confirmed the validity of the tests with the item-response theory. Discrimination analysis of the items showed that one item (item 26) was seen to be the most difficult, and two items (6 and 12) were seen to be the easiest. These three items were deleted, so only 23 items were retained on the final test. These 23 items on the reading test were also confirmed in accordance with their reliability and validity measures. Both convergent and discriminant validities were confirmed for the construct validity of the test. Furthermore, its measurement invariance was also detected to know whether or not the item bias was found between the experimental and control groups, as well as the male and female groups. The results soundly showed the measurement invariance of the test between these two groups of experimental and control. Therefore, we interpreted that the reading test was valid to measure students' reading comprehension achievement in Myanmar. The reason may be that all types of questions in the reading test are based on the prescribed reading text and its example questions for grade-9 students in Myanmar [56]. The confirmation of measurement invariance across the experimental and control groups indicates that the pre-/post-tests used in this study are equally valid and reliable for assessing reading comprehension achievement in both groups. This is important because it strengthens the internal validity of the study and supports the comparability of the results obtained from the two groups. As the measurement invariance means that the test items measure the same construct in both groups [103], the observed differences in test scores can be attributed to differences in the construct being measured, rather than to measurement bias. In this study, the test items were found to have the same meaning and value for both groups, which suggests that the test is measuring the same underlying construct of reading comprehension achievement in both groups.

In the case of the reflective questionnaire, the original reflective questionnaire (students' feedback) had 20 items. According to Pollard et al. [30], five main factors influence teachers'

reflection: planning, acting, reflecting, analysis and evaluation. However, in the analysis of the data for the reflective questionnaire, it was found that four main factors were most significant (i.e., had high factor loadings). Thus, we eliminated some inappropriate items, retaining only four main factors: planning, acting, reflecting and evaluating. Our study is in line with other studies [46–48,50], which have four main factors in this reflective teaching process. As a result, the new version of the questionnaire had only 17 items with strong reliability for measuring students' preferences for reader, strategy, text and task (for reflection). This new version was also confirmed using CFA measures, and it was also found that the reflective questionnaire had a good fit for the teacher's reflection in reading comprehension. Furthermore, we investigated its construct validity based on convergent and discriminant validity. Both convergent and discriminant validities were also confirmed based on the recommended values. Therefore, we could assume that the reflective questionnaire measures the factors (reader, strategy, text and task) that affect the students' reading events. We also analyzed the measurement invariance of the reflective questionnaire using the MICOM technique in SmartPLS, with a particular focus on gender differences (for its later use in the future). Our results suggest that the questionnaire is measurement-invariant across gender, indicating that the items are interpreted and responded to similarly by both male and female respondents. This finding is in line with the reviewed studies [24,65–67], comparing male and female groups of reading comprehension achievement. Therefore, this finding adds to the evidence supporting the reliability and validity of the questionnaire and increases confidence in the conclusions drawn from our analysis.

We conducted a quasi-experimental study to investigate the sustainable effectiveness of RTMRC on reading comprehension teaching. Apart from the investigation of its sustainable effectiveness, we ascertained that the instruments used in teaching with RTMRC are valid and useful for the investigation of RTMRC effectiveness in teaching reading comprehension. The RTMRC teaching is appropriate in the Myanmar context because it could bring some benefits to both teachers and students. Often, students in Myanmar avoid giving the answer 'no' when their teacher asks them, 'Do you understand me, the text and the questions?'. However, when the students completed an anonymous reflective questionnaire, it was found that they presented their opinions clearly. The students indicated that they did not understand the teacher's questions, they did not like individual work and they did not like their teacher's strategy. Using these opinions, the teacher had to determine what changes to make for the next steps. As the RTMRC provided students the opportunity to present their opinions of themselves (with the help of a reflective questionnaire), we determined that the RTMRC could aid the students in understanding their reading texts to a certain degree. Although the teacher who used the RTMRC may have been a bit overloaded, the instruction was still well-conducted due to the transparent and systematic nature of the RTMRC. The students were found to have positive views of reflective teaching and enjoyed using the reading text. We also investigated the impacts of teacher's reflection on students' reading comprehension achievement using a reflective questionnaire that included four factors: reader reflection, strategy reflection, text reflection and task reflection. The results of the regression analysis indicated that only two of these factors, reader reflection and strategy reflection, had significant impacts on students' reading comprehension achievement. The other two factors, text reflection and task reflection, did not show significant impacts. These findings suggest that a teacher's reflection when focused on the readers' needs and strategies can have a positive impact on students' reading comprehension achievement. The reader reflection factor showed the strongest impact, suggesting that teachers who reflect on their students' needs can better tailor their instruction to meet those needs. The strategy reflection factor also showed a significant impact, indicating that teachers who reflect on the effectiveness of their instructional strategies can identify and implement more effective approaches to support the students' reading comprehension. While the text reflection and task reflection factors did not show significant impacts on students' reading comprehension achievement in this study, it is possible that these factors could be more

influential in different contexts or with different populations of students. Further research could explore the effects of these factors more closely to better understand their impacts on students' learning. After the intervention with RTMRC teaching, both the experimental and control groups completed a post-test, and the results showed a significant difference between the two groups, with the experimental group having significantly higher achievement. The effect size was also high, thus confirming the success of the study. This finding aligns with previous reviews of reflective teaching studies conducted by Oo et al. [65], Htun et al. [66] and Oo and Habók [67]. Notably, sustainable reflective teaching practices were found to have a large effect size (Cohen's $d > 0.80$), which is a more distinct characteristic compared to the small and medium effect sizes (Cohen's $d < 0.80$) observed in the method-centered teaching strategies of reciprocal teaching [61,62], interactive teaching [63] and questioning [64] in the aforementioned reviews.

6. Limitations

Despite the promising findings of this study, there are some limitations that need to be addressed. Firstly, the study only focused on grade-9 students from Myanmar, which may limit the generalizability of the results to other populations. Secondly, the study only investigated the impact of RTMRC on students' reading comprehension achievement and did not consider other factors (e.g., motivation, well-being, thinking skills and so on) that may have contributed to their performance. Thirdly, the sample size may be relatively small for conducting Rasch analysis and structural equation modeling (SEM), which require more participants for more robust results. However, some researchers recommended that the minimum sample size is around 150 for Rasch analysis [104,105], between 100 and 150 for SEM [106] and around 200 for SEM [75]. Lastly, the study relied on self-reported measures from students, which may be subject to social desirability bias or inaccuracies in self-assessment. Thus, other reflective tools (such as tape recording, peer observation, teachers' diary writing or self-reflection notes) are also recommended for future studies.

7. Conclusions

To conclude, our study could prove the sustainable effectiveness of RTMRC on (randomly selected) grade-nine students' reading comprehension achievement in Myanmar. The practical implications of this study are significant for educators and policymakers. By validating the pre-/post-tests and the reflective questionnaire for teachers, we established reliable and valid tools for assessing the effectiveness of reflective teaching on students' reading comprehension achievement. Our findings provide evidence that sustainable reflective teaching is a promising approach to improving students' reading comprehension skills. In addition, the measurement invariance across experimental and control groups, as well as across gender, ensures that the instruments are appropriate for use in diverse settings and with different student populations. This is especially important given the increasing diversity of students in today's classrooms. The results of this study have several practical implications for educators. First, reflective teaching can be incorporated into teacher training programs to help teachers develop the necessary skills to implement this approach in their classrooms. Second, teachers can use validated instruments to evaluate the effectiveness of their teaching and to make data-driven decisions to improve students' learning outcomes. Moreover, policymakers can use the findings of this study to support the implementation of reflective teaching in schools and to advocate for the integration of reflective teaching in teacher-training programs. This can lead to empowering teachers and, ultimately, student learning outcomes.

Overall, the practical implications of this study highlight the importance of using reliable and valid instruments to evaluate the effectiveness of teaching interventions. By validating the instruments and demonstrating the effectiveness of sustainable reflective teaching on students' reading comprehension achievement, this study provides a valuable contribution to the field of education and has the potential to inform future research, teacher-training programs and policy decisions.

Author Contributions: Conceptualization, T.Z.O., A.H. and K.J.; Methodology, T.Z.O. and A.H.; Validation, T.Z.O.; Investigation, T.Z.O.; Data curation, T.Z.O.; Writing—original draft, T.Z.O.; Writing—review & editing, K.J.; Supervision, A.H.; Funding acquisition, K.J. All authors have read and agreed to the published version of the manuscript.

Funding: This study was funded by the Scientific Foundations of Education Research Program of the Hungarian Academy of Sciences and by the ICT and Societal Challenges Competence Centre of the Humanities and Social Sciences Cluster of the Centre of Excellence for Interdisciplinary Research, Development and Innovation of the University of Szeged. The authors are members of the New Tools and Techniques for Assessing Students Research Group.

Institutional Review Board Statement: This study was conducted according to the guidelines of the Declaration of Helsinki, and approved by the Institutional Review Board of the University of Szeged Doctoral School of Education (20/2019, 24 June 2019).

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study through the respective principals.

Data Availability Statement: Data are unavailable due to privacy or ethical restrictions.

Acknowledgments: The authors acknowledge the principals from two secondary schools in the Sagaing region, Myanmar (Basic Education Upper Secondary Schools 5 and 6) for their active cooperation and kind allowance for conducting this research.

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

References

- Dewey, J. *How We Think: A Restatement of the Relation of Reflective Thinking to the Educative Process*; D.C. Heath and Company: New York, NY, USA, 1933.
- Fook, J. Reflective practice and critical reflection. In *Handbook for Practice Learning in Social Work and Social Care*; Lishman, J., Ed.; Jessica Kingsley Publishers: London, UK, 2015; pp. 440–454.
- Babaei, M.; Abednia, A. Reflective teaching and self-efficacy beliefs: Exploring relationships in the context of teaching EFL in Iran. *Aust. J. Teach. Educ.* **2016**, *41*, 1–27. [[CrossRef](#)]
- Paterson, C.; Chapman, J.A. Enhancing skills of critical reflection to evidence learning in professional practice. *Phys. Ther. Sport* **2013**, *14*, 133–138. [[CrossRef](#)] [[PubMed](#)]
- Clarke, P.A. Reflective teaching model: A tool for motivation, collaboration, self-reflection, and innovation in learning. *GA Educ. Res. J.* **2008**, *5*, 1–18.
- Gheith, E.; Aljaberi, N. Reflective teaching practices in teachers and their attitudes toward professional self-development. *Int. J. Progress. Educ.* **2018**, *14*, 161–180. [[CrossRef](#)]
- McAlpine, L.; Weston, C.; Berthiaume, D.; Fairbank-Roch, G.; Owen, M. Reflection on teaching: Types and goals of reflection. *Educ. Res. Eval.* **2004**, *10*, 337–363. [[CrossRef](#)]
- Gill, G.S. The nature of reflective practice and emotional intelligence in tutorial settings. *J. Educ. Learn.* **2014**, *3*, 86–100. [[CrossRef](#)]
- Gordon, E.J. ‘The good, the bad and the ugly’: A model for reflective teaching practices in coaching pedagogy. *Strategies* **2017**, *30*, 21–27. [[CrossRef](#)]
- Ulla, M.B. Teacher training in Myanmar: Teachers’ perceptions and implications. *Int. J. Instr.* **2017**, *10*, 103–118. [[CrossRef](#)]
- Goodman, A.E. Foreword: Making history by investing in the future. A briefing paper from IIE’s Centre for International Partnerships. In *Investing in the Future: Rebuilding Higher Education in Myanmar*; Institute of International Education, Ed.; IIE Publishing: New York, NY, USA, 2013; pp. 5–8.
- Soe, T. A study of contemporary trends and challenges of English language teaching in Myanmar. In Proceedings of the International Conference on Burma/Myanmar Studies Burma/Myanmar in Transition: Connectivity, Changes and Challenges, Chiang Mai, Thailand, 24–25 July 2015; pp. 1–10.
- Lall, M. Evolving education in Myanmar: The interplay of state, business and the community. In *Dictatorship, Disorder and Decline in Myanmar*; Skidmore, M., Wilson, T., Eds.; ANUE Press: Canberra, Australia, 2008; pp. 127–150.
- Maley, A. ELF: A teacher’s perspective. *Lang. Intercult. Commun.* **2009**, *9*, 187–200. [[CrossRef](#)]
- Tin, T.B. Learning English in the periphery: A view from Myanmar (Burma). *Lang. Teach. Res.* **2014**, *18*, 95–117. [[CrossRef](#)]
- Ministry of Education. *National Education Strategic Plan 2016–2021 (Executive Summary)*; Ministry of Education in Myanmar: Napyidaw, Myanmar, 2016.
- United Nations Educational Scientific and Cultural Organization. *Strengthening Pre-Serviced Teacher Education in Myanmar (STEM): Phase II Final Narrative Report*; UNESCO Yangon Project Office: Yangon, Myanmar, 2020.
- Nourdad, N.; Asghari, R. The effect of reflective reading on reading comprehension of Iranian EFL learners. *Int. J. Appl. Linguist. Engl. Lit.* **2017**, *6*, 267–273. [[CrossRef](#)]

19. Ghilay, Y.; Ghilay, R. ISMS: A new model for improving student motivation and self-esteem in primary education. *Int. Electron. J. Elem. Educ.* **2015**, *7*, 383–398.
20. Joyce, B.M.; Weil, M.; Calhoun, E. *Models of Teaching*, 9th ed.; Pearson: Boston, MA, USA, 2015.
21. Awla, H.A. Learning styles and their relation to teaching styles. *Int. J. Lang. Linguist.* **2019**, *2*, 240–245. [[CrossRef](#)]
22. Eggen, P.; Kauchak, D. *Strategies and Models for Teachers: Teaching Content and Thinking Skills*, 6th ed.; Allyn & Bacon Publishing: Boston, MA, USA, 2012.
23. Reiser, R.A.; Dempsey, J.V. *Trends and Issues in Instructional Design and Technology*, 3rd ed.; Pearson Education, Inc.: Hoboken, NJ, USA; Allyn & Bacon: Boston, MA, USA, 2012.
24. Richards, J.C.; Lockhart, C. *Reflective Teaching in Second Language Classrooms*; Cambridge University Press: New York, NY, USA, 2007.
25. Amakyi, M.; Mensah, A.A. Reflective practice in teacher education in Ghana. *Int. J. Educ. Pract.* **2017**, *2*, 42–50. [[CrossRef](#)]
26. Aliakbari, M.; Adibpour, M. Reflective efl education in iran: Existing situation and teachers' perceived fundamental challenges. *Egit. Arast.-Eurasian J. Educ. Res.* **2018**, *77*, 129–144. [[CrossRef](#)]
27. Taggart, G.L.; Wilson, A.P. *Promoting Reflective Thinking in Teachers*; Corwin Press: Thousand Oaks, CA, USA, 2005.
28. Dennison, P. Reflective practice: The enduring influence of Kolb's experiential learning theory. *Compass J. Learn. Teach.* **2009**, *1*, 1–6. [[CrossRef](#)]
29. Kennedy-Clark, S.; Eddles-Hirsch, K.; Francis, T.; Cummins, G.; Ferantino, L.; Tichelaar, M.; Ruz, L. Developing pre-service teacher professional capabilities through action research. *Aust. J. Teach. Educ.* **2018**, *43*, 39–58. [[CrossRef](#)]
30. Pollard, A.; Black-Hawkins, K.; Hodges, G.C.; Dudley, P.; James, M.; Linklater, H.; Pollard, A.; Pollard, A.; Swaffield, S.; Swann, M.; et al. *Reflective Teaching in Schools*, 4th ed.; Bloomsbury: London, UK, 2014.
31. Hulsman, R.L.; Harmsen, A.B.; Fabriek, M. Reflective teaching of medical communication skills with DiViDU: Assessing the level of student reflection on recorded consultations with simulated patients. *Patient Educ. Couns.* **2009**, *74*, 142–149. [[CrossRef](#)]
32. Garzon, A. Unlicensed EFL teachers co-constructing knowledge and transforming curriculum through collaborative-reflective inquiry. *Profile Issues Teach. Prof. Dev.* **2018**, *20*, 73–87. [[CrossRef](#)]
33. Ratminingsih, N.M.; Artini, L.P.; Padmadewi, N.N. Incorporating self and peer assessment in reflective teaching practices. *Int. J. Instr.* **2018**, *10*, 165–184. [[CrossRef](#)]
34. Ahmada, A. The effectiveness of Jigsaw learning model in teaching reading comprehension on narrative text. *J. Darussalam J. Pendidik. Komun. Pemikir. Huk. Islam* **2019**, *10*, 258–268. [[CrossRef](#)]
35. Gilbert, J. A study of ESL students' perceptions of their digital reading. *Read. Matrix Int. Online J.* **2017**, *17*, 179–195.
36. Azmi, A. Teaching reading comprehension by using pictures. *Engl. Educ. J.* **2020**, *7*, 260–271. [[CrossRef](#)]
37. Mohamedi, A.; Rico, A. Assessment of reading comprehension in primary education: Reading processes and texts. *Leng. Mod.* **2020**, *55*, 37–52.
38. Khamis, N.Y. A new competency-based framework for English for specific engineering purposes (ESEP) practitioners. *Asian EFL J.* **2019**, *23*, 46–65.
39. Lim, C.K.; Eng, L.S.; Mohamed, A.R.; Ismail, M. Relooking at the ESL reading comprehension assessment for Malaysian primary schools. *Engl. Lang. Teach.* **2018**, *11*, 146–157. [[CrossRef](#)]
40. Connors-Tadros, L. *National and State Definitions of Reading Proficiency and Measurement of Reading Proficiency (CEELO FAST Fact)*; Center on Enhancing Early Learning Outcomes: New Brunswick, NJ, USA, 2014.
41. Afshar, H.S.; Farahani, M. Reflective thinking and reflective teaching among Iranian EFL teachers: Do gender and teaching experience make a difference? *Procedia-Soc. Behav. Sci.* **2015**, *192*, 615–620. [[CrossRef](#)]
42. Zhang, L.J. English language teaching today: Linking theory and practice. In *English Language Teaching Today: Linking Theory and Practice*; Renandya, W.A., Widodo, H.P., Eds.; Springer International Publishing: Cham, Switzerland, 2016; pp. 127–142.
43. Yusuf, Y.Q.; Fitrisia, D. Investigating metacognitive awareness of reading strategies to strengthen students' performance in reading comprehension. *Asia-Pac. J. Educ. Educ.* **2015**, *30*, 1–16.
44. Zhang, L.; Zhang, L.J. Relationships between Chinese college test takers' strategy use and EFL reading test performance: A structural equation modeling approach. *RELC J.* **2013**, *44*, 35–57. [[CrossRef](#)]
45. Snow, K. *Reading for Understanding: Toward an R & D Program in Reading Comprehension*; RAND Corporation: Santa Monica, CA, USA, 2003.
46. Staden, S.V. Reading between the lines: Contributing factors that affect grade 5 learner reading performance. *RELC J.* **2010**, *44*, 74–83.
47. Suwanto, S. The effectiveness of the paraphrasing strategy on reading comprehension in Yogyakarta city. *J. Lit. Lang. Linguist. Open Access Int. J.* **2014**, *4*, 1–7.
48. Walker, B.J. *Diagnostic Teaching of Reading: Techniques for Instruction and Assessment*, 6th ed.; Merrill Prentice Hall: Upper Saddle River, NJ, USA, 2008.
49. Widdowson, H.G. *Teaching Language as Communication*; Oxford University Press: New York, NY, USA, 2015.
50. Yang, X. Study on factors affecting learning strategies in reading comprehension. *J. Lang. Teach. Res.* **2016**, *7*, 586–590. [[CrossRef](#)]
51. Oo, T.Z.; Habók, A. The development of a reflective teaching model for reading comprehension in English Language Teaching. *Int. Electron. J. Elem. Educ.* **2020**, *13*, 127–138. [[CrossRef](#)]

52. Dollah, S.; Weda, S. Videotaping students' perception in a graduate program EFL class. *TESOL Int. J.* **2018**, *13*, 124–133. Available online: <https://www.tesol-international-journal.com/volume-13-issue-3-2018/> (accessed on 25 April 2021).
53. Gangalakshmi, C.; Naganathan, R. Reflective teaching—A source for reconstituting teaching pedagogy. *Asian EFL J.* **2019**, *23*, 16–31.
54. Zohdijalal, P.; Mohammadi, M. Enhancing professional development of Iranian EFL teachers through collaborative reflection. *Asian EFL J.* **2021**, *25*, 104–113.
55. Brookfield, S.D. *Becoming a Critically Reflective Teacher*; Jossey-Bass: San Francisco, CA, USA, 2017.
56. Bonifacci, P.; Ferrara, I.C.; Pedrinazzi, J.; Terracina, F.; Palladino, P. Literacy Acquisition trajectories in bilingual language minority children and Monolingual peers with similar or different SES: A three-year longitudinal study. *Brain Sci.* **2022**, *12*, 563. [\[CrossRef\]](#)
57. Papastefanou, T.; Marinis, T.; Powell, D. Development of reading comprehension in bilingual and Monolingual children—Effects of language exposure. *Languages* **2021**, *6*, 166. [\[CrossRef\]](#)
58. Qi, W.; Jiang, Y. Use of a graphic organiser as a pedagogical instrument for the sustainable development of EFL learners' English reading comprehension. *Sustainability* **2021**, *13*, 13748. [\[CrossRef\]](#)
59. Schurz, A.; Coumel, M.; Hüttner, J. Accuracy and fluency teaching and the role of extramural English: A tale of three countries. *Languages* **2022**, *7*, 35. [\[CrossRef\]](#)
60. Seraj, P.M.I.; Klimova, B.; Habil, H. Use of mobile phones in teaching English in Bangladesh: A systematic review (2010–2020). *Sustainability* **2021**, *13*, 5674. [\[CrossRef\]](#)
61. Hamdani, B. teaching reading through reciprocal teaching method. *Celt. J. Cult. Engl. Lang. Teach. Lit. Linguist.* **2020**, *7*, 23. [\[CrossRef\]](#)
62. Decristan, J.; Bertram, V.; Reitenbach, V.; Schneider, K.; Köhler, C.; Rauch, D.P. Linguistically responsive reciprocal teaching in primary school: Effectiveness of an intervention study on students' reading competence. *J. Multiling. Multicult. Dev.* **2022**, *11*, 1–19. [\[CrossRef\]](#)
63. Ceyhan, S.; Yıldız, M. The effect of interactive reading aloud on student reading comprehension, reading motivation and reading fluency. *Int. Electron. J. Elem. Educ.* **2020**, *13*, 421–431. [\[CrossRef\]](#)
64. Liu, H. Does questioning strategy facilitate second language (L2) reading comprehension? The effects of comprehension measures and insights from reader perception. *J. Res. Read.* **2021**, *44*, 339–359. [\[CrossRef\]](#)
65. Oo, T.Z.; Magyar, A.; Habók, A. Effectiveness of the reflection-based reciprocal teaching approach for reading comprehension achievement in upper secondary school in Myanmar. *Asia Pac. Educ. Rev.* **2021**, *4*, 1–24. [\[CrossRef\]](#)
66. Htun, N.N.; Nu, T.; Oo, T.Z. The Effectiveness of the Reflection-based interactive teaching approach on students' reading comprehension achievement. *J. Field-Based Lesson Stud.* **2023**, *4*, 85–123. [\[CrossRef\]](#)
67. Oo, T.Z.; Habók, A. Reflection-based questioning: Aspects affecting Myanmar students' reading comprehension. *Heliyon* **2022**, *8*, e09864. [\[CrossRef\]](#)
68. Basic Education Curriculum, Syllabus and Textbook Committee. *English Text: Grade 9*; The Government of the Republic of the Union of Myanmar, Ministry of Education: Yangon, Myanmar, 2019.
69. Palincsar, A.; Brown, A. Reciprocal teaching of comprehension fostering and comprehension-monitoring activities. *Cogn. Instr.* **1984**, *1*, 117–175.
70. Salari, F.; Hosseini SM, H. Competitive team-based learning vs. reciprocal teaching of reading: A study in reading classes. *J. Lang. Teach. Res.* **2019**, *10*, 489. [\[CrossRef\]](#)
71. Dew, T.P.; Swanto, S.; Pang, V. The Effectiveness of reciprocal teaching as reading comprehension intervention: A systematic review. *J. Nusantara. Stud. JONUS* **2021**, *6*, 156–184. [\[CrossRef\]](#)
72. Lin, H.; Yu, H. A collaborative summarization system based on multiagent conversation. *Int. J. Hum.-Comput. Interact.* **2017**, *33*, 563–575.
73. Zhang, X.; Duke, N.K. The role of collaborative dialogue in informational text comprehension and learning. *Read. Res. Q.* **2018**, *53*, 185–204.
74. Mafarja, N.; Zulfaidi, H. Relationship between critical thinking and academic self-concept: An experimental study of reciprocal teaching strategy. *Think. Ski. Creat.* **2022**, *45*, 101113. [\[CrossRef\]](#)
75. Gilbert, F. The reciprocal rebellion: Promoting discussion in authoritarian schools. *Chang. Engl. Stud. Cult. Educ.* **2022**, *29*, 232–250. [\[CrossRef\]](#)
76. Israa, E.A. The effectiveness of interactive teaching strategies in teaching English language. *J. Al-Farahidi's Arts* **2023**, *15*, 483–492. [\[CrossRef\]](#)
77. Alsamadani, H.A. Developing EFL students' vocabulary and reading comprehension skills within an interactive learning environment. *Int. J. Comput. Sci. Netw. Secur.* **2022**, *22*, 145–152.
78. Wagner, M.N.L.; Kupriyanova, M.E.; Maximova, O.B.; Ovezova, U.A. Improving the effectiveness of teaching: The impact of interactive methods in teaching English students as they increase their proficiency level to B2. *Theory Pract. Lang. Stud.* **2022**, *12*, 2232–2238. [\[CrossRef\]](#)
79. Lu, Q. Research on interactive teaching for English classroom in junior high school. *Educ. Linguist. Res.* **2022**, *8*, 77. [\[CrossRef\]](#)
80. Fatmawati, E.; Rahman, A.Q.; Mahmud, M. Teacher's questioning strategies in EFL classroom interaction. *Expo. J. Pendidik. Bhs. Ingg.* **2020**, *9*, 59–72. [\[CrossRef\]](#)
81. Sulaiman, R. Questioning strategies applied by lecturers in EFL class. *Linguist. Initiat.* **2021**, *1*, 41–48. [\[CrossRef\]](#)

82. Ningsih, A.R.; Erlina, D.; Amrina, R.D. English teacher's questioning strategy in EFL classroom interaction. *Jadila J. Dev. Innov. Lang. Lit. Educ.* **2021**, *1*, 478–499. [[CrossRef](#)]
83. Reflianto, S.P.; Kuswandi, D.; Widiati, U. English teachers' competency in flipped learning: Question level and questioning strategy in reading comprehension. *Int. J. Instr.* **2022**, *15*, 965–984. [[CrossRef](#)]
84. Şeker, H. In/out-of-school learning environment and SEM analyses on attitude towards school. In *Application of Structural Equation Modeling in Educational Research and Practice*; Kiene, M.S., Ed.; Sense Publishers: Rotterdam, The Netherlands; Boston, MA, USA; Taipei, Taiwan, 2013; pp. 135–167.
85. Williams, B.; Onsmann, A.; Brown, T. Exploratory factor analysis: A five-step guide for novices. *Australas. J. Paramed.* **2010**, *8*, 1–11. [[CrossRef](#)]
86. Nami, Y.I.; Koizumi, R. Structural equation modelling in educational research: Primer. In *Application of Structural Equation Modeling in Educational Research and Practice*; Khine, M.S., Ed.; Sense Publishers: Boston, MA, USA, 2013; pp. 23–51.
87. Kline, R.B. *Principles and Practice of Structural Equation Modelling*, 3rd ed.; Guilford Press: New York, NY, USA, 2011.
88. Kline, R.B. *Principles and Practice of Structural Equation Modeling*, 2nd ed.; Guilford Press: New York, NY, USA, 2005.
89. Gleaner, J.L.; Morgan, G.A.; Leech, L.N. *Research Methods in Applied Settings: An Integrated Approach to Design and Analysis*, 3rd ed.; Taylor and Francis: New York, NY, USA, 2017.
90. Fornell, C.; Larcker, D.F. Evaluating structural equation models with unobservable variables and measurement error. *J. Mark. Res.* **1981**, *18*, 39–50. [[CrossRef](#)]
91. Habók, A.; Magyar, A. Validation of a self-regulated foreign language learning strategy questionnaire through multidimensional modelling. *Front. Psychol.* **2018**, *9*, 1–11. [[CrossRef](#)]
92. Kagoya, S.M.; Tinali GZ, P.; Caine, J. A multi-group analysis of salient determinants of E-government implementation success in developing countries. A study of Uganda and Tanzania. *Univ. Salaam Libr. J.* **2022**, *16*, 119–138. [[CrossRef](#)]
93. Othman, N.H.; Othman, N.; Juhdi, N.H. Does entrepreneurship education affect pre-start-up behavior in Malaysia? A multi-group analysis approach. *Front. Psychol.* **2022**, *13*, 8729. [[CrossRef](#)]
94. Ngah, A.H.; Gabarre, S.; Han, H.; Rahi, S.; Al-Gasawneh, J.A.; Park, S.H. Intention to purchase halal cosmetics: Do males and females differ? A multigroup analysis. *Cosmetics* **2021**, *8*, 19. [[CrossRef](#)]
95. Keung, C.P.; Cheung, A.C.; Mak, B.S.; Tam, W.W. Perceptions of Hong Kong secondary school teachers on effective pedagogical practices for curriculum reform: A multi-group analysis. *Asia Pac. J. Educ.* **2021**, *41*, 1–13. [[CrossRef](#)]
96. Hair, J.E., Jr.; Black, W.C.; Babin, B.J.; Anderson, R.E. *Multivariate Data Analysis*, 7th ed.; Prentice-Hall: New Jersey, NJ, USA, 2010.
97. Nunnally, J.C.; Bernstein, I.H. *Psychometric Theory*, 3rd ed.; McGraw-Hill: New York, NY, USA, 1994.
98. Adams, D.; Mabel HJ, T.; Sumintono, B.; Oh, S.P. Blended learning engagement in higher education institutions: A differential item functioning analysis of students' backgrounds. *Malays. J. Learn. Instr.* **2020**, *17*, 133–158. [[CrossRef](#)]
99. Zwick, R.; Thayer, D.T.; Lewis, C. An empirical Bayes approach to Mantel-Haenszel DIF analysis. *J. Educ. Meas.* **1999**, *36*, 1–28. [[CrossRef](#)]
100. Cohen, J. *Statistical Power Analysis for the Behavioural Sciences*; Academic Press: New York, NY, USA, 1969.
101. Goulet-Pelletier, J.; Cousineau, D. A review of effect sizes and their confidence intervals, Part I: The Cohen's d family. *Quant. Methods Psychol.* **2018**, *14*, 242–265. [[CrossRef](#)]
102. Badia, G. Combining critical reflection and action research to improve pedagogy. *Portal Libr. Acad.* **2017**, *17*, 695–720. [[CrossRef](#)]
103. Maiolatesi, A.J.; Clark, K.A.; Pachankis, J.E. Rejection sensitivity across sex, sexual orientation, and age: Measurement invariance and latent mean differences. *Psychol. Assess.* **2022**, *34*, 431–442. [[CrossRef](#)]
104. O'Brien, T.P.; Klassen, R.M. Validating a measure of mathematics beliefs for low-income African American middle school students: Applying the Rasch Model. *J. Appl. Meas.* **2012**, *13*, 66–82.
105. van der Ark, L.A.; Sijtsma, K. A latent class approach to item analysis. *J. Educ. Behav. Stat.* **2003**, *28*, 341–362.
106. Teo, T.; Tsai, L.T.; Yang, C. Applying structural equation modeling (SEM) in educational research. In *Application of Structural Equation Modeling in Educational Research and Practice*; Khine, M.S., Ping, L.C., Cuningham, M., Eds.; Sense Publishers: Boston, MA, USA, 2013; pp. 2–21.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.