Research Paper

Effect of Recall Protocol on Iranian Undergraduate University Students' General English Achievements: Examining the Role of Gender

Mahnaz Moayeri[1], Mohammadreza Khodareza[2]*, Valeh Valipour[3], Ramin Rahimi[4]

[1]Ph.D. Candidate, Department of English Language, Tonekabon Branch, Islamic Azad University, Tonekabon, Iran

[2]*Assistant Professor, Department of English Language, Tonekabon Branch, Islamic Azad University, Tonekabon, Iran

[3]Assistant Professor, Department of English Language, Tonekabon Branch, Islamic Azad University, Tonekabon, Iran

[4]Assistant Professor, Department of English Language, Tonekabon Branch, Islamic Azad University, Tonekabon, Iran

Abstract

The multifaceted roles of corrective feedback in second language instruction continue to attract the attention of both researchers and practitioners interested in how it can most effectively be integrated into classroom interaction in ways that benefit second language development (Tedick & Lyster, 2020). Therefore, the present study intends to shed a light on the effectiveness of the recall protocol method on Iranian male and female intermediate EFL learners' general English achievements. To fulfill this study, 80 male and female undergraduate university students were divided randomly into two groups; The experimental group received the recall protocol feedback method whereas the control group was taught through the conventional feedback techniques. Consequently, the results of the independent samples T-test revealed that the experimental group outperformed the control group. Moreover, the results of two-way ANOVA showed that male students obtained significantly higher scores than female students considering the use of recall protocol. Furthermore, the pedagogical implications suggested that such findings can be colossal as a way to assist in creating efficiency, especially on the off chance that the disclosures are to be utilized by L2 teachers.

INTRODUCTION

Swain (1985, as cited in Guo, 2013, pp. 4-5) contended that it isn't sufficient to have as it were comprehensible input for second language learning to take place. It is more vital to have the openings to deliver them so that learners are able to test their theories of target language forms through output. In this way, Swain defined language learning as not as it was being uncovered to comprehensible input, but moreover as a result of the language learner's exertion to create output. Hence, language learning takes put whereas learners lock in conversational interaction, get input, and battle to reformulate and make their yield comprehensible to their interlocutors. Concluding from thinks about conversational interaction, Pica (1994, as cited in Guo, 2013, p.5) famous three commitments of intelligently negotiation claimed to be accommodating for second language acquisition. To begin with, it makes input more comprehensible; second, it gives input and advances the

*Corresponding Author's Email:
mkhodareza@yahoo.com

generation of altered output, and third, it brings learners' consideration to the L2 frame in testing their speculation of the L2. In straightforward terms, in this manner, the conversational interaction inquiries about subsumes as it develops input, output, and adjustment through interaction (Gass and Mackey, 2006a, as cited in Guo, 2013, p.5).

The effectiveness of the strategy in teaching and learning the course of general English to undergraduate students can be a matter of investigation for the present study. Table 1 outlines the immediate recall protocol procedure for both reading and listening. Bernhardt (as cited in Bernhardt and James, 1987, p. 87).

Table 1
Recall Protocol Procedure

Reading
1. Select an unglossed text of approximately 200 words.
2. Tell the students they may read the text as often as necessary and that they will be asked to write down what they recall.
3. Students should be given time to read.
4. Students are asked to write down everything they remember.
5. Students' written protocols are collected.
6. The protocols are used either for an immediate follow-up exercise or for the writing of a future lesson plan that addresses
a.) cultural features,
b.) conceptual features,
c.) grammatical/lexical features that have interfered with comprehension.

Listening
1., Select an unglossed text, with a running time of one to two minutes, at a speed of approximately 200 words per minute.
2. Tell the students that they will hear the text once, but they are not to take written notes.
3. Tell them that when they have finished you will ask them to write down everything they remember from the text.
4. After they have heard the text twice, ask the students to write down everything they remember.
5. Collect the protocols.
6. Use the protocols either for an immediate follow-up exercise or for a future lesson plan that addresses
a.) cultural features,
b.) conceptual features,
c.) grammatical/ lexical /phonological features that interfere with comprehension.

Besides using different feedback approaches, gender differences can also be considered as a determining factor in improving learners' general English achievements. For instance, Gholizade (2013) in her study "The Investigation of Differential Effects of Recast and Metalinguistic Feedback on Accuracy, Fluency, and Complexity of Speaking Performance of Male and Female EFL Learners" attempted to compare the effectiveness of two types of corrective feedback namely, recast and metalinguistic feedback considering the learners' gender. The last ten years witnessed a steady increase in the number of studies that have examined the effects of corrective feedback on L2 speaking performance. This includes

both descriptive and experimental research examining a wide range of variables (e.g., type and amount of feedback, mode of feedback, learner's language proficiency level, instructional context, and attitudes towards feedback). The results obtained from the ANCOVA and t-test showed that corrective feedback, in the form of metalinguistic, was effective in leading to speaking accuracy, fluency, and complexity. This study failed to find any significant difference between male and female participants.

Research Question
During the last decade, some changes have occurred in the procedures of providing feedback.

Classroom teacher feedback was mostly explicit negative feedback, but the emergence of communicative and content-based teaching approaches brought about some changes in the way feedback techniques are used in the classroom. There is now a shift from explicit negative feedback, which may lead to negative affective reactions on the part of the learners, to implicit negative feedback (Jafarigohar and Gharbavi, 2014, p. 696). Firstly, the current research was carried out to shed a light on the effectiveness of recall protocol approaches on Iranian undergraduate university students' general English achievements. Secondly, the present study was done to compare and contrast the male and female learners' performances in terms of their general English achievements using different feedback approaches. More importantly, the results of this study can be applied in teaching general English to university students. Thus, the present study sought to answer the following questions:

RQ1: Is there any statistically significant difference among all participant groups of the study in terms of general English achievement posttest means after being treated with recall protocol?

RQ2: Is there any statistically significant difference among male and female participants of the study in terms of general English achievement posttest means after being treated with recall protocol?

LITERATURE REVIEW
Theories that support CF
Corrective feedback. Corrective feedback (CF) is any type of response or move the teacher makes to warn the learner about the presence of an error (Kartchava, 2019, as cited in Ulrich-Verslycken, HajHamid, and Kaffafi Azar, 2021). Corrective feedback has been a key issue in language teaching and learning and language pedagogy for almost half a century. Whether CF can benefit the second language acquisition process was a highly controversial issue in the early stages of this field. However, with the development of empirical research, increasing evidence has emerged to support that CF can assist language learning by improving

learners' accuracy (Zhang, Cao, and Zheng, 2022).

The recall protocol procedure in L2 reading assessment. Working beneath a learner-based hypothesis of perusing comprehension, Bernhardt (1991) illustrated the failure of conventional, quantitative reading comprehension evaluation to look at the complex and dynamic forms included in L2 reading comprehension. She found that "in the event that a test is to satisfactorily evaluate L2 reading capacity it must recognize the status of the reader's information base," and "an effective evaluation component must be integrator in nature" (Bernhardt, 1991, p. 193). As a research-based, suitable, and reasonable elective, the review convention strategy is seen as a profoundly substantial and successful L2 reading comprehension appraisal degree that gives both subjective and quantitative data. According to Berkemeyer (1989, p.131):

It does not allow students to guess their way through the text nor does it influence students' understanding of the text. In short, the immediate recall protocol demands that the reader comprehend the text well enough to be able to recall it in a coherent and logical manner. This procedure allows misunderstandings and gaps in comprehension to surface; a feature that other methods of evaluation cannot offer. Bernhardt (1991) has long championed the utilization of the quick, free-response reading recall protocol as an especially effective degree of a learner-based L2 reading comprehension worldview.

That paradigm advances a multifaceted reading comprehension inquiry about and evaluation approach combining quantitative and subjective strategies in arrange to reach a more total picture of the comprehension handle. Clearly, the recall protocol method can be an imperative portion of that multifaceted approach. As Brisbois (1992, p. 168) noted: testing methods, such as the recall protocol procedure, need to gain wider acceptance as a measure of reading comprehension. Not only is the recall protocol more sensitive than discrete-point tests, but its sensitivity becomes more pronounced as reading proficiency increases.

Recall as an authentic assessment task. The genuineness of the recall protocol method

to evaluate L2 reading comprehension has come beneath address. Schmidt-Rinehart (1994), reports on investigations that utilized the recall protocol method for investigation. She states, "although the advantages of this comprehension measure are well documented one would rarely be asked to perform a similar task in real life" (Schmidt-Rinehart, 1994, p.186).

In reality, nothing may well be assisted from the truth as can be contended on three levels. To begin with, from a hypothetical point of view, the recall protocol method has been effectively utilized as a noninvasive procedure for understanding how people have put a message together in both L1 and L2 investigations (Bartlett, 1932). Second, from a commonsense level, the recall protocol is experienced daily. The associate at that point continues to relate his or her understanding of the article. Upon afterward review of the first piece we regularly discover that things were excluded or decorated based on the understanding made by the reader.

According to Berkmeyer (1989), the retelling of what somebody has perused is, in truth, a day-by-day, or maybe a common event- one in which human creatures lock in actually and promptly within the course of discussion. Third, for L2 reading comprehension the review convention strategy has been appeared to be an imperative portion of coordinates, numerous measures appraisal approach. Not at all like discrete point measures, it is an integrator errand where students write down everything, they keep in mind approximately what they studied and, in this way, give a wealthy test of their personal development of the text. Thus, the present study considers sets the review convention strategy as an authentic task and clear elective to utilize in conjunction with more conventional disobedient such as the numerous choice and cloze examinations. Bernhardt, (1983 a, pp.31-32) mentions:

1. The recall procedure shows where a lack of grammatical skill interferes with student/text communication.

2. The recall procedure does not influence the readers' understanding of the text.

3. The procedure stresses the importance of understanding. Students cannot simply guess

answers; they must attempt to form an understanding of the text.

The studies by Heinz (2004), Chang (2006), Brewer (2000), Riley and Lee (1996), Brantmeier (2006), Packiam Alloway and Gathercole (2005), Boyte (2016) showed that recall protocol is practical for large-scale assessment, and can lead to improved diagnostic and placement testing. Using this system as part of a multiple-measures approach, valid and reliable quantitative score information is readily available and directly linked to a qualitative database ripe for additional examination to advance L2 reading comprehension research and model development. The results showed that resources in long-term memory also play an important diagnostic role in reading and language abilities.

METHOD
Participants
In the current study, 80 Iranian undergraduate students who were studying at the University of Gilan and the Higher Education Institute of Ahrar in Rasht were selected as a sample of the study through convenience random sampling. Since the participants of the study were students of the researcher, they were selected through convenience random sampling. The students were from distinctive majors but all of them had to take a 3-credit general English course. This course was an obligatory one for all undergraduate students. Afterward, the participants were divided into two groups and each group was randomly assigned into an experimental group and a control group. Moreover, the classes were mixed gender including both male and female students, 42 male and 38 female students respectively. It's worth mentioning that the study was piloted with 24 students both male and female learners before conducting the main study.

Instrumentation
Reading for general English (developed by Pourgive, Tajalli, Sadighi, & Yamini, 2015) was used as the course book which incorporates diverse sorts of exercises such as pre-reading questions, relevant contextual clues, word formation, coordinating and matching, equivalent words, antonyms, contextualized lexicon, cloze

passages, cross-word puzzles, mapping exercises, comprehension questions.

Oxford Placement Test (OPT) was administered as both a pre and post-tests to measure and evaluate and assess the student's general English achievement and it was also administered to ensure the homogeneity of the participants before conducting the study. It consisted of six parts including vocabulary, reading comprehension, fill in the blanks, multiple choice, true or false, and grammatical exercises.

Design of the Study

The current study used a pretest-posttest comparison group design as one type of quasi-experimental design. The participants of the present study were 80 Iranian undergraduate students who were divided into two groups. The experimental group which received recall protocol as a modern feedback method and the control group which was taught through conventional feedback techniques. Since the study followed a pretest-posttest strategy, all the students of the experimental and control groups took part in a pretest. Oxford Placement Test (OPT) was utilized as the pre-test. The following step of the study was the treatment stage. The experimental group practiced the recall protocol feedback method. On the contrary, the control group was taught through traditional and conventional feedback techniques such as informal, formal, summative, student peer, student self, and constructive feedback. It's worth noting that a total of 13 sessions were held for each class. The primary session was the "Introduction" session. The second session was the pretest for both groups. Sessions 3-10 were allotted to treatment for the recall protocol class. Participants in the control group were taught through conventional feedback techniques based on a teacher-fronted strategy. Session 13 was the posttest for both groups. It is worth mentioning that the same OPT test was utilized for the pilot study. The result of the pilot study provided reliability for the current research.

Procedure of the Study

Experimental Group. Lyster (2016) recall protocol feedback framework was implemented as a modern feedback method in the experimental group. The teacher welcomes students, gives a brief preview of the course, provides the students with the objectives, and invites the students to ask their questions and suggest their ideas. Then teacher starts with brainstorming the issue and reads the reading of the unit and then explains the reading in his/her own words. At this stage the teacher lets the students review and write whatever he/she taught. Then he/she asks the students to close the book and asks the students to recall whatever they understand from the reading. The students should write what they recall on a piece of paper in their own words. Afterward, the teacher collects the papers and corrects them for the next session. The students should do the tasks at the end of a unit. Whenever they make mistake, the teacher is responsible for providing and explaining the correct form of the mentioned answer to the task. This feedback technique can include the meaning of words and grammatical problems.

Control Group. Three-and-one marking, based on the work of Clarke (2020) on formative feedback was used as a conventional feedback framework in the control group.

Stage 1. At the very beginning of new work/ assignment/inquiry, the teacher explained to the class that the students had to change the way the teacher mark their written work, in order to help them make more progress in the future.

Stage 2. The teacher read all of the student's written work through very carefully before making any annotations. Next, the teacher highlighted 3 places in the writing where the student best met the learning intention/s of the activity. Then she indicated (for example, with a star) one place where an improvement can be made to the original work.

Stage 3. She drew an arrow to a suitable place near the star and wrote a 'close the gap' prompt to support the student in making an improvement to their work. This can be provided in a variety of different forms: A) Reminder prompt - suitable for more competent students. For example, 'Say more about …' or 'Explain why you think this is …'. B) Scaffold prompt - suitable for most students as it provides more structure to improve the work. For example, a question: 'Can you explain why …?' or a directive: 'Describe some of the …' or an unfinished

sentence: 'She shows this by the way she …' C) Example prompt - particularly supportive of less able students. For example, 'Choose one of these statements and/or create your own.' (Teacher provides 2 valid statements for the students to select from or use as a model).

Stage 4: the researcher provided time in class to enable students to read and respond to the 'close the gap' comment. Then, she used this time to follow up on individual needs with 1 or 2 specific students 'face to face'. Afterward, she commented upon their improvement at the first available opportunity.

Data Analysis Method
In the present research, investigation, and examination of the adequacy and effectiveness of recall protocol on Iranian undergraduate college students' general English achievement considering gender was based on separate assessments. The primary assessment was the pilot investigation of the specified tests. The results of the pilot study gave the right premise

for applying the essential consideration. The second section of the study was a quantitative information investigation of the OPT test as a pretest for all the students. The third section of the study comprised of practicing the review convention approach on students of general English courses. Forth part was the posttest. Appropriate factual analyzes and strategies were utilized in arrange to supply an appropriate conclusion

RESULTS
Results of the Reliability Analyses (pilot study)
Before conducting the main study, a pilot study was done with 24 EFL learners and the internal consistency within the items of the tests was estimated by computing Cronbach's alpha coefficient. Straub, Boudreau, and Gefen (2004) state that reliability analysis should be performed for the instruments that are used for a study to validate them and estimate the internal consistency of the measures. The Cronbach's values for the tests are listed in Table 2.

Table 2
Results of the Reliability Analyses for the Pretest and the Posttest

1. Instrument	2. Cronbach's Alpha	3. N of sample
4. Pretest	5. .87	6. 24 EFL learners
7. posttest	8. .89	9. 24 EFL learners

As it was shown in Table 2, the values of the reliability for the pretest and the posttest were $\alpha_{pretest}$ =.87 and $\alpha_{posttest}$=.89, respectively. George and Mallery (2003) suggest the following rule of thumb to interpret the results of the reliability analyses: "$\alpha > .9$ – Excellent, $\alpha > .8$ – Good, $\alpha > .7$ – Acceptable, $\alpha > .6$ – Questionable, $\alpha > .5$ – Poor, and $\alpha < .5$ – Unacceptable" (p. 231).

Therefore, the values of the reliability estimated for the pretest and the posttest were "good" values suggesting that these tests observed the criteria of internal consistency. After conducting the pilot study and establishing the reliability of the tests, they were administered to the main sample, and the data collected were analyzed.

Table 3
Descriptive Statistics for the Pretest Scores

groups	Gender	Mean	Std. Deviation	N
recall protocol (Ex 1)	Female	13.09	2.42	22
	Male	13.44	1.78	18
	Total	13.25	2.14	40
Control (Ctrl)	Female	13.31	1.53	16
	Male	13.54	1.84	24
	Total	13.45	1.70	40

For the general English test that was administered at the beginning of the study, the mean scores for the control group came to ($M_{Ctrl. female}$=13.31), ($M_{Ctrl. male}$ = 13.54). In addition, for the experimental group, the mean values equaled ($M_{Ex.1 female}$=13.09), and ($M_{Ex.1 male}$ = 13.44), respectively. The descriptive table showed that there were simply minor differences among the means of the two groups at the beginning of the study. The male participants in the control group reported the highest mean score. In contrast, female participants in the experimental group had the lowest mean score.

Descriptive statistics for the post-test scores
Descriptive statistics were also computed for the results of the posttest of General English for the two groups. The results of the descriptive statistics for the post-test scores are presented in Table 4.

Table 4
Descriptive statistics for the post-test scores

groups	gender	Mean	Std. Deviation	N
recall protocol	female	14.90	2.59	22
	male	18.33	1.08	18
	Total	16.45	2.66	40
control	female	13.56	2.75	16
	male	13.91	2.94	24
	Total	13.77	2.84	40

For the general English test that was given at the end of the study, the mean scores for the control group came to ($M_{Ctrl. female}$=13.56), ($M_{Ctrl. male}$ = 13.91). In addition, for the experimental group, the mean values amounted to ($M_{Ex.1 female}$=14.90), and ($M_{Ex.1 male}$ 18.33), respectively. The results of the descriptive statistics indicated that there were some differences among the means of the two groups at the end of the study. The male participants in the recall protocol group reported the highest mean score (M= 18.33). In contrast, the female participants in the control group had the lowest mean score (M= 13.56). When it comes to the extent to which the scores were deviated from the mean score, the highest degree of variation of the scores was related to male participants in the control group ($SD_{Ctrl male}$= 2.94). In the posttest scores, female and male participants in the control group received the lowest mean scores. On the other hand, female and male participants in the experimental group who had received recall protocol performed significantly better than the control group in the post-test of general English.

Assessing the Assumption of Normality
The next assumption of the parametric tests is the level of measurement. This assumption was met in this study since the dependent variable (general English achievement) was measured at the interval level and was a continuous scale. Furthermore, the measurements that made up the data were independent of one another and each measurement was not influenced by any other measurement. For examining the assumption of normal distribution, the distribution of the general English test scores was checked through computing Trimmed means and running Skewness analyses. The results of the Skewness analyses and trimmed means are given in Table 5.

Table 5
Test of Normality Assumption

	Groups	Mean	95% Confidence Interval for Mean		5% Trimmed Mean	Skewness	Kurtosis
			Lower Bound	Upper Bound			
pretest	Ex 1	13.25	12.56	13.93	13.33	-.375	-.857
	ctrl	13.45	12.90	13.99	13.47	-.137	-.581
posttest	Ex 1	16.45	16.02	17.22	16.45	-.767	.526
	ctrl	13.77	12.86	14.68	13.72	.233	-.779

The 5% trimmed mean was calculated by excluding the 5% largest and 5% smallest values from the test scores and the arithmetic mean of the remaining 90% of the sample was computed. Then, the first mean values and the new trimmed means were compared to inspect the possible differences between the two means for the two sets of general English tests. The findings showed that the extreme scores did not affect the means and the new trimmed means were between the lower bound and the upper bound of 95% confidence interval for the mean. The results implied that, since the trimmed means and the mean values were nearly the same for the general English test scores, the values were not too different from the remaining distribution. Besides, the ratios of the skewness and kurtosis over their respective standard errors were within the ranges of +/- 2. According to George and Mallery (2010), the values for skewness and kurtosis between -2 and +2 are considered acceptable in order to prove normal two-way ANOVA distribution. Thus, the normality of the distribution was confirmed.

Investigating the First Research Question
After establishing the normality assumption, independent samples t-test was conducted to compare the means between experimental and control groups. The results are given in Table 6 and Table 7.

Table 6
Descriptive Statics for Experimental and Control Groups

	Group	N	Mean	Std. Deviation	Std. Error Mean
Scores	Exp	40	16.4500	2.66939	.42207
	Cont	40	13.6000	2.90711	.45965

According to table 6, the experimental group achieved a significantly higher mean score (M=16.45, SD= 2.66) than the control group (M=13.6, SD= 2.9).

Table 7
Independent Samples T-test for Experimental and Control Groups' Posttests

		Levene's Test for Equality of Variances		t-test for Equality of Means					95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Scores	Equal variances assumed	.086	.769	4.567	78	.000	2.85000	.62404	1.60764	4.09236
	Equal variances not assumed			4.567	77.439	.000	2.85000	.62404	1.60749	4.09251

Considering the Table 7, it is clear that the sig value of Levene's test was larger than the critical value (p= .769> .05), therefore the line for equal variances was considered. With (F= .086, t= 4.567, p= .000) it was confirmed that there was a significant difference between the experimental group and the control group.

Investigating the Second Research Question

To answer the second research question that if there are any statistically significant differences among male and female participants of the study in terms of general English achievement posttest means after being recalled protocol two-way ANOVA was carried out.

Table 8
Descriptive Statics for Experimental and Control Groups across Gender

Group	Gender	Mean	Std. Deviation	N
Exp	Female	14.9091	2.59870	22
	Male	18.3333	1.08465	18
	Total	16.4500	2.66939	40
Cont	Female	13.2500	2.81662	16
	Male	13.8333	3.00241	24
	Total	13.6000	2.90711	40
Total	Female	14.2105	2.78177	38
	Male	15.7619	3.25958	42
	Total	15.0250	3.12189	80

A two-way ANOVA was conducted that examined different feedback methods on learners' general English achievements. According to Table 9, there was a statistically significant interaction between different feedback methods and gender on learners' general English achievements (F $(1, 76)$ = 6.124, p= .016). Moreover, Figure 1 displays the estimated marginal means for the experimental and the control groups on the posttest of the General English test.

Table 9
The Results of Two-Way ANOVA Test

Source	Type III Sum of Squares	Df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	281.798[a]	3	93.933	14.624	.000	.366
Intercept	17736.886	1	17736.886	2.761E3	.000	.973
Group	184.886	1	184.886	28.785	.000	.275
Gender	78.277	1	78.277	12.187	.001	.138
Group * Gender	39.336	1	39.336	6.124	.016	.075
Error	488.152	76	6.423			
Total	18830.000	80				
Corrected Total	769.950	79				

a. R Squared = .366 (Adjusted R Squared = .341)

Estimated Marginal Means of Scores
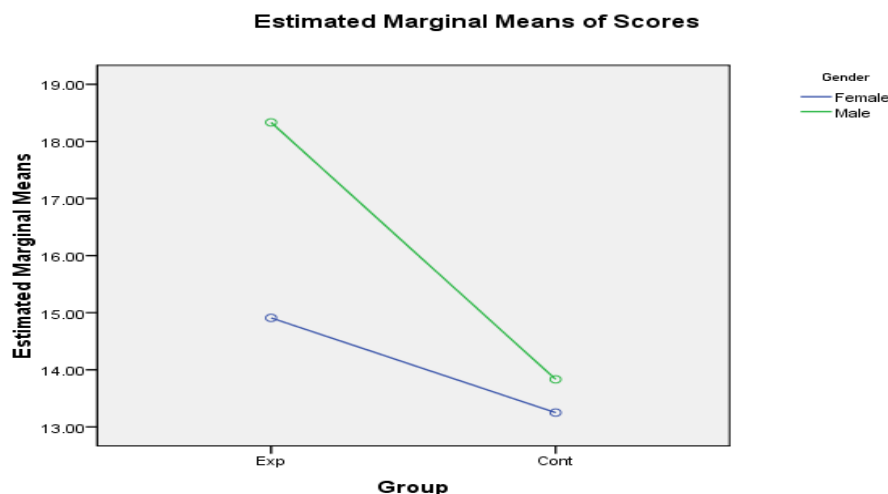
Figure 1
The plot for a dependent variable mean score in a combination of independent variables

Based on Figure 1, there was a significant interaction effect, since the lines for the independent variable were not parallel and thus, can be seen as a set of non-parallel lines. These results confirmed that there was a statistically significant difference between male and female learners in the experimental and control groups.

DISCUSSION

To compare the language learners' performance on the general English test in the recall protocol feedback group and the control group, the results of the two-way ANOVA test were documented that are as given in Table 6. The descriptive Table 4 demonstrated that the recall protocol feedback was more effective in developing the learners' General English than the conventional treatment that was given to the control group. Overall, male participants who were exposed to the recall protocol feedback reported higher mean scores than female participants. The results of two-way ANOVA revealed that the average assessment scores for the posttests were not the same across the recall protocol group and the control group at the end of the study. The interaction effect of the study groups and gender factor were also statistically significant in favor of the male participants. This meant that the treatment resulted in a significant difference between the two groups and the difference across gender was also statistically significant.

The results suggested that there were statistically significant differences between the recall protocol and the control groups at the end of the study in terms of their General English. Moreover, male participants in the experimental group outperformed the female ones Therefore, therefore the null hypothesis was rejected suggesting that implementing recall protocol feedback results in higher effectiveness on male and female participants' general English achievement as compared to the traditional types of feedback that were implemented for the control group.

Overall, for the general English test, the mean difference between the pretest and the posttest that was reported for the participants in the control group was lower than that compared to the mean difference for the participants in the experimental group. This indicated that the control group had the lowest amount of improvement from the pretest to the posttest.

The over-specified conclusions are in line with the result of the study because it found recall protocol as the dependable one which can offer assistance the students' changes in EFL courses and can be measured by means of measurable strategies, it can give the aces and cons of educating and learning prepare, and the execution of the students can be recognized by the instructor agreeing to the reaction to the criticism. This finding is in line with the study by Guinness et al. (2020) and studies by Amalia et al. (2019), Gholizade (2013), Iraji et al.

(2014) and numerous other thinks about which said remedial input as a capable device for tending to errors and cementing desires conjointly appeared contrasts between male and female students considering recall protocol adequacy from distinctive viewpoints.

CONCLUSION

The issue of this study was to explore the viability of recall protocol on making strides in the EFL learners' general English achievement. Agreeing with the findings displayed within the past parts, recall protocol is profitable for learning English as a foreign language. It can be said that recall protocol is the execution which is utilized to make strides and create the achievement of students. Within the preparation of educating and learning English as a foreign language, the educator tries to supply valuable nourishment back to the students in arrange to fortify their desires and to adjust to understudy mistakes amid lessons.

Hence, it could be an effective device utilized to progress the students' execution which can be summarized within the word "learning". Usually, the issue that can be bolstered by agreeing to the comes about and discoveries of the display proposal which appears the convenience of recall protocol and comparative strategies. The study clarified that the group treating with the recall protocol outperformed the control group. It implies that, at least in the present study, the recall protocol makes a difference for the students to form higher achievements compared to a control group. To summarize the discoveries of the study it can be said that the review convention appeared a much better accomplishment. Gender contrasts considering the utilize of recall protocol and general English achievement was the other center of the study. A solid part of gender orientation contrasts in communication has long been substantiated by sociolinguistic inquiries about. It is hence likely that gender orientation moreover plays a part in how a second language is given by instructors and peers and how it is gotten by language learners.

The findings of the study appeared that by and large, males compared to females did superior execution utilizing recall protocol as the

rectification apparatus. Familiarizing the educators with the work of the distinctive sorts of learning and evaluation devices, and the genuine application of these approaches inside the course can offer help second/foreign language learners to move forward in their language capability. Teachers can progress their understanding capacities, through giving related approaches. The eventual outcomes of the research may alter the perspectives of L2 teachers, to have a more significant look at the thought of distinctive procedures they utilize in scrutinizing classes, and it exceptionally well could be beneficial in helping teachers with checking what goes on in their classes.

This sort of investigation can be colossal as a way for assist in creating efficiency, especially on the off chance that the disclosures are to be utilized by L2 teachers. By being careful of the students' slant, the educators can deliver criticism as required by the students. Teachers got to contrast these practices by pivoting methodologies that are obviously capable which relate not only to the students' capability level however in expansion with other learning components just like the students' age, necessities, and suspicions. The discoveries suggest that mindfulness of L2 learners' cognitive styles makes a difference in L2 teaches to utilize more reasonable remedial criticism techniques that facilitate their learners' learning preferences/styles.

References

Amalia, Z. D. H., Fauziati, E., & Marmanto, S. (2019). Male and female students' preferences on the oral corrective feedback in English as a foreign language speaking classroom. *Humaniora, 10* (1), 25-33.

Bartlett, F. C. (1932). *Remembering*. Cambridge: Cambridge MA.

Berkemeyer, V. C. (1989). Recall protocol data: Some classroom implications. *Die Unterrichtspraxis*, *21*(3), 131-137.

Bernhardt, E. B. & James, C. (1987). The Teaching and Testing of Comprehension in Foreign language Learning. University of Wisconsin at Madison. Retrieved form:

https://files.eric.ed.gov/fulltext/ED2854 20.pdf

Bernhardt, E. B. & Deville, C. (1991). Testing in foreign language programs and testing programs in foreign language departments: Reflections and recommendations. In R. V. Teschner (Ed.), Issues in language program direction: *Assessing foreign language proficiency of undergraduates* (pp. 43-59). Boston, MA: Heinle & Heinle Publishers, Inc.

Boyte, K. J. (2016). Foreign Language Teachers' Attitudes Toward Written Recall Protocol as a Practice of Reading Comprehension Assessment. *TESOL International Journal, 11*(2), 108-132.

Brantmeier, C. (2006). The Effects of Language of Assessment and L2 Reading Performance on Advanced Readers' Recall. *The Reading Matrix, 6* (1), 1-17.

Brewer, D. D. (2000). Forgetting in the recall-based elicitation of personal and social networks. *Social Networks*, 22, 29–43.

Brisbois, J. E. (1992). *Do first language writing and second-language reading equal secondlanguage reading comprehension? An assessment dilemma.* Unpublished doctoral dissertation, The Ohio State University, Columbus, Ohio.

Chang, Y. (2006). On the use of the immediate recall task as a measure of second language reading comprehension. *Language Testing,* retrieve from https://www.researchgate.net/publication/249870332_On_the_use_of_the_immediate_recall_task_as_a_measure_of_second_language_reading_comprehension

Clarke, S. (2020). A Little Guide for Teachers: Formative Assessment. SAGE publications Ltd.

Gholizade, R. (2013). The Investigation of Differential Effects of Recast and Metalinguistic Feedback on Accuracy, Fluency, and Complexity of Speaking Performance of Male and Female EFL Learners. *Journal of Novel Applied Sciences. 2*(9), 417-428.

Guinness, K., Detrich, R., Keyworth, R. & States, J. (2020). Overview of Corrective Feedback. Oakland, CA: The Wing Institute. https://www.winginstitute.org/instructional-delivery-feedback

Guo, L. (2013). Effects of Recasts and Metalinguistic Feedback on Developing ESL Learners' Pragmatic Competence. Ph.D. dissertation. University of Kansas.

Heinz, P. J. (2004). Towards enhanced second language reading comprehension assessment: Computerized versus manual scoring of written recall protocols. *Reading in a Foreign Language, 16*(2).

Iraji, S., Zoghi, M., & Nemat-Tabrizi, A. (2014). Corrective Feedback and Learners' Uptake across Gender in an EFL Context. *Procedia - Social and Behavioral Sciences* (98), 680 – 687.

Lyster, R. (2016). The effects of corrective feedback on instructed L2 speech perception. *Studies in Second Language Acquisition, 38*(1), 35–64.

Jafarigohar, M. & Gharbavi, A. (2014). Recast or Prompt: Which One Does the Trick? *Procedia - Social and Behavioral Sciences*, 98, pp.695 – 703.

Packiam Alloway, T. & Gathercole, S. (2005). The role of sentence recall in reading and language skills of children with learning difficulties. *Learning and Individual Differences, 15*(4), 271-282

Riley, G. L. & Lee, J. F. (1996). A comparison of recall and summary protocols as measures of second language reading comprehension. Retrieved from: https://journals.sagepub.com/doi/10.1177/026553229601300203

Schmidt-Rinehart, B. C. (1994). The effects of topic familiarity on second-language listening comprehension. *Modern Language Journal*, 78, 179-189.

Tedick, D. J., & Lyster, R. (2020). Scaffolding language development in immersion and dual languageclassrooms. Routledge.

Ulrich-Verslycken,K., HajHamid, L., & Kaffafi Azar, F. (2021). Oral Corrective Feedback: A Synopsis for Teachers. Retrieved from: https://www.correctivefeedbackforteachers.com/blog/oral- corrective-feedback-a-synopsis-for eachers#: ~: text=Corrective%20feedback%20(CF)

%20is%20any, Oliver%20%26%20Adams%2C%202021).

Zhang, J., Cao, X., & Zheng, N. (2022). How Learners' Corrective Feedback Beliefs Modulate Their Oral Accuracy: A Comparative Study on High- and Low-Accuracy Learners of Chinese as a Second Language. Retrieved from: https://www.frontiersin.org/articles/10.3389/fpsyg.2022.869468/full.

Biodata

Mahnaz Moayeri is currently a PhD candidate in the Department of English Language at Islamic Azad University, Tonekabon Branch in Iran. She has been teaching English courses for 14 years in the university context. She studied BA and MA in the University of Guilan, Iran. She ranked first in both MA and PhD. Her research interests fall within the realms of applied linguistics, computer-assisted language learning, teaching English language methodologies, and task-based language teaching.
Email: *maah.moayeri@gmail.com*

Mohammadreza Khodareza is an assistant professor of TEFL in the Dept. of ELT at Islamic Azad University, Tonekabon Branch in Iran with 26 teaching experience. He has taught courses on English language research, teaching methods, linguistics, and English as a foreign language in ELT. He has done studies on different topics related to his field of study.
Email: *mkhodareza@yahoo.com*

Valeh Valipour is an assistant professor of TEFL in the Department of ELT at Islamic Azad University, Tonekabon Branch in Iran. With 28 years of teaching experience, she has delivered courses on discourse analysis, contrastive analysis, linguistics, and English as a foreign language in ELT. Additionally, she has conducted studies on topics related to applied linguistics and teaching English language methodologies.
Email: *v.valipour123@gmail.com*

Ramin Rahimy is an assistant professor in TEFL. He is the faculty member of the Islamic Azad University, Tonekabon Branch, Iran where he has been teaching English courses for over 20 years. He has ever published papers on different fields of second language acquisition and applied linguistics in the academic journals. He is interested in psycholinguistic aspects of L2 acquisition, translation studies, and sociolinguistics and is focusing on EFL vocabulary learning.
Email: *rahimy49@yahoo.com*