Journal of Modern Research in English Language Studies
Vol. 5, No. 4, 2018,1-26

Draft Specific Scoring and Teacher Corrective Feedback: Hearing Learners' Voice

Masoud Azizi[1*], Majid Nemati[2]

[1*]Assistant Professor, Department of Foreign Languages, Amirkabir University of Technology, Tehran, Iran, *mazizi@aut.ac.ir*
[2]Associate Professor, Department of English Language and Literature, University of Tehran, Tehran, Iran, *nematim@ut.ac.ir*

Abstract

One may not comment on the effectiveness of teacher corrective feedback (CF) before first ensuring learners' attendance. The majority of the studies carried out on teacher CF have mistakenly presupposed learners' attendance to and noticing of teacher feedback without any attempt to check or ensure them. The present study was an attempt to examine the effect of CF on learners' writing ability when it is accompanied by Draft Specific Scoring, a technique designed to maximize learners' motivation to attend to teacher feedback while minimizing the negative effect grading might have on learners' attention (Azizi, 2013; Nemati & Azizi, 2013). In so doing, 57 intermediate students of English Language Literature at University of Tehran, in the form of two groups with one receiving CF and the other one receiving CF plus Draft-Specific Scoring (DSS), were studied. The results of the Split-plot ANOVA between the two groups' pretest and posttest indicated that the treatment group could significantly outperform the control group in overall writing proficiency as well as the four components assessed in IELTS writing task 2. In addition, learners' motivation, attendance, and attitudes were explored into using a questionnaire and a written interview. The participants experiencing DSS reported a high level of motivation and attendance. They also held a very positive attitude toward the technique they had undergone. The results indicate that it is possible to make teacher corrective feedback work if the intervening variables, more particularly motivation, are taken care of.

***Keywords**:* Draft Specific Scoring (DSS), IELTS, Learners' Perception, Motivation, Teacher Corrective Feedback, Writing Ability

Received 23 April 2019	Accepted 08 June 2019
Available online 08 July 2019	DOI: 10.30479/jmrels.2019.10708.1340

1. Introduction

Talking about the effectiveness or ineffectiveness of teacher corrective feedback does not sound reasonable if learners do not attend to teacher feedback. The myriad of conflicting results in favor or against the effectiveness of teacher CF could be attributed to the fact that feedback provision has always been equated with teacher feedback provision plus students' attendance and implementation of that. However, as Ferris (1999), Guenette (2007) and many others have emphasized, there exist a variety of individual differences, one of the most important of which being motivation (e.g., Bruton, 2009, 2010; Ferris, 1999; Guenette, 2007; Lee, 2008), that need to be taken into account if one is to comment on the effectiveness or ineffectiveness of teacher corrective feedback.

The trend in the studies available in the literature suggests that what is often presupposed in most debates on the effectiveness of corrective feedback is that the simple act of feedback provision on the part of the teacher entails learners' attendance and subsequently the application of the feedback to their future writing samples. In other words, the teacher feedback in the literature is equated with 'teacher feedback plus students' attendance to and implementation of it', which cannot be warranted if not checked. The myriad of conflicting results in the literature regarding the effectiveness of corrective feedback (Bitchener & Ferris, 2012; Lee, 2014; Mawlawi-Diab, 2015; Zheng & Yu, 2018) could be regarded as evidence for the fact that learners' attendance has not been ensured or at least checked in the majority of studies on corrective feedback so far. In addition, learners are often reported not to be motivated enough to attend to teacher feedback (Lee, 2008; Truscott, 1996). There are also a number of other factors which can divert learners' attention away from teacher feedback. Therefore, only when learners' attendance to teacher feedback and its implementation in their writing practices are ensured, one can reasonably comment on the effectiveness or ineffectiveness of teacher corrective feedback (Nemati & Azizi, 2013).

2. Literature Review

Truscott's questioning of the value grammar correction in writing classes in 1996 triggered a series of reactions and debates on the validity of his claim. Truscott claimed that grammar correction not only does not aid students but it also hinders their learning process. He believes that learners are not motivated enough to attend to teacher feedback and even when they do, they are not motivated enough to apply it to their writings. He also states that those who do not receive any corrective feedback have a more positive attitude toward writing. To him, it is preferable for teachers not to correct learners' errors because grammar correction has no place in writing instruction and should be abandoned as a result. Based on a meta-analysis he

did on CF in 2007, Truscott concluded that learners receiving teacher corrective feedback are more likely to shorten or simplify their writings in order to avoid situations in which they are more likely to make mistakes and consequently be corrected. To him where learners' scores in overall accuracy improve, it may simply be due to their learning how to avoid structures they are not very sure of.

Ferris (1999), calling Truscott's anti-correction position "premature and overly strong" (p. 2), believes that for having an effective grammar correction and instruction, there are many variables one need to take into account including learners' first language background, their English proficiency level, and experience with grammar instruction. She holds that it is of utmost importance for teachers to raise learners' motivation and help them develop independent self-editing skills. According to Ferris (1999) surveys indicate that learners highly value and demand teacher feedback. To Ferris (2004) and Guenette (2007), the existence of conflicting results about the effectiveness of CF in the literature, which is due to the large variation in the research design and methodology of such studies, has made it very difficult to draw any conclusion about the effectiveness or ineffectiveness of teacher corrective feedback. Ferris (1999) believes that only when we have adequately researched the individual student variables affecting learners' willingness and ability to benefit from error correction and have identified the methods, techniques, and approaches to error correction which can lead to short or long-term student improvement, can we definitely support or refute Truscott's thesis.

According to Bruton (2009), it is both logical and intuitive that more evidence, no matter positive and negative, results in improvement in learners' level of correctness. Bruton (2010) also emphasizes on the link between motivation and effort to improve. He believes that variables such as instruction, tasks, and grades are so important that they can affect learners' success or failure and as a result cannot be overlooked. He asserts that often the participants are not given any purpose or objective for what they are supposed to do, and sometimes no feedback on content or encouragement is given to students in L2 writing research. In other cases, no grades are provided or if grades are given, no reference is made to content, which encourages avoidance. All these can demotivate learners. He believes that students need to have a reason for trying to improve their accuracy level. To Bruton (2010), the climate of the responses as well as grades is so important if CF is supposed to work. However, he fails to present a study in which motivation was present and correction was found helpful (Truscott, 2010).

While Bruton's (2010) belief in the role of grades seems intriguing, the literature indicates that grading learners' writing samples has its own flaws. Grades can divert learners' attention away from teacher feedback.

Students have frequently been observed ignoring teacher feedback when they see a grade on their paper (Lee, 2009). Lee observed that "teachers award scores/grades to student writing although they are almost certain that marks/grades draw student attention away from teacher feedback" (p. 16). Based on the feedback analysis she did, she concludes that all the teachers give their students' writings a score although they do not believe much in their usefulness because they think scores and grades divert learners' attention away from teacher feedback to the extent that some students may even ignore them particularly when they are not required to revise and resubmit their drafts for better grades. One teacher remarked, "the majority of students do not pay attention to the comments." Another teacher even said, "for students, they only look at the scores" (p. 17).

Li and Barnard (2011), interviewing tutors responding to and commenting on students' writings, sought the extent to which their participants attached importance to the awarding of grades when giving feedback. All participants considered awarding a grade as an integral part of the feedback. One interviewee remarked that he gave feedback because it would help students get a better score. Another one said that written feedback can explain how and why a student got a certain grade. Li and Barnard (2011, p. 146) argue that according to their findings, tutors' main reason in providing learners with feedback was "less that of seeking to improve the students' writing skills and more that of justifying – to themselves, to their students, and to their academic superiors – the award of a specific grade for the assignments to hand."

One may wonder why teachers do not stop grading or scoring student writing if they are aware of the harm it does. Lee (2009), quoting the same teachers, argues that grading is necessary for summative purposes, and this summative evaluation is often what most educational institutes require their teachers to provide. One teacher in the follow-up interviews emphasized the importance of grading by saying that he believes that compositions, except identifying students' difficulties in writing, serve another function, i.e., they serve for teachers to hand over score sheets. As such it seems that "the summative function of feedback has made teachers use scores/grades although they are fully aware of the harm that can be done to students" (p. 17).

Moreover, learner engagement with teacher corrective feedback has been found to be dynamic and vary across individuals (Zheng & Yu, 2018), which is affected and mediated by both learner factors and contextual factors simultaneously (Han, 2019). Learners' beliefs can have a tremendous effect on their engagement with teacher CF, for instance (Han, 2017). "Teachers should consider students' beliefs when providing WCF, and foster the development of learner beliefs conducive to deep engagement with WFC"

(Han, 2017, p. 133). Students have been reported demanding their teachers to assess their writing by assigning it a grade (Lee, 2008) mostly because its interpretation is much easier for them in comparison with the sometimes vague or excessive amount of comments written in the paper margins. Lee (2008), studying both high proficient (HP) and low proficient (LP) students of English during an academic year, examined their preference for the type of feedback they received. She observed that 72.2 percent of HP students and 40.9 percent of LP students chose the option 'mark/grade + error feedback + written comments.'

In addition, one should not overlook learners' feelings when being engaged with teacher corrective feedback as they can affect the way they interact with or attend to teacher feedback. The studies so far have reported different emotional reactions to teacher feedback on the part of the learners (Han & Hyland, 2019; Zhang & Hyland, 2018). Some students have been reported to feel proud (Ferris, Liu, Sinha, & Senna, 2013) and self-confident (Storch & Wigglesworth, 2010), others were frustrated (Zheng & Yu, 2018), indifferent, relieved, or even excited (Han & Hyland, 2015). This indicates that whatever method of feedback provision we adopt; we need to be considerate of the feelings it may trigger in our students.

No matter what conclusion research studies come up with, language teachers seem to continue providing their learners with corrective feedback mostly because they think they should. Leki (1990) asserts that although written comments on students' writings are time consuming, teachers still continue to provide them with these comments because they believe that it will help the writers improve. He also believes that teachers do so because their job not only requires them to evaluate students' writings, but it also needs them to justify their evaluation. However, learners are reported not paying attention to teacher feedback, which seems to be due to their lack of motivation to do so. If one wishes to examine the effectiveness or ineffectiveness of teacher corrective feedback, he or she needs to first ensure that learners do attend to the feedback in the process of instruction and data collection. This could be a very challenging task as learners are often not motivated enough. Moreover, due to organizational obligations and also learners' demands, teachers often provide learners with a grade on their writing samples in addition to the feedback they provide.

The literature indicates that these grades further divert learners' attention from teacher feedback. As a result, what is needed is solution which not only motivates students to attend to teacher feedback but also satisfy the need for assessing learners' writing while not adversely affecting learners' attendance. In other words, having been confronted with all such contradictions, we need to find a middle ground compromising all such challenges. We need a way to motivate learners' to attend to teacher feedback

while providing them with grades that can satisfy teachers' sense of obligation in having summative evaluation and learners' sense of need for such an evaluative feedback without jeopardizing learners' attention to teacher feedback. A solution which not only does not divert learners' attention from teacher feedback, but it also gives them, at least for the majority of learners, a reason and the needed motive to attend to that.

The solution we came up with in Azizi (2013) and Nemati and Azizi (2013) was a simple technique we called Draft Specific Scoring, based on which learners are provided with both corrective feedback and a grade representing the teacher's general evaluation of that piece of work. However, the grades learners receive are not fixed or final. They may improve based on the quality of the revisions students make. Learners can improve their grades by applying teacher feedback to their writings. This improvement may also be initiated by the learner herself as a result of her reflection on the way she could improve her writing in terms of both structure and content. Students are given two opportunities to go through this procedure of redrafting and revising and as a result improving their score. Learners' final score in the course would be the mean score of all the grades they have received in the final version or edition of each one of their assignments during the course.

As a result, the present study was an attempt to examine the effect of this technique on learners' overall writing proficiency as assessed using IELTS writing task 2 scoring rubric as well as the four components present in that rubric namely, task response, cohesion and coherence, lexical resources, and grammatical range and accuracy. In addition, it was attempted to explore into learners' attitudes and opinion toward this type of instruction as well as their level of attendance and motivation to teacher feedback in a course taught using Draft Specific Scoring technique. In so doing, the following research questions were stated:

1. Do students receiving CF plus DSS significantly differ from students receiving only CF in their overal scores in the IELTS writing task 2 as well as the four components examined in the related scoring rubric namely, task response, cohesion and coherence, lexical resources, and grammatical range and accuracy?

2. What do students undergoing DSS think and feel about the technique they experienced?

The present study can be regarded as one of few studies in which it was tried to have both motivation and teacher feedback present and then assess the effectiveness of teacher feedback, a study Truscott (2010) accuses Bruton (2010) of not being able to present an example of.

3. Method

The present study was an attempt to check the effect of a newly-developed technique named Draft Specific Scoring (DSS) on the extent to which it could help learners improve their writing ability. In the first phase of the study, this effect was examined in terms of changes in learners' scores in IELTS writing task 2 as well as their scores in the four components examined in the rubric. In addition, in the second phase of the study, participants' opinion about the instruction they went through and the extent to which they attended to teacher feedback was examined.

3.1. Participants

For the purpose of examining the effect of DSS on participants' scores in IELTS writing task, two intact groups were used. The treatment group consisted of 26 participants (10 male and 16 female participants) with the age range of 22 to 25. There were also 31 participants in the control group consisting of 12 male and 19 female students whose age ranged between 21 and 27. They were all undergraduate students of English Language Literature at the University of Tehran taking the 'Advanced Writing' course as part of their curriculum. There were also two raters, both male, with at least 10 years of experience in teaching English writing. They were both university instructors specializing in L2 writing.

3.2. Instrumentation

Learners' writing samples were collected in three different sessions, later considered as the pretest, mid-test, and posttest. In order to rate these samples, IELTS Scoring Rubric for Writing Task 2 was used. For the purpose of the second research question, part of the data collection was done using a researcher-made questionnaire with 25 items in a 5 point Likert scale format ranging from 'completely agree' to 'completely disagree.' Three main subscales were identified after data collection, Motivation, Attendance, and Attitude. In addition, learners' opinion regarding the technique used was probed into using a written interview. Learners in the treatment group were asked to write an essay expressing their opinion and feelings regarding their experience with DSS.

3.3. Data Collection Procedure

The two groups were checked regarding their proficiency level using Oxford Quick Placement Test and only the ones who were classified as intermediate were selected to be included in the study. The rest of the participants, although not excluded from the course, were not included in the data analysis.

During the first three sessions, the preliminaries of writing were taught to both groups, and using model essays, different parts and components of an essay were discussed and instructed. The base of the evaluation carried out by the instructor was TOEFL iBT independent Task in writing which is very similar to IELTS task 2 in writing. As such, learners were informed of the criteria based on which their writing samples were evaluated and scored. In the fourth session, samples of students' writing were collected as the pretest in the class. Participants in both groups were given 80 minutes to plan and write about the given topic. The samples were scored and returned to the participants with teacher comments on them. The scores learners received on their returned drafts were not the ones given by expert raters because the papers were rated for the purpose of the study much later while the pretest samples were commented and returned to students the following session. They received scores given by their instructor based on the quality and the general impression of their writings. The two sets of scores given by expert raters were later contrasted for making sure that the participants in both groups were comparable in their writing proficiency. No significant difference was found between the two groups at the pretest t (55) = .92, p = .36.

To prevent Halo and Hawthorne effects, both groups were kept blind to the fact that they were being studied. During the class time, some of the learners' writing samples were chosen and discussed with the whole class, and their weaknesses and strengths were pointed out. Each session, learners' essays were collected, taken home, scored, and commented on by one of the teacher researchers. At the end of each session, the participants were assigned a new topic to write about for the following session. Their essays had to be at least 150 words long, typed and printed in an A4 paper. Learners' essays were read by the researcher, and for the grammatical mistakes, learners were provided with indirect corrective feedback, i.e., the errors were underlined but not corrected. To keep the conditions, the same for all, no explicit feedback was given in the samples for the problems they had with the style of writing and issues such as topic development, topic relevance, coherence, and cohesion. Instead some of those samples with such problems were identified and discussed with the whole class during the class time. However, for all essays, if necessary, it was commented that they needed to be improved stylistically in terms of topic development, for instance. The participants were required to revise the drafts they had submitted based on the feedback they had received and return them to the teacher the following session. The two groups were told that their final score would be the average score for all the scores they had received for their assignments during the course. Both groups wrote 10 assignments during the course including the pretest, mid-test, and the posttest. Their final exam was regarded as their posttest. Four weeks before their final exam, the mid test was administered. Unlike the

assignments students had to write at home, the samples written at the three tests had to be at least 250 words long.

Up to this point the procedure followed was the same for both the control and treatment groups. However, the two differed in one major aspect. The scores given to the essays written by learners in the control group were fixed, that is, they did not change after the revisions were made by learners, but in the case of the treatment group, learners could improve their scores by the revisions they made. For example, a learner who had received 14 out of 20 for the draft she had submitted could revise her sample based on the feedback she had received and improve her score. She could receive 16, or 18 or any other score based on the quality of her revised sample. She could even receive the same score in case the revisions were not satisfactory or she had attempted to avoid the use of the structure. The revised samples were again commented on and returned. The learners had one more opportunity to revise their returned samples and undergo the same procedure. This is what we called *Draft Specific Scoring* (Azizi, 2013; Nemati & Azizi, 2013).

Both groups received a sample of the score profile in which the instructor would record their scores in order to come up with their final score at the end of the semester. Their final score would be the mean of all the scores they received on their assignments during the semester. For the treatment group, the final score they received on the last revision they submitted was taken into account while for the control group the single score they received for each assignment was used to calculate their final score.

In order to control for the handwriting effect on raters (Klein & Taub, 2005; Russell, 2002), all essays written by both groups in pretest, mid-test and the posttest were typed first. All the mistakes, regardless of their type, were typed exactly as they were written by participants. All typed essays, 171 in total, were coded by numbers so that it was impossible for the raters to identify which essay belonged to which group or which test. All essays were given to two experienced raters to be rated based on IELTS writing scoring rubric for task 2. All essays were shuffled and given to raters for each rating at once so that the time factor could be controlled for.

For the purpose of the second research question, a 25 item researcher-made questionnaire was prepared and distributed among the learners in the treatment group almost at the end of the course checking for their motivation, feelings and attitude toward the course and the grading system, and their attendance to teacher feedback. This questionnaire was not piloted because it was designed to check for learners' motivation and beliefs in a course taught using DSS. Therefore, only those who had taken part in such a course could respond to that, whose number was very low. As such, it was some type of descriptive analysis rather than a well-designed survey used to check for

learners' opinion. However, in order to make the task of presentation easier, after doing a factor analysis, three components were identified namely, Motivation, Attendance, and Attitude. It is worth mentioning that doing a factor analysis with such a low number of participants is not recommended. However, since the purpose of using this questionnaire was only descriptive, it does not seem that doing so may do any harm. It was simply for the ease of presentation. Moreover, learners were also asked to write about what they liked and disliked about the grading system in the Advanced Writing course, the results of which together with the results of the above-mentioned questionnaire were used in order to answer the second research question.

3.4. Data Analysis

For checking the effect of Draft Specific Scoring on learners' overall writing proficiency and the writing components examined in the IELTS scoring rubric for task 2, a number of mixed between-within-subjects ANOVAs, also called Split-plot ANOVA or SPANOVA, were used for data analysis. However, the second question was descriptive in nature. Therefore, the mean score for each item and scale was calculated and finally learners' opinion expressed in their writing sample was examined.

4. Results and Discussion

4.1. Results

The first research question was to examine the effect of DSS in comparison with more traditional methods of CF provision on learners' writing proficiency as assessed through the use of IELTS Writing Task 2 Scoring Rubric. Table 1 summarizes the rater reliability indices for the IELTS four components. For the data analysis, the scores given by the rater with higher intra rater reliability were used. There is no rater reliability index for IELTS holistic score as it is the mean of the four scores awarded for the four components in the scoring rubric. If this mean score ends up in .25 or .75, it is rounded up. If it is smaller than that, it is rounded to the lower half or complete band score.

4.1.1. Learners' IELTS Writing Scores

Regarding the participants' writing scores in IELTS, as evident in Table 2, while both groups started the course almost at the same level, their improvement over time seems to be different. In the mid-test, the treatment group undergoing DSS appears to outperform the control group by one band score, which changed to 1.5 band score at posttest. However, the control group could only show an improvement of less than a half band score from pretest to mid-test. It could finally reach a half band score at posttest.

Table 1

The Rater Reliability Indices for IELTS Components

IELTS Components	Inter-rater Reliability	Intra-rater Reliability (Rater I)	Intra-rater Reliability (Rater II)
1.Task Response	.89	.88	.93
2.Coherence & Cohesion	.91	.89	.91
3.Lexical Resources	.89	.86	.92
4.Grammatical Range & Accuracy	.86	.85	.89

Table 2

Learners' IELTS Writing Scores

	Group	N	Minimum	Maximum	Mean	S.D
Pretest	Treatment	26	4.5	8.0	5.67	1.03
	Control	31	4.0	8.0	5.42	1.04
Mid-test	Treatment	26	5.0	9.0	6.67	1.10
	Control	31	4.5	8.0	5.87	.91
Posttest	Treatment	26	5.0	9.0	7.25	1.27
	Control	31	4.0	8.5	6.02	1.06

A SPANOVA was performed for the two groups across the three time periods to examine the effect of the intervention. There was a significant interaction between Time and Group, Wilks' Lambda = .71, $F (2, 54) = 11.37$, $p < .0005$, partial eta squared = .30. There was also a substantial main effect for Time, Wilks' Lambda = .31, $F (2, 54) = 59.34$, $p < .0005$, partial eta squared = .69. In addition, the main effect for Group was found statistically significant, $F (1, 55) = 8.54$, $p = .00$, partial eta squared = .13, suggesting a benefit for the treatment group receiving DSS over the control group receiving corrective feedback without DSS. The pairwise comparisons across time for each group showed a significant difference in all cases but for the control group from mid-test to posttest.

4.1.2. Task Response Component

Regarding the first writing component in IELTS Writing Task 2 Rubric, i.e., Task Response, a very similar pattern of results to the one observed in the case of learners' mean scores as explained above was obtained. Table 3 summarizes the descriptive statistics for the two groups across time.

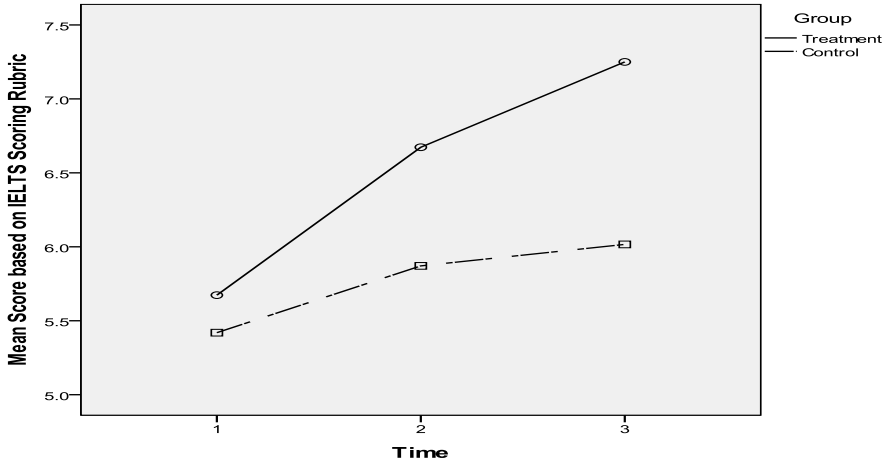

Figure 1. Groups' Mean Scores across Time

Table 3

The Two Groups' Scores in IELTS Task Response Component

	Group	N	Minimum	Maximum	Mean	S.D
Pretest	Treatment	26	4.0	8.0	5.42	1.19
	Control	31	4.0	8.0	5.24	1.07
Mid-test	Treatment	26	5.5	9.0	6.83	1.09
	Control	31	4.0	8.0	5.76	1.01
Posttest	Treatment	26	5.0	9.0	7.19	1.36
	Control	31	4.0	8.0	5.95	1.10

In the case of Task Response component, a significant interaction between Time and Group was observed, Wilks' Lambda = .74, F (2, 54) = 9.27, p < .0005, partial eta squared = .26. There was a substantial main effect for Time, Wilks' Lambda = .36, F (2, 54) = 47.99, p < .0005, partial eta squared = .64, and the main effect for Group, which compares the effect of the intervention on the two groups, was also found statistically significant, F (1, 55) = 9.69, *p* = .00, partial eta squared = .15. As such, it can be concluded that the treatment group had a better performance in comparison with the control group due to undergoing DSS. The pairwise comparison for each group across time showed a significant difference in all cases except for each group's improvement from mid-test to posttest.

4.1.3. Coherence and Cohesion Component

While the treatment group could demonstrate a gain of more than 1.5 band score in this IELTS component, the control group could only improve by about half a band score from pretest to posttest (see Table 4).

Table 4

The Two Groups' Scores in IELTS Coherence & Cohesion Component

	Group	N	Minimum	Maximum	Mean	S.D
Pretest	Treatment	26	4.0	8.0	5.50	1.09
	Control	31	3.5	7.5	5.27	1.02
Mid-test	Treatment	26	5.0	9.0	6.67	1.19
	Control	31	4.0	8.0	5.81	1.05
Posttest	Treatment	26	5.0	9.0	7.32	1.27
	Control	31	4.0	8.5	5.86	1.13

Examining the effect of DSS on learners' scores in the second component of IELTS writing scoring rubric, namely Coherence and Cohesion, a significant interaction between Time and Group was observed, Wilks' Lambda = .67, F (2, 54) = 13.33, $p < .0005$, partial eta squared = .33. There was a substantial main effect for Time, Wilks' Lambda = .35, F (2, 54) = 50.75, $p < .0005$, partial eta squared = .65. In addition, the main effect for Group was found statistically significant, F (1, 55) = 10.13, p = .00, partial eta squared = .16, suggesting a benefit for the treatment group. All the mean differences for each group from one test time to next were statistically significant but for the change from mid-test to posttest in the case of the control group.

4.1.4. Lexical Resources Component

Table 5 presents the descriptive statistics for the two groups' scores in the third component of IELTS writing rubric across time. It seems that the pattern of change in the two groups' scores has been to some extent similar to the previous components although the change has been less in the case of the control group.

A significant interaction between Time and Group was observed, Wilks' Lambda = .56, F (2, 54) = 20.94, $p < .0005$, partial eta squared = .44, with a substantial main effect for Time, Wilks' Lambda = .37, F (2, 54) = 47.07, $p < .0005$, partial eta squared = .64. The main effect for Group, which examines the effect of the intervention on the two groups, was also found statistically significant, F (1, 55) = 6.72, p = .01, partial eta squared = .11, suggesting a benefit for the treatment group again. Like the previous component, the pairwise comparisons across time for each group showed

only one non-significant mean difference which was for the control group from mid-test to posttest.

Table 5

The Two Groups' Scores in IELTS Lexical Resources Component

	Group	N	Minimum	Maximum	Mean	S.D
Pretest	Treatment	26	4.0	9.0	5.75	1.17
	Control	31	4.0	8.5	5.63	1.02
Mid-test	Treatment	26	5.0	8.5	6.46	1.04
	Control	31	4.5	8.0	5.87	.95
Posttest	Treatment	26	5.0	9.0	7.27	1.29
	Control	31	4.0	8.5	5.94	1.01

4.1.5. Grammatical Range and Accuracy Component

Table 6 summarizes the descriptive statistics for the two groups across time for the last writing component examined in IELTS task 2 writing rubric. Although the results seem to be similar to the previous ones, the range of change appears to be more limited.

Table 6

The Two Groups' Scores in IELTS Grammatical Range & Accuracy Component

	Group	N	Minimum	Maximum	Mean	S.D
Pretest	Treatment	26	4.5	8.0	5.71	1.10
	Control	31	4.0	8.0	5.37	1.09
Mid-test	Treatment	26	4.5	9.0	6.56	1.27
	Control	31	4.5	9.0	5.79	1.04
Posttest	Treatment	26	5.0	9.0	7.04	1.39
	Control	31	4.0	9.0	5.94	1.23

The SPANOVA performed on the two groups' scores across the three time periods revealed a significant interaction between Time and Group, Wilks' Lambda = .81, $F (2, 54) = 6.23$, $p = .00$, partial eta squared = .19. There was a substantial main effect for Time, Wilks' Lambda = .41, $F (2, 54) = 39.73$, $p < .0005$, partial eta squared = .60, and the main effect for Group was found statistically significant, $F (1, 55) = 6.42$, $p = .01$, partial eta squared = .11, again suggesting an advantage for the treatment group over the control group. The mean differences across different testing sessions were statistically significant in all cases but for both groups' improvement from mid-test to posttest.

4.1.6. Participants' Viewpoints

The second research question intended to examine what learners in the treatment group thought and felt about the grading system they went through. To answer that question, a 25 item questionnaire was developed by the researchers. This questionnaire could not be piloted as the participants who could respond to it were limited to those who had experienced the instruction using DSS. Identifying components in the questionnaire using a factor analysis was not recommended due to the low number of participants. There were 26 participants in the treatment group, one of whom was absent on the day of data collection. However, a factor analysis was run to help the researcher classify the items into different categories for the ease of presentation. Based on the results of the factor analysis, the reliability analysis, and the wordings of the questions, three components were identified: Motivation, Attendance, and Feelings and Attitudes. Learners could choose from among 'Completely agree' (5 points), 'Agree' (4 points), 'No idea' (3 points), 'Disagree' (2 points), and 'Completely disagree' (1 point). The Cronbach's Alpha reliability index for the whole questionnaire was .93.

The Motivation Component. Table 7 presents the items which fell into this category. The Cronbach's Alpha for this category was observed to be .87.

Table 7

Items Examining Learners' Motivation

Item No.	Item wording	Mean score
1	This system of grading does not motivate me to revise my first drafts.*	4.00
2	Improving my score motivates me to revise my first drafts.	4.40
3	This grading system gave me a reason to attend to teacher feedback.	3.84
4	This course was successful in motivating me to attend to my mistakes in writing.	4.12
5	I hate it when grading is used to make me study.*	3.56
	Mean Score for Motivation Component	3.98

* Notice that this item needs to be reversed. The reported mean scores are based on reversed items.

The Attendance Component. Table 8 presents the items which fell into this category. The Cronbach's Alpha for this category was .81. The high mean obtained shows a high level of attendance to teacher feedback on the part of the learners.

Table 8

Items Examining Learners' Attendance

Item No.	Item wording	Mean score
1	During this course, I tried to learn from my mistakes.	4.40
2	I tried to correct all the errors I committed for my next assignment.	4.20
3	I tried to learn why what I wrote was wrong so that I would not repeat them in my new assignments.	4.16
4	I tried to learn the rules for the correct use of a point in grammar so that I could use it correctly in my future writings.	4.08
5	I read all the comments written by the teacher on my writings.	4.28
6	I paid attention to the comments the teacher provided on my first and mid drafts.	4.12
7	When seeing my score, I was eager to check the teacher comments to see why I lost points.	4.44
	Mean Score for Attendance Component	4.24

In addition to the above items, the extent to which each group attended to teacher feedback can be reasonably inferred from the rate of revisions each group made for each draft of each assignment. The rates at which each group handed in its first draft on each assignment was not that different from each other mostly due to the fact that their final score was the mean of the scores each group received on their assignments. As such, if an individual missed an assignment, his or her final score could heavily suffer. The treatment group handed their first draft of their assignment at the rate of 98.07%. This rate was 93.60% for the control group. However, for the second draft and the third draft, the rates changed dramatically. For the second draft, the revision rate was 73.56% for the treatment group. Although the control group was strongly recommended to revise their first drafts, this rate was found to be only 5.64%. Finally, for the third draft, while the revision rate was 54.96% for the treatment group, it was only 1.83% for the control group, which shows the control group's lack of motivation and attendance to teacher feedback.

However, one point needs to be noted. Although a great amount of difference was observed between the two groups in their revision rates, individual differences cannot and should not be ignored. There were individuals in the treatment group who preferred to write their first draft only and were not interested in revisions at all. There were also learners who revised their drafts for the fourth time but were asked not to due to lack of time on the part of the teacher. Similarly, there was a participant in the control group who was observed to be self-motivated for revising her writing samples. As such, when we examine the collective effect of this approach, we should also be aware of individual differences.

The Feeling and Attitude Component. Table 9 presents the items which fell into this category. The Cronbach's Alpha for this category was .86. The mean score for items in this category indicates learners' very positive attitude toward the type of instruction they received.

Table 9

Items Examining Learners' Feelings and Attitude

Item No.	Item wording	Mean score
1	I hate the way we had to revise our writings.*	4.36
2	When I saw my score, I did not want to go through & review my mistakes.*	4.00
3	I am happy with my score on the first draft & I hate being required to revise them to receive higher scores.*	4.04
4	I did not feel stressed when writing because I knew I had the chance to revise it later.	3.24
5	I think receiving comments on our first and mid drafts is very important.	4.48
6	This system of grading could not help me write better in new assignments.*	3.80
7	I think I could not learn that much in this course.	4.24
8	I felt positive about the way we were required to revise our writings.	4.08
9	I think the comments I received were not helpful for me.*	3.80
10	I found the process of revising enjoyable.	3.68
11	I feel my efforts could produce desirable outcomes.	4.12
12	Although it was difficult, I feel satisfied with the effort I had to make in revising my first drafts.	4.20
13	I think I could not learn that much in this course.*	4.12
	Mean Score for Feelings and Attitude Component	4.01

* Notice that this item needs to be reversed. The reported mean scores are based on reversed items.

Treatment Group's Opinion. In addition to the questionnaire, learners in the treatment group were asked to state their opinion and feeling about the whole course in the form of an essay. They were asked to share their opinion as frankly and honestly as possible about what they liked and disliked about the program and the grading system they underwent. They were ensured that the content of their essays would not affect their scores at all. The results were very intriguing. Learners' comments confirmed their responses in the previous section. They confirmed learners' high motivation to attend to teacher feedback. What follows is the selection of learners' comments on the course. The paragraphs quoted are exactly the same as what learners have written; the errors have not been corrected.

> *The advantage of this method in comparison with other writing classes was that in my last classes whenever I got my paper and saw my mistakes, the only thing was that it made me disappointed. I did not focus on my mistakes carefully, whereas in this class I used to*

look at my mistakes and try to find the correct form in in order to prepare a revision. I knew that I will learn more, also get a better score… it was very motivating. What I liked a lot about this system was that I became very meticulous in my latter writings. … I may perform the same method of teaching if I become a teacher as well. (Code 101)

The grading system acted as a subliminal devise of motivation and a relatively necessary 'candy,' rewarding the good boy for his accomplishment. Even though I study for my own personal gain, the fact that my improvement was evinced in the form of gradually increasing grades was heart-warming and essentially subtly encouraging. (Code 103)

The grading system incorporated in this course compelled me to improve my awful grammar skills, and as a result, I grew fond of writing. This new grading technique was the most important reason for my success … Sometimes when I did not receive a good score, and when I also did not have enough time to revise my writing, I knew that I had a second chance to hand in my assignment; therefore, I could feel much more relaxed. (Code 104)

This procedure of revising and receiving comments made students more willing toward writing and lessened their apprehension about their grades since they knew they could revise it afterwards. (Code 106)

The first thing that I liked about this grading system was that I was motivated to revise my writings and try to improve them in order to improve my grade. I was sure that since I can revise my writings and get better grades, I will finally have a high score in this course. So I never really felt any stress about my final score, and I could only focus on improving my writing style. Another thing that I liked about this system was that I could learn from mistakes (Code 109)

The grading system for this course was new and interesting; the thought that we had two chances to revise each of our writings was a very reassuring one. No other teacher had provided us with this opportunity, and this made us less anxious about the grade itself so we could solely focus on improving our writing skills, and trying to correct our mistakes. Every time our papers were given back to us, I would first read the comments at the end of my writing, then go through the whole text and see what my mistakes had been, and I would immediately start revising it in my head. (Code 110)

I appreciate your grading system, and I hope that our education system moves toward seeking an effective method for teaching. (Code 111)

I noticed that in this class, most people didn't get good marks in the beginning, but later if they try to learn from mistakes, their marks will be raised. This is far better than giving the same students always the same marks. (Code 113)

All in all, I find your grading system very encouraging and useful, and writing, which has always been an ordeal for me, has become much easier and much more interesting. (Code 114)

Another great feature of this method was that our final score relied solely on ourselves. As opposed to all other courses, where the professor's personal evaluation determines a student's fate, here our score relied almost entirely on ourselves. Because we had the opportunity to revise our writings, we were essentially the main decisive agent in this equation. Thanks to this new approach, I finished this course on a high note. I learned a lot about grammar, and I also learned to write systematically, and academically, two important traits that will guide me in my academic life. (Code 118)

In the beginning, this way of revising paragraphs bored me. I thought my writings were the best I could produce, and that it was useless even to make the effort to change them. However, by starting to get better and better in every revised draft, I figured out I can actually produce better works if I worked harder. This thought increased my self-confidence as well as my perseverance. (Code 119)

As a matter of fact, I am really a fan of this kind of grading system that we experienced during this course. The possibility of being able to correct and edit each writing for two times was really useful. I myself learnt a lot of my mistakes. (Code 122)

I like the grading system our teacher uses, because his evaluation is not just based on our final exam… this was one of the most important motivations that encouraged us to revise them to have a better grade and correct our mistakes in our previous compositions. (Code 124)

4.2. Discussion

Regarding the effect of DSS on learners' writing ability as assessed by the means of IELTS writing scoring rubric, it was observed that both groups could significantly improve over time with the DSS group outperforming the control group. Both groups started the program at a similar level, and they could both improve by the end of program. However, the

extent to which each group improved was significantly different from each other. While the improvement for the control group was very steady and slow over time, the treatment group's improvement was more eye-catching. Both groups started the program at a score around 5.5. By the end of the program, while the control group could improve only by .6 band score, the treatment group had an improvement of about 1.5 band score, which was about one band score more than the control group. The effect size for this difference was very large (.13). The same patterns of results were observed in the case of all four IELTS components namely, Task Response, Coherence and Cohesion, Lexical Resources, and Grammatical Range and Accuracy all with large effect sizes.

The most noticeable component in this regard is the last one, i.e. Grammatical Range and Accuracy, as it is related to the most controversial part of teacher corrective feedback debate. Truscott (1999) emphasized that he argued against grammar correction rather than teacher feedback. He does not reject the possible positive effect of teacher feedback. However, he does reject the effectiveness of grammar correction. The fourth IELTS component shows evidence not in favor of Truscott's thesis. While both groups improved in accuracy, the treatment group outperformed the control group by a large difference. One might argue that, as Truscott does in many cases, since there was no group receiving no corrective feedback, this study is simply another study comparing different types of feedback rather than correction vs. no correction. He has a point there. However, if we assume that corrective feedback is ineffective as Truscott claims, then there must be no difference between a control group only receiving no corrective feedback and a control group receiving corrective feedback when it comes to evaluating the effect of DSS. In any case the superiority of this type of instruction is evident.

The fact that even the control group could improve over time contradicts Truscott's thesis regarding the ineffectiveness of CF and is consistent with a number of other studies in the literature (e.g. Ashwell, 200; Chandler, 2003, 2004). However, as mentioned above, since there was no group receiving no feedback at all, one cannot safely comment on the implications of the observed results regarding the effectiveness or ineffectiveness of teacher corrective feedback. On the other hand, one should not ignore the fact that in the present study, there could have been a high level of motivation on the part of the control group as well. Learners in the control group too had to work hard if they wished to improve their final score since even for this group, learners' final score in the writing course they were going through as part of their university curriculum was the mean of all the scores they had received on the assignments. The only difference was that they did not have the opportunity to improve each single score as a result of the revisions they could make. In other words, it could be assumed that even

the learners in the control group were more motivated than participants in most other studies on CF in the literature. It seems that it can reasonably be concluded that if motivation is ensured on the part of the learners, teacher corrective feedback works. This confirms Ferris (1999), Guenette (2007), Chandler (2004), and Bruton's (2010) belief in the important role of the intervening individual variables in the debate on the effectiveness of teacher CF, more particularly the role of motivation to attend to teacher feedback.

This does not refute Truscott's thesis all together as he claims and admits at the same time that teacher corrective feedback does not work because learners are not motivated enough to attend to teacher feedback or apply it to their writings. He could be right. In fact, the plethora of conflicting results regarding the effectiveness of CF could be due to the fact that in almost all those studies, teacher CF has been equated with provision of feedback by teachers and learners' attendance to and implementation of that. As a result, where the latter elements were present for any reason, CF has been found to be effective, and where they were absent, CF was reported to be ineffective.

The results indicate that draft specific scoring can be a fruitful technique for ensuring learners' attendance to teacher feedback by giving them a good reason or motive to do so. It not only helps learners improve their writing proficiency over time, but it also solves a second problem in the field. It can neutralize the negative effect scoring learners' writing samples may have on their attention and attendance to teacher feedback. The literature indicates that grades divert learners' attention away from teacher feedback, but still teachers continue to accompany their comments with a grad mostly because they feel they need to due to organizational obligations. Learners also demand receiving such a grade, which makes the situation more complex. Therefore, not assessing learners' writing does not seem to be an option at least in most contexts. DSS, responding to all such concerns, can be the solution. In DSS, grades are used as motivators rather than distractors. Students, having a good reason to attend to teacher feedback, see grades both as rewards and a yardstick to help them recognize the extent to which they need to work on their writing samples in order to reach their objectives.

Examining learners' opinion about this system of grading also suggests a positive attitude towards teacher feedback by language learners. The part examining learners' motivation to revise their writing samples had a mean of almost 4, which means that learners agreed that this system of grading managed to motivate them to attend to teacher feedback and revise their writings accordingly. Learners' mean score for the items checking their attendance to their mistakes was much higher (4.24 out of 5). In addition, learners' feelings and attitudes toward DSS were also quite positive with a mean score of 4.01. Achieving this level of attendance with a positive attitude

and feeling is what is lacking in the literature on corrective feedback and is exactly what we should be looking for, and it is the superiority of DSS over other similar feedback provision methods. This indicates that DSS has been successful in pursuing the predefined objectives regarding motivating learners to attend to teacher feedback.

Learners' opinions regarding DSS not only confirms the results of the questionnaire but also show a clear pattern common in most of the writings. Most learners believed that this grading system could motivate them to attend to teacher feedback. This attendance was not simply for revision purposes, but it was for learning from mistakes in order to write better in future writings. Learners reported not throwing their papers in the waste basket as soon as they saw their grades. Instead they reported going through their mistakes to see how they could improve their scores.

Their account of the story also indicates that they had a quite positive attitude toward the new grading system. They said they were eager to make revisions and felt positive about the process of revision based on teacher feedback. They did not feel stressed when writing their assignments because they knew they would have other chances to compensate for their flaws in writing. In addition, although grading is the integral part of this system, decreasing the weighting of the final exam and distributing its share among the assignments during the semester could help learners feel more responsible for the final score they received and feel less stressed.

Another drawback about CF which is heavily invested in by Truscott (2007) is the effect corrective feedback may have on the fluency, grammatical complexity, and accuracy of texts learners write. Truscott believes that corrected students write shorter and simpler texts. He claims that even when accuracy scores are found to have improved, it could be due to the fact that learners have learned how to avoid situations in which they were more likely to make mistakes and be corrected as a result. However, as Azizi (2013) and Nemati and Azizi (2013) demonstrated, learners receiving CF using DSS improved in fluency. Regarding the grammatical complexity of texts which they wrote, learners both in the control and DSS groups were observed to have significantly improve over time measured via the number of dependent clauses they used. Checking grammatical complexity by the means of another measure, i.e. the ratio of clauses to T-units, the pattern of results was different. While the control group had significantly written less grammatically complex texts, the treatment group receiving CF with DSS had demonstrated no significant change, suggesting that even if DSS does not help increase the grammatical complexity of learners' texts, it will not adversely affect it. Finally, regarding the accuracy of texts written by learners, Nemati and Azizi (2013) observed that while the DSS group could significantly improve over time, the control group failed to do so.

All these considered, it seems plausible to conclude that teacher corrective feedback works if the intervening variables are taken care of. DSS as a technique to ensure learners' attendance to teacher feedback can be helpful in making CF work.

5. Conclusion and Implications

It appears that what is important is not whether we provide our students with corrective feedback or not, but it is whether we ensure their attendance to it. Learners need to notice the input we provide them with no matter if it is in the form of positive or negative evidence. We need to provide learners with a reason and motive to pay attention to the feedback they receive on their writing samples. Simultaneously we need to avoid practices which can jeopardize learners' attention and attendance to teacher feedback. One such a practice is grading learners' writing samples. However, for a number of reasons it seems that not assessing student writing is not an option for most teachers. As a result, what is needed is a middle ground comprising among all these challenges. The solution we came up with was a technique we called Draft Specific Scoring (Azizi, 2013; Nematic & Azizi, 2013). DSS allows teachers to continue their preferred practices while minimizing the negative effect of grading and changing its weak point to strength. It uses grading as a motivating factor which not only does not divert learners' attention from teacher feedback but also ensures their attendance to it.

DSS showed its effectiveness in helping learners improve in writing proficiency while maximizing their motivation to attend to teacher feedback and learn from their mistakes. It helps learners keep a positive attitude toward what they are doing. Learners feel relaxed and more confident when DSS is used as part of their instruction. They have a positive attitude toward it and enjoy writing when assessment is accompanied with DSS. In addition, DSS not only does not negatively affect features such as the fluency, grammatical complexity, accuracy of the texts written by learners, but it may also help improve them.

DSS also addresses Hamp-Lyons' (2007) concern. She believes that in most contexts, writing assessment is taking over writing instruction. As a result, grading and scoring student writing is receiving increasingly more attention. DSS combines assessment with instruction without omitting any of them. It keeps both assessment and instruction in one go.

Experiencing DSS in writing instruction programs, learners will not throw their writing samples in the waste basket as soon as they see the grade on them. Instead they will go through their mistakes to find out the reason why they had made such mistakes and how they can correct them to improve

the scores they had received. In addition, teachers will have a profile of learners' scores to easily come up with a final score and satisfy learners' demand for grading as well as the institutional demands for such an evaluation. Keeping such a profile for each student can also help teachers keep track of their learners' improvement over time.

In case teachers intend to achieve their objectives in a writing program, they need to be aware of the very important role motivation plays in learners' attendance to the feedback they provide. If not motivated, learners will not pay attention to the comments they are provided with, and will repeat the same mistakes in the following assignments. Therefore, before adopting any method of feedback provision or any type of feedback, instructors need to think of a way to motivate them to attend to and apply the teacher feedback.

The present study indicates that there could be more intervening variables between teacher feedback and their effect on learners' new pieces of writing. Motivation was one of them. There could be much more which need to be looked for. As Bruton (2009) states, teacher feedback must work and when it does not, one should look for what it is that hinders it.

References

Ashwell, T. (2000). Patterns of teacher response to student writing in a multi-draft composition classroom: Is content feedback followed by form feedback the best method? *Journal of Second Language Writing, 9(3)*, 227–257.

Azizi, M. (2013). *Draft-Specific-Scoring: A technique to ensure learners' attendance to teacher feedback in L2 writing* (Unpublished doctoral dissertation). University of Tehran, Iran.

Bitchener, J., & Ferris, D. R. (2012). *Written corrective feedback in second language acquisition and writing*. New York, NY: Routledge.

Bruton, A. (2009). Improving accuracy is not the only reason for writing, and even if it were. *System, 37*, 600-613.

Bruton, A. (2010). Another reply to Truscott on error correction: Improved situated designs over statistics. *System, 38*, 491-498.

Chandler, J. (2003). The efficacy of various kinds of error feedback for improvement in the accuracy and fluency of L2 student writing. *Journal of Second Language Writing, 12*, 267–296.

Chandler, J. (2004). Dialogue: A response to Truscott. *Journal of Second Language Writing, 13*, 345-348.

Ferris, D.R. (1999). The case for grammar correction in L2 writing classes: A response to Truscott (1996). *Journal of Second Language Writing, 8*, 1–10.

Ferris, D.R. (2004). The "grammar correction" debate in L2 writing: Where are we, and where do we go from here? (and what do we do in the meantime?). *Journal of Second Language Writing, 13*, 49-62.

Ferris, D. R., Liu, H., Sinha, A., & Senna, M. (2013). Written corrective feedback for individual L2 writers. *Journal of Second Language Writing*, 22, 307-329.

Guenette, D. (2007). Is feedback pedagogically correct? Research design issues in studies of feedback on writing. *Journal of Second Language Writing, 16*, 40–53.

Hamp-Lyons, L. (2007). Editorial. *Assessing Writing, 12* (1), 1–9.

Han, Y. (2017). Mediating and being mediated: Learner beliefs and learner engagement with written corrective feedback. *System,* 69, 133-142.

Han, Y. (2019). Written corrective feedback from an ecological perspective: The interaction between the context and individual learners. *System*, 80, 288-303.

Han, Y., & Hyland, F. (2015). Exploring learner engagement with written corrective feedback in a Chinese tertiary EFL classroom. *Journal of Second Language Writing, 30*, 31-44.

Han, Y., & Hyland, F. (2019). Academic emotions in written corrective feedback situations. *Journal of English for Academic Purposes, 38*, 1-13.

Klein, J. & Taub, D. (2005). The effect of variation in handwriting and print on evaluation of student essays. *Writing Assessment, 10*, 134-148.

Lee, I. (2008). Student reactions to teacher feedback in two Hong Kong secondary classrooms. *Journal of Second Language Writing, 17*, 144-146.

Lee, I. (2009). Ten mismatches between teachers' beliefs and written feedback practice. *ELT Journal, 63*, 13–22.

Lee, I. (2014). Feedback in writing: Issues and challenges. *Assessing Writing, 19*, 1–5.

Leki, I. (1990). Coaching from the margins: issues in written response. In B. Kroll (Ed.), *Second language writing: Research insights for the classroom* (pp. 57– 68). Cambridge: Cambridge University Press.

Li, J., & Barnard, R. (2011). Academic tutors' beliefs about and practices of giving feedback on students' written assignments: A New Zealand case study. *Assessing Writing, 16*, 137-148.

Mawlawi-Diab, N. (2015). Effectiveness of written corrective feedback: Does type of error and type of correction matter? *Assessing Writing, 24*, 16–34.

Nemati, M. & Azizi, M. (2013). Grading, no longer an obstacle to learners' attendance to teacher feedback. *Applied Research on English Language, 2*(2), 129-143.

Russell, M. (2002). *The influence of computer print on rater scores*. Technology and Assessment Study Collaborative. CSTEEP, Boston College.

Storch, N., & Wigglesworth, G. (2010). Learners' processing, uptake, and retention of corrective feedback on writing. *Studies in Second Language Acquisition, 32*, 303-334.

Truscott, J. (1996). The case against grammar correction in L2 writing classes. *Language Learning, 46,* 327–369.

Truscott, J. (2007). The effect of error correction on learners' ability to write accurately. *Journal of Second Language Writing, 16*, 255–272.

Truscott, J. (2010). Further thoughts on Anthony Bruton's critique of the correction debate. *System, 38*, 626-633.

Zhang, Z., & Hyland, K. (2018). Student engagement with teacher and automated feedback on L2 writing. *Assessing Writing, 36*, 90-102.

Zheng, Y., & Yu, S. (2018). Student engagement with teacher written corrective feedback in EFL writing: A case study of Chinese lower-proficiency students. *Assessing Writing, 37*, 13-24.

Bibliographic information of this paper for citing:

Azizi, M., & Nemati, M. (2018). Draft specific scoring and teacher corrective feedback: hearing learners' voice. *Journal of Modern Research in English Language Studies*, *5*(4), 1-26.