Diagnostic Assessment of Writing through Dynamic Self-Assessment

Siamak Mazloomi[1] & Mona Khabiri[1]

[1] Department of Foreign Languages and Literature, Islamic Azad University, Science and Research Branch, Tehran, Iran

Correspondence: Siamak Mazloomi, Department of Foreign Languages and Literature, Islamic Azad University, Science and Research Branch, Tehran, Iran. Tel: 98-912-185-9527. E-mail: siamakmaz@gmail.com or monakhabiri@yahoo.com

Abstract

Deeply rooted in the sociocultural theory of mind by Vygotsky, Dynamic assessment (DA) asserts that mediation is essential for online diagnosis in the classroom. One of the major challenges facing language teachers is the assessment of the learners' Zone of Proximal Development (ZPD) level or diagnosing the amount of mediation or scaffolding they require to achieve their potential level. Ongoing assessment of the learner's ZPD and the tailoring of mediation to fit the learning environment seems to be a vital stage. Dynamic self-assessment (DSA) can be applied for diagnostic purposes in writing classes. In this research, it is assumed that the analysis and comparison of teacher's assessment and DSA will not only indicate their ZPD level or the amount of mediation the learners require but also diagnose their weaknesses and strengths in writing. A quasi-experimental research on 60 sophomore English Translation students in essay writing classes in Islamshahr Azad University revealed that DSA not only significantly affects the EFL learners' writing ability, but also it is incrementally correlated with teacher's assessment through 8 weeks of treatment, and the analysis of DSAs reveals the leaner's' weaknesses and the areas which should be emphasized.

Keywords: diagnostic language testing, dynamic assessment, self-assessment, zone of proximal development

1. Introduction

Despite the fact that the main goal of Diagnostic language testing is "to help test-takers better understand their own linguistic strengths and weaknesses so as to aid language learning" (Harding, Alderson, & Brufaut, 2015, p. 324), few truly diagnostic language tests have been developed so far, perhaps the best example being the DIALANG Project (Alderson, 2005). Moreover, most of these Diagnostic tests deal with reading, writing, and listening, especially through the use of computers and internet. Such summative approach these diagnostic language tests have applied does not provide teachers and learners with adequate constant assessment of the learners' weaknesses and strengths and may not be of much help for pedagogical purposes. Rather, a dynamic means of assessment seems more advantageous in language learning and teaching environments for diagnostic purposes.

To develop a tentative framework for a theory of diagnosis in second or foreign language assessment, Alderson, Brunfaut, & Harding (2014) recently conducted a research on how diagnosis is accomplished across a range of professions and proposed a set of five broad principles, encompassing the whole enterprise of diagnostic assessment (pp. 21-22):

1)   It is the user of the test who diagnoses, not the test. This principle emphasizes the fact that diagnosis should be done either by the teacher or the learners in the classroom, not the test. DSA has got the potential to reveal such diagnosis to both the teacher and the learners, raising their awareness of the amount and kind of mediation, required to compensate for the learning weaknesses.

2)   "User-friendly, targeted, discrete and efficient" instruments should be designed in order to help the teacher make a diagnosis. DSAs, in form of checklists, rubrics, scales, can-do lists, etc., should meticulously be developed according to the classroom syllabi to check the intended objectives achieved by the learners.

3)   Various stakeholder views should be accounted for in the diagnostic assessment process. Alderson et al. (2014) specifically name learners' SAs as one of the main viewpoints towards the diagnosis.

4) Diagnostic assessment should be incorporated within a system that gives way to all four diagnostic stages: "(1) listening/observation, (2) initial assessment, (3) use of tools, tests, expert help, and (4) decision-making." DSA will allow for all four stages: through the teacher-student negotiations and interactions, the first two stages are accomplished; different forms of DSA are available for the third stage; and decision-making is done by teachers and students in a learner-centered approach.

5) Future treatment and goals are sought as a result of the diagnosis through DSA.

Therefore, stake-holder involvement (including learners themselves) (principle 3), targeted, purpose-built diagnostic tools, selected from a bank according to purpose (principle 2), rich and detailed feedback (principle 2), and treatment or intervention to address specific problems which have been identified (principle 5) construct an ideal diagnostic assessment (Harding et al., 2015), which can be achieved through DSA. A skilled "diagnostician" (a well-trained, skilled teacher) should perform all of these requirements (principle 1).

Having been abstracted from descriptions of diagnostic practices across a range of professions (including such fields as medicine, education, information technology and mechanics), such a process has a firm theoretical basis; however, its application to the field of language assessment remains untested (Alderson et al., 2014). A dynamic assessment can provide both teachers and learners with the rich feedback they require in order to set future goals and overcome the learners' weaknesses at different stages of learning through scaffolding and mediating. Although a summative/performance-based diagnostic assessment (such as DIALANG) might be helpful for administrative purposes and placement functions, a dynamic/development-based diagnostic assessment seems more practical and realistic in the online classroom environment giving the stakeholders (teachers and learners) the kind of awareness they need to set their mutual and individual goals for future improvement from one ZPD level to another through adequate scaffolding and mediating (Lantolf, 2000).

Much classroom writing assessment is restricted to the teachers' feedback given to the topical essays the learners are assigned to write by themselves. Little guidance and goal-orienting mediation (based on diagnosis) have been provided to the learners who are in dire need of writing skills especially in an academic setting. On the other hand, the modern humanistic and collaborative approach to education in general is concerned with the question of how students can develop a more active and responsible role in their own learning and assessment (Dewey, 2012; Dyke, 2006). In spite of the trend to find alternative forms of assessment to increase the validity and reliability of assessments, as well as to increase formative and dynamic aspects of learning, self- and peer-assessment have not been much practiced by students and teachers at any level (Taras, 2002). In Iran, for instance, both students and teachers seem to have little previous experience of such forms of alternatives in assessment in the language classrooms, as assessment has traditionally been the teachers' sole prerogative and obligation.

While self-assessment (SA) and other alternative forms of assessment are not much practiced in real academic and learning environments, they have been widely researched and consensually emphasized by many language teachers and scholars (Ross, 1998). The huge bulk of research in the field has indicated that SA is crucially effective and significant in developing the various language skills and learning strategies and raising the awareness and motivation necessary for language learning (Birjandi & Hadidi, 2010). Therefore, SA specifically seems apt to be included in any language learning syllabus as a part of formative/dynamic assessment.

In this research, it is believed that SA as a sort of dynamic/formative assessment can be introduced as a task for scaffolding and for diagnostic purposes since it can fill up the gap between the learners' knowledge and that of their teachers' criteria for their writing assessment. It can also help them develop their own criteria for success in their writing classes. Although DSA itself demands scaffolding (training), the process can be seen as mediating between what the learners already know or can monitor on their own and what they are expected to do. As a result, investigating students' assessment of their own EFL learning skills is important for a deeper understanding of the students' role in assessment, as well as a clearer elaboration of assessment procedures.

The aim of this paper, then, is to explore the implications of the diagnostic process proposed by Alderson et al. (2014) for the assessment of writing through the application of DSA. Accordingly, based on his/her experience and observation, the teacher can hypothetically assess the learners' current ZPD, their probable weaknesses and their future demands, consulting students about their perceptions of the course requirements and goals as well as their weaknesses. To check his/her hypothesis, the teacher should construct self-assessment checklists and rubrics based on the goals of the intended course and the criteria he/she expects the learners to observe in their writing. These self-assessment checklists had better be approved by other teachers in the field as well. Then, the teacher can administer these self-assessment checklists and rubrics in the classroom after each assigned writing task so that the learners can assess their own writing performances, if trained and consulted by the teacher. The

teacher then can formulate his/her diagnostic decisions through these self-assessment tools as well as his own holistic diagnostic assessment after several sessions of instruction and assessment.

The feedback which is provided by the teacher through scaffolding and DSA training would help learners understand their own learning and writing strategies (Harding et al., 2015). The feedback should encourage learners to monitor their own learning and development (metacognitive strategies) and to set their own goals after they are made aware of the criteria for their writing assessment. The feedback provided by computer-based diagnostic systems, such as DELTA and DIALANG are rather general and pre-programmed (Harding et al., 2015), while in a classroom setting, more detailed and individualized feedback is required so that the learners are aware of their own specific weaknesses and strengths. "The more specific the diagnosis can be, the more likely it is that useful teaching and learning materials can be devised." (Harding et al., 2015, p. 326) The immediate feedback these computer-based diagnostic systems can provide the learners with can also be bestowed through self-assessment checklists right after they have performed specific tasks in the classroom while teacher-learner interaction is constantly held through mediation and scaffolding in such a dynamic assessment (Poehner & Lantolf, 2013).

On the other hand, Vygotsky described the Zone of Proximal Development (ZPD) as the difference between the actual development level as determined by individual problem solving and the level of potential development as determined through problem solving under adult guidance or collaboration with more knowledgeable peers. Ellis & Barkhuizen (2005) point to three effective mechanisms in ZPD. The first one is that the intervention should be graduated starting with implicit help and becomes more specific until the appropriate level is reached. Second, the help should be contingent; i.e., the help should be offered only when it is needed and withdrawn as soon as the novice shows signs of self-control and ability to function independently. The third is collaborative interaction, ongoing assessment of the learner's needs and abilities and the tailoring of help to fit these. It requires the learner's ZPD measurement in a dialogic manner or so-called scaffolding, promoting self-regulation, which can be done dynamically through self-assessment.

There are many researches concerning the use of self-assessment (SA) to improve and measure learners' writing ability (Birjandi & Siyyari, 2011; Black, 1998; Dragemark, 2009; Janssen-van, 1992; Naeini, 2011; Rea-Dickins, 2006; Sadler, 1989; Sullivan & Lindgren, 2002; Taras, 2001, 2002, 2003; Wang Yucui, 2007). They have all emphasized that incorporation of a writing test, especially SA, into the present testing system will improve the learners' writing ability as well as their self-perception and self-regulation, raising their awareness of their weaknesses and strengths. Dragemark (2009) also indicated that SA can affect learners' language proficiency and their lifelong learning strategies in a writing class. Yet none of them have applied SA in a dynamic approach for diagnostic purposes.

2. Research Questions

The following research questions have been posed to investigate whether DSA can be used to measure EFL learners' ZPD and the diagnosis they require in English writing classrooms:

1)   Does the application of DSA significantly affect the EFL learners' writing abilities?

2)   Is there any significant relationship between the teachers' assessment and the EFL learners' DSA in writing classes?

3)   Does the analysis of the teachers' assessment and the learners' DSA the amount of feedback the learners require and their weaknesses?

3. Method

3.1 Participants

Two groups of English Translation sophomores, a total of 60, participated in the present study who were selected out of 76 participants based on their scores on a PET. They had all passed the first year of English Translation successfully and were at lower intermediate level. The two groups were randomly assigned into experimental and control groups, consisting of 30 students in each group.

The experimental group attended eight sessions of instruction, through which they were administered the SA checklists and rubrics to evaluate their performances in expository writing for each section of instruction—based on the units of their textbook. The participants' performances were rated by two EFL teachers, holding MA in TEFL, and each having at least ten years of experience in language teaching.

The control group also attended similar classes with the same teacher, materials, and instructions, yet they were not asked to self-assess their writing skill; instead, they did the regular exercises and tasks assigned in their

textbooks and were assessed by their teacher.

3.2 Instrumentation

All the participants attended a sample Preliminary Language Test (PET) to measure their proficiency level and homogenize them. Both the experimental and control groups also took the following pre-tests and post-tests and were evaluated based on the scoring scales which are described hereunder.

To assess the participants' writing skill, they were required to write a 250-word composition on a PET/IELTS writing task in expository genre both at the beginning and the end of the term; then, the ESL composition profile by Jacobs et al. (1981, as cited in Hughes, 2003) was employed by two raters to score the participants' compositions, and the participants of the experimental group used the scale for the purpose of SA. The five criteria to evaluate a composition in Jacobs' scale are: 1. Content, 2. Organization, 3. Vocabulary, 4. Language Use, and 5. Mechanics. A checklist (Figure 1 illustrates the kind of questions asked on each criterion) was also developed based on these criteria suited for the level of students and based on the teacher's expectations. Several questions on each criterion—a total of 40—were posed to the learners in the checklist in a Likert scale (Yes, No, or Not Sure) so that they could assess their own writing ability generally and specifically. More emphasis was put on the content, organization and mechanics of writing, and fewer questions were posed on the less observable criteria, such as vocabulary and language use.

	Yes	No	NS
Content			
1. Have you understood the topic you have written about?			
2. Have you written your thesis statement first?			
3. Have you written an outline of your essay before you start writing?			
And some other questions on the content of the essay.			
Organization	Yes	No	NS
4. Have you written an introduction to your essay?			
5. Have you written a motivator (e.g. a question, quotation, anecdote, statistics or a controversial idea) at the beginning of your introduction?			
6. Have you written a narrowed-down thesis statement?			
And some other questions on the Organization of the essay.			
Grammar & Mechanics	Yes	No	NS
7. Have you used transitions to join sentences and paragraphs?			
8. Do all your sentences have a subject and a verb?			
9. Have you applied the tenses in your sentences correctly?			
And some other questions on the grammar and mechanics of their writing.			
Vocabulary & Language Use	Yes	No	NS
10. Are all your sentences meaningful and logical?			
11. Are all your words correctly used?			
12. Are you sure about all the words you used in your writing?			

Figure 1. Self-assessment of five-paragraph essays checklist

At the end of every session of instruction, the participants were asked to write a five-paragraph essay on a specific topic. They were also required to keep their assigned essays to keep track of their own development and their goals achieved. The DSA of the assigned essays revealed whether the students could recognize their own weaknesses and set future goals for their learning (metacognitive awareness). The participants were also requested to keep their journals and to write their opinions, goals, attitudes and assessments of their essay writing skills. They were later analyzed by the researchers qualitatively. The writing rubrics/ scales (Jacob's) were introduced and distributed at the beginning of the course to give the learners the rationale for the use of DSAs. Yet the checklists were the focus of this research since the participants found them easier and more precise to use and score.

Moreover, through a semi-structured interview at the end of the experiment, all the students in the experimental group (30) could express their attitudes and beliefs towards the impact of DSA on their language skill development and their goal-setting. The interview, which lasted 150 minutes, revealed if the use of DSA had made the learners feel more responsible for their own learning and had led them to develop better criteria for their own assessment.

3.3 Procedure

One week after the beginning of the winter semester in 2015, the sophomores of English Translation learning English essay writing at the Islamic Azad University of Islamshahr took a PET. Sixty learners whose scores fell one SD above and below the mean were selected as the prime homogeneous participants of the study. The researcher did not exclude the learners who had not met the criteria; nevertheless, he mainly focused on those who had met the criteria for the study.

The selected learners received a pre-test of IELTS writing task (2) as well to assess how well they were familiar with L2 writing before they experienced the treatment. The researcher conducted the treatment throughout the winter semester which took three months (each week one session of writing according to their syllabus, each session lasting for 90 minutes). Both groups of learners received similar instruction for their ordinary writing courses by the same teacher and materials; meanwhile, the experimental group were asked to self-assess their essays in the checklists based on the ESL composition profile by Jacobs et al. (1981, as cited in Hughes, 2003), divided into five sections: (1). Content, (2). Organization, (3). Vocabulary, (4). Language Use, and (5). Mechanics.

The experimental group was also individually interviewed after the treatment and their opinions and attitudes towards SA were collected qualitatively and quantitatively.

The following Figure 2 illustrates the whole procedure of the experiment:
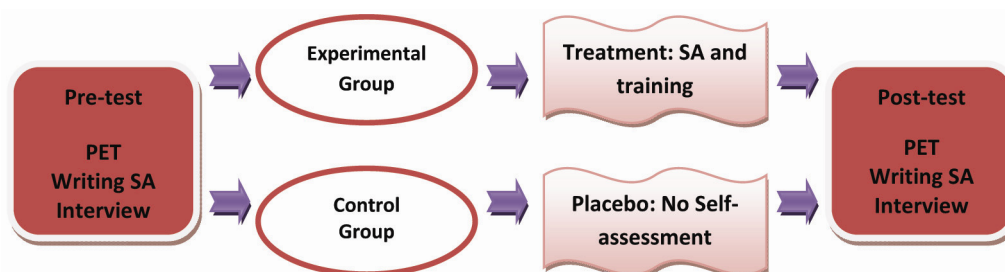

Figure 2. Illustrating the procedure of the research

In the writing classes, both experimental and control groups received similar materials, a course book entitled as "*Practical Writer with Readings*" by Baily & Powel (1989). The learners were instructed on the general style of five-paragraph essays, including introductions, body paragraphs and conclusions as well as different kinds of expository genre, such as cause-effect and comparison and contrast essays. They were also assigned to write a five-paragraph essay on each topic discussed in the classroom every session.

Both classes followed process writing method of teaching in which the teacher is involved and gives feedback on every step of writing, like outlining, drafting, redrafting and revising. That is, each session the organization of a five paragraph essay they were supposed to write was explicitly explained and a sample essay was read aloud from their book. Then, a topic was introduced to the participants to write on, followed by a warm-up activity through brainstorming. An outline was then developed based on the most appropriate subtopics they could write about in each paragraph. They were given some time to provide the first draft of their writings while the teacher gave feedback whenever needed. Some of the essays were randomly read aloud and corrected as samples. At the end of each session a different topic on the same theme was introduced to the participants to write about as assignment. These writings were then collected and rated by the teacher. Yet the experimental group was also asked to self-assess their writings based on the checklists provided by the teacher based on the ESL composition profile by Jacobs et al. (1981, as cited in Hughes, 2003).

The same procedure was followed in the control group, but they did not have any practice in DSA. Instead, they were asked to revise their essays and were assessed by their teacher. An independent samples *t*-test run between the writing post-tests of the experimental and control groups demonstrated the impact of DSA practices on their writing skill.

4. Results

4.1 Testing Assumptions

The present data were analyzed through the parametric independent samples *t*-test which is based on two main

assumptions of normality and homogeneity of variances. The first assumption was met. As displayed in Table 1 below, the ratios of skewness and kurtosis over their respective standard errors were within the ranges of +/- 1.96 (Field, 2009).

Table 1. Testing normality assumption

Group		N	Skewness			Kurtosis		
		Statistic	Statistic	Std. Error	Ratio	Statistic	Std. Error	Ratio
Control Group	PETPre	30	-.547	.427	-1.28	.265	.833	0.32
	PreWR	30	.531	.427	1.24	-.688	.833	-0.83
	PostWR	30	.568	.427	1.33	-.108	.833	-0.13
Experimental Group	PETPre	30	-.307	.427	-0.72	.793	.833	0.95
	PreWR	30	.320	.427	0.75	.714	.833	0.86
	PostWR	30	-.256	.427	-0.60	-1.093	.833	-1.31

Note. Pre = Pre-test, Post = Post-test, WR = Writing

4.2 PET and Participant-Selection Statistics

The PET test was administered to 76 participants. Based on the mean plus and minus one SD, 60 participants were selected in two groups: experimental and control groups.

Table 2. Descriptive statistics; PET

	N	Mean	Std. Deviation	Variance
PET	76	52.19	16.193	262.211
Valid N (listwise)	76			

4.3 Investigating the First Null Hypothesis

Based on the first research question, the first null hypothesis was set as: *the use of DSA in writing does not have a significant effect on EFL learners' writing.*

An independent samples *t*-test was run to compare the experimental and control groups' mean scores on post-tests of writing in order to probe the effect of DSA on the improvement of the writing ability of the Iranian EFL learners. As displayed in Table 3, the mean scores for experimental and control groups on post-tests of writing were 84.17 and 74.93, respectively.

Table 3. Descriptive statistics; post-tests of writing by groups

	Group	N	Mean	Std. Deviation	Std. Error Mean
PostWR	Experimental	30	84.17	6.524	1.191
	Control	30	74.93	5.139	.938

The results of the independent samples *t*-test (t (58) = 6.09, $p < .05$, r = .62, representing a large effect size) indicated that there was a significant difference between experimental and control groups' mean scores on the post-tests of writing. Since all the other variables in the two groups were controlled, it can be concluded that the DSA significantly improved the writing ability of the Iranian EFL learners. Thus the first null-hypothesis **was rejected**. It can be concluded that the application of SA in the experimental group has significantly affected their writing ability.

Table 4. Independent samples *t*-test; post-tests of writing by groups

	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
Equal variances assumed	4.580	.037	6.090	58	.000	9.233	1.516	6.198	12.268
Equal variances not assumed			6.090	54.985	.000	9.233	1.516	6.195	12.272

It should be noted that the assumption of homogeneity of variances was not met (Levene's F = 4.58, *p* < .05). However, in case the groups have equal sample sizes, there is no need to worry about the violation of this assumption (Bachman, 2005).

4.4 Investigating the Second Null Hypothesis

The second null hypothesis was considered as: *there is no significant relationship between the teacher's assessment and the EFL learners' DSA in writing.*

To probe the second research question, the Pearson's Product-moment correlation coefficient is calculated between the EFL learners' DSAs and their teacher's assessment of their performances both at the pre-tests and post-tests as well as throughout the eight treatment sessions.

A: Correlation coefficient between writing pre-test in experimental group and their pre-test DSA

The results of the Pearson correlation (r (28) = .34, *p* > .05, representing a non-significant effect size) indicated that there was a non-significant relationship between pre-test of writing and pre-test of writing DSA.

Table 5. Pearson correlation; pre-test of writing and DSA

		PreWR
SAWRPre	Pearson Correlation	.342
	Sig. (2-tailed)	.065
	N	30

B: Correlation coefficient between writing post-test in experimental group and their post-test SA

The results of the Pearson correlation (r (28) = .61, *p* < .05, representing a moderate effect size) indicated that there was a significant moderate relationship between post-test of writing and post-test of writing DSA.

Table 6. Pearson correlation; post-test of writing and DSA

		PostWR
SAWRPost	Pearson Correlation	.612
	Sig. (2-tailed)	.000
	N	30

Since the Pearson correlation coefficients run between post-tests of writing and post-tests of writing DSA represented a significant relationship, it can be concluded that the second null-hypothesis **was rejected**. That is, the DSA ability of the students in the experimental group has improved and approached the teachers' assessment throughout the term, leading to more valid results in their DSA. Such an improvement indicates that the EFL learners were able to internalize those criteria mentioned in the checklists by the teacher and practiced through teacher's intervention and mediation throughout the treatment.

The following diagram (Figure 3) shows the gradual increase in the correlation coefficient between teachers' assessments and learners' DSAs through the eight sessions in which the learners submitted their assignments as well as their DSAs.
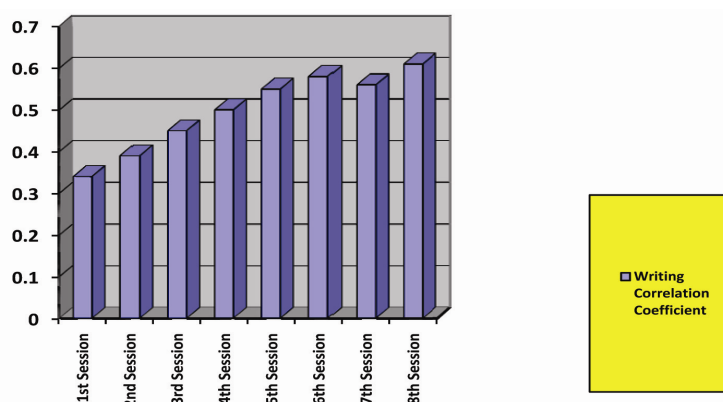
Figure 3. Incremental increase in the correlation coefficient between teacher's assessments and learners' DSAs through the eight sessions of treatment

4.5 Investigating the Third Null Hypothesis

The third null hypothesis based on the third research question was: *the analysis of the teachers' assessment and the learners' DSA does not represent the EFL learners' weaknesses and the amount of feedback the learners require*?

The significant correlation coefficient between the teacher's assessment and learners' DSAs indicate that the learners have been able to internalize the appropriate criteria for their DSAs and have improved their writing ability.

An analysis of the learner's DSAs and the teacher's assessments especially on their pre-test and post-test tasks was conducted in five different sections:

1)   Content: The content of the learners' writings has been developed highly throughout the term, and they have been able to achieve the appropriate criteria to choose better content through the teacher's feedback and intervention. They have been able to write more narrowed-down thesis statements and support them more appropriately through the body paragraphs, using more precise topic sentences and more relevant supporting sentences.

2)   Organization: The learners have been able to apply the adequate organization for a five-paragraph essay after eight sessions of practice, observing the criteria the teacher has emphasized throughout the term to write introductions, body paragraphs and conclusions. The organization of such an essay seems to be the easiest part of the essay writing criteria the learners have been able to acquire right from the beginning.

3)   Vocabulary: Probably the most problematic criterion for the EFL learners in their writing skill is vocabulary use. They have got serious problems, especially at the lower intermediate level they are in, due to their small vocabulary size and their inadequate use of collocations and lexical items. It seems that the learners require more practice to develop their vocabulary size and lexical use. They are highly influenced by their native language and negatively transfer some words and collocations from their native language into their second language writing.

4)   Language Use: Although the learners have been able to develop their language usage significantly, observing the basic grammatical rules, they seem to have problems in expressing their thoughts clearly and appropriately. They absolutely need more practice on the language use and functions as they move from the lower intermediate level to higher language proficiency levels.

5)   Mechanics: As mentioned, the language learners at this stage have been able to use the mechanics of writing correctly, but at times, the language use was not appropriate. They still have problems with more complex and compound sentences, adverb and adjective clauses, indirect quotations, embedded questions, etc. They still need more practice in using punctuations.

Such an analysis of the learners' DSAs and its comparison with the teacher's assessment of their performances will provide both the teachers and learners (stakeholders) invaluable information on how much mediation and feedback they need in order to reach their potential level and what their weaknesses and strengths are. That is, the difference and analysis between learners' DSAs and teacher's assessment will indicate the learners' ZPD and the kind of instruction/intervention they require to achieve their potential level in later remedial sessions or in further instructions.

4.6 Investigating the Interviews

To investigate the learners' and their teacher's attitudes towards DSA, the researcher interviewed every individual participant in the experimental. Four main questions were asked in this semi-structured interview, and the results were gathered and reported both quantitatively and qualitatively:

1)   Do you think DSA is a useful way of assessing learners' achievements?

2)   Do you think DSA is a profitable method of orienting learners and raising their awareness of learning goals and procedures?

3)   Which DSA technique do you prefer: checklists, rubrics or journals?

4)   What are the advantages of DSA application in the EFL writing classes?

Answering the first question, 78% of the learners found DSA a useful way of assessing their own achievements although 22% of the learners and the teacher involved in the experiment were doubtful if they could rely on the learners' DSA as a sole instrument for assessment. Although most studies have found SAs quite reliable in the literature and correlated to the teacher's assessment (Falchikov & Boud, 1989; Oscarson, 1989; Shrauger & Osberg, 1981; Sullivan & Hall, 1997; Ross, 1998), there are still doubts about using it as a mere means of assessment, except maybe in the learner-centered syllabi.

However, 89% of the learners and the teacher found DSA a useful method of orienting learners and raising their awareness of learning goals and procedures. Eleven percent found it difficult to become aware of their own weaknesses and deficiencies merely through DSA. They demanded more teacher feedback and intervention to come up with crystal clear results.

Three techniques of DSA were introduced and practiced in the present study in the experimental group: journals, checklists and scales (or rubrics). Eighty five percent of the EFL learners as well as the teacher found checklists quite appropriate and easy to use. Ten percent preferred scales/rubrics and only five percent of them were in favor of journals. The majority of the learners found scales/rubrics and journals somehow confusing and difficult to use and interpret.

In spite of such general tendency towards checklists, an investigation of the correlations between teacher's and student's assessment using various techniques of DSA revealed that the students could assess their general competence holistically in a reasonably accurate way, especially while applying scales/rubrics or journals. However, the learners' ability to assess their specific detailed competence using checklists revealed varying degrees of correlation with that of their teacher's. That is, the checklists did not reveal a constant correlation between teacher's assessment and those of the learners, demonstrating that teachers and learners may not correlate in their assessment while dealing with specific skills necessary for essay-writing, such as grammatical points, cohesive devices, or coherence. Hence, subjective, holistic SA seems more valid and although objective DSAs (checklists) sound easier to use and score and more reliable but is less valid. Yet generally, in a learner-centered approach, the reliability and validity of DSAs are not important; rather, DSA is used as a learning task and as a means of teaching. Accordingly, checklists provide the learners with a more precise account of their weaknesses and the points they have to improve, giving them the chance to focus more on the details of a skill.

The majority of the learners (75%) mentioned the revision of their own efforts as the main advantage of DSA, leading to an understanding of their own weaknesses. The DSAs of their writing helped the learners distinguish their lack of awareness of mechanics of writing, punctuations, cohesion, coherence, and grammar points. They had also become more aware of the organization of their essays and the genre structures they needed. Moreover, they found DSAs less face-threatening and less stressful than teachers' assessment and felt more responsible for their own language learning (Kavaliauskiene, 2004; Skolverket, 2001). The teacher also agreed that the learners were able to develop some metacognitive strategies and awareness to the criteria they had set. That is, they found out how to deal with a given writing task under time restrictions, how to brainstorm various topics, how to choose a specific essay format and organization for a given topic, how to outline the organization of their essays, how to take care of their coherence and cohesion and grammatical structures, how to start and end their paragraphs and support their main ideas in each paragraph, and how to develop and use their vocabulary knowledge on each topic to write accurately and appropriately. However, they need to improve their vocabulary size and use and their language use and functions.

Surprisingly, they did not have much difficulty with the organizational demands of essay writing, perhaps because those genre structures follow a universal structure and are quite comprehensible for the EFL learners. They could easily cope with the necessities of writing a paragraph in English, such as topic sentences and

supporting sentences, and even essay writing, such as writing introductions and conclusions. The learners were more concerned with sentence-level problems, such as the use of appropriate lexis, grammatical mistakes, spelling mistakes, and mechanics of writing in general.

5. Discussions

The findings of the current research demonstrate that firstly, the application of DSA in EFL classes significantly affects the improvement of their writing skills. Secondly, through DSA training and practice, the learners' DSAs gain a closer correlation with those of the raters and teachers, adding up to the validity of their self-ratings. And finally, the use of DSA indicates and diagnoses the learners' weaknesses in their writing skill.

With regard to the first hypothesis of the research, the findings of the present study (t (58) = .40, p > .05) reveals that the application of DSA in writing classes significantly affects the EFL learners' writing ability, which is quite in line with the results found by Birjandi & Siyyari (2011), Dragemark (2009), Javaherbakhsh (2010), Khodadady & Khodabakhshzade (2012), Naeini (2011), Ross (1998), Sadler (1989), Sullivan & Lindgren (2002), Taras (2001, 2002, 2003), and Wang (2007). Like (2003), this study also supports the positive effect of teacher's feedback and intervention on the learners' DSA and has revealed that such a help by the teacher will lead to a higher level of self-awareness and a better ability to detect their errors. In their interviews, the learners also expressed the fact that through their teacher's mediations they could make sure whether they were on the right track assessing their works or not.

At the same time, Wang (2007) also confirms that DSA is an effective way to improve learners' writing ability, but he declares that it is more helpful in self-revising than in timed-essay writing. Brown (2005) also applied some communicatively oriented criteria as a means of DSA which she finds "both reliable and useful" (p. 174) for student DSA of writing, as well as for learning specific language skills.

The results of the present study also suggests that learners' ability of DSA can be improved through practice and training, being in line with Sadler (1989) and Taras (2001; 2002; 2003) who think that feedback is important in the DSA process. The gradual increase in the correlation coefficient between teachers' and students' assessment not only contributes to the incremental validity of the learners' DSAs but also indicates that learners gain the kind of skills and expertise they require to assess their performances through time and practice. In this regard, teachers' feedback and intervention are of outmost importance. Here, the learners, who received their teachers' feedback directly and explicitly after each production and DSA, were able to obtain the criteria needed for their own assessments.

The second null hypothesis compared the correlation coefficient between teachers' and students' assessments between the pre-test and post-test of the experimental group. The results indicated that the correlation coefficient between teachers' and students' assessments got closer throughout the term from 0.34 to 0.61, definitely due to the training and feedback the learners received. Moreover, the incremental increase in the relationships between teachers' and learners' assessments approved the validity and reliability of DSA. Although in a learner-centered approach and autonomous humanistic learning environment, the reliability and validity of DSAs do not matter much as far as they are applied for the sake of learning, these measurements reveal the fact that students have been able to achieve and internalize those criteria provided and mediated by the teacher as the norms of the target discourse community.

The third hypothesis, checked through the analysis of the learner's DSAs and the teacher's assessments, indicated that the EFL learners were able to successfully achieve the skills to provide appropriate content, organization and mechanics of writing through the eight sessions of instruction and intervention by the teacher. However, such mediation by the teacher at their lower intermediate level could not be helpful to improve their vocabulary and language use, leaving them in need of extending their vocabulary size and developing their language and vocabulary use.

Such an analysis of the learners' DSAs and its comparison with the teacher's assessment of their performances demonstrates how much mediation and feedback the learners need in order to reach their desirable potential level and what their weaknesses and strengths are. That is, the analysis of learners' DSAs and teacher's assessment will diagnose the kind of instruction/ intervention they require to achieve their potential level in later remedial sessions or in further instructions.

The investigation of the interviews with the language learners in the experimental group revealed that the majority of the learners (89%) found the utilization of DSA in their language learning classes fruitful, and that it led to revising their own learning procedures and setting goals for further improvement. As one of the interviewees mentioned, "SA gave us the opportunity to look at our own works more meticulously." Most of the

learners (78%) preferred checklists to other forms of DSA due to their ease of use, finding them "to the point and easy to handle and less time-consuming." Moreover, they found DSAs less face-threatening and less stressful than teachers' assessment and felt more responsible for their own language learning (Kavaliauskiene, 2004; Skolverket, 2001). They developed a positive attitude towards DSA, similar to what Andrade & Du (2007) found. Some of them even suggested, "Why don't the teachers let us assess our papers by ourselves and find out how much we have learned so far?"

This investigation also revealed that grammatical mistakes were the learners' main concern, maybe due to their own knowledge of grammar which acted as an affective filter. They mainly set their goals to overcome their grammatical errors, such as using "-s" at the end of the verbs for third-person singular, using modals for WH-questions, not using those modal verbs for embedded questions and indirect quotations, etc. Surprisingly, the organizational demands of writing were not considered by the learners as problematic as other aspects, such as grammar, spelling and punctuation since those genre structures perhaps follow a universal logical pattern and are quite comprehensible and adoptable for the EFL learners. Maybe due to the techniques of brainstorming and outlining regularly practiced for each topic, they found it easy to follow their own logical plans in their essays.

In a nutshell, it can be concluded that DSA can positively affect EFL learners' writing skill. It is as reliable and valid as any other formative/dynamic assessments and can be applied as a means of self-regulation and self-management to make learners more aware of their learning process and more involved in their own goal-setting and autonomous learning. DSA can also diagnose the learners' weaknesses and strengths from one ZPD level to another and distinguish the amount of feedback and mediation they need to overcome their problems.

6. Conclusion

Based on the results of the present study, it can be concluded that the DSA significantly improves the writing ability of the Iranian EFL learners. However, the improvement is achieved if the learners receive appropriate feedback and training by the teachers. In that case, they will be able to develop their own criteria for the assessment of their own performance, being aware of the details often left implicit in the normal classes. Moreover, it was revealed that learners were more concerned with the grammatical errors and mechanics of writing than the organizational structure of a five-paragraph essay. Surprisingly, the organization of writing was most easily handled by the Iranian EFL learners, maybe due to the universal logical patterns that exist within such genres.

On the other hand, the results of the Pearson correlation between the teachers' assessment and the learners' DSA in the writing class (R (28) = .61, $p > .05$ and R (28) = .706, $p > .05$) showed that there was a significant moderate relationship between post-tests of writing assessed by the teachers and post-tests of writing DSAs. That is, after the implementation of DSA and its training, learners' DSA would develop and gradually result in closer correlation to teachers' assessment, ultimately leading to more reliable and valid DSAs. This indicates that teachers' feedback and training play a vital role in developing students' ability to assess themselves appropriately. Moreover, the teachers can diagnose the points they have to emphasize in the classroom so that the learners can achieve their goals.

An analysis of the learner's DSAs and the teacher's assessments indicated that the EFL learners successfully improved the content, organization and mechanics of their essay writing through the eight sessions of instruction and intervention by the teacher. However, such mediation by the teacher at their lower intermediate level could not be helpful to improve their vocabulary and language use, leaving them in need of extending their vocabulary size and developing their language and vocabulary use. That is, such analysis could precisely diagnose the learners' weaknesses and their need for especial remedial instructions.

The interviews with the language learners in the experimental group revealed that the majority of the learners (89%) found the utilization of DSA in their language learning classes fruitful, leading to revising their own learning procedures and setting goals for further improvement. Most of the learners (78%) preferred checklists to other forms of DSA, such as journals and rubrics, maybe due to their ease of use. Moreover, they have found DSAs less face-threatening and less stressful than teachers' assessment and have felt more responsible for their own language learning.

As a result, DSA can be applied as a means of self-orienting and self-regulating tasks and activities in the writing classrooms and seems to be quite effective in doing so (Black et al., 2003). For instance, in a writing class, students can be assigned to write an essay on a particular topic and analyze and assess it through presented scales/rubrics or checklists. These self-orienting activities will help the learners develop their own learning goals diagnosing their weak points and improve their self-confidence and motivation in learning. They will feel more

responsible for their own studies and will have a say in their own assessment if assessment is taken for learning (*AfL*) (Blanche & Merino, 1989; Brown, 2004; Kavaliauskiene, 2004; Liang, 2006; Matsuno, 2009; Oscarson, 1989; Skolvertket, 2001; Sullivan & Hall, 1997).

On one hand, EFL teachers should be trained and made aware of the standard criteria they have to use in order to rate EFL learners' performances in writing and diagnose their deficiencies. This could be done through Teacher Training Classes and by providing them with the necessary scales and rubrics for assessment. They also have to be trained on how to implement DSA techniques in the classroom and how to give feedback on the learners' productions since it is the training and the feedback by the teachers which lead to higher accuracy of learners' DSAs and better perception of the assessment criteria. As Blanche & Merino (1989) state, there will not be any significant correlation between teachers' assessment and students' DSA without appropriate training and guidance.

On the other hand, teachers should be trained and apply those DSA techniques and tasks in the classroom environment while giving feedback and mediating whenever it is necessary. They can provide the learners with some scales/rubrics or checklists to self-assess themselves or use them for peer assessment, which is equally effective according to Birjandi & Siyyari (2011). They should also give feedback to the learners on their DSA directly or indirectly. They can spend some time after each unit of instruction to let the learners self-assess their achievement and set their goals for further development. They should train and scaffold the learners on how to assess themselves explaining the rubrics/checklists and giving feedback in case needed.

To sum up, it is suggested to include DSA checklists and scales for every curriculum on language teaching, especially writing, to help students gain the kind of awareness they need to self-assess and self-regulate their improvement and diagnose their own problems.

References

Alderson, J. C. (2005). *Diagnosing foreign language proficiency: The interface between learning and assessment*. New York, NY: Continuum.

Alderson, J. C., Brunfaut, T., & Harding, L. (2014). Towards a theory of diagnosis in second and foreign language assessment: Insights from professional practice across diverse fields. *Applied Linguistics, 36*(2), 236-260

Andrede, H., & Du, Y. (2007). Student responses to criteria-referenced self-assessment. *Assessment and Evaluation in Higher Education, 32*(2), 159-181. http://dx.doi.org/10.1080/02602930600801928

Bailey, E. P., & Powell, P. A. (2009). *The practical writer with readings* (7th ed.). New York: Thomson Wadsworth.

Birjandi, P., & Hadidi, T. N. (2010). The role of self-assessment in promoting Iranian EFL learners' motivation. *ELT, 3*(3), 211-220. http://dx.doi.org/10.5539/elt.v3n3p211

Birjandi, P., & Siyyari, M. (2011). Self-assessment and peer-assessment: A comparative study of their effect on writing performance and rating accuracy. *Iranian Journal of Applied Linguistics, 13*(1), 23-45.

Black, P., Harrison, Ch., Lee, C., Marshall, B., & Wiliam, D. (2003). *Assessment for learning: Putting it into practice*. Maidenhead: Open University Press.

Blanche, P., & Merino, B. J. (1989). Self-assessment of foreign language skills: Implications for teachers and researchers. *Language Learning, 39*, 313-340. http://dx.doi.org/10.1111/j.1467-1770.1989.tb00595.x

Brown, A. (2005). Self-assessment of writing in independent language learning programs: The value of annotated samples. *Assessing Writing, 10*, 174-191. http://dx.doi.org/10.1016/j.asw.2005.06.001

Brown, H. D. (2004). *Language assessment: Principles and classroom practices*. New York: Longman.

Dewey, J. (2012). *Democracy and Education*. Pensylvania: Electronic Classics Series. DIALANG. Retrieved from www.lancaster.ac.uk/researchenterprise/dialang/about.htm/

Dragemark, O. A. (2009). *Self-assessment of writing in learning English as a foreign language: A study at the upper secondary school level*. Goteborg: Geson Hylte Tryck.

Dyke, M. (2006). The role of the "Other" in reflection, knowledge formation and action in a late modernity. *International Journal of Lifelong Education, 25*(2), 105-123. http://dx.doi.org/10.1080/02601370500510728

Falchikov, N., & Boud, D. (1989). Student self-assessment in higher education: A meta-analysis. *Review of Educational Research, 59*(4), 395-430. http://dx.doi.org/10.3102/00346543059004395

Field, A. (2009). *Discovering statistics using SPSS* (3rd ed.). London: SAGE Publications.

Harding, L., Alderson, J. C., & Brunfaut, T. (2015). Diagnostic assessment of reading and listening in a second or foreign language: Elaborating on diagnostic principles. *Language Testing, 32*(3), 317-336. http://dx.doi.org/10.1177/0265532214564505

Hughes, A. (2003). *Testing for language teachers*. Cambridge: CUP.

Javaherbakhsh, M. R. (2010). The impact of self-assessment on Iranian EFL learners' writing skill. *English Language Teaching, 3*(2). http://dx.doi.org/10.5539/elt.v3n2p213

Kavaliauskiene, G. (2004). Quality Assessment in Teaching English for Specific Purposes. *ESP World*. Retrieved from http://esp-world.info/Articles

Khodadady, E., & Khodabakhshzade, H. (2012). The effect of portfolio and self-assessment on writing ability and autonomy. *Journal of Language Teaching and Research, 3*(3), 518-524.

Liang, J. (2006). Overview of self-assessment in the second language writing classroom. Paper presented at the 2006 TESOL Convention, Tampa, Florida. Retrieved from http://secondlanguagewriting.com/documents/overview.doc

Matsuno, S. (2009). Self-, peer-, and teacher- assessments in Japanese University EFL writing classrooms. *Language Testing, 29*(1), 75-100. http://dx.doi.org/10.1177/0265532208097337

Naeini, J. (2011). Self-assessment and the impact on language skills. *Educational Research, 2*(6), 1225-1231. Retrieved from http://www.interesjournals.org/ER

Oscarson, M. (1989). Self-assessment of language proficiency: Rationale and applications. *Language Testing, 6*(1), 1-13. http://dx.doi.org/10.1177/026553228900600103

Ross, S. (1998). Self-assessment in second language testing: A meta-analysis and analysis of experiential factors. *Language Testing, 15*(1), 1-11.

Sadler, D. R. (1989). Formative assessment and the design of instructional systems. *Instructional Science, 18*, 119-144. http://dx.doi.org/10.1007/BF00117714

Shrauger, S. J., & Osberg, T. M. (1981). The relative accuracy of self-predictions and judgments by others in psychological assessment. *Psychological Bulletin, 90*(2), 322-351. http://dx.doi.org/10.1037/0033-2909.90.2.322

Skolverket. (2001). *Assessment and grading*. Stockholm: Elanders Gotab.

Sullivan, K., & Hall, C. (1997). Introducing Students to Self-assessment. *Assessment & Evaluation in Higher Education, 22*(3), 1-15. http://dx.doi.org/10.1080/0260293970220303

Sullivan, K., & Lindgren, E. (2002). Self-assessment in autonomous computer-aided second language writing. *ELT Journal, 56*(3), 258-266. http://dx.doi.org/10.1093/elt/56.3.258

Taras, M. (2001). The use of tutor feedback and student self-assessment in summative assessment tasks: Towards transparency for students and for tutors. *Assessment & Evaluation in Higher Education, 26*(6), 606-614. http://dx.doi.org/10.1080/02602930120093922

Taras, M. (2002). Using assessment for learning and learning from assessment. *Assessment & Evaluation in Higher Education, 27*(6), 501. http://dx.doi.org/10.1080/0260293022000020273

Taras, M. (2003). To feedback or not to feedback in student self-assessment. *Assessment & Evaluation in Higher Education, 28*(5), 549. http://dx.doi.org/10.1080/02602930301678

Wang, Y. H. (2007). The addition of an affect test and self-assessment into ESL writing assessment: Process and effect. A case study in a non-English major postgraduate ESL writing class. *Professional Teaching Articles, 20*, Article 1.

Copyrights