The Impact of Dynamic Assessment on Iranian EFL Students' Writing Self-Assessment

Minoo Alemi[1]

Assistant Professor, Islamic Azad University, Tehran-west Branch

Abstract

The rise of sociocultural theory has pushed up the value of dynamic assessment as one of the alternatives to traditional testing. The purpose of this study was to investigate (a) Iranian EFL (English as a Foreign Language) learners' self-assessment and self-rating of their writing ability and the effect of dynamic-assessment-based course on their accuracy, and (b) the interrelationships among teacher rating, self-assessment, and self-rating in the writing performance of twenty two engineering students taking a writing course as part of a general English course. Conducting descriptive statistics, correlational analyses, and *t*-tests revealed that the participants overrated their writing ability as measured against teacher rating before their exposure to the dynamic-assessment-based course. While they got close to each other after dynamic assessment and high correlations among all participants' self-rating, self-assessment, and teacher ratings proved the impact of dynamic assessment. The results suggest that dynamic assessment could help Iranian EFL learners to get a better awareness of their criteria for writing evaluation, and subsequently they become more accurate in assessing their own writing ability, and this is a step forward in education.

Keywords: dynamic assessment, self-assessment, self-rating, teacher rating, EFL writing, self-rating criteria

---

[1] Corresponding Author: minooalemi2000@yahoo.com

The Impact of Dynamic Assessment on Iranian EFL

1. Introduction
Assessment, as defined by Lynch (2001), is a range of procedures that includes measurement and testing but is not restricted to these forms. It is the systematic information we gather in order to make decisions about individuals, resulting from tests or other measurement procedures. The main purpose of assessment is to support the teaching/learning process. According to Gipps (1994), assessment is undergoing a paradigm shift from a psychometric to a broader model of educational assessment. Dynamic Assessment posits a qualitatively different way of thinking about assessment from how it has traditionally been understood by classroom teachers and researchers. The pedagogical approach of assessment, understanding learners' abilities, instruction, and supporting learner development are a dialectically integrated activity called dynamic assessment (Poehner, 2008).

The advent of Dynamic Assessment is under the result of the Sociocultural Theory of Mind (SCT), proposed by Vygotsky. Nowadays, educational settings are dependent on a dialogically integrated activity of assessment and instruction (Poehner, 2005, 2008; Lantolf and Poehner, 2004; Lantolf and Thorne, 2006).

One of the prominent studies about the effect of Dynamic Assessment on EFL writing was carried out by Xiaoxiao and Yan (2010). Despite the large amount of contributions of Dynamic Assessment to L_2 pedagogy, very few studies have been done in EFL writing (e.g. Donato, 2000; Lantolf and Thorne, 2006; Murphy and Maree, 2006). Therefore, the target of Dynamic Assessment (DA) could be achieved through interaction and mediation to facilitate the development of writers.

However, there are lots of research literatures on DA in psychology and general education, the approach is highly new in L_2 writing. Studies of DA's implications for problems particular

Alemi

to the development of L2 writing have been paid less attention, and there is a lot of room to work on it.

2. Review of Literature

Considering the underlying premises of the Sociocultural Theory (SCT), two aspects are given a pivotal role: the social environment as an impetus to drive individuals forward and the cultural artifacts to trigger the dynamics of thinking and acting. In this regard, individuals should be provided with ample opportunities to expand their knowledge and capabilities while engaged in social mediation and dialogic interaction. A recurrently mentioned account of Vygotsky's formulation of cultural development that is extrinsically bound to the concept of Zone of Proximal Development (ZPD) is almost prevalent throughout the mainstream of Dynamic Assessment. The genesis of cultural development is on the two distinctive social and psychological planes. At the interpsychological phase, social interaction and supportive intervention lead individuals to the higher level of attainment. At the intrapsychological phase, the external regulatory structures are gradually being transformed through the process of internalization to the internal cognitive and motor functions with resort to automatic processing (Vygotsky, 1978, 1981, cited in Lantolf and Thorne, 2006). Additionally, such a notion of assisted performance offers a favorable moment to individuals to further develop their abilities and to progress from other-regulation to self-regulation. Hence, Lantolf and Poehner (2004) put a high emphasis on the unification of assessment (the systematic representation of individuals' achievement) and instruction (the provision of a finely tuned mediation). Such a movement ultimately results in the establishment of a diagnostic tool to investigate the causes of poor performance on the one hand and a prognostic tool to enable the individual to go beyond the current level of functioning with increasing reliance on the external support on the other.

The Impact of Dynamic Assessment on Iranian EFL

In line with the fundamental properties that basically underlie Vygotsky's conceptual understanding of development, it should be noted that the solo performance of a person renders an inadequate account of the individuals' mastery level. In a similar vein, just a cursory look at the substantial procedures of Dynamic Assessment indicates that the results of an intervention that leads individuals to the subsequent developmental levels are part of the different formats of assessment (Sternberg and Grigorenko, 2002; cited in Poehner, 2005, 2008). Regarding the sandwich format, the instruction or, in other terms, the mediation phase is "sandwiched" between the pretest and posttest to bring the final results; however, the cake format which interestingly resembles the layers on a cake forms the unified whole as concurrently linked to the graded series of hints to remove the successive impediments. In this case, assessment is as an instructional intervention n a systematic manner. In unpredictable circumstances, the successive hints or the "graduated prompt approach" proposed by Campione and Brown (1985, 1987; cited in Wang, 2010) should be implemented within the dynamics of learning and teaching. The implementation of the series of hints is systematically in accordance with the specific requirements of the context of occurrence. At the outset, the presentation of "general hints" is normally intervening in the administration procedures of assessment and then gradually "specific hints" should be provided to modify the incorrect responses rather explicitly.

Two rudimentary models of Dynamic Assessment are of paramount significance. The interactionist account depicts the relation between the learner and the mediator with reference to the dialogic cooperation in a roughly qualitative fashion, while those static psychometric properties of assessment are taken into consideration with the application of the easily quantifiable account of intervention to meet the "predetermined endpoints" (Lantolf and Poehner, 2004; Poehner, 2005, 2008). Importantly,

Alemi

the ontological perspectives on human abilities are heavily contingent upon the above-mentioned theoretical constructs that are inherently derived from Vygotsky's conceptualization of the mediated mind. Following the Sociocultural Theory (SCT) as a departure point that causes a dichotomy between widely-held practices and traditional approaches, a number of controversial issues should be accounted. First and foremost, the unification of mental and physical possessions as the representative of the individual-social boundary is implied. One of the central points that requires certain degree of attention is a dialectical perspective that relies notably on an organic and inseparable wholeness to elucidate individually-oriented functions (attention-retention- intention) and socially constructed interplays (Lantolf & Poehner, 2007). On the whole, real life embedded in contextual and symbolic representations is also dialectically organized to bring an organic unity. Regarding the notion of the mediated mind, the physical and symbolic tools which are regularly the byproduct of the cultural development also set the ground for changing the events. The contributions of Sociocultural Theory (SCT) to L_2 development as an interpretive tool to broaden the awareness of educators and to provide learners with assistance to reach the actual developmental level have been investigated from a variety of perspectives.

What casts light on the merits of alternatives in assessment have been partly numerated by Brown and Hudson (1998) and subsequently reiterated by a host of other influential figures (e.g., Huerta-Macias, 2002; Hyland, 2003; Brown, 2004). On the whole, the defining characteristics of alternatives are quite conductive through giving learners an opportunity to control their own procedural and declarative knowledge and alleviating the burden of assessment from the teacher (Dickinson, 1987, cited in Oscarson, 1989; Harris, 1997). Alternatives are crucial in delineating individuals' performances and skills in a rather

The Impact of Dynamic Assessment on Iranian EFL

qualitative mode with the application of certain measurement tools, including portfolios, journals, diaries, conferences, observations, self-assessment, peer-assessment and the like (Brown, 2004) . It's important to keep abreast of the latest developments in recent studies. In this regard, reliability and practicality that have been conventionally considered as the influential characteristics of a comprehensive test should be studied in depth with regard to assessment procedures. However, alternative techniques offer greater washback and authenticity in a discerning way as the hallmarks of such a formative assessment. Under such circumstances, the implementations of the performance-based alternatives which engage learners in higher-order thinking and cyclic evaluation essentially comprise the main bulk of L_2 pedagogy (O'Malley and Valdez Pierce, 1996; cited in Brown, 2004).

What counts primarily is the elicitation of authentic performance within a supportive climate and the enrichment of evaluative procedures with employment of the predominant learner-driven assessment. Initially, what is relevant to the mutual responsibility of learners and teachers is the maintenance of learners' performance profiles and the specification of objectives in advance. Therefore, the endpoints should be clearly specified to prevent learners from being involved in unsolicited tasks and aimless chatter. In other words, learners should be aware of the purpose behind the activity and simultaneously they should be intended to satisfy the requirements of the pedagogical tasks. When learners consciously shoulder the responsibility for their own learning and assessing their own performances in a dynamic fashion, a number of issues may be put forth. A collaborative interaction between teachers and learners should not be limited to the instructional level. Obviously, learners are being interactively involved in monitoring, evaluating and planning their own progress at the assessment level. In contrast with the statutory

end-of-course exam, learners try to render an exhaustive account of their own achievement during the course of instruction for summative and in some cases placement purposes (e.g., Heilenman, 1991; LeBlanc and Painchaud, 1985; all cited in Butler and Lee, 2010). As a result, they are given the role of assessor to self-report their own weaknesses and strengths, learning styles and degree of understanding or mastery level from their perspective through self reflection.

Meanwhile, a clearer picture of students' growth is interactively being drawn during the gradual process of learning. In addition to the compilation of students' performances which is known as portfolio, self-assessment as the impetus of learner autonomy (Harris, 1997) and the medium of reflection emphasizes affective, cognitive, and metacognitive states of learners (Oscarson, 1989). According to Chamot and O'Malley (1994, cited in Harris, 1997), the implementation of self-rating not only provides feedback to the student but also directs the future learning. Under proper guidance, learners should be initially given insights into the notion of monitoring and evaluating others' performance through dialogic interaction and social mediation which prepare the way for the rise of "scaffolded assessment." Then, they gradually develop their abilities with the assistance of professional experts and peers to direct their future development autonomously. In this sense, they should be trained to deal with rating criteria to assess their own performance independently within highly specified rubrics while they are seizing an opportunity to reflect on their own written or spoken recordings during the course of instruction and look at their own products from their viewpoints.

Generally, it should be noted that three issues of validity, reliability and objectivity of alternative assessment that have a close affinity with the standardization of a test should be exhaustively investigated due to their importance (Huerta-Macias, 2002). In line with Bachman and Palmer (1989), self-

ratings are valid and reliable measures that indicate the true ability (construct) being measured. Regarding the challenging issue of objectivity, the fact is that the possibility of running an unbiased test has always been one of the central issues of administration. Contrary to our expectation, a standardized test is no more objective than an alternative assessment instrument (Huerta-Macias, 2002). The documentation and evaluation of students' products can be eventually integrated with students' journey through the learning process which is the result of the dialogic interaction between teachers, peers, and learners. Self-assessment as a facilitative and above all evaluative instrument has been set to enrich the teacher-driven and the peer-driven assessments.

The effects of self-assessment as an integral part of productive skills are particularly on the non-linear and generative writing process. It includes planning, sequencing, drafting, revising, and editing which have been hotly debated with the aim of fostering the awareness and stimulating the motivation among writers.

In a similar fashion, Matsuno (2009) stated that the idiosyncratic attributes of individuals are patently manifested in the process of self-rating, while peer-assessors rate the writing products almost leniently. Notably, reflective teachers promote writers to be reflective in a way of taking account of their own modicum of success. Ultimately, the desired objective of all writing pedagogies should be directed toward weaning writers away from instructors while training them to develop their own abilities to fulfill the future tasks competently.

At the pre-task phase, nomination of topic and activation of background knowledge establish a baseline for the subsequent stages. Next, the relevant ideas should be generated to structure the basic format of writing. At the post-task phase the overall structure, layout, and evidential supports should be thoroughly evaluated. It must be kept in mind that goal-setting and

Alemi

responding to students' writing are substantial components of writing. The purpose behind shaping feedback heavily relies on raising the awareness of learners and fostering their improvement. Correcting errors is a demanding task on the part of the teacher, peer, and learner and in the case of misapplications may bring the counterproductive results. The assessors who are making a sound decision in terms of "whens," "whos," "whichs," and "hows" of error feedback should employ a series of diverse approaches embedded in the pedagogical contexts. Hyland (2003) states that some experts believe that errors should be totally prohibited due to the fossilization phenomenon and bad habit formation (e.g., Higgs and Clifford, 1982; Lalande, 1982). However, others give a priority to the selective type of error correction (Bates et al., 1993; Ferris, 1995c; Henddrickson, 1978). In this case, those errors that impede communication of meaning and cause misunderstandings and misinterpretations should be eliminated. In a radical view, some may have a strongly held belief in tolerance of errors instead of correcting malformed pieces of information while hampering students' progress in writing due to the debilitating effect of such a direct error correction (Cook, 1991; Corder, 1981; Krashen, 1984; Selinker, 1992; Truscott, 1996). One of the salient features of the effect of raters' (peer-raters, self-raters) training on their evaluation of the writing ability is predominantly relevant to calling attention to errors. Direct means of error correction may have a variety of forms, including crossing out redundant parts and interpolating the correct forms that are especially applicable at the lower proficiency level. However, in some cases, indirect ways of correction may be desirable, specifying the location of errors or labeling specific errors with regard to the particular features of violation that are exclusively more challenging at the higher level of mastery (Kroll, 2001; Hyland, 2003). Learners may erroneously equate progress with an increase in accuracy,

The Impact of Dynamic Assessment on Iranian EFL

despite the fact that the acquisition of new knowledge during the transitory stages may proportionally increase the possibility of making errors confronting ambiguities and uncertainties. On the contrary, the possibility of making errors may be minimized during the learning plateau at the cost of blockage of communication Yule, Damico and Hoffman (1987 cited in Harris, 1997).

Eventually, the inseparability of instruction and assessment as a distinguishing feature of Dynamic Assessment in education has been in vogue recently. Therefore, there is a need of studies on DA in the development of L2 writing in education system.

3. Purpose of the Study
Studies of DA's implications for problems particular to the development of L_2 writing have been paid less attention, and there is a lot of room to work on it. Thus, this study was designed to investigate the relationship among students' self scores, teacher scores, and self assessment in pre-DA, post-DA, and between pre-DA and post-DA of this study. Thus, this study was designed to investigate the following research questions:

1. Is there any relationship among students' self scores (students rate their writings), teacher scores, and self-assessment (students assess their writing ability in a questionnaire) in the pre-DA phase?
2. Is there any relationship among students' self scores, teacher scores, and self-assessment in the post-DA phase?
3. Is there any relationship between students' pre-DA self-scores, teacher's scores, and self-assessment and those of their post-DA?

4. Methodology
4.1 Participants
This study was conducted on 22 freshman engineering students at one of the state universities in Iran. The participants consisted of 11 males and 11 females whose age ranged from 18 to 20.

The course was General English, a three- credit course that met for three classroom hours per week in a fourteen-week semester.

4.2 Instruments

Two instruments have been used in this study, namely, asking the students to write about seven topics such as globalization, corporal punishment, and co-education (as pre-test, treatment, and post-test) and a self-assessment questionnaire which was adapted from Bandura's (1995) Self-Efficacy Scales (with reliability of .86) and then modified (with reliability of .88). The questionnaire included 11 items about students' judgments of their abilities in academic writing regarding composition, grammar, usage, and mechanical skills. The students responded on a five-point likert scale ranging from 1 (no confidence at all) to 5 (completely confident).

4.3 Data Collection Procedure

Data collection procedure is as follows:

   Pre-DA/Pre-test Phase: In the first session, students took a pre-test writing a paragraph in half an hour in English about why they chose their university and their major. The reason of choosing this topic was due to their ability of writing about it spontaneously. They were not allowed to use Dictionaries; therefore, they had to rely for word meaning on strategies such as making predictions. And then they scored their own papers for self-rating and filled out a self-assessment questionnaire (on the 5-point scale ranging from 1 ("poor") to 5 ("excellent") .The instrument took about 5 minutes to complete, and it was a record of their skills at the beginning of the course.

   Treatment/Mediation Phase: In the second session, the researcher corrected their papers based on criteria of error categories which were consulted with other error analysis specialists. The criteria were writing organization, and language errors, such as word choice, verb tense, verb form, word form, articles, singular-plural, pronouns, fragments, punctuation,

The Impact of Dynamic Assessment on Iranian EFL

spelling, sentence structure, idioms, and subject-verb agreement. Then the students' scores were compared with that of the researcher in order to raise consciousness among them about the difference between their scores. Given feedbacks were continuously negotiated between the researcher and the students to encourage the learners to function at their potential level of ability. Next, the researcher assigned another topic based on assigned readings of their textbook. For seven sessions the students wrote seven draft essays on varied topics and submitted their assignments in courseware (an academic forum at the university site). Each time the researcher and the students scored the papers and negotiated all feedbacks to facilitate the L2 writer's accurate use of English writing.

Post-DA/post test Phase: At the last session, the researcher asked the students to write the last topic (what they had learned from their English class during the term) as the post-test in the class. And then they scored themselves for self-rating based on the feedbacks they received for assessing their writing process during the semester, and filled out a self-assessment questionnaire. At the same time, the researcher scored the last papers.

4.4 Data Analysis

Cronbach's Alpha was run to calculate the reliability of the self-assessment questionnaire and the result was .66 in pretest and .79 in posttest, respectively. It indicates that the uniformity increased among students' writing after the treatment. Descriptive statistics (mean and standard deviation) of the self-assessment questionnaire, students' self-rating, and teacher rating was also analyzes. Then, the researcher examined the relationship between students' self-rating, teacher scores, and students' self-assessment in the pretest and posttest. Moreover, correlation was conducted to examine the relationship between students' self-rating in the pretest and posttest, students' self-assessment in the pretest and posttest. Finally, t-tests between

Alemi

teacher rating in the pretest phase and that in the posttest phase, students' pre self-rating scores and post self-scores were run to see if there were any significant differences. It is worth to mention that gender has not been taken into account in this study.

4.5 Results
The results of the pre-course phase and post-course phase in order to answer our three research questions are reported in this section.

4.6 Pre-course Phase: The Relationship among Self-Assessment, Self-rating, and Teacher Rating
First, descriptive statistics on self-assessment, students' self-rating, and teacher rating are given in order to investigate the relationship among the variables. As Table 1 depicts, students' general self-assessment of their writing ability was an average of 3.28 on a 5-point Likert scale.

Table 1
Descriptive statistics for pre-course self-assessment

	Mean	N	Std. Deviation	Std. Error Mean
Pre-course Self-assessment	3.28	22	.397	.085

The result of students' self-assessment of the individual components showing their writing ability (Table 2) reveals that they evaluated themselves as having the highest ability in Item 4 (M=3.68), as shown below:

Item 4: I can write simple sentences with good grammar.

However, they ranked themselves as lowest in Item 7 (M=2.18), which was concerned with the use of idioms and expressions:

Item 7: I can correctly use idioms and expressions in my composition.

The Impact of Dynamic Assessment on Iranian EFL

Table 2

Descriptive Statistics for Pre-Course Self-Assessment Items

		N	Mean	Std. Deviation
	O1	22	3.59	.503
	Q2	22	2.86	.710
	Q3	22	3.59	.666
	Q4	22	3.68	.839
	Q5	22	3.64	.790
Pre-course Self-assessment Items	Q6	22	3.23	.752
	Q7	22	2.18	.733
	Q8	22	3.55	.739
	Q9	22	3.55	.671
	Q10	22	2.95	.653
	Q11	22	3.23	.528

Item 4, among high-ability items, is within the Language Section of the self-assessment questionnaire. This selection shows that general English students are highly confident in writing simple sentences in English since writing individual sentences with correct grammar is one of the important points in high school English education and in many of the general English courses for university students. However, the lowest rank belonged to correct use of idioms and expressions in compositions. This self-assessment is in line with the low ability of general English students in using idioms and expressions due to the lack of attention to this part in their previous English education in high school and general English courses in universities which are basically focused on grammatical knowledge and non-idiomatic lexical knowledge. Self-rating was the second variable in the pre-course phase. In fact, they were asked to write a composition and then rate it on a 5-point scale of 1 (very poor) to 5 (excellent) in the self-rating part. Table 3 shows that they gave the average of 3.36 to their writing.

Alemi

Table 3
Descriptive Statistics for Pre-Course Self-Rating

	Mean	N	Std. Deviation	Std. Error Mean
Pre-course Self-rating	3.36	22	1.177	.251

The average of their ratings (3.36) is high revealing that they had high opinion of themselves regarding their writing performance in their pre-test. It also showed consistency in their evaluation of their general writing ability prior to writing task performance which was nearly the same average, i.e. 3.27. The similarity of the averages shows that the students cannot change their general perception of their writing ability expressed through self-assessment. However, just similarity of the averages may be misleading, and the correlational study shows whether the two sets of scores are interrelated.

Comparing students' self-assessment and self-rating with teacher rating examines the accuracy of the first two against teacher rating. Table 4 shows the finding related to teacher rating, and the average score the teacher gave to students' writing task in the pre-course phase was 2.36. This reveals that the teacher's evaluation of students' writing performance is much lower than students' self-assessment and self-rating, and it proves students' over-evaluation of their language skills, which is a feature of students at the low-proficiency level.

Table 4
Descriptive Statistics of Pre-Course Teacher Rating

	Mean	N	Std. Deviation	Std. Error Mean
Pre-course Teacher Rating	2.36	22	.790	.168

In order to answer the first research question completely, beside these descriptive findings, the calculation of the correlations among the three variables of self-assessment, self-rating, and teacher rating is required. As shown in Table 5, the correlation between students' self-assessment and their task-specific self-

The Impact of Dynamic Assessment on Iranian EFL

rating is .23 , and it's not significant at the *p*<.05. The correlation between students' self-assessment and teacher rating (.31) was not significant either. However, a significant correlation (.62[**]) was realized between students' self-rating and teacher rating.

Table 5
Correlations Among Pre-course Self-assessment, Self-rating, and Teacher Rating

		Pre-course Teacher Rating	Pre-course Self-assessment
Pre-course Self-rating	Pearson Correlation	.620[**]	.237
	Sig. (2-tailed)	.002	.287
	N	22	22
Pre-course Self-assessment	Pearson Correlation	.312	.276
	Sig. (2-tailed)	.157	.214
	N	22	22

The reasons behind lack of significant correlation between students' self-assessment and their task-specific self-rating on the one hand and self-assessment and teacher rating on the other may lend support to the fact that first, students in the Iranian ELT context have a false understanding of their writing ability in their self-assessment. Second, ELT students don't have a proper understanding of self-rating and self-assessment scales.

However, the correlation between students' self-rating and teacher rating is statistically significant (r=.62, *p*<.05). This finding shows that when the students rate a writing task they have performed, they evaluate their writing ability in their piece of writing more logically. Consequently, their rating approaches

Alemi

to teacher rating, and subsequently task-based ratings are more likely to correlate with each other.

4.7 Post-course Phase: The Relationship among Self-assessment, Self-rating, and Teacher Rating

The second objective of this study was to examine the relationship among students' self-assessment, students' self-rating, and teacher rating. Therefore, first averages of these three ratings are reported, and then the correlational indexes show the relationships among these three variables in the post-course phase of the study. During the term, there was a dynamic assessment of their writing, and there was a dialogic interaction between the teacher and the students on the accuracy of their self-ratings against teacher ratings.

As shown in Table 6, the self-assessment average at the end of DA-based writing course reached to 3.31. Unfortunately, **c**ompared with the average of 3.27 for pre-course self-assessment, there was not a big change in students' self-assessment of their writing ability in spite of receiving training.

Table 6
Descriptive Statistics of Post-course Self-Assessment

	Mean	N	Std. Deviation	Std. Error Mean
Post-course Self-assessment	3.31	22	.347	.085

An investigation of students' self-assessment of their writing ability, as revealed through the self-assessment questionnaire, depict that the students' highest evaluation of their writing ability is on Item 4 (M = 3.64), evaluating their ability to write simple sentences with good grammar**.** Therefore, the participants considered themselves to have the highest ability in writing simple sentences with good grammar because of grammar-based English language teaching and testing in Iran.

The Impact of Dynamic Assessment on Iranian EFL

Table 7

Descriptive Statistics for Post-Course Self-assessment Items

	A1	22	3.59	.666
	A2	22	3.14	.774
	A3	22	3.55	.739
	A4	22	3.64	.790
	A5	22	3.59	.666
Post-course Self-assessment Items	A6	22	3.36	.727
	A7	22	2.36	.790
	A8	22	3.23	.813
	A9	22	3.59	.666
	A10	22	3.23	.813
	A11	22	3.18	.501

Item 7 which relates to the ability to correctly use idioms and expressions in compositions received the lowest with the average of 2.36. Since it was the lowest rank in the pre-course phase of self-assessment as well, it shows it's the most difficult part for EFL students. The reasons could be the very low weight given to idioms and expressions in high school and university general English courses in one hand and linguistic variation in the idiomatic expression of concepts in Persian and English on the other.

The average of Students' self-rating of their writing ability is 3.45 (Table 8). This rate is higher than that of their self-assessment in the post-course phase. As shown in Table 9, the average of teacher score is 2.77, and the self-rating average shows a much higher rate. The lower average score of teacher shows that EFL students are not aware of complicated nature of writing ability and subsequently tend to overrate their writing performance.

Alemi

Table 8

Descriptive Statistics of Post-Course Self-Rating

	Mean	N	Std. Deviation	Std. Error Mean
l	3.45	22	.800	.171

Table 9

Descriptive Statistics of Pre-Course Teacher Rating

	Mean	N	Std. Deviation	Std. Error Mean
Post-course Teacher Rating	2.77	22	.973	.207

A correlational analysis was done to find the relationships among the three variables in the post-course phase. As shown in Table 10, the results show the correlational index of .540 between students' post-test self-assessment and their post-test self-rating of their writing ability which is statistically significant at $p<.05$. This finding reveals that the students' general assessment of their writing ability went hand in hand with their task-specific rating of their writing ability after taking the DA-based writing course. This is highly prominent because it shows that the same criteria were involved in general assessment and the task-based assessment of their writing ability. In spite of the match or mismatch between these two ratings and teacher rating, this finding proves the effect of dynamic assessment on the students' assessment ability in both off-task and on-task situations because they gain insights into their writing ability throughout the DA-based course.

The Impact of Dynamic Assessment on Iranian EFL

Table 10

Correlations Among Post-Course Self-assessment, Self-Rating, and Teacher Rating

		Pre-course Teacher Rating	Pre-course Self-assessment
Post-course Self-rating	Pearson Correlation	.384	.540[**]
	Sig. (2-tailed)	.078	.010
	N	22	22
Post-course Self-assessment	Pearson Correlation		.472[*]
	Sig. (2-tailed)		.027
	N		22

Furthermore, the correlation between students' self-assessment and the teacher rating of their writing ability was significant (.472). This finding again is the evidence to the effect of the DA-based writing course on students' awareness of their writing ability and subsequently the closer approximation of their assessment to teacher rating.

In contrast to the mentioned significant correlations, the correlation between students' self-rating and teacher rating was not significant (.384). The reasons are multiple for this unexpectedly low correlation although the students were expected to gain the insights through the course. The first reason may be due to different scaling in the self-assessment questionnaire and that of self-rating which was based on Un-informed assessment, and that's why it had a lower correlation with teacher rating. The next reason could be the conscious application of the criteria for teacher's rating while the students might have failed to apply rating criteria instructed to them through the DA-based course.

Alemi

4.8 Pre-self-assessment and Post-self-assessment Phases: The Impact of Dynamic Assessment on Self-assessment and Self-rating

In order to examine the effects of a DA-based writing course on the accuracy of EFL students' self-assessment and self-rating and the match of these two with teacher rating, a number of t-tests were run for these three variables in the pre-course and post-course. As shown in Table 11, t-test (t = -.44, df = 21) reveals that there is not a significant difference between the mean scores of pre-course (M = 3.27) and that of the post-course (M = 3.31) for students' self-assessment of their writing ability.

These small differences between the two means are reasons to argue for and against the advantages of dynamic assessment. The positive part for dynamic assessment is that the participants gained a better evaluation of their writing ability. While they over-evaluated their ability in the pre-course phase (M= 3.36), which was very far from teacher rating (M = 2.36), they became more accurate in their self-assessment at the end of the course. In fact, they assessed their post-course ability as 3.31 which was close to the pre-course average of 3.27 despite their improvement in writing throughout the course. They also improved by decreasing the pre-test mismatch between self-assessment (3.27) and teacher rating (2.36) to a smaller difference in rating from 3.31 for self-assessment to 2.77 for teacher rating. The second explanation for benefit of the DA-based course accounts for the decrease in the difference between self-assessment and self-rating as a result of the course. It's worth mentioning that before the start of the course, the students were not aware of the evaluation criteria for good writing. It was the reason of overrating and variation in assessment. Consequently, a mismatch between off-task general self-assessment and on-task self-rating as well as non-significant correlation between the two variables in the pre-test, i.e. before the DA-based instructional course (r = .23) were observed.

The Impact of Dynamic Assessment on Iranian EFL

However, they gained insight into assessment criteria through regular self-assessment and the dialogic discussion between the teacher and the students on the accuracy of their self-assessment at the end of the course.

Table 11
T-tests for Self-Assessment, Self-rating and Teacher Rating

	Paired Differences							
	Mean	SD	Std. Error Mean	95% Confidence Interval of the		t	df	Sig. (2-
				Lower	Upper			
Pre-course and Post-course Self-assessment	-.037	.389	.083	-.209	.135	-.448	21	.659
Preteacher score – Postteacher score	-.409	.796	.170	-.762	-.056	-2.409	21	.025
Preselfscore-Postselfscore	-.174	.650	.136	-.455	.107	-1.283	22	.213

The argument against the efficacy of the DA-based writing course on the students' writing awareness is due to the low correlation between self-rating and teacher rating. It reveals that the DA-based course could not provide the students with sufficient insight into writing assessment criteria.

6. Conclusion

The main purpose of this study was to examine general English students' evaluation of their writing ability before and after taking a DA-based writing course. The findings were evidence of the consistent self-rating and teacher-scaffolded dynamic assessment provided general English students with insights into their writing ability which led to their more accurate assessment of their writing ability. This study has a number of advantages for EFL students and teachers. First, the teachers can aware L2 students of the criteria involved in the evaluation of writing. Next, L_2 students need to get involved in self-rating to evaluate their own writing ability and become autonomous. Finally,

Alemi

since low-level L2 students are to over-assess their writing ability, a DA-based course can help them to get more accuracy in self-assessment.

The next purpose of the study was to find the correlation between teacher rating and student self-rating of writing ability. The results proved the impact of dynamic assessment on the lowering of mismatch between the two variables. When the students received teacher scaffolding and learned the criteria the teacher applied   to the evaluation of their writing, their self-rating was more logical. This study suggests that dialogical interaction, dynamic assessment, and teacher scaffolding of the self-rating process can improve the match between teacher rating and students' self-rating.

References

Bachman, L. F., & Palmer, A. (1989). The construct validation of self-ratings of communicative language ability. *Language Testing, 6*, 14-25.

Bandura, A. (1995). *Self-efficacy in changing societies*. New York: Cambridge University Press.

Bates, L., Lane, J., & Lange, E. (1993). *Writing clearly: Responding to ESL compositions*. Boston, MA: Heinle & Heinle.

Brown, H. D. (2004). *Language assessment: Principles and classroom practice*. White Plains, NY: Pearson Education.

Brown, J. D., & Hudson, T. (1998). The alternatives in language assessment. *TESOL Quarterly, 32*(4), 653-674.

Butler, Y. G., & Lee, J. (2010). The effects of self-assessment among young learners of English. *Language Testing, 27*(1), 5-31.

Cook, V. J. (1991). *Second language learning and language teaching*. London: Edward Arnold.

Corder, S. P. (1981). *Error analysis and interlanguage*. Oxford: Oxford University Press.

Donato, R. (2000). Sociocultural contributions to understanding the foreign and second language classroom. In J. P. Lantolf (Ed.), *Sociocultural theory and second language learning*. Oxford: Oxford University Press.

Ferris, D. R. (1995). Teaching ESL composition students to become independent self-editors. *TESOL Journal, 4*(4), 18-22.

Gipps, C. (1994). *Beyond testing: Towards a theory of educational assessment.* London. Falmer Press.

Harris, M. (1997). Self-assessment of language learning in formal settings. *ELT Journal*, *51*(1), 12-19.

Haywood, H. C., & Tzuriel, D. (2002). Applications and challenges in dynamic assessment. *Peabody Journal of Education*, *77*(2), 40-63.

Hendrickson, J. M. (1978). Error correction in foreign language teaching: Recent theory, research, and practice. *The Modern Language Journal, 62*, 387-398.

Higgs, T., & Clifford, R. (1982). The push toward communication. In T. Higgs (Ed.), *Curriculum, competence, and the foreign language teacher* (pp. 57-79). Skokie, IL: National Textbook Company.

Huerta-Macias, A. (2002). Alternative Assessment: Response to commonly asked questions. In J.C. Richards & W. A. Renandya (Eds.), *Methodology in language teaching: An anthology of current practice.* Cambridge: Cambridge University Press.

Hyland, K. (2003). *Second language writing*. Cambridge: Cambridge University Press.

Hyland, K., & Hyland, F. (2006). *Feedback in second language writing.* Cambridge: Cambridge University Press.

Krashen, S. D. (1984). *Writing: Research, theory and application.* Oxford: Pergamon Press.

Kroll, B. (2001). Consideration for teaching an ESL/EFL writing course. In M. Celce-Murcia (Ed.), *Teaching English as a second or foreign language* (3rd ed.). Boston, MA: Heinle & Heinle.

Lalande, J. F. (1982). Reducing composition errors: An experiment. *The Modern Language Journal, 75*, 305-313.

Lantolf, J. P., & Poehner, M. E. (2004). *Dynamic assessment in the language classroom.* CALPER, PA: The Pennsylvania State University.

Lantolf, J. P., & Poehner, M. E. (2007). *Language proficiency or symbolic capability: A dialectical perspective*. CALPER, (9), PA: The Pennsylvania State University.

Alemi

Lantolf, J. P., & Thorne, S. L. (2006). *Sociocultural theory and the genesis of second language development*. Oxford: Oxford University Press.

Levykh, M. G. (2008). The affective establishment and maintenance of Vygotsky's Zone of proximal development. *Educational Theory*, *58*(1), 83-101.

Lynch, B. K. (2001). Rethinking assessment from a critical perspective. *Language Testing*, *18*(4), 351-72.

Matsuno, S. (2009). Self-, peer-, and teacher-assessments in Japanese university EFL writing classrooms. *Language Testing*, *26*(1), 77-100.

Murphy, R. and Maree, D.J.F. (2006). A review of South African research in the field of dynamic Assessment. *South African Journal of Psychology, 36*(1), 168-191.

Oscarson, M. (1989). Self-assessment of language proficiency: Rationale and applications. *Language Testing*, *6(1), 1-13.*

*Poehner, M .E. & J. P. Lantolf. (200*5). Dynamic assessment in the language classroom. *Language Teaching Research, 9*(3), 1-33.

Poehner, M. E. (2008). *Dynamic assessment: A Vygotskian approach to understanding and promoting L2 development*. Chicago, IL: Springer.

Selinker, L. (1992). *Rediscovering interlanguage.* London: Longman.

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

Wang, T. H. (2010). Web-based dynamic assessment: Taking assessment as teaching and learning strategy for improving students' e-learning effectiveness. *Computers and Education*, *54*, 1157-1166.

Xiaoxiao, L. & Yan, L. (2010). A case study of dynamic assessment in EFL process writing. *Chinese Journal of Applied Linguistics*, *33*(1), 24-40.