

Alternative Assessment Procedures in Iranian EFL Writing Classes: The Washback Effect and Learners' Attitudes

Farrokhlagha Heidari

University of Sistan and Baluchestan, Iran

Heidari.f@english.usb.ac.ir

Abstract

This study aimed to investigate the washback effect of three alternative assessment procedures, namely oral-conferences, portfolios, and corpora-based feedback on the writing achievement and attitudes of Iranian EFL university learners. The participants taking part in the study were 156 native Iranian students in a mixed Reading-Writing course. Through the two three-month semesters, the learners were studied in terms of three experimental and one control groups. It should be mentioned that revision was a major component of the pedagogy and learners were asked to incorporate feedback into their revision processes. The findings revealed that the aforementioned assessment procedures led to a significantly better performance of the participants and created a considerable change in their attitudes.

Keywords: Washback, Alternative Assessment, Oral-conferences, Portfolio, Corpora-based Feedback

Received: January 2010; Accepted: December 2010

1. Introduction

A long-standing debate in language education has been the interaction between assessment and learning. Saif (2006, p. 1) defines “the effect of testing, assessment, and assessment procedures on course content, teaching, learning, and classroom activities” as washback. A glance at the literature on language teaching and testing reveals that there is considerable variation in the way authors have theoretically portrayed this phenomenon. While some authors consider tests as having nothing but negative consequences for teaching methodology and syllabus content (Vernon, 1956; Wiseman, 1961); others look at tests more positively with important implications for curriculum and as potential instruments for educational reform (Alderson, 1986; Black and William, 1998; Hughes, 1989, among others).

Self-directed learning and learner autonomy have become increasingly popular in language learning in recent years. In line with this development, assessment practices have also started to change so that learners are guided to assume responsibility not only for their learning but also for the assessment of their learning processes and products (Benson, 2001). Advances have been made with alternatives in assessment in particular, with the distinction that alternative assessment tends to emphasize self-reflection and students' involvement in assessment.

1.1. Significance of the Study

Recently, the idea of what writing ought to be has been changing and more attention has been upon the communicative aspect of writing. Writing is now a meaningful, social act and a problem-solving activity (Kroll, 1990). Even though such a shift of image has taken place, the revolution in language classes

Alternative Assessment Procedures in Iranian...

is very slow, particularly so in writing classes. This can be attributed to a number of factors. Most often students feel they are poor writers and when asked about the skill they are best at, none select writing (Dauod, 1998; Jing, 1998). It is not the case solely with the students. Teachers find teaching writing challenging, in Jing's (1998) words, "writing is the most constrained and problematic, though time- and energy consuming, it is the least rewarding" (p. 30). Furthermore, there seem to be some problems adherent to the significance of feedback and assessment procedures. In spite of the popularity of a variety of alternative assessment techniques, e.g. checklists, portfolios, self assessment, peer assessment, conferences, writing workshops, and computer-delivered feedback, teachers in general resort to traditional assessment based on their own quantitative evaluation and the teacher written response continues to play a central role in most L2 and foreign language (FL) writing classes.

The paradigm shift, from product-oriented to process-oriented approach towards writing, demands teachers to look for new and alternative assessment procedures. To encounter the unfavorable attitudes on the part of learners, on the one hand, and to give due attention to assessing this skill and providing students with appropriate and effective feedback in the curriculum, on the other, there is a need for a better and deeper understanding of the whole process. To that end, different variables have been identified and examined by researchers. The present study addresses the basic question of washback effect of alternatives in assessment on EFL students' writing achievement. In this study, these assessment tools are seen as feedback providing tools and negotiation fostering strategies at the service of students' better learning and ultimately better performance on writing tasks. Furthermore, students' attitudes as one potential factor affecting the efficacy of these alternative assessments are also taken into account.

1.2. Research Questions

Q1. Do alternative assessment procedures (oral-conferences, portfolios, and corpora-based feedback) have any impact on Iranian EFL students' writing achievement?

One broad question to be answered in 3 specific questions:

Q2. Is there any difference between learners' attitudes towards oral-conferences prior to and after experiencing them?

Q3. Is there any difference between learners' attitudes prior to and after construction of portfolios?

Q4. Is there any difference between learners' attitudes prior to and after integrating concordancing software into the writing process?

2. Background of the Study

Over the last 30 years, an enormous shift has occurred in the way teachers and researchers look at writing and writing instruction. They no longer believe, if they ever did, that "writing is writing". They are much more inclined to think of writing as an extended process that occurs over time and that draws upon different approaches to thinking and expression at different points in that process. The result of these changes is a lack of fit between current models for teaching and learning, on the one hand, and traditional models for writing assessment, on the other. Fortunately, the models for assessment are now changing. In this brave new world, assessment may in fact be based on work that engages students over a period of time inside and outside the classroom, and on teachers' evaluations of the work. Clearly, in the new view of assessment, it would be difficult to imagine a conversation about writing assessment in which no one mentioned alternative assessment (Nelson, 2000).

Alternative Assessment Procedures in Iranian...

Broadly speaking, alternative assessment is characterized as an alternative to standardized, norm-referenced, multiple-choice testing and typically claims the following features: the students' involvement in setting goals and criteria for assessment; use of higher-level thinking and/or problem solving skills; measuring metacognitive, collaborative and intrapersonal skills as well as intellectual products; and contextualization in real world applications (Linn and Baker, 1996; Maclellan, 2004).

2.1. Oral-conferences

While it is widely accepted that oral interaction has an important role to play in the planning, writing or revision stages of producing a text in L1 contexts (Bruffee, 1984), the scope and extent of its contribution are still unclear, especially to L2 writers.

Some issues remain only partially answered, especially the effectiveness of oral conferences for improving students' writing. It has been pointed out that some L2 learners have cultural or social inhibitions about engaging informally with authority figures such as teachers, let alone questioning them, as this can result in students passively and unreflectively incorporating teacher's suggestions into their work (Goldstein and Conrad, 1990). Goldstein and Conrad found that only those students who negotiated meaning successfully in conferences were able to carry out extensive and better revisions to their writing. This finding was supported by Williams (2004) who found greater uptake of tutor advice in terms of revisions when tutor suggestions were explicit, when students actively participated and negotiated in the conferences, and when they wrote down their plans during their sessions with tutors. Williams noted that negotiation was important especially for higher level text-

based revisions, although her study suggested that the majority of revisions linked to conferences were, in fact, surface level ones.

2.2. Portfolios

Barrett (2000) defined a portfolio as a purposeful representative collection of the works of the student that exhibits his/her independent efforts, progress, and achievement. Process and product synthesized, writing portfolios are for writers and composition instructors alike, not only a product- a body of writing to be assessed- but ideally, evidence of the process by which that writing is created, shaped, revised, selected, and presented (Burch, 2000).

Tiwari and Tang (2003), attempting to describe how students perceive their experience of preparing portfolios reported that the students favored the use of portfolio assessment. An unexpected finding was the fact that those students who lacked motivation came to be interested in collaborative learning and their eagerness for learning increased during the process of preparing portfolios. Conducting a quantitative project, Song and August (2002) were interested to study the potential impact of portfolio assessment on passing writing assessment test (WAT) in an ESL context. They discovered that portfolio assessment is as valid as any standardized test in predicting the students' success in an English course. They found that non-native English students are likely to pass their English courses when they are evaluated through the portfolio evaluation scheme than when they are required to pass their standardized final written test. Song and August (2002) believed that the use of the portfolio assessment seems to be a more appropriate evaluation alternative for the ESL students. The results indicated that the portfolio group outperformed the non-portfolio group. However, the authors mentioned that "both groups generally had difficulties passing a timed impromptu test" (p. 61).

Alternative Assessment Procedures in Iranian...

Tanner, Longayroux, Beijaared, and Verloop (2000) found that using portfolios as an instrument for professional development during a one-year pre-service teacher education course for language students demonstrate a trainee's learning process overtime. It also illustrates an individual's development by strengthening the student teachers' reflective skills.

Portfolios do have their own weak points, of course. To begin with, the researchers contended that portfolios are indeed less reliable than more quantitative evaluations such as test scores (Sewell, Marczak, and Horn, 2000).

And like any other qualitative measure, data from portfolio assessments are also difficult to analyze. Sewell et al. (2000) add that if goals and criteria are not clearly defined for the students in advance, the portfolio will change to just a haphazard collection that would show no pattern of growth.

2.3. Corpora-based Feedback

In recent years, there has been growing interest in applying the results of corpus research to second language pedagogy. The general aim of corpus research is to provide a rich collection, or corpus covering as wide a range of words and word combinations as possible, and to provide information about the frequency of use of these word combinations.

Integrating concordancing software into the writing process can give students direct access to corpus information as they write. This offers L2 novice writers a cognitive support tool with which to check their evolving language hypotheses. Watt (1998) maintains that a Web Concordance incorporates interactive elements which assist independent learning. In this way, concordances are linked to discovery learning (Robinson, 1997) and raising language awareness (Wichmann, 1995; Willis, 1998). Todd (2001), for example, reports a study in Thailand where students were able to make inductive

decisions about language use and to self-correct their common lexical errors after they had made small concordances of these items on the internet. Gaskell and Cobb (2004) also report learner self-correction of grammatical errors after tutors pasted in links to an online concordance to facilitate feedback and reflection. In Thurstun and Candlin's pilot study (1998), for example, participants reacted positively toward this innovation in vocabulary teaching. However, they also reported some negative reactions, such as, "some students were puzzled by the cut-off sentences of the one-line concordances and daunted by the difficulty of the authentic academic texts" (p. 271). Koosha and Jafarpour (2006) wanted to see if concordancing materials presented through data-driven learning (DDL) had any effect in the teaching/learning collocation of prepositions. The obtained results showed that there was a significant difference between the performance of the subjects in the conventional group and the DDL group.

While the previous research did not thoroughly investigate student responses, Sun's (2000) study was aimed at exploring student reactions to a corpus-based lesson. Using a questionnaire approach, he examined Taiwanese EFL students' feedback toward web-based concordancing. The majority of the students were positive toward the web-based concordancing, mostly because it allowed them to encounter authentic language use. Among its specific benefits, they felt the approach was most helpful in acquiring knowledge of the actual usage of individual words as well as phrases, and in reading comprehension.

The least effective areas, in their view, were writing proficiency, application of grammatical rules, and understanding the meaning of vocabulary. The students also expressed concern about the slow speed of Internet connections and the time involved in conducting an analysis of concordance data. Qualitative data obtained through open-ended questions likewise showed that

students perceived data analysis of concordance output as problematic due to the huge amount of data available and difficulties in adjusting to the inductive learning style fostered by corpus pedagogy.

Research is sure to continue in a range of areas related to computer feedback as technology changes our conceptions of instruction, assessment, and L2 literacy and as teachers continue to integrate technology into their writing classrooms. While the impact of these tools is an area for further investigation, they appear to offer considerable promise for supporting teacher feedback with authentic language data.

3. Research Method and Procedure

3.1. Participants

The participants taking part in the study were 156 native Iranian students aged between 20 and 24 majoring in English Literature or Translation at University of Sistan and Baluchestan and Azad University in Zahedan during the autumn and spring semesters of 2009-2010. All participants were freshmen enrolling for Reading II and the study continued with the same students as sophomores taking Reading III, a required course for the second-year students and actually a mixed Reading-Writing course. It means that the course was basically a Reading one providing the students with a chance to write about a variety of issues related to the topics of each reading section of their textbook.

Due to certain limitations, the subjects were not randomly selected; in fact all subjects enrolling for Reading II were included in the study. McBurney calls such samples 'convenience samples', "a random sample that is chosen for practical reasons" (2001, p. 246). This kind of sample, he later adds, is quite acceptable.

3.2. Instrumentation

Several means of data sources were used to assure the validity of findings:

3.2.1. Questionnaire Surveys

Questionnaires have been used for a variety of purposes: to investigate students' reactions to materials, their teachers, courses, tasks, etc. Questionnaires often reveal publicly acceptable beliefs rather than the true beliefs or actual behaviors. Because the researcher aimed at examining students attitudes regarding a set of variables, she decided to include a set of questions for each variable. Before designing the questionnaires, some informal meetings were held with subjects. In this way, some information which would ease item writing process could be gathered. The initial pool of items was pre-tested and the necessary modifications were made. It should be mentioned that an anterior examination of content validity (Weir, 1990) was preferred in this study. Anterior validity mainly focuses on expert judgment on the construct of the items.

Two questionnaire surveys were respectively administered to the students at the beginning (the first session of Reading II) and at the end of the course (week 14 of Reading III): pre-course and post-course questionnaires. The questionnaires were in two sections, the first part yielding personal information as well as students' evaluation of their own language skills, with particular attention to their writing skills. The second section included 25 statements on a Likert scale in which the respondents were asked to indicate their degree of agreement on a scale (strongly agree, agree, neutral, disagree and strongly disagree) which were rated from 5 to 1 respectively.

Alternative Assessment Procedures in Iranian...

In order to eradicate the probability of misunderstanding items, the questionnaires were developed in Persian. Furthermore, at the end of the questionnaires, an open-ended question was included 'to discover something not anticipated by the designer' (McBurney, 2001, p. 238). Therefore, the subjects were invited to add their comments if they were not covered in the questionnaires.

3.2.2. Formative Writing Tests

Eighteen formative tests based on the topics presented in the thought-provoking 'Discussion and Writing Section' at the end of each chapter of the course book were used as feedback providing tools in this study. The writing sections provided students with a chance to broaden their view on the topics of the readings, to address more global issues and concerns, and to write about a variety of issues.

3.2.3. Pre-test and Post-test Essay Writing

At the beginning (the second session of Reading II) and at the end of the course (week 13 of the Reading III), the students were provided with opportunities to write two short essays functioning as pre-test and post-test of this study. These helped the researcher to investigate the impact of the treatments shown through the performance of the control and experimental groups. As mentioned before, the students were asked to write on the topics presented in the 'Discussion and Writing Section' of their books.

3.2.4. The Concordancing Software

Mark My Words (Milton's 2006) is sophisticated concordancing software employed in this study. It allows teachers to insert customizable comments in any language in the student's electronic document and to link the comments to the online resources including corpus data. The program can identify shortlist suggested comments when the teacher right-clicks the mouse, so minimizing the time needed to insert comments while guiding teachers in the cases that they are unsure in selecting appropriate comments.

3.2.5. Students' Course Book

'Select Readings' (Intermediate and Upper-intermediate) by Lee and Gundersen (2001) was the proposed course book which provides learners of English with high-interest reading passages from authentic sources. These passages act as springboards into reading comprehension activities, vocabulary building, grammatical analysis, and thought-provoking discussion and writing. The main focus of the course was on reading passages, reading skills developments, and writing sections.

3.3. Procedure

The following steps were followed in the process of conducting the research: The first session of the course was devoted to administering the pre-course questionnaire to the experimental groups and it was also a training session lasting for one hour. The main purpose of the training session was to acquaint the participants with the nature, purpose, and the design of the specific alternative assessment which was going to be utilized for that group.

Alternative Assessment Procedures in Iranian...

During the second session of the course, the students in our three experimental groups as well as the participants of the control group were asked to write an essay, functioning as the pre-test in this study, on the topic presented at the end of chapter one of their textbook. The Multiple Trait Scoring proposed by Hamp-Lyons (1992) was used for assessing the students' essays. Hamp-Lyons (1992) states that the Multiple Trait Scoring implies giving separate scores for more than one facet or trait on any single essay. She adds that this approach is very different from the old analytic scoring which focused on relatively trivial features of text (grammar, spelling, handwriting) and which did indeed reduce writing to an activity apparently composed of countable units stung together. In other words, Multiple Trait Scoring is an approach to the whole writing assessment and not only the scoring (Hamp-Lyons, 1992). To develop scoring criteria and standards based on the Hamp-Lyons' (1991) Multiple Trait Assessment, a group rather than a single expert were employed to take into account the specific context and a range of levels appropriate to the context. The readers decided on six criteria (the nature of ideas, reasonable context, development of specifics, text structure, control of the language use, and communication effectiveness), all of which were both scored and ideally reported. The actual scoring involved two readers for each text, with a third reader if those two disagreed. Then two/three readers' scores were summed or averaged to arrive at the final, single-number score for research use. In fact, the information yielded from the Multiple Trait Scoring can be used for decision making and it can also be communicated to students in the form of diagnostic feedback. Hamp-Lyons (1995) believes that the students' essays maybe adequately represented by a single score for research use.

Then throughout the two three-month semesters each consisting of 24 sessions, all subjects were required to write eighteen essays on the topics, extracted from 'Discussion and Writing Section' at the end of each chapter of their course book.

The students of different groups received different feedback methods on the essays they had written on each of the above topics. It should be mentioned that revision was a major component of the pedagogy in this research and learners were required to incorporate feedback into their own revision processes. In the control group, the students' essays were gathered by their teacher and the students received feedback in the form of some written comments accompanied by a single-number score. The students were provided with an opportunity to revise their texts but they necessarily had to submit their revised essays the next session.

The key difference between the control and the three experimental groups was that in the experimental groups there was no talking about grades, instead, the students' essays were read in order to get a general feel of their writing ability and to provide them with some different types of feedback to revise their writings. In fact process-oriented and collaborative models of instruction and assessment which feature multi-draft assignments, teacher-student conferences, peer response, and setting out to assess students' writings as works in progress were the tenets in these three groups. It may be such a relief for our students knowing that the writing class is there to help them improve rather than evaluate and categorize them. In fact, in the experimental groups writing was indeed a form of learning.

In the oral conference group, before each writing conference, each student wrote a draft of an essay on a topic assigned by the instructor, and, during revision talk, the instructor and student identified problem areas in the

Alternative Assessment Procedures in Iranian...

student's writing, talked about ways to improve the writing, and revised the essay. The instructor displayed a preference for having the students identify problems themselves and self-correcting them. In other words, the students had to take active roles by identifying problems, by providing explanations for revisions, and by writing revisions without waiting for the instructor's directive to do so.

In the class utilizing portfolio assessment, each student's work was collected, like an archivist's collection of a writer's entire oeuvre, into a portfolio. Students in this class produced the same certain number of written documents during the term, along with the self-evaluations in which they discussed the strengths and weaknesses of each paper. It was agreed in advance that they had to provide a record of that record through self-reflection and select just six of those documents, one out of every three, to be assessed by a group of students. The students in the portfolio group were allowed unlimited opportunities to rewrite their six selected essays. It means that the students were given the opportunity to continue working on them and they were never done until the student decided they were or until the end of the course. It is believed that when students know that they can improve as writers through extra effort in revising and through the selection process available to them, they may become more responsible and more independent; in today's terminology, they may become 'empowered'.

Finally, in the corpora-based or concordancing group, the students' electronically submitted essays along with the teacher's comments were linked directly to a concordance file. In other words, there are corpus-linked programs which can help teachers provide resource-assisted feedback. In this group, encouraging the students to submit their writings electronically allowed teacher to hyperlink errors in the essays directly to the concordance file where

students could examine language use and the context and collocations of the words. Based on the authentic language use the students self-corrected their errors. Like our other two experimental groups, the students in the concordance group were also required to revise and try to improve their first drafts.

Finally, at the end of the course (week 13 of Reading III) all subjects were asked to write an essay, functioning as the post-test, to investigate the effect of alternatives in assessment on their writing proficiency. The scoring procedure for the post-test was quite similar to that of the pre-test. Control and experimental groups were compared on the basis of their mean scores.

As the last step, the post-course questionnaire was administered to participants of the experimental groups (in week 14 of Reading III) to investigate the potential changes in their attitudes in the light of alternatives in assessment.

4. Findings of the Study

The present study was carried out on the basis of a number of null hypotheses. To test each hypothesis, the relevant data were collected. Now in what follows the results of the application of statistical procedures to test the various hypotheses are presented.

H0₁. Alternative assessment procedures (oral-conferences, portfolios, and corpora-based feedback) do not have any impact on Iranian EFL students' writing achievement.

After randomly assigning the subjects into four groups, a One-Way ANOVA was performed to check the homogeneity of our four groups at the outset of the experiment. Table 1 reveals the results.

Table 1. ANOVA results for Pre-test scores

	Sum of Squares	df	Mean Squares	F	Sig.
Between Groups	31.57	3	10.52	.405	.777
Within Groups	3846.55	152	25.3		
Total	3878.12	155			

The F-ratio (Sig.= .777) shows that there is no significant difference in terms of the four groups' (each consisting of different levels of language proficiency) performance on the pre-test at the beginning of the study. Thus, it can be safely concluded that the four groups participating in the study met the condition of homogeneity.

After the two three-month semesters each consisting of 24 sessions, the treatments were completely carried out and then a post-test (essay writing) was administered at the end of the course.

For examining the first null hypothesis of this research concerning the washback effect of the alternative assessment procedures on subjects' writing proficiency, the statistical procedure of One-way ANOVA was carried out on the post-test scores (Table 2).

Table 2. ANOVA results of post-test scores

	Sum of Squares	df	Mean Squares	F	Sig.
Between Groups	1012.28	3	337.42	11.6	.000
Within Groups	4421.59	152	29.08		
Total	5433.87	155			

An F-ratio of 11.6 reveals that the first null hypothesis of the study is strongly rejected due to a significant difference observed among the groups. This can possibly be attributed to the effectiveness of the treatments.

In order to find the location of differences, a Post-hoc Scheffe test was performed. The results of which are presented in Table 3.

Table 3. Post-hoc Scheffe test multiple comparisons for the Post-test

(I) Groups of students	(J) Groups of students	Mean differences	Sig.
Corpora-based Feedback	Oral-conference	-.77	.76
	Portfolio	-.25	.987
	Control	2.85*	.003
Oral-conference	Corpora-based Feedback	.77	.76
	Portfolio	.23	.991
	Control	3.62*	.000
Portfolio	Corpora-based Feedback	.25	.987
	Oral-conference	-.23	.991
	Control	3.39*	.000

* The mean differences are significant at the .05 level.

The obtained results revealed that there is a significant difference between the performance of the three experimental groups involved in the study and that of the control group. The mean differences reported in Table 3 indicate that the most significant difference lies between the performance of the oral-conference group and that of the control group. Based on the mean differences between the control group and the three experimental groups, the experimental groups can be ordered in the following way based on their improvement: (1) oral-conference group (2) portfolio group (3) corpora-based feedback group. The control group revealed the least of all improvements that can be attributed to the lack of appropriate feedback and passiveness of the students.

As mentioned earlier, two questionnaire surveys were administered at the beginning and at the end of the course to the subjects in the three experimental groups to evaluate their attitudes towards each of these alternative procedures. A pilot test was conducted on the initial pool of items to help establish a preliminary version of the questionnaires for further refinement and evaluation. Cronbach's alpha for attitude indicated a satisfactory level of

Alternative Assessment Procedures in Iranian...

internal consistency for all the questionnaires ($\alpha > .60$). In addition to reliability coefficients, confirmatory factor analysis (CFA) was employed to confirm the anticipated subscale structure (attitude) of the questionnaires. The CFA supported that questionnaires had appropriate validity. The Rotated Factor Patterns for the three questionnaires reveal a satisfactory level of validity for the three utilized questionnaires.

The Rotated Component Matrix results for the three questionnaires appear in Table 4.

Table 4. The Rotated Component Matrix

	Component 1
Portfolio Questionnaire	.79
Corpora-based Questionnaire	.86
Oral-conference Questionnaire	.81

Extraction Method: Principal Component Analysis, 1 component extracted from the questionnaires.

H0₂. There is no difference between learners' attitudes towards oral-conferences prior to and after experiencing them.

In comparison to their perceptions before experiencing oral-conferences ($M=3.7$, $SD=1.1$) in their writing classes, the subjects scored higher after really being involved in them ($M= 5.21$, $SD=.75$). A paired samples t-test, $t_{(38)}=8.17$ revealed that the oral-conferences worked with the subjects and influenced their attitudes.

H0₃. There is no difference between learners' attitudes prior to and after construction of portfolios.

To test the third null hypothesis and to investigate the probable difference between the subjects' attitude towards portfolios before and after the writing course, a paired samples t-test was conducted. The subjects' interest for

portfolios at the outset of the experiment ($M= 4.91$, $SD= .62$) was less than their interest after experiencing the task ($M= 6.16$, $SD= .83$). Having rejected the third null hypothesis, given the $t_{(38)} = 8.12$; it can be safely concluded that the students participating in the portfolio group welcomed and enjoyed the experience.

H0₄. There is no difference between learners' attitudes prior to and after integrating concordancing software into the writing process.

Conducting a paired samples t-test for evaluating the subjects' enthusiasm for integrating corpora-based feedback and concordancing software into the writing process revealed that the subjects' scores after the experiment ($M= 3.91$, $SD= .52$) was higher than their scores ($M= 3.11$, $SD= .4$) at the outset of the experiment. A paired samples t-test, $t_{(38)}= 7.11$, allowed us to reject the fourth null hypothesis and to conclude that the writing courses improved the students' attitude for corpora-based and concordancing feedback.

5. Conclusion

Now putting together the small pieces, we try to interpret the big picture as clearly as possible. This study mainly explored the washback effect of alternative assessment procedures on Iranian EFL university students' writing achievement throughout two semesters each consisting of 24 sessions. The findings indicate significant effects of alternative assessments on Iranian EFL writing achievement and point to the importance of students' active engagement in the learning process and assessing their learning. The better performance of the three experimental groups provides support for the promotion of learner-centered approaches in which learners' autonomy and collaboration are the two major promises.

Alternative Assessment Procedures in Iranian...

Furthermore, different strategies of feedback incorporation throughout the students' composing process, to the extent that feedback can be couched in terms more meaningful than a single score, seem to be responsible for the out-performance of the experimental groups. John Harris (1986, p. 23), an expert in assessment in higher education, believes that "improvements in instruction begin with feedback on student achievement. Such feedback is dependent on assessment, and the occasional use of outside tests is not enough." At first some students had troubles adjusting to a no-grade class, but later they viewed the no-grade method as less stifling. One student wrote: "*I like the idea that I can use what I learn on all my papers before they are graded.*" He added that "*my first papers could have ruined my opportunity to get a good grade, going back and revising all my papers improved my writing to a great extent.*" In fact, students believed that these writing courses had definitely lowered their grade obsession.

Besides other things, the out-performance of the oral-conference group provides support for the beneficial effects of negotiation and face-to-face interaction that provide students with opportunities to assume a more active role in their own learning. It can be concluded that since in oral-conferences teachers and students have face-to-face interaction, there is no ambiguity in terms of teacher's comments; therefore, tutor's suggestions are explicit to the students who actively participate and negotiate in the conferences.

The survey findings indicate that in general, the students of the three experimental groups reported positive feelings about alternative assessment procedures at the end of the experiment and that there was a significant difference between their pre-course and post-course attitudes. The students' positive attitudes towards these assessment procedures can be attributed to their increased opportunities for actual in-class use of them. The students'

negative attitudes towards these alternatives in assessment prior to the study may be due to their lack of exposure and not experiencing them before. However, the continued use through the course allowed the students to gradually become familiar, and therefore comfortable with them. In other words, having received more direct information, training and real practice seemed to position the students in a situation to develop more favorable feelings about the assessment procedures. This is consistent with the literature that has argued for the need to train learner for successful use of alternatives in assessment and letting these procedures speak for themselves (Kennedy and Miceli, 2001).

What is more interesting is that most of the students reported that the oral-conferences and the portfolio classes had helped them develop the other three basic skills (speaking, reading, and listening) to a great extent. In all experimental groups, the students' attitudes improved and their writing became more worth reading than ever before.

However, it should be noted that there is a danger of the enthusiasm for alternatives in assessment being inflated to such an extent that they are seen as a sort of language teaching panacea. Instead of being promoted as a panacea, they should be incorporated appropriately into the teacher's battery of reference and teaching resources as the useful additional teaching and learning tool that they undoubtedly are.

References

- Alderson, J. C. (1986). "Innovations in language testing", in Portal, M. (ed.), *Innovations in Language Testing*, NFER Nelson, pp. 93-105.
- Barrett, H. C. (2000). Electronic teaching portfolios: Multimedia skills+ portfolio development= powerful professional development. Proceedings of the Society for Information Technology & Teacher Training (SITE) Annual Conference. Montreal Canada.
- Benson, P. (2001). *Teaching and Researching Autonomy in Language Learning*, Harlow: Longman.
- Black, P., and William, D. (1998). "Assessment and classroom learning", *Assessment in Education*, 5 (1), pp. 7-74.
- Bruffee, K. (1984). "Peer tutoring and the conversation of mankind", in Olsen, G. (ed.), *Writing Centers: Theory and Administration*, Urbana, IL: NCTE, pp. 3-14.
- Burch, C. B. (2000). "Inside the portfolio experience: The students' perspective", *Trends and Issues in Post-secondary English Studies*, NCTE, pp. 148-164.
- Dauod, S. A. (1998). "How to motivate EFL learning and teaching of academic writing by cross-cultural exchanges", *English for Specific Purposes*, 17(4), pp. 391-412.
- Gaskell, D., and Cobb, T. (2004). "Can learners use concordance feedback for writing errors?", *System*, 32(3), pp. 301-319.
- Goldstein, L., and Conrad, S. (1990). "Student input and negotiation of meaning in ESL writing conferences", *TESOL Quarterly*, 24(3), pp.443-460.
- Hamp-Lyons, L. (1991). "Scoring procedures for ESL contexts", in Hamp-Lyons, L. (ed.), *Assessing Second Language Writing in Academic Contexts*, Norwood, NJ: Ablex, pp. 241-276.
- Hamp-Lyons, L. (1992). Holistic Writing Assessment for LEP Students. Proceeding of the Second National Research Symposium on Limited English

- Proficient Issues: Focus on Evaluation and Measurement. OBEMLA, available at: <http://www.ncela.gwu.edu/pubs/symposia/vol2/holistic-ref.htm>.
- Hamp-Lyons, L. (1995). "Rating non-native writing: The trouble with holistic scoring", *TESOL Quarterly*, 29 (4), pp. 759-765.
- Harris, J. (1986). "Assessing outcomes in higher education", in Adelman, C. (ed.), *Assessment in American Higher Education*, Washington, D. C: GPO, pp. 19-34.
- Hughes, A. (1989). *Testing for Language Teachers*, Cambridge University Press.
- Jing, H. (1998). "English corner activities", *English Journal*, 72 (8), pp. 45-46.
- Koosha, M., and Jafarpour, A. (2006). "Data-driven learning and teaching collocation of prepositions: The case of Iranian EFL Adult Learners", *Asian EFL Journal*, 8 (8), pp. 192-209.
- Kroll, B. (1990). *Second Language Writing: Research Insights for the Classroom*, Cambridge: Cambridge University Press.
- Lee, L., and Gunderson, E. (2001). *Select Readings*, Oxford: Oxford University Press.
- Linn, R., and Baker, E. (1996). "Can performance-based student assessments be psychometrically sound?", in Baron, J. and Wolf, D. (eds.), *Performance-Based Student Assessment: Challenges and Possibilities*, Ninety-fifth Yearbook (Chicago, IL, National Society for the Study of Education).
- Macellan, E. (2004). "How convincing is alternative assessment for use in higher education?", *Assessment & Evaluation in Higher Education*, 29 (3), pp. 311-321.
- McBurney, D. H. (2001). *Research Methods*, Belmont: Wadsworth Thomas Learning Incorporation.
- Milton, J. (2006). "Resource-rich web-based feedback: Helping learners become independent writers", in Hyland, K. and Hyland, F. (eds.), *Feedback in Second Language Writing: Contexts and Issues*, New York: Cambridge University Press, pp. 123-139.

Alternative Assessment Procedures in Iranian...

- Nelson, B. (2000). *Identifying and Language Learning*, Harlow: Pearson Education.
- Robinson, P. (1997). "Individual differences and the fundamental similarity of implicit and explicit adult second language learning", *Language Learning*, 47, pp. 45-99.
- Saif, S. (2006). "Aiming for positive washback: a case study of international teaching assistants", *Language Testing*, 23 (1), pp. 1-34.
- Sewell, M., Marczak, M., and Horn, M. (2000). "The use of portfolio assessment in evaluation", *Cyber Net Evaluation*, Tuscan, Arisona: University of Arizona Press.
- Song, B., and August, B. (2002). "Using portfolios to assess the writing of ESL students: A powerful alternative?", *Journal of Second Language Writing*, 11(1), pp. 49-72.
- Sun, Y. C. (2000). Using on-line corpus to facilitate language learning. Paper presented at the Annual Meeting of the Teachers of English to Speakers of Other Languages, British Columbia, Canada.
- Tanner, R., Longayroux, D., Beijaared, D., and Verloop, N. (2000). "Piloting portfolios: Using portfolio in pre-service teacher education", *ELT Journal*, 54(1), pp. 20-30.
- Thurstun, J., and Candlin, C. (1998). "Concordancing and the teaching of the vocabulary of academic English", *English for Specific Purposes*, 17(3), pp. 267-280.
- Tiwari, A., and Tang, C. (2003). "From process to outcome: The effect of portfolio assessment on student learning", *Nurse Education Today*, 23 (4), pp. 269-277.
- Todd, R. (2001). "Induction from self-selected concordances and self-correction", *System*, 29 (1), pp. 91-102.
- Vernon, P. E. (1956). *The Measurement of Abilities*, University of London Press.
- Watt, R. J. C. (1998). Technology Applications Program, available at: <http://www.dundee.ac.uk/English/wics/wics.htm>.

- Wichmann, A. (1995). "Using concordances for the teaching of modern languages in higher education", *Language Learning Journal*, 11, pp. 61-63.
- Williams, J. (2004). "Tutoring and revision: Second language writers in the writing center", *Journal of Second Language Writing*, 13 (3), pp. 173-201.
- Willis, J. (1998). "Concordances in the classroom without a computer", in Tomlinson, B. (ed.), *Materials Development in Language Teaching*, Cambridge: Cambridge University Press.
- Wiseman, S. (1961). "The efficiency of examinations", in Wiseman, S. (ed.), *Examinations and English Education*, Manchester University Press.

