Tabaran Institute of Higher Education
Iranian Journal of Language Testing
Vol. 5, No. 1, March 2015
ISSN 2251-7324

On the Effectiveness of Quizzes on L2 Idioms Learning

Abbas Ali Zarei[1]

Abstract

The purpose of the present study was to investigate the effect of different quiz frequencies on Iranian EFL learners' comprehension and production of English idioms. Furthermore, the study compared students' attitudes toward frequent quizzes before and after the course and also investigated their opinions as to the most popular quiz frequency in the comprehension and production of idioms. To this end, 120 male and female language learners at intermediate level of proficiency were selected in four groups. Each group was randomly assigned to one of the treatment conditions. The first group received quizzes every week; the second group received quizzes biweekly; the third group was administered a quiz once a month; and the fourth group (the comparison group) received no quiz during the instructional period. The collected data were analyzed using two one-way ANOVAs, three Wilcoxon signed ranks tests, and two Chi-squares. The results of ANOVAs indicated that frequent quizzes had a significant effect on the comprehension and production of English idioms. The results of the Wilcoxon signed ranks tests revealed that the participants' attitudes changed positively towards frequent testing in general and its effect on the comprehension and production of idioms in particular. The results of the Chi-squares revealed that bi-weekly quizzing was the most popular quizzing frequency both in the comprehension and production of idioms. The findings of the present study may have implication for teachers, learners as well as syllabus designers.

Key words: *idiom, idiom comprehension, idiom production, frequent quizzes*

1. Introduction

In recent years, lexical knowledge has aroused much enthusiasm in second language teaching and learning. However, to speak fluently and naturally does not mean that one has to learn just single words; rather, there are many things like collocations, phrasal verbs and idioms one can make use of in order to get one's message across naturally. In fact, "there is a general consensus that the vocabulary of a language is much more than a list of individual words" (Zyzik, 2009, p. 1). A speaker's lexicon also contains multiword units such as phrasal verbs, collocations, and idioms (Carter, 1998).

Idioms - the focus of attention in this study - are a part of natural language which native speakers of language frequently use in their communication (Cooper, 1999; Irujo, 1986a;

1 *English Department, Faculty of Humanities, Imam Khomeini International University, Qazvin, Iran.*
E-mail: a.zarei@hum.ikiu.ac.ir , aazarei@yahoo.com

Concklin & Schmit, 2008). Not being able to grasp the meaning of idioms in real communication might lead to feelings of embarrassment. Moreover, since the use of idioms is a characteristic of native speakers and advanced EFL learners, the absence of this competence indicates that one belongs to the non-native or foreigner camp (Çelik-Yazici, 2004). For these and other reasons, they deserve much attention in language programs and should be included in syllabus design and not be ignored in EFL environments (Cooper, 1999).

Regarding the above statements, teachers are responsible for providing opportunities for students to learn idiomatic expressions and help them develop competence in idioms. There are many techniques to improve students' learning. One such technique involves giving quizzes. According to Farhady, Jafarpur, and Birjandi (1994), classroom testing or quizzes increase the efficiency of teaching and encourage students to study and review more. By receiving feedback from quizzes, students find areas of strength and weakness. Quizzes also help teachers recognize the problematic areas in the teaching process, which results in improving their performance (Bangert-Drowns, Kulik, & Kulik, 1991). On the other hand, although most scholars consider quizzes as a facilitator of learning, other scholars argue that the delivery, construction, and scoring processes take time from instruction (Wilder, Flood, & Stromsnes, 2001).

This study intends to investigate the effect of different quiz frequencies on Iranian EFL learners' comprehension and production of English idioms. Furthermore, the study compares students' attitudes towards frequent quizzes before and after the course and also investigates their opinion as to the most popular quiz frequency. It intends to answer the following research questions:

1. Does quiz frequency have any significant effect on Iranian learners' comprehension of English idioms?
2- Does quiz frequency have any significant effect on Iranian learners' production of English idioms?
3- Are there any significant differences in the perceptions of learners as to the effectiveness of frequent quizzes before and after the course?
4- Which quiz frequency is the most popular in learners' view?

2. Review of related literature

Since vocabulary has a substantial influence on the way students produce and comprehend language (Gathercole, 2006), it can be claimed that learning a second language largely means learning its vocabulary (Gass, 1999). However, the vocabulary of a given language is much more than a repository of single words (Carter, 1998). A speaker's lexicon also includes larger lexical items or multiple word units such as metaphors, similes, proverbs and idioms (Hillert & Swinney, 1999). Generally speaking, idioms – the focus of attention in this study – are the most frequent type of multiword units (Grant, 2007) that "have at least one element with figurative meaning" (Zyzik, 2009, p. 1).

Idioms are not well defined. Despite a great deal of research on the properties of idioms in the linguistic literature, scholars do not agree on a single definition (Moon, 1998; Cook, Fazly, and Stevenson, 2008; Grant & Bauer, 2004; Cullen, 2004). The definition of idiom varies remarkably among scholars and depends largely on the context of use (Liu, 2003). According to

Irujo (1986, p. 288) "an idiom is a conventionalized expression whose meaning cannot be determined from the meaning of its parts". Cutting and Block (1977) hold that "idioms are sometimes viewed as a unitized phrase with interpretations that are independent of the literal meanings of their individual words" (p. 57). According to Ifill (2002), "the individual words in an idiom cannot be replaced by synonyms and still retain the idiomatic reading of the phrase. This is what qualifies them as fixed forms" (p. 8).

Simpson and Mendis (2003) are of the opinion that "the word idiom conjures up the language that is thought to be entertaining, engaging, charming, colourful, and memorable" (p.419). But such a description is not enough incentive to learn idioms and put them in the EFL curriculum. However, there are good reasons for focusing on idioms. For one thing, a large proportion of language is composed of figurative features (Zyzik, 2009; Concklin & Schmitt, 2008; Celik-Yazici, 2004). Erman and Warren (2000) found that multi-word expressions form 50 percent of a language. They also estimated that 58.6 percent of the spoken discourse and 52.3 percent of the written discourse are composed of multiword expressions. Moreover, they found that most English speakers use about 4.08 idioms per minute. In another study, cooper (1999) transcribed idioms from 3 hours of taped television program and verified that most English speakers use about three idioms per minute. Therefore, the first reason for the importance of idioms is that they are pervasive (Bortfeld, 2003). In addition, idioms are not only frequently used in language, but also play an important role in discourse. What makes idioms a critical part of any language is that they have processing privilege since they are mentally stored and retrieved like a single word (Abel, 2003; Androu & Galantomas, 2008; Wray, 2000, 2002).

At the same time, idioms are widely recognized as a stumbling block in the acquisition of a foreign language. They have been proven to be one of the most difficult phenomena of language for native speakers (Nippold, 1991) and ESL/EFL students (Adkins, 1968; Irujo, 1986b; Cedar, 2008). According to Nippold (1991), "there is no clear point in human development when it can be said that idioms have been mastered" (p. 101). The main reason is that idioms are not literal.

The proper use of idioms needs considerable effort. Unfortunately, English/American idiom dictionaries are poor in providing adequate information about the correct use of idioms, that is to say, when and in which context they should be used (Cedar, 2008). Language learners need clarification and feedback on their performance for acquiring to use idioms correctly. As a result, idiom learning is limited to classroom setting.

2.1. Frequent Testing and Quiz

The term frequent testing has been subject to many different interpretations and definitions. Some define frequent testing as a kind of examination which is carried out weekly (Keys, 1934), others as a kind of assessment which is performed on a daily basis (Dineen, Taylor, & Stephens 1989), while others define it on a monthly basis (Kling, Miller, & Reardon, 2005).

The relevant literature on frequent quizzing has witnessed a substantial amount of controversy. On the one hand, the proponents of quizzes believe that quizzes have undeniable advantages, some of which include the following. Frequent quizzes help students to retain the material for longer periods of time or make them ready for high stakes exams (Johnson & Kiviniemi, 2009). Taking into consideration the kinds of nation-wide and high stakes tests

students are required to take, frequent testing has an important role to play in preparing students for these exams (Johnson & Kiviniemi, 2009). In addition, many studies have shown that frequent testing increases students' classroom attendance (Clump, Bauer, & Alex, 2003; Jones, 1984; Wilder et al., 2001). Frequent testing is also beneficial because it can provide the school, teachers, parents, and students with useful feedback on student performance (Bangert-Drowns et al., 1986; Standlee & Popham, 1960).

One of the other important advantages of frequent testing is that it creates extrinsic motivation for the students; since students want to obtain good grades in the course, they try hard and spend a lot of time preparing for the quizzes (Dustin, 1971). There is another view that getting good grades on quizzes motivates students. This has a circular effect in that students prepare more for the quizzes since they are sources of motivation (Zarei, 2008).

Another reason why frequent testing is beneficial to students' learning is that frequent testing covers small amounts of materials. Therefore, they are processed more deeply and meticulously (Standlee & Popham, 1960). Moreover, Selakovich (1962) believes that frequent testing even results in more classroom discussion of the content or material covered in the same class.

As far as stress and anxiety are concerned, Dustin (1971) believes that stress is reduced through frequent testing. Teachers can also make sure that students are doing the required readings and assignments in the class through frequent testing (Connor-Greene, 2000; Weinstein & Wu, 2009).

On the other hand, there are also a number of issues raised against frequent testing. Administering and scoring tests are really time-consuming and it may take the class time away from efficient instruction. Frequent testing might also become tedious for students and decrease students' interest in the materials and learning in general (Kulik, et. al, 1991). Marshall (2007) thinks that too much testing does not lead to fruitful and lifelong learning because teachers put their focus only on the tests and teach to the test, providing their students only with the amount of information they need to do well on the tests.

Many studies have been done regarding the effect of frequent testing on students' learning. The earliest study, to the researchers' best knowledge, was conducted by Turney (1931) in an educational psychology course. The participants of the study were classified into two groups based on their performance on the pretest. The experimental group received weekly quizzes and the control group received just one mid-term. It was found that frequency of testing was a soure of motivation and resulted in higher performance.

The next study in chronological order was done by Keys (1934). The students were divided into two groups in which the frequency of testing differed in a way that the experimental group was tested on a weekly basis, while the control group was given tests once a month. The instructor, the content of the course, and the assignments were the same. The experimental group outperformed the control group, suggesting that frequent testing leads to efficient learning.

In another study by Dustin (1971), the effects of frequent quizzes were investigated. The experimental group received the tests every week, while the control group received the tests on a monthly basis. Results showed that the students in the experimental group had significantly higher scores than those in the control group.

Martin and Srikameswaran (1974) investigated the effects of frequent testing on long-term retention of content in a Chemistry class. The study included two groups, experimental and

control; the only difference between the groups came with the experimental group taking tests every week. The students in the experimental group outperformed those in the control group since they were motivated to do some extra work due to frequent quizzes.

In another study, Fulkerson and Martin (1981) found that frequent testing through short, objective tests resulted in better performance than the longer ones given less frequently to the students. The experimental group in this study was given eight tests, each consisting of 25 questions, while the students in the comparison group received four tests each consisting of 50 questions. The findings were in line with those of Keys (1934) and Dustin (1971). Another study was carried out by Dineen et al., (1989) to see the effect of daily frequent testing versus weekly testing on students' performance. It was found that "frequent testing was more effective for the weaker students than the stronger students" (Dineen et al., 1989, p. 200).

A study was conducted by Kika, McLaughlin, and Dixon (1992) to see the effects of weekly testing versus biweekly testing upon students' learning in an algebra course. The experimental group took tests every week, while the control one took them once a fortnight. The findings revealed that the experimental group performed better than the other group. Another study was conducted by Grover, Becker, and Davis (1989) to see if unit (four-chapter) testing and chapter testing had any beneficial effect on students' performance. The experimental group was tested at the end of each chapter, while the control group was given tests after each unit was covered. The study found that there was no significant difference between the two groups. However, students in the experimental group had a higher mean in most of the tests compared with their counterparts in the control group.

Although daily testing, as the previous study found, did not prove to benefit all kinds of students, changing the frequency of testing can result in better information retention, test performance, and student achievement. Beaulieu and Frost (1989) carried out a study with college students of management. There were three groups in the study. The first group was given three tests during the semester, the second group received seven tests, and the third group was given thirteen tests during the semester. Although the group with thirteen tests outperformed the other two, the differences among the groups were not statistical.

All the studies reviewed so far have been carried out on a relatively small scale including no more than a few hundred students. There was a very large scale study done by Khalaf and Hanna (1992), in which nearly 2000 biology students participated. The participants were divided into two groups, experimental and control. The students in the experimental group were given tests every two weeks, while the members of the control group received a test on a monthly basis. The results revealed that frequent testing had a beneficial effect on students' achievement and retention of information. Similarly, Geist and Soehren (1997) carried out a study with dental students dividing them into two groups of weekly and no quiz. They found that frequent quizzing had a positive effect on students' performance.

Clump et al., (2003) conducted a study on frequent testing with four groups. The first group took no test, the second group took one test, the third group took two tests, and the fourth one took three tests. He found that the fourth group significantly outperformed the other three groups indicating that the higher the frequency of quizzes, the higher the scores. Kling et al., (2005) carried out another study with marketing students. The participants were divided into two groups, experimental and control. The experimental group was given twelve tests during the semester, while the control group was tested three times. The results were somehow

contradictory. Although the experimental group outperformed the other one in the final exam, the control group outperformed the experimental group in all tests given during the semester. In another study conducted by kamuche (2005), it was found that students who took tests weekly performed better than those who did not take any quizzes.

As to the effect of frequent testing on students' retention of information, a study was conducted by Roediger and Karpicke (2006). The participants of the study consisted of undergraduate university students. The results of the study showed that those students who were tested frequently during the course remembered information better than those who were not given tests frequently.

Kamuche (2007) intended to determine the effect of unannounced versus announced quizzes on student learning and achievement at university. The study took place over two semesters and the difference between the experimental, unannounced quiz, and control group, announced, was sought on the basis of the final scores. The results of the analysis showed a significant difference between the two groups as far as academic performance was concerned. The unannounced group outperformed the announced group in academic achievement and learning.

Zarei (2008) examined the effect of frequent testing on Iranian English students' performance and classroom attendance. The results revealed that the more frequently students were exposed to quizzes, the better their performance appeared to be. Moreover, the results showed that the administration of frequent quizzes had a positive correlation with classroom attendance.

Marcell (2008) carried out a study on online frequent testing. In his study, he compared a group of learners who took quizzes online on the basis of daily readings with another group not given any quizzes, whether traditional or online. The results revealed that students tested online came to the class with more preparation and raised more questions and made more comments in the class.

A meta-analysis was conducted by Basol and Johanson (2009) on the effect of different testing frequencies on student learning and exam performance. To this end, the authors made use of 78 studies. The studies were classified into three frequency types: high, medium, and low frequency. The findings of the meta-analysis revealed that although there were no statistically significant differences among the three groups, frequent testing was beneficial to student learning and academic achievement.

However, the findings of a study reported by Zgraggen (2009) contradict those of the above studies. The study compared the effect of weekly versus bi-weekly testing on student learning and retention of information. In this study, the participants were divided into two groups. The experimental group was given the tests every week, and the control group received the tests on a bi-weekly basis. There was a significant difference between the two groups, but contrary to previous findings, the control group outperformed the experimental group on both the final exam and the retention test, which was administered one month after the final exam. This indicates that the controversy surrounding the issue of frequent quizzes is not fully resolved by empirical evidence.

Of course, sight must not be lost of the fact that the effectiveness of quizzes depend, to a large extent, on how the learners view them. In fact, the way students perceive of frequent testing determines the amount of information they learn and retain. To study students' perceptions about

frequent testing, Keys (1934) gave students a questionnaire to compare two types of frequent testing: weekly versus monthly testing. The participants of the study believed that weekly testing brought more educational benefits than monthly tests.

In a study carried out by Fulkerson and Martin (1981) the students showed a positive attitude towards frequent testing and believed that frequent testing provided them with valuable feedback and facilitated their learning progress. In a meta-analysis by Bangert-Drowns et al., (1986), it was reported that students highly favored more frequent testing and believed that frequent testing gave them a sense of motivation and long term retention of materials covered in the classroom.

Wilder et al., (2001) carried out a study on the students' perceptions about frequent testing or quiz system. 94% of the participants were reported to welcome the quiz system. About 60% of them said that due to the quiz system or frequent testing, they had to attend classes and stick to course requirements. The findings of a study carried out by Kling et al., (2005), which is consistent with Fulkerson and Martin's (1981) study, revealed that students taking more frequent tests expressed more positive attitudes towards the class and teacher's methodology compared to the control group, who took tests on a less frequent basis.

One cannot find any study conducted on the effect of frequent testing on students' idiom learning in an English learning classroom, to the researchers' best knowledge. To fill this gap, the current study intends to investigate the effect of frequent testing on students' idiom learning.

3. Methodology

3.1. Participants

This study was conducted with 120 male and female language learners at intermediate level of proficiency who studied English in private institutes in Lahijan, Rasht, and Langerood, in Iran. They ranged from 17 to 25 years of age. The participants were selected from among native speakers of Persian in four groups of 30 members each. Randomly, one class served as the control group and the other three acted as the experimental groups to receive different quiz frequencies. Group 1 received quizzes every week; group 2 received quizzes biweekly; group 3 received quizzes once a month, and group 4 received no quiz.

3.2. Instruments

In the present study, the following materials and instruments were used: To homogenize the participants, the 35-item vocabulary subtest of a Michigan general proficiency test was used. In addition, to minimize the effect of the participants' prior knowledge of the target idioms, a pretest was also administered. The pretest contained the idioms which were to be presented during the semester. It included 200 sentences each containing one selected idiom which students were required to translate into Persian.

The main course book at intermediate level introduced by the institutes was " New Interchange". Since it was difficult to find a course specifically devoted to idioms, each session, 15-20 minutes of class time were allocated to teaching idioms. The materials presented to the participants contained 28 chapters of the idiom book entitled "*English idioms in use*" by McCarthy and O'Dell (2002). The book is designed for intermediate level learners.

The post tests of the study were of two kinds: to measure the participants' productive knowledge of idioms, a 30-item fill-in-the blanks test was used. The English definitions of the

idioms were given in parentheses as a hint to help the students fill the blanks. A 30-item multiple choice test was also used to measure the participants' receptive knowledge of idioms.

Another instrument was a five-item questionnaire which was given to students in order to investigate their perceptions toward the effectiveness of frequent quizzes.

3.3. Procedures and data analysis

To begin with, each group of participants was randomly assigned to one of the four treatment conditions. The multiple-choice vocabulary subtest of the Michigan general proficiency test was used to homogenize the participants. Data from those students who scored more than one standard deviation above or below the mean were excluded from all subsequent analyses. Initially, there were 147 participants. After excluding heterogeneous learners, there remained 133 learners, two groups of 30 members each, one group of 31 and one group of 32 members. To be able to use balanced ANOVA procedures, data from three extra students in the larger groups were randomly excluded from analysis. To minimize the effect of the participants' prior knowledge of the idioms to be taught, a pretest was administered. Those idioms to which students responded correctly were not included in the post tests. Every session 10 idioms were presented to participants. 200 idioms were presented over 20 sessions spanning a whole semester. Each idiom was used in a sentence in bold face with its Persian meaning. The first group received quizzes every week; the second group received quizzes biweekly; the third group was administered a quiz once a month; and the comparison group received no quiz during the instructional period. At the end of the instructional period, an idiom comprehension post test (in multiple choice format) and an idiom production post test (in fill-in-the blank format) were administered. A five-item questionnaire was also administered to investigate the students' perceptions toward the effectiveness of frequent quizzes and to find out which quiz frequency the students liked best. The questionnaire was administered twice: once before the course started, and once, after the course.

Since the idiom recognition and production post tests were constructed by the researchers based on the idioms which were presented in classes, their content validity was taken for granted. The KR-21 method was used to estimate the reliability of the tests. The reliability index of the receptive and productive tests turned out to be (.72) and (.78), respectively. The reliability of the constructed questionnaire was also checked and Chronbach alpha turned out to be .84.

To analyze the scores of the participants on the post tests, two separate one-way ANOVA procedures were used. One ANOVA procedure was used to investigate the effects of quiz frequency on idiom comprehension. The same procedure was repeated to compare the scores of the participants on the test of idiom production. Moreover, three separate Wilcoxon Signed Ranks tests were used to answer the first three questions of the questionnaire, seeking to compare the learners' perceptions toward the quiz system, the effectiveness of quiz frequency on learners' idiom comprehension, and their perceptions as to the effectiveness of frequent quizzes on learners' idiom production before and after the course.

Two separate Chi-square procedures were also used to analyze the scores obtained from the fourth and fifth questions of the questionnaire to obtain the learners' perceptions about which type of quiz frequency has the most effect on their comprehension and production of idioms.

4. Results

4.1. Investigation of the first question

The first question sought to investigate the effect of quiz frequency on EFL learners' idiom comprehension. A one-way ANOVA procedure was used to investigate the result of the participants' post-test. Descriptive and test statistics are summarized in Table 1.

Table 1: Descriptive and test statistics for the ANOVA on idiom comprehension

Groups	N	Mean	Std. Deviation	Std. Error
Weekly	30	25.10	5.62	1.02
Biweekly	30	23.63	5.08	.92
Monthly	30	21.36	5.91	1.08
no quiz	30	19.93	6.86	1.25
	F = 4.55		Sig. = .005	

As it can be seen in Table 1, the weekly quiz group has the highest mean, followed closely by the biweekly quiz group, and the monthly quiz group. The control group has the lowest mean. Moreover, we can safely claim that there are significant differences in the performances of the groups on the idiom comprehension post test ($F_{(3, 16)} = 4.55$, $p < 0.05$). To locate the differences among the means, a post-hoc Scheffe' test procedure was run, which yielded the following results.

Table 2: Multiple comparisons of means for the learners' idiom comprehension

frequency		**Mean Difference**	**Std. Error**	**Sig.**
Weekly	Biweekly	1.46	1.52	.820
	monthly	3.73	1.52	.118
	no quiz	5.16*	1.52	.012
Biweekly	monthly	2.26	1.52	.533
	no quiz	3.70	1.52	.124
Monthly	no quiz	1.43	1.52	.830

A look at Table 2 makes it clear that the only significant difference is that between the weekly and the comparison group.

4.2. Investigation of the second question

The aim of the second question was to investigate the effect of quiz frequency on EFL learners' idiom production. To this end, another one-way ANOVA procedure was used. Descriptive statistics and test results are summarized in Table 3.

Table 3: Descriptive statistics and test results for the ANOVA on idiom production

Groups	N	Mean	Std. Deviation	Std. Error
weekly	30	23.16	5.83	1.06
biweekly	30	21.56	6.37	1.16

 ISSN 2251-7324

monthly	30	18.56	7.28	1.32
no quiz	30	16.53	6.99	1.27
		F =6.014	Sig. .001	

Table 3 shows that the weekly group has the highest mean, followed closely by the biweekly group, and the monthly group. The control group has the lowest mean. In addition, based on Table 3, we can safely claim that there are significant differences in the performances of the groups on the idiom production post test ($F_{(3, 116)} = 6.01$, $p < 0.05$). To locate the differences among the means, another post-hoc Scheffe' test procedure was run, which yielded the results summarized.

A look at Table 4 makes it clear that the weekly and biweekly groups are both significantly better than the comparison group; the other differences are not statistically significant.

Table 4: Multiple comparisons of means for the learners' idiom production

(I) frequency	(J) frequency	Mean Difference (I-J)	Std. Error	Sig.
weekly	Biweekly	1.60	1.71	.833
	Monthly	4.60	1.71	.072
	no quiz	6.63*	1.71	.003
Biweekly	Monthly	3.00	1.71	.387
	no quiz	5.03*	1.71	.039
Monthly	no quiz	2.03	1.71	.705

4.3. Investigation of the third question

The third question sought to investigate if there were any significant differences in the learners' perceptions as to the effectiveness of frequent quizzes before and after the course. For this purpose, three Wilcoxon Signed Ranks tests were used. The first question of the questionnaire attempted to obtain the learners' general perceptions about the quiz system before and after the course. The Results of the Wilcoxon Signed Ranks Test are summarized in Table 5.

Table 5: The results of the Wilcoxon Signed Ranks Test

		N	Mean	Std. Deviation
Before		120	2.23	.764
After		120	3.49	.925
		N	Mean Rank	Sum of Ranks
after – before	Negative Ranks	5	37.40	187.00
	Positive Ranks	91	49.11	4469.00
	Ties	24		
		Z = 7.991	Sig. = .001	

ISSN 2251-7324

Based on Table 5, since the Z-value is statistically significant (Z = 7.991, p < .01), we can safely claim that there is a significant difference between the means of the groups. So, the learners' general perceptions have changed positively after the course.

The same procedure was gone through for the second question of the questionnaire, which sought to investigate the learners' perceptions as to the effectiveness of frequent quizzes on learners' idiom comprehension. Results are given in Table 6. Based on Table 6, since the Z-value of is statistically significant (Z = 8.199, p < .01), we can safely claim that the means of the groups are significantly different, and that the learners' perceptions have changed positively after the course.

Table 6: The Wilcoxon Signed Ranks Test results on Idiom Comprehension

		N	Mean	Std. Deviation
Before		120	2.22	.76
After		120	3.73	.99
		N	Mean Rank	Sum of Ranks
after – before	Negative Ranks	8	25.88	207.00
	Positive Ranks	94	53.68	5046.00
	Ties	18		
		Z = 8.199	Sig. = .001	

The aim of the third question of the questionnaire was to investigate whether there are any significant differences in the learners' perceptions as to the effectiveness of frequent quizzes on their idiom production. To this end, once again, the same procedure was run, yielding the following results:

Table 7: Descriptive and test statistics for the third Wilcoxon Signed Ranks Test

		N	Mean	Std. Deviation
Before		120	2.09	.721
After		120	3.93	.857
		N	Mean Rank	Sum of Ranks
after – before	Negative Ranks	4	15.50	62.00
	Positive Ranks	105	56.50	5933.00
	Ties	11		
		Z = 8.992	Sig. = .001	

Since the Z-value is statistically significant (Z = 8.992, p < .01), we can safely claim that the means of the groups are meaningfully different. So, the learners' perceptions have changed positively after the course.

4.4. Investigation of the Fourth Question

The fourth question sought to investigate which type of quiz frequency EFL learners liked best. For this purpose, two Chi-Square procedures were used separately to answer the fourth and fifth questions of the questionnaire.

 ISSN 2251-7324

The first Chi-Square was used to obtain the learners' perceptions about which quiz frequency has the most effect on their comprehension of idioms. Table 7 contains the summary of the frequencies of the Chi-Square including the observed and expected values. Based on Table 7, it can be seen that the most popular quiz frequency is biweekly (N=55), followed by monthly (N=29), weekly (N=26), and midterm (N=8) for the comprehension of idioms. It can also be seen that a significant majority of the learners expressed a preference for biweekly quiz frequency.

Table 7: Chi-square results for the fourth question of the questionnaire

	Observed N	Expected N	Residual
no quiz	2	24.0	-22.0
midterm	8	24.0	-16.0
monthly	29	24.0	5.0
biweekly	55	24.0	31.0
weekly	26	24.0	2.0
		Quiz frequency mode	
Chi-Square		72.083	
df		4	
Asymp. Sig.		.000	

Another Chi-Square was used to obtain the learners' perceptions about which quiz frequency has the most effect on their production of idioms. Table 8 contains the descriptive and Chi-Square results.

Table 8: Chi-square results for the fifth question of the questionnaire

	Observed N	Expected N	Residual
no quiz	2	24.0	-22.0
midterm	8	24.0	-16.0
monthly	22	24.0	-2.0
biweekly	52	24.0	28.0
weekly	36	24.0	12.0
		Quiz frequency mode	
Chi-Square		69.667	
Df		4	
Asymp. Sig.		.000	

Once more, it can be seen that a significant majority of the learners expressed a preference for biweekly quiz frequency.

5. Discussions and conclusion

As to the first and second research questions, the results of the analyses revealed that the differences among the weekly, biweekly and monthly groups were not statistically significant on both the comprehension and production post tests. This finding is consistent with the study done by Basol and Johnson (2009) in which there were no significant differences among high, medium, and low frequencies.

The present study also showed that the performance of the weekly quiz group was significantly better than that of the control group both on the comprehension and production post tests indicating that the weekly quiz had a significant effect on the students' comprehension and production of idioms. Along the same line, Geist and Soehren (1997) and Ballard and Johnson (2004) found evidence in favor of weekly quizzes compared with no quiz indicating that weekly quizzes enhance students' performance. But they did not specify whether by learners' performance they meant comprehension, production, or both. The finding of the present study is also consistent with studies such as Martin and Srikameswaran (1974), Graham (1999) and Kamuche (2005), who confirmed that students who received weekly quizzes outscored students who received no quiz during the course. Surprisingly enough, in contrast with the aforementioned findings, Haberyan (2003), in his study, found that there was no significant difference between the weekly quiz group and the no-quiz control group as to the students' performance in the class.

In the present study, it was shown that the weekly-quiz group performed significantly better than the control group in idiom comprehension, and the weekly and biweekly groups were both significantly better than the control group in idiom production. Although the weekly quiz group performed better than the biweekly quiz group on idiom comprehension and production tests, the differences between the two groups were not significant. This finding is contrary to that of Martin and Srikameswaran (1974) in which there was a significant difference between the weekly and biweekly groups. It was found that students who were tested on a weekly basis significantly outscored their bi-weekly tested counterparts. Similarly, in another study carried out by Kika et al., (1992), a significant difference was found between the two groups with the weekly group outperforming the biweekly group. Contrary to both the present study and the above-mentioned studies, Zgraggen's (2009) study showed that the biweekly quizzes were more effective than the weekly quizzes.

It is clear from the present study that there is no significant difference between the performance of the weekly and monthly groups. This finding is contrary to the study conducted by Keys (1934) in which it was shown that the group which received weekly quizzes outperformed those who were given quizzes every month. In another study, Dustin (1971) found that the weekly frequency was more effective than the monthly one.

The finding of the present study with regard to the absence of a significant difference between the performance of the biweekly group and the monthly one is not consistent with the large-scale study conducted by Khalaf and Hanna (1992). In the latter study, concerning the effect of quiz frequency on the students' learning and retention of the material, the authors concluded that students in the biweekly group outperformed the monthly group.

There are several reasons why more frequent testing when compared with infrequent testing has a more influential impact on students' learning, in this work comprehension and production of idioms. One of the reasons behind the success of weekly quizzes may be attributed to class attendance; weekly quizzes make students come to class consistently. As previous

studies have shown, there is a positive relationship between frequent testing and students' attendance in the classroom (Wilder et al., 2001; Zarei, 2008) which consequently results in a positive effect on overall course grades (Wilder et al., 2001; Clump et al., 2003) since it provides more opportunities for learning in the classroom. This can particularly explain why monthly testing did not turn out to be significantly effective. The reason might be that in the present study monthly frequency did not make the students attend the class on a regular basis.

The reason behind the success of the weekly and biweekly groups in comparison with the unquizzed group in this study may be self evident. It may as well be that frequent testing encourages students to study and review more (Farhady et al., 1994) because frequent quizzing makes students come into contact with the materials of the class more evenly. Furthermore, it creates more opportunities for students to spread out their study time on a more regular basis. In fact one of explanations for the effect of bi-weekly quizzes on the production of idioms might be the processing time which was available for students.

As frequent testing motivates students to do extra work in the class (Martin & Srikameswaran, 1974), it makes the long-term retention of the materials possible. The short nature of frequent quizzes facilitates and reinforces the learning of materials in a systematic way, idioms in this case, because everything is being tested and then stored in small chunks systematically (Fulkerson & Martin, 1981). Furthermore, frequent testing makes students come to class with preparation (Dustin, 1971; Standlee & Popham, 1960). Through frequent testing, small amounts of materials are tested. Therefore, these materials are processed more deeply and meticulously and lend themselves to more efficient learning (Standlee & Popham, 1960). Furthermore, frequent testing produces a lot of small discussions between the teacher and students, which again helps learners retain the material for a longer period of time (Selakovich, 1962; Farhady et al., 1994). Fitch, Drucker, and Norton (1951) also found that students who were tested on a weekly basis were engaged in more discussions and interaction in the class.

As frequent quizzes expose students to the materials covered in the class more regularly, there is the probability that students become more familiar with the instructional expectations of the teacher and the methodology. In other words, they become more test-wise and detect the kinds of questions to be included in the final exam, in this study the post-test (Farhady et al., 1994).

Still another reason for the better performance of the weekly and biweekly quiz groups might be that quizzes create a lot of extrinsic motivation for the students since students want to obtain good grades in the course; therefore, they try hard and spend a lot of time preparing for the quizzes (Dustin, 1971; Standlee & Popham, 1960). Although, motivation and interest are the incentives for learning, Zarei (2008) argues that "motivation is not always the cause of good grades; it may well be the result of them. Quizzes increase course grades by supplying motivation, and motivate students to study by improving their grades" (pp.5-6). In addition, when students take tests regularly, they become accustomed to the tests, and this reduces their sense of test anxiety. Therefore, when taking the final examination or post-test they experience lower levels of debilitative test anxiety compared with those who take tests less frequently. The lower levels of debilitative test anxiety may, in turn, boost the learners' test performance.

As to the third research question, the results of the analyses revealed that students' attitudes changed positively and meaningfully after the course with regard to frequent testing and its effect on the comprehension and production of idioms. This finding lends strong support to

the research conducted by Keys (1934). In Keys' study, two types of frequent testing were employed: weekly and monthly. The participants expressed their positive attitudes towards more frequent testing. In a number of studies, the same findings were replicated (Fulkerson & Martin, 1981; Bangert-Drowns et al., 1986; Wilder et al., 2001; Kling et al., 2005). In these studies, students expressed their positive attitudes towards frequent testing and believed that their attitudes changed significantly near the end of the course because frequent testing injected into them a sense of motivation and brought for them the long retention of materials.

Although in this study the researcher did not ask students to give reasons for their answers in the questionnaire, some reasons might be mentioned. They expressed positive attitudes perhaps because quizzes provided them with a sense of motivation and learning (Fulkerson & Martin, 1981; Bangert-Drowns et al., 1986; Kling et al., 2005), and gave them fruitful feedback on their process of learning (Bangert-Drowns et al., 1986; Standlee & Popham, 1960). Another explanation for the significant difference in students' attitudes before and after the course is that students harbor a tendency for procrastination; they do not like to be tested frequently, but after taking the quizzes, they begin to appreciate the positive effect quizzes had on their learning.

As to the fourth research question, although this study confirmed that weekly quizzes have significant effect on the comprehension and production of idioms, the students preferred biweekly quizzes. This preference may be partially due to their procrastination habit, or more logically time limitations. Another reason might be the fact that the students might feel less worried and stressed in this frequency compared with the weekly one. In other words, the learners have come to realize the benefits of frequent quizzes by the end of the course. However, they may still fear that they may not be able to cope with the pressure of weekly quizzes, possibly due to the fact that they may have other assignments in other courses. On the other hand, the reason why the monthly and midterm quiz frequencies did not appear to be popular might be that the students did not think these two frequencies could provide them with enough feedback, motivation, and progress check as the bi-weekly frequency did. To cut the long story short, one can say that the students preferred the more balanced way of frequent testing, not the two sides of a pole.

The findings of the present study may have implications for teachers, learners, and curriculum designers. The knowledge of how various testing frequencies influence students' learning can help teachers make more informed decisions as to how to provide feedback and how to assess learners' comprehension and production of idioms. The wash back effect of frequent testing methods will, in turn, influence students' learning of idioms.

As to students' preferences, one can claim that taking their perceptions into account may result in a pedagogy or curriculum which is geared more towards students' abilities. In other words, the kind of curriculum at work will be more or less student-centered. Moreover, having enough knowledge of students' perceptions and preferences makes the testing process more individualized, thus leading to more efficient learning.

The findings of this study suggest that frequent testing should be treated more fairly in the field and be looked upon as a valuable source for all stakeholders in English language education including policy makers, teachers, students, and parents. Another theoretical promise of this study is that both this study and the current literature on frequent testing suggest that

 ISSN 2251-7324

students highly welcome frequent quizzes because of different reasons they harbour; therefore, it is time to call for a renewal of the role of frequent testing in the field.

References

Abel, B. (2003). English idioms in the first language and second language lexicon: a dual representation approach. *Second Language Research*. 19(4), 329-358.

Adkins, P.G. (1968). Teaching idioms and figures of speech to non-native speakers. *Modern Language Journal*, 52(3), 148-152.

Androu, G. & Galantomas, L. (2008), Designing a conceptual syllabus for teaching metaphors and idioms in a foreign language context. *Portra Linguarum*, 9, 69-77.

Ballard, C. L., & Johnson, M. F. (2004). Basic Math Skills and Performance in an Introductory Economics Class. *Journal of Economic Education*, 35(1), 3-24.

Basol, G., & Johnson, G., (2009). Effectiveness of frequent testing over achievement: A meta analysis study. *International Journal of Human Sciences*, 6(2), 99-120.

Bangert-Drowns, R.L., Kulik, J.A., & Kulik, C.C. (1991). Effects of frequent classroom testing. *Journal of Educational Research*, *85*(2), 89-99.

Beaulieu, R. P., & Frost, B. F. (1989). Impact of examination frequency on achievement. *Journal of Instructional Psychology*, *16*(3), 145-150.

Bortfeld, H. (2003). Comprehending Idioms Cross-Linguistically. Experimental Psychology, 50(3), 212-230.

Carter, R. (1998). *Vocabulary. Applied Linguistics perspectives*. London & New York: Routledge.

Cedar, P. (2008). Learners' Recognition of Thai-English idiom counterparts. *Linguistic Journal*, 3(3), 1-13.

Çelik-Yazici, I. (2004). Idiom processing in L23: Through rose-colored glasses. *The Reading Matrix*, 4(2), 105-116.

Clump, M. A., Bauer, H., & Alex, W. (2003). *To attend or not to attend : Is that a good question*? The Internet TESOL Journal, 8(12), Retrieved March 12, 2010, from http://findarticles.com/p/articles/mi

Concklin, K. & Schmitt, N. (2008) Formulaic sequences: Are they processed more quickly than non-formulaic language by native and nonnative speakers. *Applied Linguistics,* 29(1), 72-89.

Connor-Greene, P. A. (2000). Assessing and Promoting Student Learning: Blurring the Line Between Teaching and Testing. *Teaching of Psychology*, 27(2), 84-88.

Cook, P. Fazly, A., Stevenson, S. (2008). Unsupervised Type and Token Identification of Idiomatic Expressions. *Association for Computational linguistics,* 29(2), 1-44

Cooper, T. C. (1999). Processing of idioms by L2 learners of English. *TESOL quarterly*, 33(2), 233-262.

Cullen, D. M. (2004). *Should Idioms be Taught to Second Language Learners?* Doctoral dissertation. McGill University.

Cutting, J. C. & Block, K. (1997). That's the way the cookie bounces: Syntactic and semantic components of experimentally elicited idiom blends. *Memory and Cognitionو* 25(1), 57-71.

Dineen, P., Taylor, J., & Stephens, L. (1989). The effect of testing frequency upon the achievement of high school mathematics course. *School Science and Mathematics*, *89*(3), 197-200.

Dustin, D. S. (1971). Some effects of exam frequency. *The Psychological Record*, *21*(3), 409-414.

Erman, B. and Warren, B. (2000). The idiom principle and the open choice principle. *Text,* 20(1), 29-62.

Farhady, H., Jafarpur, A., & Birjandi, P. (1994). *Testing language skills: From theory to practice*, Tehran: SAMT Publication.

Fitch, M. L., Drucker, A. J., & Norton, J. R. (1951). Frequent testing as a motivating factor in large lecture classes. *The Journal of Educational Psychology*, *42*(1), 1-20.

Fulkerson, F. F., & Martin, G. (1981). Effects of exam frequency on student performance, evaluations of instructor, and test anxiety. *Teaching of Psychology,* 8(2), 90-93.

Gass, S. (1999). Incidental Vocabulary Learning. SSLA, 21, 319-333 Gibbs, R. (1992). What do idioms really mean? *Journal of Memory and Language*, 31, 485-506.

Gathercole, S. (2006). Complexities and Constraints in non-word repetition and word learning. *Applied Linguistics*, 27(4), 599-613.

Geist, J. R. and Soehren, S. E. (1997). The Effect of Frequent Quizzes on Short- and Long-Term Academic Performance. *Journal of Dental Education,* 61(4), 339-345.

Grant, L.E. (2007). In a manner of speaking: Assessing frequent spoken figurative idioms to assist ESL/EFL teachers, *system*, 35, 169-181.

Grant, L. & Bauer, L. (2004). Criteria for re-defining idioms: Are we barking up the wrong tree? *Applied Linguistics, 25*(1), 38-61.

Grover, C. A., Becker, A. H., & Davis, S. F. (1989). Chapters and units: Frequent versus infrequent testing revisited*. Teaching of Psychology*, 16(4), 192-194.

Haberyan, K. A. (2003). Do Weekly Quizzes Improve Student Performance on General Biology Exams? *American Biology Teacher,* 65(2), 110-114.

Hillert, D. & Swinney, D. (1999). The Processing of Fixed Expressions during Sentence Comprehension. *Conceptual Structure,* Discourse, *and Language*, 4, 1-16.

Ifil, T. (2002). *College Seeking the Nature of Idioms: A Study in Idiomatic Structure*. M. A. thesis. Haverford college.

Irujo, S. (1986a). Don't put your leg in your mouth: Transfer in the acquisition of idioms in a second language. *TESOL Quarterly, 20*(2), 111-128.

Irujo, S. (1986b). A piece of cake: Learning and teaching idioms. *English Language Teaching,* 40, 236-242.

Johnson, B. C. Kiviniemi, M. T. (2009). The effect of online chapter Quizzes on Exam performance in an undergraduate Social psychology course. *Teach psychology*, 36(1), 33-37.

Jones, C.H. (1984). Interaction of absences and grades in a college course. *Journal of Psychology*, 116, 133-136.

Kamuche, F. U. (2005). Do Weekly Quizzes Improve Student Performance? *Academic Exchange Quarterly*, 9, 3, 188-193.

Kamuche, F. U. (2007). The effect of Unannounced Quizzes On Students Performance: Further Evidence. *College Teaching Methods & Styles Journal*. 3(2), 21-25.

Keys, N. (1934). The influence on learning and retention of weekly as opposed to monthly tests. *Journal of Educational Psychology*, *25*, 427–436.

 ISSN 2251-7324

Khalaf, A. S. S., & Hanna, G. S. (1992). The impact of classroom testing frequency on high school students' achievement. *Contemporary Educational Psychology. 17(1),* 71-77.

Kika, F. M., McLaughlin, T. F., & Dixon, J. (1992). Effects of frequent testing of secondary algebra students. *Journal of Educational Research*, 85(3), 159-162.

Kling, N., Miller, C., & Reardon, J. (2005). The impact of testing frequency on student performance in a marketing course. *Journal of Education for Business,* 81(2), 67-72.

Liu, D. (2003). The most frequently used spoken American English idioms: A corpus analysis and its implications. *TESOL Quarterly, 37*(4), 671-700.

Marcell, M. (2008). Effectiveness of regular on-line quizzing in increasing class participation and preparation. *International Journal for the Scholarship of Teaching and Learning, 2(1), 1-9.*

Marshall, B. (2007). A crisis for efficacy? *Education Review, 20(1),* 29-35.

Martin, R. R., & Srikameswaran, K. (1974). Correlation between frequent testing and student performance. *Journal of Chemical Education.* 51(7), 485-486.

Moon, R. (1998): *Fixed Expressions and Idioms in English. A Corpus-Based Approach*, Oxford: OUP.

Nippold, M. A. (1991). Evaluating and enhancing idiom comprehension in language-disordered students. *Language, Speech ,and Hearing Services in Schools*,2 2, 100-106.

Roediger, H. L. III, & Karpicke, J. D. (2006). Test-enhanced learning: taking memory tests improves long-term retention. *Psychological Science.* 17(3), 249-255.

Selakovich, D. (1962). An experiment attempting to determine the effectiveness of frequent testing as an aid to learning in beginning college courses in American government. *The Journal of Educational Research*, 55(4), 178-180.

Simpson, R., & Mendis, D. (2003). A corpus-based study of idioms in academic speech. *TESOL Quarterly, 37*(3), 419-441.

Standlee, L.S., & Popham, W. J. (1960). Quizzes' contribution to learning. *Journal of Educational Psychology*, *51*(6), 322-325.

Turney, A. H. (1931). The effect of frequent short objective tests upon the achievement of college students in educational psychology. *School and Society*, 33(858), 760-762.

Weinstein, S. E., Wu, SH.W. (2009). Readiness Assessment Tests versus Frequent Quizzes: Student Preferences. *International Journal of Teaching and Learning in Higher Education.* 21(2), 181-186.

Wilder, D. A., Flood, W. A. & Stromsnes, W. (2001). The use of random extra credit quizzes to increase student attendance. *Journal of Instructional Psychology*. Retrieved March 7, 2010, From http://findarticles.com/p/articles/mi

Wray, A. (2000). Formulaic sequences in second language teaching: Principle and Practice. *Applied Linguistics*, 21, 463-489

Wray, A. (2002). *Formulaic language and the lexicon*. Cambridge: Cambridge University Press.

Zarei, A. A. (2008). On the Learnability of three categories of Idioms by Iranian EFL learners: *Journal of Humanities of the University of Kerman*. 2 (2), 82-100.

Zgraggen, F. D. (2009). *The Effects of Frequent Testing in the Mathematics' Classroom*. Doctoral dissertation. University of Wisconsin-Stout.

Zyzik, E. (2009). Teaching and Learning Idioms: The Big Picture. *Clear News*. 13(2), 4-8.