



EVALUATING PNU GENERAL ENGLISH TEXTBOOK FROM TWO PERSPECTIVES: SCIENCE VS. HUMANITIES STUDENTS

Fereidoon Vahdany¹

¹English Department, Payame-Noor University, Tehran, Iran

ABSTRACT

This study aimed at evaluating the General English textbook currently in use at Payame Noor University (PNU) of Iran. In fact, it investigated the differences between students majoring in humanities and students majoring in science and engineering in terms of their attitudes and achievement after taking the corresponding course and studying the General English textbook for one semester. The study was conducted in two stages: administering attitude questionnaires, and comparing students' grades at final examination of the general English course. About 150 students took part in the questionnaire including 94 students majoring in humanities and 56 ones majoring in science and engineering. Additionally, the final exam grades of 520 students, 260 from the faculty of science and 260 from the faculty of humanities, were compared in order to find any differences in their achievement in English. The results demonstrated that the textbook falls short of students' expectations. It was also revealed that while students majoring in humanities showed more positive attitudes towards the textbook, students majoring in science and engineering achieved higher grades in the general English final exam at PNU.

© 2015 AESS Publications. All Rights Reserved.

Keywords: EFL, ESL, General English, Textbook evaluation, PNU, Humanities, Science, Engineering.

Contribution/ Originality

This study is one of very few studies which have investigated the use of general English books to serve target students majoring in different courses. Actually, the assumption of the authors of the book is that general English should be the same for students of different academic fields. However, this study reveals that such books may not be appropriate either in terms of difficulty or topics covered for different majors in college.

1. INTRODUCTION

1.1. Background

English language instruction has many important components but textbooks play a fundamental role in ESL/EFL classrooms and programs. As [Richards \(2005\)](#) suggests textbooks are a key component in most language programs. In some situations they serve as the basis for much of the language input learners receive and the language practice that occurs in the classrooms. They may provide the basis for the content of the lessons, the balance of skills taught and the kinds of language practice the students take part in. In other situations, the textbook may serve primarily to supplement the teacher's instruction. For learners, the textbook may provide the major source of contact they have with the language apart from input provided by the teacher. In the case of inexperienced teachers, textbooks may also serve as a form of teacher training; they provide ideas on how to plan and teach lessons as well as formats that teachers can use. Much of the language teaching that occurs throughout the world today could not take place without the extensive use of commercial textbooks. Learning how to use and adapt textbooks is hence an important part of a teacher's professional knowledge.

[Azizifar et al. \(2010\)](#) stated that "Textbooks play a role in actualization of the plans and decisions into interesting and useful materials, tasks and activities" (cited in [Amiryousefi and Ketabi \(2011\)](#)). [Sheldon \(1988\)](#) suggests that textbooks "represent the visible heart of any ELT program" and have more advantages for both the learner and the teacher, when they are being utilized in the ESL/EFL classroom. According to [Sheldon \(1988\)](#) published materials have more validity than teacher-generated or in-house materials. Sometimes teacher-generated or in house-materials are more effective than published materials because textbooks are often too limited or too general. [O'Neill \(1982\)](#) asserted that textbooks are generally, susceptible to student's needs, though they are not planned specifically for them, they are also cheaper and more effective in term of time, and also provide beginner teachers with security, guidance and support. Although some theorists such as [O'Neill \(1982\)](#) and [Williams \(1983\)](#) have implied to inherent danger of the novice teacher who may use a textbook as a pedagogic crutch, such an over-reliance may really have the reverse effect of saving students from a teacher defects.

[Hutchinson and Torres \(1994\)](#) stated that textbooks may have fundamental role in innovation. They believe that textbooks can aid teachers through potentially disturbing and threatening change processes, introduce new and/or untried methodologies, indicate change gradually, and develop scaffolding upon which teachers can construct a more creative methodology of their own.

While many of the aforementioned theorists implied some advantages of using EFL/ESL textbooks, there are many other researchers who have, however, questioned the actual role of textbooks in EFL/ESL classes. Textbooks can also hand false confidence to teachers that they are developed by abstemious people. Other researchers have determined their criticism against the socio-cultural biases embedded in textbooks, such as [Prodomou \(1988\)](#) for instance argue that target language culture as a vehicle for teaching the language in textbooks and suggest that teaching a language is not possible without embedding it in its culture base. [Prodomou \(1988\)](#) is also critical, but emphasizes more on transferring effects of such materials on students, and how they can have disengagement role regarding learning.

Others such as Carrell and Korwitz (1994) and Renner (1997) believe that many EFL/ESL textbooks display gender bias, sexism, and stereotyping. Gray (2000) also has supported the socio-cultural aspects of many textbooks. He advocates the need to recognize the course book's status as cultural artifact, and suggests that critical engagement with cultural content makes both cross-cultural and educational sense. Clearly, there is no consensus on this issue, and some degree of caution should be taken into account regarding using these types of materials in certain teaching and learning contexts.

As far as evaluation of textbook is concerned, there are a few important ideas to review in this section. Mukundan (2007) argues about two types of evaluation: predictive evaluation and retrospective evaluation. He discusses that predictive evaluation is done for the purpose of the selection of books while retrospective evaluation involves persistent evaluation of the textbook after it is selected and while it is used. He states that there are two types of predictive book selection, implicit model and explicit model. The implicit model which is based on teacher impressions and intuitions, also known as the impressionistic model. This model can be considered as an impressive way of evaluating textbooks, particularly if done by experienced teachers. However, the disadvantage of impressionistic evaluation is that team evaluation cannot be carried out this way.

He claims that the explicit model is more efficient alternative to the implicit model, from the advantages of explicit model are that this kind of evaluation is effective institutions where many teachers teach specific level and team evaluation are necessary. He states that explicit evaluation can be done through an evaluation instrument which is normally in the form of a checklist. In general, Mukundan (2007), argues about the importance of retrospective evaluation in countries (like in Malaysia) where book selection is carried out by a central authority within the Ministry of Education. Similarly, Ellis (1997) distinguishes two types of materials evaluation: a predictive evaluation designed to make a decision concerning what materials to use; he continues that there are two basic ways in which teachers can do this kind of evaluation. One is to rely upon evaluations carried out by 'expert' reviewers. Alternatively, teachers can do their own predictive evaluations; there are numerous checklists and guidelines existent to help them. And a retrospective evaluation designed to test materials that have really been used, such an evaluation is suitable for teacher to determine whether it is valuable to use materials again, which activities 'work' and which do not, and how to improve the materials to make them more efficient for future use.

A retrospective evaluation also can be used for 'testing' the validity of a predictive evaluation. Ellis (1997) also argues that retrospective evaluation can be impressionistic or empirical but most teachers prefer to use impressionistic evaluation and empirical evaluation are less usual because they are time-consuming. According to Ellis (1997) there are two main approaches that have a pivotal role for the purpose of selecting, improving and modifying materials to meet the needs of students and instructors in a specific teaching and learning context. These approaches are named as micro approach and macro approach to evaluation of materials. He asserted that one way in which an empirical evaluation can be made more treatable is via micro-evaluation. In a micro evaluation the focus is on the efficiency of the tasks and a macro-evaluation focuses on an general assessment of whether an entire set of materials has worked in relation to needs identified.

English curriculum in Payame Noor University of Iran which has been organized in two courses is a case in point. First, all students regardless of their majors have to take a General English course in which they study a common textbook. After passing this course, students will take an ESP course with different textbooks depending on their corresponding fields of study. The efficiency of a common textbook in the General English course for students majoring in both science and humanities in Payame Noor University has been subject to criticism by some scholars. They argue that students majoring in humanities have a different set of needs, attitudes, and abilities from those majoring in science and engineering, and therefore need to study a separate textbook which takes account of these differences.

Meanwhile, the strengths and weaknesses of the General English course which is supposed to prepare PNU students for a more advanced ESP course are of utmost importance for curriculum developers, university administrators, and tertiary education policy makers. Without a thorough evaluation of the current textbook, the success rate of students in the later ESP course will be at risk. The purpose of this paper is to evaluate the General English textbook currently used at PNU written by [Alimohammadi and Khalili \(2006\)](#) to find out whether it meets the standards of a general English course and the needs of students majoring in science versus humanities. Meanwhile, this research investigates the differences between the attitudes of science and humanities students towards these textbooks and compares their grades after taking part in a General English course in which this textbook has been taught by instructors.

1.2. Research on Textbook Evaluation in Iran

In Iran several projects have been carried out to evaluate textbooks, among which [Mal Amiri \(2008\)](#), [Fathi \(2010\)](#), [Karamouzian \(2010\)](#), [Rahimpour and Hashemi \(2011\)](#), [Hashemnezhad and Maftoon \(2011\)](#), [Nemati \(2009\)](#), [Karamoozian and Riazi \(2008\)](#), and [Zohrabi and Sabouri \(2012\)](#), are typical examples. [Mal Amiri \(2008\)](#) conducted a study in order to evaluate ESP programs at both MS/A and Ph.D. levels at the faculties of Science and Research Campus, three sets of questionnaires were devised to be responded by students, instructors, and heads of departments. The results of the study indicated that the MS/A and Ph.D. students have certain English language needs which are not thoroughly met in the ESP classes they attend and the majority of MS/A students, instructors, and heads of departments evaluated their ESP courses as moderate and unsuccessful. According to results of this study the three groups of respondents complained of problems in their ESP courses and asked for modifications and changes.

[Fathi \(2010\)](#) chose three texts to analyze: an architectural text, a medical text, and an ordinary text. The focus was on architectural text. Two kinds of questionnaires were prepared to be answered by both the students and the teachers about the content of the textbook "English for the Students of Engineering" that is taught at different technical courses in Azad University of Tabriz. The results of analysis of students' responses showed that most of the students had the same idea that the texts in the textbook weren't so useful for them all. They believed that useless materials and shortage of time were the main reasons for their not being successful in learning English. They also stated that the texts in the textbook covered all branches of engineering and that is a disadvantage. However, the results of analysis of teachers' responses indicated that most of the teachers are not

satisfied with the texts of the book. They suggested that the contents of the texts should be richer than this with more relevant technical terms. She came to the conclusion that there is the pressing need for preparing an appropriate and independent textbook for architecture students at the college level in Iran.

In a study conducted by Karamouzian (2010) the checklist method of textbook evaluation was selected for evaluating reading comprehension textbooks. The checklist consisted of five sections: 1) general information, 2) overall impression, 3) organization, 4) content, and 5) overall consideration. To select the textbooks, the books with the highest frequency were chosen from the whole country. The " *Reading through Interaction*" textbooks 1, 2, and 3 edited by Mirhassani and Farhady (2010), and published by Zabankadeh Publications in Iran. The findings of the study suggested that the total average scores-including overall impression, organization, content, and overall considerations are close and the overall quality of all books is convenient. However, the lack of information on grammar and pronunciation should be covered by appropriate supplementary materials. Furthermore, she recommended that the quality of the books can be better improved by adapting those parts that have several weak points.

Rahimpour and Hashemi (2011) evaluated the three English language textbooks currently taught at high schools throughout the country. The findings of this study offered convincing evidence that the English language textbooks that were currently taught at high schools in Iran did not meet the teachers' expectations.

Hashemnezhad and Maftoon (2011) evaluated the book " *English Grammar for College Students 1 and 2* (Azabdaftari, 2003). This book is intended for the higher learning across the country. Both the questionnaire and the checklist were used in this study. The questionnaire that was distributed to twenty EFL lecturers included five different aspects of subject matter, vocabulary and structure, exercises, illustrations, and physical makeup of the book and checklist included three cognitive, affective, and psychomotor domains. The findings showed that there is only correlation between subject matter and physical makeup of the book and there is no correlation between the other aspects. According to the results of the study the importance of affective and psychomotor domains were ignored in the book and the emphasis is on cognitive domain. They finally concluded that in order to prepare learners to a fully-fledge acquisition of language, the writer of the book as well as the syllabus designers should try to make full use of all three domains : cognitive, affective, and psychomotor.

Nemati (2009) conducted a study in which she evaluated second Pre-University English book under the title of 'enable' from two different aspects and two different steps: first, evaluation of the book against the available and common criterion, and second, evaluation of vocabularies of the book systematically. The results of data from questionnaire revealed that the book is acceptable from the viewpoint of the majority of English teachers who are familiar with the book. However, the findings of vocabulary analysis showed that it is better to change and do some modification in the sequence of presenting the text so the texts with more frequent vocabularies and as a result easier ones come first and texts with rare and much more difficult vocabularies occurs texts that appear to the end of the book.

Karamoozian and Riazi (2008) conducted a study in which they tried to explore several well-known checklists in terms of their content and specifically their practicality. The checklists were examined in terms of their formats, scope, applied terms and concepts, weighting/rating systems, guidelines, and whether they were piloted by the developers. The results of this review revealed that although the reviewed checklists have several strong points specifically regarding their format and scope, they mostly fail in terms of other features that lead to practicality. Furthermore, the available checklists are mostly designed to evaluate general English textbooks while they are not generalizable enough to be adapted to evaluate other English language textbooks.

Zohrabi and Sabouri (2012) evaluated the merits and demerits of Iranian first year high school English textbook. In this study, data were obtained by quantitative (questionnaire) and qualitative methods (interviews). The findings of this study showed that English book I is structure-based and it cannot meet the curriculum goals and students' needs. They concluded that in this book only reading skill is emphasized and communicative role of the language is ignored. Meanwhile, there is no focus on the culture of foreign language countries. Amiryousefi and Ketabi (2011) examined the validity of anti-textbook views in EAP courses. Results of their study indicate that current textbooks prepared for EAP classes are not very satisfactory and fail to meet students' needs and interests. This study, therefore, supports the weak anti-textbook view presented by Harwood (2005) claiming that textbooks can help both teachers and students but they should contain what is needed. It is, therefore, suggested that a need analysis should be carried out before any attempts are made to develop textbooks, and teachers should become sensitive to their students' needs and interests.

1.3. Research Questions

This study aims at evaluating the General English textbook prepared for students majoring in science and humanities at PNU, and to see whether this textbook meets the needs and demands of the two faculty groups. It seeks to find answers to the following questions:

1. Is there any difference between science vs. humanities students in terms of their attitude towards the appropriateness of the General English textbook used at PNU?
2. Is there any difference between science vs. humanities students in terms of their achievement in the general English course at PNU?

2. METHODOLOGY

2.1. The General Design of the Study

The purpose of this research was evaluating the "General English" textbook written by Alimohammadi and Khalili (2006), and investigating its efficiency for students majoring in humanities compared with students of science at PNU. Based on two research questions presented in chapter one, the process of evaluating the General English textbook was designed at two levels. A questionnaire was prepared to elicit the views of a representative sample of PNU students about different aspects of the textbook. Students were presented with a self-administered questionnaire in which they were asked to answer questions related to characteristics of textbook they studied. Secondly, the final exam scores of two groups of PNU students majoring in humanities and science

were compared in order to discover which group achieved higher proficiency after studying the shared textbook for one semester. In fact, the second question intends to find out if the book in focus is appropriate in terms of its difficulty level for the two groups.

2.2. Participants

For the questionnaire part of the study 150 students (94 students from faculty of humanities and 56 ones from faculty of science) who were studying General English at Payame Noor University of Rasht took part in the investigation and shared their views on different items on the questionnaire. They were between the ages of 18 to 30 years old and were selected randomly. This sample included both male and female students without any intentional control on gender factor.

With regard to the second research question, the end-of-term General English scores of 520 students out of which 260 ones were selected from the faculty of science and 260 ones from the faculty of humanities were collected in order to undergo inferential statistics. The participants were selected based on convenience sampling strategy from Payame Noor University of Rasht, and were within the age group of 18-32 years and the sample included both gender group.

2.3. Research Instruments and Data Collection Procedures

Two data collection instruments were chosen in this study in order to evaluate The General English textbook at PNU:

2.3.1. Questionnaire (Appendix I)

The items on the questionnaire explored students' attitudes with regard to four major aspects of the textbook:

a) The Language Content

This category consisted of 11 items (number= 2, 3, 4, 5, 6, 7, 8, 9, 11, 15, 21) asking students about the grammar, vocabulary and language components which were covered in the textbook.

b) Students' Needs

This category included 7 items (number=1, 12, 16, 22, 24, 26, 27) which focused on the learner needs and expectations.

c) Study Skills

This part consisted of 5 items (number =13, 14, 17, 23, 25) focusing on the incorporation of materials for independent learning, learning techniques and study skills promoted by the textbook.

d) Content Variety

This section incorporated four items (number=10, 18, 19, 20) which focused on the variety of the contents and subjects and the extent to which the textbook' contents are up-to-date.

2.3.2. PNU Final Exam of the General English Course

Actually this test is end-of-the term exam from the bank of question in the PNU. This test includes 30 multiple choice items covering grammar, reading and vocabulary and all the items are based on the book. The fact that the test was constructed by university professors there and had been reviewed strictly and was administered under uniform condition makes it dependable measure of achievement. However, the reliability and validity of the test cannot be highly guaranteed.

2.4. Data Analysis

After the data collection procedures were completed, the collected data was coded and analyzed using inferential statistics in order to discern any possible difference between students majoring in humanities and science and engineering. With regard to the questionnaire part of this study, which was meant to compare the attitudes of the two groups towards the General English textbook, a Mann-Whitney U test was run to find out whether there is a significant difference between the two groups of students. The reason for this choice of inferential statistics was that the data gathered through the questionnaire were of ordinal type, and the two groups of raters were independent. Therefore, non-parametric statistics were selected for data analysis. SPSS software was used for running the test. A decision level of $\alpha=0.05$ was chosen to investigate the significance of the difference. All of the data collected through questionnaires were precisely checked in terms of any irregularity or atypical behavior, and no single response was spotted as outlier within the questionnaires rated by the participants. Therefore, all participants' responses were included in the final Mann-Whitney U test performed by SPSS software.

With regard to the analysis of the final exam part, an independent samples t-test was run to the scores of the both humanities and science students by using SPSS software to answer the third research question finding the possible difference between the two groups in terms of their achievement after studying the General English textbook for one semester at Payame Noor University. Again, all the scores were checked for possible irregularities and all the students' scores were included for the final t-test analysis.

3. RESULTS AND DISCUSSION

3.1. Questionnaire Results

After conducting the questionnaire part of the study with the participation of humanities and science students at Payame Noor University of Rasht, the results were properly coded and entered into a spread sheet in the SPSS 16.0 software. As mentioned before, a Mann-Whitney U test was run using SPSS in order to investigate the significance of any difference between the views of the two groups of students with regard to the effectiveness of their common General English textbook. All the data gathered through questionnaires were included in the final analysis because no single data showed any irregularity or atypical behavior.

As Table 3.1 demonstrates, students from the faculty of humanities at PNU had more positive attitudes towards their textbook than those from the faculty of science and engineering. According to the results, students majoring in humanities perceived the General English textbook as more effective ($M=84.53$) in all categories of language content, students' needs, study skills, and content

validity than their counterparts majoring in science and engineering ($M=60.35$). The difference between the mean ranks is relatively high ($=24.18$) which enables us to predict a significant difference between the two groups even before calculating Mann-Whitney statistics (figure 3.1). This is especially the case when the sum of the ranks for students of humanities ($S=7945.50$) is almost twice as large as the sum of the ranks for students of science and engineering ($S=3379.50$). However, we cannot jump into any interpretation before taking into account the significance level of these findings.

There is also another note of caution regarding the interpretation of the ranks table here: Mann-Whitney U tests are based on ranks of the original values and not on the values themselves, that is to say, each case is ranked without regard to group membership and after ranking the cases, the ranks are summed within groups. This fact must also be taken into account in order to avoid any misinterpretation of the results.

Table-3.1. Summary of students' ranks regarding attitude towards General English textbook

	Faculty	N	Mean Rank	Sum of Ranks
Evaluation of Textbook	Humanities	94	84.53	7945.50
	Science	56	60.35	3379.50
	Total	150		

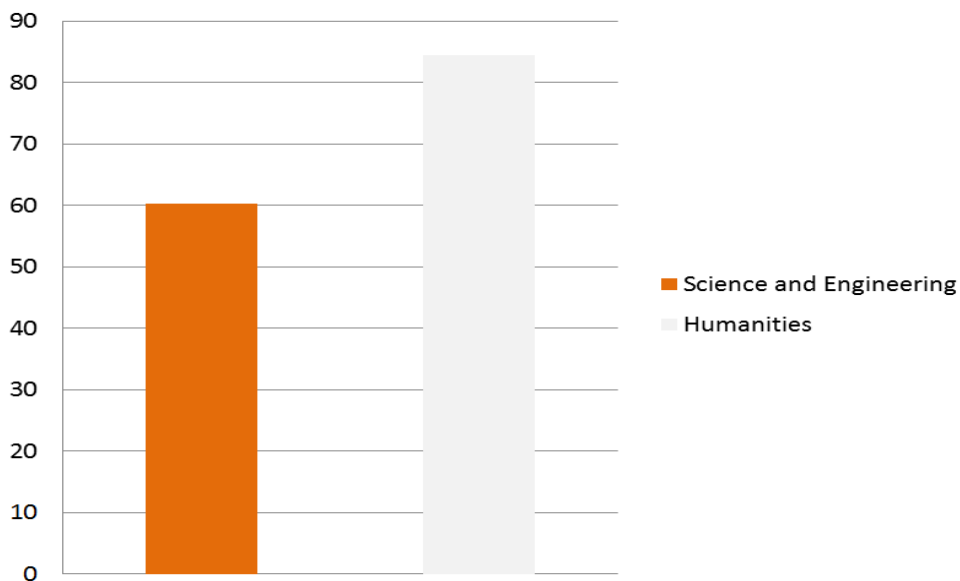


Figure-3.1. Mean rank of students regarding attitude towards General English textbook

Table 3.2. shows the Mann-Whitney U test statistics of the same data which was gathered through questionnaire from the two group of students. Because the test variables assumed to be ordinal, the Mann-Whitney testis based on ranks of the original values and not on the values themselves. Mann-Whitney U test reveals that with a decision level of $p=0.05$, the difference between the attitudes of students majoring in humanities and those of students majoring in science and engineering was significant (Asymp. Sig. $=0.001$). This means we can conclude with 95%

confidence that as far as all four categories of language content, students' needs, study skills, and content validity are concerned, humanities students perceived the General English textbook as more effective compared to students of science and engineering.

Table-3.2. Mann-Whitney U test statistics of students' ranks regarding General English textbook at PNU

	Evaluation of Textbook
Mann-Whitney U	1783.500
Wilcoxon W	3379.500
Z	-3.298
Asymp. Sig. (2-tailed)	.001

3.2. Final Exam Results

One of the most common ways to evaluate a textbook is to look at the performance of the students in the final exams. Although students' grades cannot be considered as the only source of judgment towards the value of a textbook, it can definitely provide the instructor with a picture of the textbook's efficiency. This is especially the case if the evaluation is used for a comparison between the achievements of two groups after studying a textbook in class.

Table 3.3 demonstrates the descriptive statistics of the final-exam performance of both groups of students majoring in humanities and science respectively. The science group gained a higher mean (M=12.6) than the humanities group (M=10.8) at the General English final exam. On the other hand, the relatively lower standard deviation of 2.6 for students of science and engineering compared to that of 2.9 for humanities students, it can suggest that their grades were more homogeneously distributed around the mean than humanities students.

Table-3.3. Descriptive statistics of students' grades in General English final exam at PNU

faculty	N	Mean	Std. Deviation	Std. Error Mean
Science	260	12.6435	2.62922	.16306
Humanities	260	10.8504	2.96507	.18389

In order to scientifically approve the significance of the difference between means, an independent samples t-test was run the results of which are shown in table 4.4. According to this table, the significance value of Levene's test is 0.112 which is greater than 0.1, and therefore, one of the assumptions of using a t-test has been met here. The t column displays the observed t statistic for each sample, calculated as the ratio of the difference between sample means divided by the standard error of the difference. The degree of freedom for this test is 518.

Table-3.4. T-test statistics of students' grades in General English final exam at PNU

Levene's Test for Equality of Variances		t-test for Equality of Means						
F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
							Lower	Upper
2.533	.112	7.296	518	0.000000000001	1.79308	.24577	1.31025	2.27590

If we look at the significance of the observed difference between the two group (Sig. 2-tailed=0.000000000001), we can be 95% sure that students majoring in science and engineering performed better in the final exam compared to students majoring in humanities. In other words, with a decision level of $p=0.05$, there is only 5% chance that the observed difference between the mean performance of the two group of students is purely random.

The 95% Confidence Interval of the Difference provides an estimate of the boundaries between which the true mean difference lies in 95% of all possible random samples of 520 students. Since the significance value of the test is less than 0.05, we can safely conclude that the average of 1.79 more scores gained by science students studying the General English textbook for one semester compared to humanities students is not due to chance alone.

3.3. Discussion of Questionnaire and Final Exam Results

According to the findings of the questionnaire and after running a Man-Whitney U test, it can be inferred that the difference between the attitudes of humanities students and science and engineering students is significant (Asymp. Sig. =0.001). This means that in all categories of language content, students' needs, study skills, content variety which were included in the questionnaire, students majoring in humanities regarded the General English textbook as more effective and more in line with their needs compared to students majoring in science and engineering. This finding can have several explanations. The most likely one is that humanities students might generally have a different set of language needs and may perceive English not a major priority for their future studies. However, students majoring in science and engineering view learning English as a skill which might improve the prospect of their future studies and careers and therefore are more sensitive to the quality of the General English textbooks. This hypothesis is further proved when we take into account the number of science and engineering students seeking academic degrees abroad which has been relatively higher than their counterparts majoring in humanities and arts. Such assumptions, however, should be confirmed by comprehensive investigation of the needs, motivation, and priorities of the two groups with regard to their future careers and studies.

Another possible explanation would be the background English proficiency and academic capacity of the students. It is believed that most students who graduate from high schools in Iran with higher academic achievement prefer to pursue their education in science and engineering. These students might have better aptitude in learning English and have experienced learning

English at an earlier age. Consequently, science students are more likely to compare the quality of their textbook with those English textbooks which they have studied in those language institutes and rate the general English textbook at PNU as outdated or less fulfilling and effective compared to students majoring in humanities.

The findings also suggest that both groups were not completely satisfied with the content of the General English textbook.

As far as the diversity of topics and material covered in the textbook is concerned, neither groups found the variety of topics not quite suitable to their needs and expectations and more than one third of the Literature students and almost one quarter of the science students maintained that the choice of topics in the General English textbook was relatively poor.

Another striking finding was that almost half of each group believed that the General English textbook did not stimulate creativity and learner autonomy. Additionally, about half of both humanities and science students approved the contention that the General English textbook did not include different study techniques and around one third of them believed that the textbook did not promote autonomous learning. A note of caution is worth mentioning here: due to the nature of educational practices at PNU in which class attendance is not obligatory and students are supposed to rely on their textbooks for learning and passing the exams, and ignore the specific characteristics and limitations which a general English course usually has. Therefore, we cannot rush into any generalization regarding the extent to which the textbook promotes autonomy and learner independence from the learners' point of view. As far as the final exam results are concerned, several possible explanations for the significant superiority of science students in comparison to humanities students might be possible to discern. One theory could be that students majoring in science and engineering performed better in the final exam because the General English text book has better met their needs and expectations from a language textbook. However, by looking at the questionnaire result, this reason seems not to be true, since students majoring in humanities expressed more positive attitudes towards the textbook than their counterparts in science and engineering fields of study. Although the questionnaire results were used to reject this theory, we need to be cautious in our arguments, and as mentioned before, textbook evaluations must be supplemented by a retrospective study to explore teachers and students' views on the efficiency of the book in meeting the needs and to satisfying the set objectives.

It is also possible to conclude that science and engineering students generally have a higher language aptitude and therefore performed better than their counterparts in the final exam. This fact can be further approved if we take into account the tendency of most students with higher grades at high school to continue their studies in science and engineering. If this statement is true, then we cannot necessarily conclude that science students reached higher levels of achievement merely because the General English textbook is more suitable for meeting their needs and expectations. Nevertheless, any judgment regarding the language aptitude of both groups of students requires further research, and it is not reasonable to identify language aptitude with success or failure in other areas of knowledge.

Another possible explanation of the significant supremacy of science and engineering students might be their previous experience in learning English. It might be the case that due to the seeming

vital role of learning English for these students' future careers, they are more likely to have already taken part in language classes and therefore, have already gained a better command of English than students majoring in humanities. If this contention is true, the observed higher scores of science students might not have anything to do with the quality of the General English textbook. While such an argument seems quite compelling, it is by no means reasonable to say science students have such a higher English background than students majoring in humanities and any generalization in this regard needs a thorough investigation of language proficiency of the two groups of students at the freshman year.

Due to the multiple-choice format of final exams in PNU, the degree of familiarity of students with this format might also have affected the results. One might argue that science and engineering students in Iran spend more time and energy preparing themselves for the university examinations which is also held in multiple-choice format than humanities students, and therefore will probably perform better on such exams. Although such conclusion seems intuitively true, it has not been scientifically approved that science students perform better on multiple choice examinations and we have to postpone this contention after conducting further research and investigations.

One relatively important point that we need to bear in mind with regard to interpreting the final exam results of the general English course is that we should not mistake students' final scores on a multiple choice test with their language proficiency. This test does not provide any information about students' communicative language abilities, listening and speaking abilities. Therefore, although humanities students did not perform as good as science students in this multiple choice test, their speaking and listening skills might not be so different from them. It is highly advisable to not only evaluate the effectiveness of the General English at PNU, but also the efficiency of the final exam format for the course in order to find out to what degree the current textbook promotes communicative language learning in the class and to what extent the exam measures language skills such as listening and speaking.

4. CONCLUSION

This study was conducted in two phases. The results of the questionnaire and final exam grades both demonstrated considerable differences between students majoring in science and humanities. As far as students' attitudes are concerned, humanities students expressed more favorable attitudes towards the General English textbook and found it more effective in improving their English proficiency as well as their study skills. It was argued that this finding might be due to the compatibility of the textbook objectives with the humanities students' needs and expectations. However, this more positive attitude might be related to their language aptitude, background knowledge and different motivational variables in comparison to students majoring in science and engineering who are more likely to continue their studies in English speaking countries. Further investigation of both groups' motivational factors might provide a better picture on the issue.

Students majoring in science and engineering, on the other hand, gained better grades in the final exam of the General English textbook compared to humanities students. It was argued that this finding demonstrates that these students achieved higher levels of English proficiency after studying the textbook. Nevertheless, we should not rush into any conclusion without taking into

account these students' background knowledge, language aptitude, and their experience in multiple-choice examinations.

Overall, evaluation of a textbook is by no means a simple undertaking and conducting one study in order to investigate the effectiveness of a textbook is definitely not sufficient. This is especially the case in evaluating language textbooks at universities where students begin their course with varying needs, expectations, and levels of linguistic knowledge. Therefore, interpreting the results of such evaluations is not an easy task and requires viewing the issue from multiple perspectives. It also involves investigating students' motivational and personal variables which might severely affect their achievement in General English courses at PNU.

REFERENCES

- Alimohammadi, M. and H. Khalili, 2006. General English. Tehran: Payame Noor University.
- Amiryousefi, M. and S. Ketabi, 2011. Anti-textbook arguments revisited: A case study from Iran. *Procedia Social and Behavioral Science*, 15(2): 215-220.
- Azabdaftari, B., 2003. English language grammar for college students. Tehran: SAMT.
- Azizifar, A., M. Koosha and A.R. Lotfi, 2010. An analytical evaluation of locally produced Iranian high school ELT textbooks from 1970 to the present. *English Language Teaching*, 3(4): 132-141.
- Carrell, D. and J. Korwitz, 1994. Using concordancing techniques to study gender stereotyping in ELT textbooks. In J. Sunderland (Eds). *Exploring gender: Questions and implications for english language education*. Prentice Hall International.
- Ellis, R., 1997. The empirical evaluation of language teaching materials. *ELT Journal*, 51(1): 36-42.
- Fathi, V., 2010. The importance of materials design in ESP and EST. Available from <http://www.faculty.ksu.edu> [Accessed Oct 29, 2012].
- Gray, J., 2000. The ELT course book as cultural artifact: How teachers censor and adapt. *ELT Journal*, 54(3): 274-283.
- Harwood, N., 2005. What do we want EAP teaching materials for? *Journal of English for Academic Purposes*, 4: 149-161.
- Hashemnezhad, H. and P. Maftoon, 2011. An evaluation of english language grammar for college students 1&2: An EAP course book evaluation. Available from <http://www.ccsenet.org> [Accessed Oct29, 2012].
- Hutchinson and Torres, 1994. The textbook as agent of change. *ELT Journal*, 48(4): 315-328.
- Karamoozian, F.M. and A. Riazi, 2008. Development of a new checklist for evaluating reading comprehension textbooks. *ESP World*, 7(3): 19-39.
- Karamoozian, F.T., 2010. A post-use evaluation of current reading comprehension textbooks used in TEFL programs. *The Iranian EFL Journal*, 6(4): 24-62.
- Mal Amiri, B., 2008. A program evaluation of ESP education at MS/A and Ph.D. Levels at science and research campus, Islamic Azad University. *Human Sciences*, 56(4): 1-28.
- Mirhassani, S.A. and H. Farhady, 2010. *Reading through interaction*. Tehran: Zabankade Publication.
- Mukundan, J., 2007. Evaluation of english language textbooks: Some important issues for consideration. *Journal of NELTA*, 12(1&2): 80-84.

- Nemati, A., 2009. Evaluation of an ESL english course book: A step toward systematic vocabulary evaluation. *Journal of Social Sciences*, 20(2): 91-99.
- O'Neill, R., 1982. Why use textbooks. *ELT Journal*, 36(2): 104-111.
- Prodomou, L., 1988. English as cultural action. *ELT Journal*, 42(2): 15-25.
- Rahimpour, M. and R. Hashemi, 2011. Textbook selection and evaluation in EFL context. *World Journal of Education*, 1(2): 62-68.
- Renner, C., 1997. Women are busy, tall, and beautiful: Looking at sexism in EFL materials. From: Annual Meeting of the Teachers of English to Speakers of Other Languages. Orlando FL. pp: 11-15.
- Richards, J.C., 2005. The role of textbooks in a language program. Available from <http://www.professorjackrichards.com> [Accessed 29 November 2011].
- Sheldon, L., 1988. Evaluating ELT textbooks and materials. *ELT Journal*, 42(4): 237-246.
- Williams, D., 1983. Developing criteria for textbook evaluation. *ELT Journal*, 37(3): 251-255.
- Zohrabi, M. and H. Sabouri, 2012. An assessment of strengths and weaknesses of Iranian first year high school english course book using evaluation checklist. *English Language and Literature Studies*, 2(2): 213-222.

APPENDIX-1. Questionnaire Textbook Evaluation Checklist (Sheldon (1988) model for textbook evaluation)

	Rating			
	1=very low	2=to some extent	3= moderate	4=very high
1) What grammar items are included? Do they correspond to students' needs?				
2) To what extent is the number of grammatical points appropriate and how appropriate is their sequence?				
3) To what extent is there an even distribution of grammatical and vocabulary material among the chapters?				
4) To what extent are new structures controlled to be presented and explained before they appear in drills, dialogs, or reading material?				
5) To what extent does the presentation, practice, and recycling of new linguistic items to be appropriate for the level of language mastery of the students?				
6) Is there any principled basis for selection of vocabulary?				
7) Are learners sensitized to the structure of the lexicon through vocabulary learning exercises based on semantic relationships, formal relationships, or collocations?				
8) To what extent the number of new words introduced every lesson seems to be reasonable for the students of that level?				
9) To what extent is vocabulary introduced in appropriate contexts?				
				<i>Continue</i>

10) To what extent does the presentation of vocabulary and structure move gradually from the simple to the more complex?				
11) To what extent are new vocabulary and structures recycled in subsequent lessons for reinforcement?				
12) Does the course book use authentic reading material at an appropriate level?				
13) Is the book sensitive to what students need in order to learn well?				
14) To what extent do the aims and objectives of the text correspond to the needs and goals of the students?				
15) Are learners encouraged to use language creatively?				
16) Is there any reflection on study techniques?				
17) Are students encouraged to take some degree of responsibility for their learning?				
18) Are there any materials for independent work?				
19) To what extent do the exercises promote meaningful communication by referring to realistic activities and situations?				
20) To what extent does the subject matter cover a variety of topics suitable to the interests of the intended audience, as determined by age, sex, environment and cultural orientation?				
21) To what extent does the sentence length and syntactic complexity seem reasonable for the students of that level?				
22) To what extent does the presentation, practice, and recycling of new linguistic items to be appropriate for the level of language mastery of the students?				
23) To what extent does the text's level of difficulty match that of the class?				
24) To what extent are visual approaches such as picture, table, figure, and etc. used for optimized transferring of the meaning of the words?				
25) To what extent do the syntactic exercises cover a variety of tests (completion, open-ended, multiple-choice, true or false)?				
26) To what extent are the word explanations at the beginning of each lesson meaningful for students?				
27) To what extent are the synonyms at the beginning of each lesson helpful for vocabulary learning?				
28) To what extent is the appropriate relationship between syntactic section and reading quantitatively?				
29) To what extent are reading texts appropriate for a general English textbook?				
30) To what extent is it provided for students to have contact with the book authors through communicational channels such as fax, e-mail, site, cell phone, postbox and etc.?				

Views and opinions expressed in this article are the views and opinions of the authors, International Journal of Asian Social Science shall not be responsible or answerable for any loss, damage or liability etc. caused in relation to/arising out of the use of the content.