The Effect of Washback on Reading Comprehension Achievement of Medical Students in English for Specific Purposes Classes

Background: Testing and teaching are intertwined in education process; using various kinds of tests, teachers are able to anticipate the strong and weak points of learning in students, their progress and their accomplishment. The influence of test on teaching and learning is commonly referred to as washback. Testing washback is a twisted concept that becomes even more complex under a variety of interpretations of the washback phenomenon on teaching and learning. This study aimed to describe the effects of washback on behaviors of students, in the low stakes testing in English for Specific Purpose (ESP) environment in two-fold.

Methods: The effects of different formative tests on the medical students’ English reading comprehension were assessed using Michigan test. The effects of washback on the students’ attitudes toward English reading comprehension were measured by the English Reading Attitudes Questionnaire (ERAQ). Data were analyzed by paired t-test, t-test, and Mann-Whitney U test.

Results: Formative tests did not significantly affect students’ reading comprehension achievement in an ESP environment with (Mean = 7.17 ± 1.23 for experimental group vs. 6.59 ± 1.03 for control group; p > 0.05). But they showed significant effects on students’ English reading attitudes (Mean = 63.25 ± 7.6 for experimental group vs. 54.13 ± 8.7 for control group; p < 0.001).

Conclusions: Findings from this study indicate that although formative tests do not significantly affect students’ reading comprehension achievement in English for Specific Purpose environment, they have significant effects on students’ English reading attitudes.

Keywords: Washback; Formative Test; Reading Attitude; English for Specific Purpose.
INTRODUCTION

Testing has penetrated in all aspects of our life. It also influences on all grades of education, i.e. elementary, secondary, high school and university. Wherever there is teaching, there is also testing. On the other hand, the relationship between teaching and testing has long been concerned with both educational and applied linguistics literature. The fact that tests attract classroom teaching and the syllabus and make the teachers teach for examinations, and tempt students to do the exercises which are necessary to pass the examination is usual. Tests are generally used to make, among other things, inferences about examinees’ abilities, predictions about their future success, and decisions (such as employment, placement, selection, etc.) about the test-takers. The uses made of test results thus imply values and have an impact on, the society and the educational system in general, and the students in particular(1).

In applied linguistics research, this influence of testing on teaching and learning has been referred to as ‘washback’. The following definitions of ‘washback’ represent its meaning as used in the paper. Buck(2) defined it as the influence of the test on the classroom. This washback effect can be either beneficial or harmful. Messick(3) declares the extent to which the test acts upon language teachers and learners to do things that they would not necessarily otherwise do, and Bailey(4) defined it as the influence of testing on teaching and learning.

The test becomes a part of educational process in which changes in the educational system take place based on feedback obtained from the test. Public examinations that concern to high stakes influence more learners and teachers than class examinations that concern to low stakes, and consequently public examinations will have more difficult statistical analysis and interpretation than class examinations that are limited to students of one or two classes. However, the results of low stakes examinations are more explicit; they represent the students’ learning condition, their strong and weak points.

It has long been believed that tests directly affect the educational processes in different ways. One assumption is that teachers will be influenced by the knowledge that their students need to take a certain test and will adapt their instruction and lesson content to cover the test. By test washback teachers learn more about students learning and their own teaching. Washback has been defined differently by different researchers, based on different aspects they study. Some researchers focused on learner and learning; some studied its effects on teacher and teaching, material, curriculum, school and even educational organization and society.

Buck(2) in his definition emphasizes the impact of a test on what teachers and students do in classrooms. One of the most common definitions referred to as the influence of testing on teaching and learning (e.g., Alderson & Wall(5); Cheng & Curtis(6)). Similarly, Brown(7) refers to washback as the effect of test on teaching and learning. Shohamy(8), focuses on washback in terms of language learners as test-takers when she explains the use of external language tests to influence foreign language learning in the school context. In another article, Shohamy, Donitza-Schmidt, and Ferman(9), define washback as the connections between testing and learning. Bailey(4) also takes a broad view. He describes washback in terms of how tests influence educational practices and beliefs. More recently, Bachman and Palmer(1) define washback as a subset of a test’s impact on society, educational systems, and individuals. They mention that test impact operates at two different levels: the micro level (i.e., the effect of the test on students and teachers) and the macro level (the effect on society and its educational systems). Cheng(10) also conducted research in the context of secondary school exams in Hong Kong. She explains washback as “an active direction and function of intended curriculum change by means of the change of public examinations, as is the case under study in Hong Kong”(10). Valette in (11) stated that “washback occurs when it is the testing instrument rather than the statement of desired learner outcomes that determines the nature of the curriculum and the course of instruction”; and Saville in (12) defines it as the influence of testing on teaching and learning. Also Alderson and Wall(5) claimed that “the extent to which the test influences language teachers and learners to do things they would not otherwise necessarily do” (p. 117). Messick(3), noting that washback can have either harmful or positive effects, expands it as: We emphasize first the need to establish valid evidential grounds for trustworthy inferences about tested language proficiency to provide a basis for distinguishing test-linked positive washback from good teaching regardless of the quality of the test and negative washback from poor teaching. (3)

It is worth mentioning that in recent studies, washback encounters a broader definition, taking the view that tests can have more far-reaching influences within the field of education. For example, Andrews (13), explores the relationship between washback and curricular innovation, beyond the classroom, and depict the effects of tests on teaching, learning, the educational process, and different stakeholders in the educational system. He uses washback to refer to: “the effects of tests on teaching and learning, the educational system, and the various stakeholders in the education process” (13).

The general objective of the present study was to describe the washback effect of the English Language Tests on reading comprehension in an ESP environment for medical students. In fact, this study had three main objectives:

1. to investigate the effect of different formative tests on reading comprehension in English for Specific Purpose (ESP) environment;
2. to investigate the effect of washback on attitudes of learners; and
3. To suggest areas for further research on washback.

METHODS

Participants

The participants of this study were four classes consisting of 130 students in total, 64 students in experimental group and 66 students in control group that had enrolled in the
To know whether these formative tests had any effect on experimental students’ attitudes, another instrument was used. The quantitative data of this part of the study was collected via the English Reading Attitude Questionnaire (ERAQ) developed by Hung and Chiang (12) a master thesis to explore participants’ attitudes toward EFL reading. Students received this questionnaire one month after their post-test. It consisted of twenty-one items all were closed questions designed using four-point Likert scale for easier analysis of the results. Its points start from Strongly Agree to Strongly Disagree as follows:

- SA = 4, Strongly Agree;
- A = 3, Agree;
- D = 2, Disagree;
- SD = 1, Strongly Disagree.

The questionnaire was anonymous in order to encourage honest responses and the students were informed of this fact before it was distributed, it was also mentioned that it has no effect on their scores. Informed consent was taken from participants. Data were analyzed by paired t-test, t-test, and Mann-Whitney U test.

The research project is registered in Clinical Trials.gov (Code NCT016423).

**RESULTS**

Participants aged from 18 to 21 years old, the average age was 19. The first section of study deals with the students’ performance on pre- and post-tests. And the second section reports the findings of the students’ attitudes towards reading in English as measured by the English Reading Attitudes Questionnaire (ERAQ).

**Within Group Pre- vs. Post-Test Analysis**

The experimental group’ Michigan test score increased from 62.7±9.39 (Mean±SD) in the pre-test to 71.76±12.31 (Mean±SD) in the post-test (t = -9.6) and the critical value of t (t_{crit} = 2.045).

The control group took both the pre- and post-tests as the experimental group did. The experimental group’ Michigan test score increased from 63.37±9.15 (Mean±SD) in the pre-test to 67.58±10.28 (Mean±SD) in the post-test (t = -8.86) and the critical value of t (t_{crit} = 2.048).

**Between Groups Analysis**

The two-tailed test of significance for pre-test scores by the experimental and control groups indicates that there was no significant difference in pre-test scores between experimental and control groups (p>0.05). Therefore, it is assumed that at the beginning of the study two groups must have come from the same population (Mean Difference = -0.67, Std. Error Difference= 2.41, t(57) = -0.281, Sig. 2-tailed=0.78, p > 0.05).

As for the post-tests there was no significant difference between experimental and control groups as well (p>0.05). Therefore, it is assumed that improvement of reading comprehension in both experimental and control groups is similar. In other word, the post-tests in experimental and control groups must come from the same population because no significant differences exist (Mean Difference = 4.18, Std. Error Difference = 2.95, t(57) = 1.41, Sig. 2-tailed=0.16 p > 0.05).

**Questionnaire Findings**

The results of twenty one items of English Reading Attitude
The Effect of Washback on Reading Comprehension

Results of this study show that there is a significant difference between pre-test and post-test scores obtained by the students in both experimental and control group. There was no significant difference between the experimental and control group at the beginning of the study. There was no significant difference in post-test scores between experimental and control groups as well. The two-tailed test of significance for post-test in experimental group and control group is more than .05. To put it in other words, there was no significant difference between the two groups, and it means that washback of formative tests did not affect reading comprehension in experimental group significantly. Therefore, the first null hypothesis was accepted.

To find the influence of washback on attitudes towards reading comprehension in English, students in the control and experimental groups received a questionnaire at the end of the course. Analyzing the results of students' questionnaire scores display that with t (57) = 4.296; p < .05, there was significant difference between students' attitudes in two different groups. So the null hypothesis was rejected, and its alternative hypothesis was accepted.

In terms of pedagogical implications, the findings of this study suggest that, first, formative tests influence teachers' and learners' behavior, and help them to find their weak points, and to improve their understanding of unknown parts of the course. These tests may not improve their reading comprehension significantly, but they have significant effects on students' attitudes. Second, formative tests affect on the way and the content of teaching. By knowing the weak areas of students' language knowledge, the focus of teaching will be changed to those areas, and the teacher spends more time on them.

In terms of research implications, investigation of the data reveals that it is better to extend the categories of study and analysis beyond the scope of tests in class. The researcher should take the advantage of using different teachers in his or her study, so the results will be more convenient.

Although large strides have been made in identifying factors that might influence washback since Alderson and Wall's (5) call for more empirical research, only recently have studies begun pinpointing instances that may have had a direct impact on the amount of washback observed in specific situations, which is necessary if generalizations are ever to be made. Therefore, more case studies identifying factors affecting washback need to be carried out. Most of the washback research is confined to teachers, but all teachers and researchers believe in the fact that learners are the main determination of any learning process.

Future research should gather evidence from different populations so that the phenomenon of washback can be understood more thoroughly.

Replication of the study on different subjects and populations in order to increase generalizability is also recommended.

ACKNOWLEDGEMENT

Research committee approval and financial support: This research was financially supported by Mashhad University of Medical Sciences. The research project is also registered in Clinical Trials.gov (Code NCT016423).

Conflict of Interest: Authors declare no conflict of interest.

REFERENCES

1. Bachman LF; Palmer AS. Language testing in practice: Designing and developing useful language tests. USA: Oxford University, 1996.
12. Chiang CI. The effects of graphic organizers on Taiwanese tertiary students' EFL reading comprehension and attitudes towards reading in English. Australia: Australian Catholic University, 2005.