realities, to argue that test influence (washback), though accepted to be present to some extent, is very minimal in our situation as far as language learning is concerned. This claim emanates particularly from the defects of our system. To elaborate of this, just imagine that the goal of our language program is to achieve "language grammar" (linguistic competence), as this has been inherited to us. And imagine further that all the involved elements in the program, including teachers, learners, materials, etc., are consistent with the main goal. As a consequence, we expect, e.g., at least one character in this game - learner - across the country to be influenced by such tests, and as a result, learn "grammar" very well. This kind of grammar syllabus has been the top priority in our language plans for a long time, but what is quite disappointing is that very few of the learners have managed to master grammar, let alone other aspects. This short reasoning is to conclude that even if we make innovations in our tests in line with new language theories, we are not able, as a rule of thumb, to make profound changes in our plans. The policy of pressure from below for the changes above sounds inefficient. By this virtu, we may be able to lead our characters to the water but not able to make them drink. Thus we need to care for the education system in general and language learning in particular, in case we feel responsible to our children.

References

Alderson, J.C. (2002). Language Testing and Assessment. *Language Teaching*, 79-113.

Alderson, J.C. and D. Wall. (1993). Does Washback Exist? *Applied Linguistics*, 14, 115-29.

Bachman, L.F. (1995). *Fundamental Considerations in Language Testing*. New York:OUP.

Bailey, K.M. (1996). Working for Washback: A Review of the Washback Concept in Language Testing. *Language Testing*, 13, 257-79.

Brown, D.H. (2004). *Language Assessment: Principles and Classroom Practices*. Pearson Education, Inc.

Canale, M. and M. Swain (1980). Theoretical Bases of Communicative Approaches to Second Language Teaching and Testing. *Applied Linguistics*, 1, 1-47.

Carroll, B.J. (1980). *Testing Communicative Performance*. Oxford: Pergamon Press.

Davies, A. (2000). The Logic of Testing Languages for Specific Purposes. *Language Testing*, 18(2), 133-147.

Hart, D., S. Lapkin and M. Swain. (1987). Communicative Language Tests: Perks and Perils. *Evaluation and Research in Education*, 1,83-93.

Hughes, A. (1993). Introducing a Need - Based Test of English Language Proficiency in A. Hughes (Ed.), *Testing English for University Study*. ELT Document ≠ 127, Modern English Publications.

Hughes, A. (2003). *Testing for Language Teachers*. Second Edition. Cambridge:CUP.

Hymes, D.H. (1972). On *Communicative Competence*. In J.B. Pride and J. Holmes (eds.): Sociolinguistics. Harmonds Worth: Penguin: 269-93.

McNamara, T. (2000). *Language Testing*. Oxford: OUP.

Messick, S. (1996) Validity and Washback in Language Testing. *Language Testing*, 13,241-256.

Morrow, K. (1991). *Evaluating Communicative Tests*. In S. Anivan (Ed.), Current Developments in Language Testing (pp.111-118). Singapore: Regional Language Center.

Oller, W.J. (1979). *Language Tests at School: A Pragmatic Approach*. London: Longman.

Prodromou, K. (1998). The Washback Effect: from Testing to Testing. *ELT Journal*, 49/1,32-45.

Weir, C.J. (2005). *Language Testing and Validation*. New York: Prentice

Widdowson, H.G. (1978). *Teaching Language as Communication* Oxford: OUP.

Widdowson, H.G. (1983). *Learning Purpose and Language Use*. London:OUP.

situation is inevitably artificial and idealized and it cannot replicate the realities involved in the real language. Davies (2000) believes that those who are after authenticity are fruitlessly pursuing the chimera of authenticity. Messick (1996) observed that any kind of test is an exercise in sampling and from this sample an attempt is made to infer student's capabilities in relation to their performance in general. But because we have to make stable predictions about large universe of tasks, the fact conflicts immediately with the demands of test efficiency. Moreover, the sampling of communicative task as to involve what and how of the test needs meticulous analysis of the prospective situations, while taking the risk of losing some information, and it is almost an inefficient job to do. It seems unjustified on practical grounds to adopt such an approach to generate positive influence on the part of the learners / teachers. To say least, the time seems not yet ripe to include such tests in its scientific sense in the language syllabi for possible washback. It follows that such tests are not yet scientific - psychometrically established. On the other hand, the whole language learning system in Iran, including administrators, teachers, learners, facilities, etc. all need to be livened up before embarking on making innovations in testing. It sounds inappropriate to start off making new tests for the changes in the plans.

Another important challenge against the idea of changes in tests for the better is in line with a more utilitarian view of education. Test in education system is no more than an instrument. Though it is vital to make an effective instrument for the fulfillment of the goal, it does not mean to take a means for an end (McNamara, 2000). I believe too much emphasis on the means relegates the end to a secondary position. In such cases, the learner will be kept in fetters of test, without any heed for optimal pragmatism expected from any education system. This is precisely what is happening in our society.

Conclusion

I would like, on the basis of the existing

show "functions" rather than the formal language patterns of usage (Hymes, 1972; Widdoson, 1978 and 1983). And it is also believed that a test will promote beneficial washback to the extent that it is based on sound theoretical principles which are widely accepted in the field (Bailey, 1996). Thus the old way of testing based on a limited view of what language is and how it works is no longer productive. With respect to the non-communicative (traditional) testing, Iranian schools testing system seems to be strongly set and fossilized and this is quite obvious from a quick glimpse of the available tests, both at school and university level. They merely focus on the linguistic elements; however, the recent definition of language incorporates more than linguistic competence. For example, the communicative view of language proposed by Bachman (1995) defines the communicative language ability as the ability consisting of language competence, strategic competence and psycho-physiological mechanisms.

In line with such thinking, Morrow (1991) suggested that communicative testing have the following five characteristics:

1. Single skills: the examinations consist of a suite of free - standing modules in writing, reading, listening and speaking.

2. Tests of performance: the tests are designed to be direct measures of performance. In a communicative classroom the focus of activities is "doing something" through the language.

3. Task-based: communication through language involves the participants in carrying out tasks in the production or comprehension of language in order to arrive at shared understanding.

4. Criterion-referenced: the essential question, which a communicative test must answer, is whether or how well a candidate can use language to communicate meaning.

5. To reflect and encourage good classroom practice: this is a major concern underlying the design of test; indeed in many ways the tests themselves have grown on good classroom practice in an attempt to disseminate to other classrooms.

Morrow (1991) advocates feedback between teaching and testing as vital for educational development. Other authors like Canale and Swain (1980), Green (1985), and Hart, Lapkin, and Swain (1987) suggest the development of communicative language tests.

Communicative testing immaturity and washback insufficiency

Though almost all teachers and experts have in unison voiced the influential role attributed to the tests, I think that the significance of the test has been exaggerated to superfluity. Attempts in this respect look to the changes consistent with the communicative perspective of the language.

A major criticism against communicative tests sticks to the idea of the generalizability of the results. Since such tests seek the specifications of the practical activities conducted by the learners, then this fact confounds the elicited results and militates against the generalizability of the tasks. To circumvent the problem, experts usually propose the precise species specification of the objectives. But as it is quite evident, the identification of relevant communicative tasks underlying skills for a special group of learners will encounter the problems of sampling, practicality, relability, and validity (McNamara, 2000).

As Weir (2005) claims, the more the tasks are specified through needs analysis, the more the possibility of decreasing genererlizability. Next criticism aims at the authenticity or real life subject proposed in this domain. Since much communication is non-linguistic, hence a test
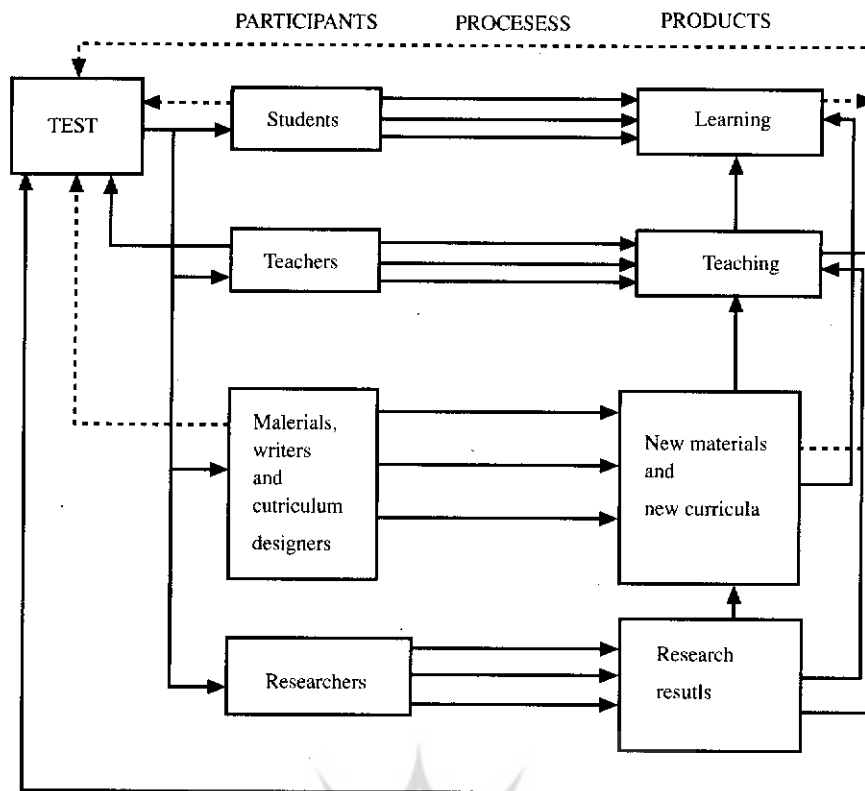
A basic model of washback adopted from Hughes (1993)

In an article entitled "Does washback exist?", Alderson and Wall (1993) propose some possible washback hypothesis. These hypotheses also incorporate the extensive coverage of the washback over the constituents of the program. They are as follows:

1. A test will influence teaching.
2. A test will influence learning.
3. A test will influence what teachers teach.
4. A test will influence how teachers teach.
5. A test will influence how learners learn.
6. A test will influence how learners learn.
7. A test will influence the rate and sequence of teaching.
8. A test will influence the rate and sequence of learning.
9. A test will influence the degree and depth of teaching.
10. A test will influence the degree and depth of learning.
11. A test will influence the attitudes to the content, method, etc. of teaching and learning.
12. Tests that have important consequences will have washback and conversely.
13. Tests that do not have important consequences will have no washback.
14. Tests will have washback on all teachers and learners.
15. Tests will have washback effect for some teachers and some learners, but not for others.

Such views by Hughes, Alderson and Wall reveal an important role a test can play in a language program. Such a status that testing has necessitates a particular attention because tests can echo a wide feedback to different parts of the system, so we need to take as much care of the test as possible. By the same token, we have to make such tests that could provide the most out of their influences.

Communicative testing and washback effect

Upon criticizing the traditional tests to be usage-based and hence unable to assess learner's communicative ability (Morrow, 1991; Carroll, 1980), the emphasis on testing has been shifted onto the "use" and the concern that testing should

activities to the demands of the test, especially when the test is very important to the future of the students, and the pass rates are used as a measure of teacher's success". This influence of test on the classroom referred to, as washback is of course very important. And the effect can be beneficial or harmful. No doubt most of the testing experts, methodologists and language teachers endorse the influential role of test on language plans and their implementation. However, some of them express some caveats as to how the tests are to be put to use. Due to the importance of tests, they believe, any deficiency in the structure of the test would infect the whole body of the language pedagogy. In this connection we are warned to care for the "negative washback" (Hughes, 2003). In contrast to negative washback comes positive washback that is the one desired for language pedagogy. Messick (1996), however, takes washback as "the extent to which the test influences language teachers and learners to do things they would not otherwise necessarily do". Now the question is how positive washback is to be demarcated.

We have to say that positive washback is a relative term. It all depends on the theories of language and learning we adhere to. For example, at the time faculty psychology was fashionable, kinds of tests requiring translation and memorization were considered to have positive washback. Later, psychometric- structuralism era demanded testing discrete linguistic points. They are still widely used and propagated. The positive washback this era generated was the sheer concentration on linguistic forms. This typs of testing was also criticized on the grounds of breaking the elements of language apart and giving no attention to the way those elements interacted in a larger context of communication. Next generation of tests called integrative tests were supposed to put the language back into its totality.

They included cloze and dictation tests (Oller, 1979). Therefore, as far as one believes in a special theory of what language is, he considers the related types of tests to produce positive washback since those tests would aim at the target as it is defined in the accepted paradigm. The positive washback that we seek is the one which is specifically treated in the light of the changes language teaching and learning have recently undergone, namely, the changes of trend towards communicative language testing.

Role of washback in language program

Washback as the test influence has both overt and covert implications for language teaching and learning (Prodromou, 1998). Such implication can be beneficial to the teachers, learners and authorities; Hughes (1993:2) pinpoints the dimensions of the language plan, which can come under the tailoring mechanisms of the test. In a program, he believes testing exerts influence on three elements; participants, process, and product. Participants include learners, teachers, administrators, material developers, publishers, and all those who are affected by the test. He defines process as "any actions taken by the participants who may contribute to the process of learning". The processes include material development, syllabus design, changes in teaching methodology, and the use of learning or test-taking strategies. And he also refers to the product as "what is learned (facts, skills, etc.) and the quality of learning (fluency, etc)".

Hughes (1993) elaborates on the idea by saying that test a puts up a special attitude and perception on the part of the participants leading to a particular kind of activity, e.g., practicing the kind of items to be found in the test, which in turn forms the learning outcome. Hughes then offers a basic model of washback representing this views stated above. This model is illustrated below:

positive and negative effects that tests may have on teaching, and want to know what they should do to ensure that these effects are beneficial. Some ELT specialists have offered advice on how to produce positive washback, calling for attention to test design and the communication between testers and teachers. In accordance with the newly announced need for modifications in types of tests in Iran for making changes in the language program, this article attempts to illuminate the notion of the test influence on language teaching and learning referred as 'washback effect'. It then reviews the recent innovations in language testing for the purpose of bringing about changes in teaching and learning language. Finally, it argues that new trends in testing are not yet well-established and that washback *per se* cannot fulfill the assumed goal.

Key Words: washback, notion, test, innovations, communicative.

Introduction

Tests, needless to say, constitute an essential part of any education system. They make an invaluable contribution to the assessment of learner's progress, proficiency, and are employed as diagnostic tools to uncover learner's errors, and interlanguage processes involved in language learning. They may also give rise to motivation, and discipline both on the part of the teacher and learner (Prodromou, 1998).

Close to the test is the concept of teaching. They have mutual bearings on each other. Despite the apparent reciprocity involved, many educationalists deem the direction to be from testing to teaching. (e.g., Alderson, 2002; Bailey, 1996; Brown, 2004; Davies, 2000; Hughes, 2003; McNamara, 2000; Weir, 2005). An analogy may cast light over the position of the test. Test can be taken as an "event" which requires special preparation for the encounter. The more the event is known, the better one can face it. The known influential dimensions of any event would thus instigate the would – be victims to make preparations for all the specified outcomes. Through this analogy we also attribute the prime influential path to move from testing to teaching. Tests, therefore, reflect the "demands" to teaching–

to the classroom, waiting for the "supply". In testing literature the idea of test affecting teaching has been called "washback effect" (Bachman, 1995). Alderson and Wall (1993) also argue for innovations in the language testing to make modifications to the curriculum.

As due emphasis is always directed to testing as a permanent influence exerting element on teaching in the education realm, and because in Iran the trend is directed to make necessary changes in tests for the sake of prospective influences on teaching/learning, this article attempts to explore the notion of the test influence on the language teaching/ learning referred to as "washback effect". It then reviews the recent innovations in the language testing for the purpose of bringing about changes in teaching / learning the language. And finally it argues that new trends in testing are not yet well - established and that "washback" or test influence alone cannot fulfill the assumed goal.

Washback or test influence

As stated before, test influence of washback is the direct and indirect effect testing can have on language teaching / learning. Brown (2004) defines washback as" ... A natural tendency for both teachers and students to tailor their classroom

Introducing Novel Tests for Meaningful Changes in English Language Teaching: The Washback Effect Revisited

Parviz Ahmadi Darani: M.A.(TEFL)
ahmadi 1906@yahoo.com
Islamic Azad University of Shahrekord

چکیده

«اثر آزمون» (washback) واژه‌ای آشنا در میان متخصصان آموزش و آزمون‌سازی است . زیرا آزمون‌ها از عناصر تعیین‌کننده‌ی نظام آموزشی هستند و این متخصصان از اثرات بالقوه مثبت و منفی آزمون‌ها بر آموزش نگران‌اند و به دنبال رویکردهایی هستند که بتوانند، اثرات مثبت آزمون‌ها را به حداکثر برسانند . برخی از آنان توصیه‌هایی برای ایجاد اثرات مثبت آزمون‌ها دارند و بر طراحی آزمون‌ها و ارتباط نزدیک بین معلمان و متخصصان آزمون‌سازی تأکید فراوان می‌کنند.

در پاسخ به نیازهای اعلام شده‌ی جدید برای تغییرات در انواع آزمون‌ها در ایران و به‌منظور اصلاحات بنیادین در برنامه‌های آموزش زبان ، این مقاله سعی دارد مفهوم «اثر آزمون» بر آموزش و یادگیری را توضیح دهد . سپس به بررسی نوآوری‌های اخیر در آزمون‌سازی در راستای نیل به تغییرات معنادار در فرایند آموزش و یادگیری زبان بپردازد و روش آزمون‌سازی ارتباطی ، آرا و نظریات متخصصان درباره‌ی این روش ، و کمی و کاستی‌های این روش را شرح دهد . در خاتمه، مقاله به این نکته اشاره دارد که باوجود پیشرفت‌های چشم‌گیر در عرصه‌ی آزمون‌سازی ، اثر آزمون نتوانسته است به اهداف تعیین شده در ایران نائل شود . این وضعیت ناشی از برنامه‌های آموزشی سنتی است که هنوز از اولویت خاصی در کشورمان برخوردار هستند . بنابراین ، تا زمانی که این نقایص در سیستم آموزشی زبان اصلاح نشود ، لحاظ کردن نوآوری‌های جدید در آزمون‌ها راه به جایی نخواهد برد.

کلید واژه‌ها : اثر آزمون، مفهوم، آزمون، نوآوری‌ها، ارتباطی

Abstract

The notion of 'washback' is common in language teaching and testing literature, and tests are held to be powerful determiners of what happens in classrooms. Educators are concerned about the potentially