

Koláčková & Šikolová, 2017

Volume 3 Issue 2, pp. 465 - 471

Date of Publication: 08<sup>th</sup> September, 2017

DOI-<https://dx.doi.org/10.20319/pijss.2017.32.465471>

This paper can be cited as: Koláčková, L., & Šikolová, M. (2017). *Language Tests and Their Role in Society*. PEOPLE: International Journal of Social Sciences, 3(2), 465-471.

This work is licensed under the Creative Commons Attribution-Non-commercial 4.0 International License. To view a copy of this license, visit <http://creativecommons.org/licenses/by-nc/4.0/> or send a letter to Creative Commons, PO Box 1866, Mountain View, CA 94042, USA.

## **LANGUAGE TESTS AND THEIR ROLE IN SOCIETY**

**Ludmila Koláčková**

*University of Defence, Language Centre, Brno, Czech Republic*  
[ludmila.kolackova@unob.cz](mailto:ludmila.kolackova@unob.cz)

**Mária Šikolová**

*University of Defence, Language Centre, Brno, Czech Republic*  
[maria.sikolova@unob.cz](mailto:maria.sikolova@unob.cz)

---

### **Abstract**

*Being continually subjected to various forms of assessment remains a certainty for most of us, and the results of those evaluations often have a significant impact on our lives, be it positive or negative.*

*Whether we have in mind aptitude tests, language tests, entrance examinations or either high school or a university, driving tests, their results are invariably of profound importance for us of profound importance for us. It is automatically taken for granted by the public is that the requirements for passing the tests are appropriate, and the instruments for the measurement of the required knowledge, abilities, performance, skills or competencies have been constructed in such a way that they reliably and accurately measure what they are intended to measure. Nevertheless, we often do not realize that this ideal is not always the case and hence, the tests frequently become a potentially dangerous instrument in the hands of the powerful in order to impose their policies. The objective of the contribution is to raise awareness and focus attention*

on the importance of tests, their design and quality control. It will illustrate the pitfalls in the area of language testing both from professional literature and authors' experience.

### **Keywords**

Language testing, Reliability, Validity, Test design

---

## **1. Introduction**

Motto:

*Tests should be labeled just like dangerous drugs: 'Use with care.'*

(Spolsky, in Bachman, 2013)

High-stakes tests generally and high-stakes language tests particularly are of profound importance, since their results can influence the study opportunities, careers, and thus the whole lives of test takers. Moreover, the tests often affect students and teachers in the educational process, and they could have a broader influence on society and educational system as such. For better understanding, we will first define the main terms we are going to use.

## **2. Definitions of basic terms**

Test *impact* can be defined as a phenomenon which can be observed on two levels, namely on the micro and the macro levels. The impact on the micro level means the influence of test on students and teachers while the impact on the macro level is represented by the influence of test on society and educational system (Bachman & Palmer, 2013).

Another term connected with tests and their effects is *washback* or *backwash* used interchangeably. Basically, its meaning could be summarized as the influence of testing on teaching and learning (Alderson, Clapham & Wall, 2006; Hughes, 2013). Another definition offered by Alderson and Wall is a bit more concrete: it is "a phenomenon by which teachers and learners do things which they would not do if there were no test" (Alderson, Clapham & Wall, 2006).

A view on *washback* from a slightly different perspective is given by Shohamy. She defines it as an intentional exercise of power over educational institutions with the objective of controlling the behavior of teachers and students (Shohamy, 2016).

From the above-mentioned definitions, it is apparent that the influence of tests can be either positive or negative. Alderson and Wall (Alderson & Wall, 2006) also state that the term

may refer both to (intended) positive (Bachman & Palmer; 2013 Davies, 2006) or beneficial (Buck, 2011; Hughes, 2013) effects and to (unintended) harmful (Buck, 2011) or negative effects (Bachman & Palmer, 2013; Davies, 2006; Hughes, 2013).

However, Valette (1977), and Spolsky (2012) argue that it represents only unforeseen and deleterious effects.

An important factor to be taken into consideration is that the quality of *washback* can be independent of test quality. Any test, whether good or bad, can have its beneficial or harmful influence (Alderson, Clapham & Wall, 2006).

Another rather concrete and real-life commentary on *washback* is that “it is the phenomenon which we all know from our student years: ‘Is it going to be in the test?’ ” (Shawcross, n. d.).

Cheng (2013), on the other hand, sees the *washback* in association with the curriculum in such a way that the intended curriculum changes are associated with a testing innovation. This leads us to the term of *measurement-driven instruction* which means that the test should manage teaching, and thus also learning (Popham, 1990). Popham sees the tests in a positive way as he thinks that introducing tests means “the most efficient way to improve the quality of public education” (Popham, 1990). In addition, he advocates the proper use of tests stating: “If the tests are created and used in an appropriate way, then focusing teaching on what is being assessed is a useful activity” (Popham, 1990).

Sometimes the tests are based on the curriculum or they correspond to it, or it could be put the opposite way around that the curriculum is based on tests and corresponds to them. This relationship is encapsulated in the term of *curriculum alignment* which means the alignment between testing and teaching (Shepard, 1989). In this respect, Pearson (Pearson in Cheng, 2013) claims that exams are commonly used as levers for change (Pearson in Cheng, 2013). The importance of the roles of the language teachers in this process is emphasized by Alimemaj (Alimemaj, 2015).

### **3. Power of Tests**

The authors have borrowed the name of this part of the paper from Elana Shohamy’s book. What we would like to point out to is the fact that the public often does not fully realize what great responsibility and power are hidden in tests and especially in their results. As E. Shohamy emphasizes: “Doing well on tests can take a person to the best university and open the

way to an excellent education, doing poorly can send a person to a low-level university and block the possibility of higher education, resulting in a poor education. Doing well on a test may lead to a person entering a desired profession, while doing poor may lead to an unwanted job. Doing well on a test may mean that a person is given permission to migrate to a new country and start a new life, while doing poorly may force a person to stay somewhere he or she does not wish to be.“ (Shohamy, 2016).

In the authors’ country of origin, test results did not use to be questioned until very recently. As we think that this could well be the case in most countries all over the world, except the US and UK, we have decided to address this issue in our article.

Since the high-stakes tests profoundly affect the lives of test-takers, it is almost automatically assumed that exams should meet a lot of requirements, namely fairness, transparency, accountability and the right to verify the accuracy of the results or measurement and recourse against them.

The well-established and influential testing institutions address these requirements in many respects. They put a lot of emphasis on quality assurance and the Code of Conduct often precisely specifies the desired ways of behavior of the testers.

To demonstrate an aspect of testing that is not widely comprehended and dealt with, i.e. its broader consequences, we will present some examples of different purposes for which the tests have been used, as well as their either positive or negative impact.

#### **4. Real Life Examples of Test Use**

Shohamy’s ideas which we find particularly appealing are those mentioned in the chapter titled Temptations. In this chapter, she lists and comments on possible situations in which tests are or can be used as various instruments of power. One of these situations is that “tests are effective for control and for redefining knowledge”. She further elaborates on this idea and says that test takers will try to learn and acquire the knowledge that is being tested. Another fact she mentions is that “tests allow cost-effective and efficient policy making”. She develops this idea in such a way that instead of implementing a complex educational reform including new teaching materials, curricula and teacher training, it is much simpler and more effective to introduce an exam. In her words, “tests, therefore, provide policy makers with the opportunity to create policy in the shortest time.” (Shohamy, 2016).

The following cases demonstrate how tests can be used and what kind of consequences they may result in.

As an example of test use for ethnic discrimination, the Latvian exam can be mentioned. This goes back to the nineties when Latvia gained its independence. The government had introduced the Latvian exam for those applying for Latvian citizenship which significantly affected Russians who historically had lived in Latvia before it became an independent state. The language exam was quite demanding and its impact was not difficult to guess – in fact, in 1996 the number of Russians has dropped from 52 % to 35 % (Shohamy, 2016).

To illustrate also the better side of using the high-stakes language examinations, let us mention the introduction of the Arabic Exam in Israel in 1988. The primary goal of introducing the exam was to improve the perception of the Arabic language, in other words the introduction of the exam was a method of directed change in order to gain respect, as well as change the teaching methods. Subsequently, three studies were conducted, one immediately before the test was administered, the second one after three years and the third one after eight years. To sum up the results of the studies, the test resulted in changing the behavior of both teachers and learners, as well as the whole domain of teaching Arabic. However, the goals reaching behind teaching, i.e. promoting the status of Arabic, were not achieved (Shohamy, 2016).

As both authors have been working for the Czech military, they have been given the opportunity to observe profound changes in language learning and teaching in the Czech Armed Forces as a consequence of the political changes after 1989 Velvet Revolution. Following the path of democracy, the Czech Republic was preparing to join NATO in the mid-nineties of the last century which entailed language learning as an integral part. At the first sight, it might seem simple, but it covered a broad spectrum of changes not only in terms of curriculum, teaching materials and teacher training, but also internal changes in attitudes and general perception of language abilities. In this respect, the implementation of NATO STANAG 6001 exam has resulted in profound changes in teaching. In comparison with the previous predominantly grammar-translation method, communicative approaches started to be exerted, since the mentioned standardized examination is skill-focused.

The exam has been of high importance ever since; its passing often has been a prerequisite for promotion or for sending abroad (NATO structures, military mission, etc.), which has made the exam a powerful tool. Moreover, in 2016 a new law for state service was passed

and in connection with this, a regulation was enforced concerning language requirements and their meeting. The language requirements for the soldiers were set without any deeper understanding of the testing system as a whole and the interpretation of exam descriptors and test results while language specialists did not have a say in this process. On top of that, this regulation requires the military to meet the requirements until the end of 2019 and failing to meet them in four-year time will result in dismissal from the Czech Armed Forces.

There have recently been rumors concerning softening these requirements, but nothing has been officially declared so far. However, if the conditions are not eased, it may have serious detrimental effects on the Czech Armed Forces.

## **5. Conclusion**

To sum up, the concept of impact and/or washback in high-stakes testing is of considerable importance for all stakeholders. Nevertheless, the research in the field is still not sufficient and there is still a need for raising public awareness of it. High-stakes test designers and responsible managers have to bear in mind this importance and hence, be fully aware of and observe their accountability. Since it is apparent that high-stakes tests influence, either positively or negatively, the learning – teaching processes, the emphasis should be put on the communication between teachers and testers, as well as other stakeholders. Ideally, the communication should result in teaching and testing alignment. And last but not least, students and/ or test candidates should also have their say in the whole process expressed in either formal or less formal ways.

## **REFERENCES**

- Alderson, J. C., Clapham, C., & Wall, D. (2006). *Language test construction and evaluation*. Cambridge: Cambridge University Press.
- Alimemaj, Z. (2015). *Language test taking in Albanian schools*. Retrieved from <http://eujournal.org/index.php/esj/article/view/6380>.
- Bachman, L. F., & Palmer, A. S. (2013). *Language testing in practice: designing and developing useful language tests*. Oxford: Oxford University Press.
- Buck, G. (2011). *Assessing listening*. Cambridge: Cambridge University Press.

- Cheng, L. (2013). *Language classroom assessment*. Alexandria, VA: TESOL International Association.
- Davies, A. (2006). *Dictionary of language testing*. Cambridge: Cambridge University Press.
- Hughes, A. (2013). *Testing for language teachers*. Cambridge: Cambridge University Press.
- Popham, W. J. (1990). *Modern educational measurement: a practitioner's perspective*. Englewood Cliffs, NJ: Prentice Hall.
- Shawcross, P. (n.d.) *What do we mean by the "washback effect" of testing?*. Retrieved from [https://www.researchgate.net/publication/242129949\\_What\\_do\\_we\\_mean\\_by\\_the\\_'washback\\_effect'\\_of\\_testing](https://www.researchgate.net/publication/242129949_What_do_we_mean_by_the_'washback_effect'_of_testing).
- Shepard, L. A. (1989). Why we need better assessments. *Educational Leadership*, 46, 4-9.
- Shohamy, E. (2016). *Power of tests: a critical perspective on the uses of language tests*. Place of publication not identified: Routledge.
- Spolsky, B. (2012). *The Cambridge handbook of language policy*. Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9780511979026>
- Valette, R. M. (1977). *Modern Language Testing*. New York: Harcourt Brace Jovanovich.