Taking account of test takers' personal characteristics in language test development process

Svetlana Baluyan[*]

Southern Federal University, 347900 Taganrog, Russia

Abstract. The article presents the results of a study aimed at solving the problem of improving the quality of a language test. It was found that taking into account test takers' individual characteristics that may affect their test performance, namely their personal characteristics, general and specific knowledge, language ability level and the emotional component at the design stage allows to increase reliability, interactiveness, impact, authenticity and even practicality of the test, which, in turn, contribute in its overall effectiveness.

The main problem in designing and developing a pedagogical test is to ensure its quality, or usefulness. The quality of a language test is affected by several different characteristics that are interconnected, and their combination provides the overall quality of the test.

It is known that such characteristics as reliability and validity are unanimously recognized as the main ones in language testing, and they are used as criteria in identifying the overall quality of the test. Other characteristics of the usefulness of tests, such as objectivity, practicality, effectiveness and others, vary in different language testing schools.

Representatives of one of the most advanced testing schools, American language testers L. Bachman and E. Palmer, adhere to the point of view according to which six characteristics are responsible for the quality of tests, namely reliability, construct validity, authenticity, interactiveness, impact and practicality. Since some of these characteristics depend on each other and are even in conflict, the authors insist on the need to try to maximize them in the given test, while maintaining an appropriate balance between them [1].

The overall quality of the test can be significantly improved if we take into account the personal characteristics of the test takers, having them in mind at the stage of test design and development, thereby improving such test qualities as reliability, impact and interactiveness. For example, determining the test takers' approximate level of foreign language ability and adapting the test to this level when designing it can help to improve its reliability. The collection of data on the personal characteristics of test takers helps to determine the degree of impact of the test on them in the form of a phenomenon known in language testing as "washback effect" or "backwash effect".

The personal characteristics of the test takers and considering them when creating pedagogical tests were the object of interest of both foreign and local researchers. At different

---

[*] Corresponding author: sbaluian@gmail.com

times this problem was dealt with by H. D. Brown, M. A. Christison, N. Entwistle, U. Esser, S. Krashen, W. Rivers, P. Skehan, Alkhazishvili A.A., Vaysburd M.L., Vitt N.V., Zimnyaya I.A., Kabardov M.K., Kuzovlev V.P., Markova A.K., Nikolayeva S. Yu., Tsetlin V. S. Their work led to the conclusion that the individuality of each person is unique and is formed from a complex of an infinite number of different properties. Obviously, when designing a test, it is not possible to take into account all these properties and it is necessary to highlight a small number of them that are significant for language test construction purposes. An analysis of several works in this area [1,2,3,4] led to the conclusion that four categories of test-takers' characteristics can affect the test results, namely their personal characteristics, their empirical, academic or topical knowledge, their general level of language ability and their possible emotional reaction to the test.

At first glance, the level of proficiency in a foreign language is in no way connected with the personal characteristics of the test takers, but in reality, personal qualities can in some way affect their test performance. Native language(s), age, sex, experience of participating in language testing, amount of preparation, level of general education - taking all these characteristics into account is considered to improve the quality of the language test.

It is equally important to determine the level of topical knowledge of the test takers and keep it in mind when designing a test. For purposes of test development, the testees can be grouped into two categories: those with a wide range of specific knowledge and those with relatively homogenious specific topical knowledge. For the first category, there is no need to take this feature of their personal characteristics into account when designing the test. The information base for developing a quality test for the homogenious group with specific topical knowledge can be the common area of knowledge using which they will demonstrate their language ability. Often, the development of such tests requires the consultation of a specialist in the field. Tests of this kind are most often required for language proficiency testing for special purposes. Taking into account the factor of test takers' topical knowledge can increase the overall quality of the language test by increasing its validity, impact, interactiveness and authenticity.

Considering test takers' general level of language ability also allows to maximize the overall quality of the test. Knowing how successfully they can solve different communicative tasks in foreign language in practice, language test designers include difficult but feasible tasks for this level in the test. In some cases, when developing a test, the level of competency in different language skills are taken into account. For example, as a result of preliminary collection of information, it became known that in a particular group of potential testees language competencies in reading are developed much better than in listening. Replacing the oral stimulus of the task in the achievement test in grammar with a written one allowed to improve the overall quality of the test.

In the second half of the 20th century, pedagogical testing took a big step forward. The emergence of Item Response Theory (IRT) made it possible to solve the most important problem - to adapt the level of difficulty of test tasks to the test takers' practical level of language ability, which can noticeably reduce time spent on testing, increase the objectivity of assessment, and, in general, improve the quality of language tests. In adaptive tests, the difficulty and the number of tasks directly depend on the test taker's response to the previous task or several tasks. The selection of the next task is carried out in such a way that its complexity in the best way allows to elicit the language level. In adaptive tests, there is no predetermined number of tasks. Upon reaching sufficient accuracy in assessing the level of language ability, the testing ends, therefore, such tests require less time to complete than traditional ones. The practicality of adaptive tests is also higher because they are cheaper than traditional ones. Practical, easy-to-use and perspective computer-based adaptive tests based on IRT best solve the problem of taking into account the general level of language ability, thereby also improving their overall quality.

In Russia, theoretical studies on adaptive testing are not as intensive as in the western countries. Among the specialists working in this field are Avanesov V.S., Vasilyev V. I., Nardyuzhev V. I., Tyagunova T. N., Chelyshkova M. B., Shmelev A. G. However, there is no doubt that the prospects for the development of testing in linguodidactics are connected with computer adaptive testing.

The study of test takers' personal characteristics is also necessary to identify their presumptive emotional reaction to the test tasks, the testing process and the test environment. In this case, information about their experience in participating in such testing, their general knowledge, and the level of language ability can be useful. For example, if a person has already taken part in such testing and is familiar with the test setting and types of tasks, then we can assume that his reaction to the test will be positive. Usually a positive emotional reaction is also observed when the test is compiled in accordance with the requirements for the correspondence of test tasks to communicative tasks in the target language use domain. Students with a lower level of language proficiency or specific knowledge may have a negative reaction.

Various methods are used to identify the personal characteristics of potential testees, such as interviews, self-reports, questionnaires, observations, and other reliable sources of information.

Thus, taking into account the test takers' individual characteristics, namely their personal characteristics, general and specific knowledge, language ability level and emotional component, can improve the reliability, interactiveness, impact, authenticity and even practicality of the language test being developed, which, in turn, in turn, contribute in its overall effectiveness.

References

1. L.F. Bachman, A.S. Palmer, *Language Testing in Practice: Designing and Developing Useful Language Tests* (Oxford, 1996)
2. H.D. Brown. *Principles of Language Learning and Teaching. Englewood Cliffs* (Prentice-Hall Regents, New Jersey, 1994)
3. M.A. Christison, Journal of the Imagination in Language Learning. **3**, 8-12 (1996)
4. P. Skehan, *Individual Differences in Second-language Learning* (Edward Arnold, London, 1989)