

THE INTERACTIVE TESTING IN THE ENGLISH FOR SPECIFIC PURPOSES TEACHING IN NON-LINGUISTIC UNIVERSITIES

Olena Hryshyna

National Technical University of Ukraine "Kyiv Polytechnic Institute", Kyiv, Ukraine
gryshyna@gmail.com

In the paper the author gives a thorough analysis of the effective tools for the control organization in terms of the recent researches in the English for specific purposes testing methodology. The writer proposes the effective technology of the reading competency measurement covering the methods and techniques of the test administration including the criteria and quality characteristics definition. The research gives the sound articulation of the test quality characteristics alongside with the outlined test task design requirements. The examples of the original communicative professional test tasks give an overall picture of the testing domain to be developed. The objectives outlined give the reader an understanding of the test structure evaluation and processing the output data. The research gives a possibility to obtain a practically ideal tool for measuring the study achievements in reading competency acquisition of the technical university students. Moreover, the research highlights the learning outcomes to be envisaged beforehand as the target skills the students-future engineers are to acquire. The paper investigates the impact of the test characteristics and qualities on the test task construction in terms of the learning and testing purposes interaction. The research focuses on the test task development requirements as a dominant creative force in a testing development, organization and administration. Being sufficiently useful, criteria bring a sound functionality in the test task and test items construction. The structure of the communicative real life situations in foreign language reading where the future technical specialists are to apply their certain reading skills for their academic purposes is elaborated and introduced in the paper.

Keywords: test quality characteristics; test construct; test item development; test task requirements, types of test task; measurement and evaluation tools; reading skills; reading competency; real-to-life domain; authenticity profile.

Introduction. The demand for successful testing is rapidly growing due to the content structure requirements in the framework of the specific criteria and an adequate assessment needed within the foreign language professional education. Many scientists in the sphere of the teaching of the English language for foreigners are engaged into the investigation and research of ideal tools to evaluate the study achievements of the students. To find a solution, it is obligatory to envisage the learning outcomes as the leading target skills the students are to acquire and the reading skills being purposes in communicative testing.

Primarily, a test is a means for gathering the data performed by the learner while reading in order to elicit responses for various reasons stated as an objective of a task. Moreover, such tasks are presented in the test items covering the main set of purposes caused by reading (Alderson 1995, Fulcher, 2007). However, the test construction itself remains the burning question in the modern testing methodology. Thus, the construct domain evokes the interaction of the testing and teaching content in terms of the long/short range objectives. Consequently, the design of a test item is subjected to the real life or communicative domain of testing. To be more scientifically and methodologically grounded, the need for the definite instruments of the test design or the test domain profile construct organization has been urged.

Test Quality Characteristics. The effectiveness and objectivity of testing measurement is needed. Therefore, to get any speech reading skill evaluated, namely, a test is considered to be such a tool. Thus, such a tool must be produced in accordance with the *quality characteristics* of a test→subtest→test task. There is a crucial concept of a test or degree measuring the ability of test→subtest→test task that determines how well and effectively, regarding the circumstances when test measures the competency of a certain skill.

A number of scientists propose to take into account the following quality characteristics: *validity, reliability, practicality, authenticity, interactivity, impact and significance of instruction (instructional value)* (Heaton, 1990, Hughes, 1996). Moreover, the communicative quality is not pronounced as a definite category (Petraschuk, 2003).

Reading through the researches on evaluation and assessment instruments can distinguish the quality characteristics of a test as follows: **validity, reliability, practicality, efficiency, item difficulty, authenticity, interactivity** (Hughes, 1996, Weir, 2005, Kokkota, 1989). Actually, quality

characteristics like the impact and significance of instruction (instructional value) are implicit in the interactive quality context. Undoubtedly, to have a thorough understanding of all the qualities and how they work as a whole it is obligatory to analyse them briefly.

Firstly, test **validity** is defined as the test orientation towards the purpose of the test. Validity demonstrates the efficiency measurement. In addition, validity shows the test suitability to a greater or lesser extent for being applied in a certain situation with certain goals. In other words, validity interacts with the objectives of the test and plays a crucial role as the supporter of the major characteristics of the test (Clapman, 1996). L. Bahman considers adequate implementation of the validity procedure as follows: firstly, to provide a clear theoretical definition and definition of the skills to be necessary measured; secondly, to specify the conditions and actions where certain language skills are to be applied and, thirdly, to obtain possible quantitative indicators for the sound measurement results interpretation (Bachman, 2001, Clapman, 1996).

Consequently, we understand the importance of such a characteristic as the validity of the test, which determines what the test measures and how well it does that. In other words, validity shows how the test is used for certain purposes. We distinguish the following methods for measuring the validity: 1) the method based on the analyses of the test content; and 2) the method lies in the correlation of the calculated coefficient that is a numerical test description. Some sources in the field of research reveal the following methods: a method based on various experimental procedures or peer review procedures, like in case the numerical characteristics cannot be obtained by the results, are the basis for the test validation procedure; a method based on a series of theoretical and experimental studies such as, at first, test structure is outlined; the hypotheses of this structure functioning is empirically formulated and tested, however, like the previous method mentioned before, numerical index is not a characteristic of the methodology validity, but only the possibility to pronounce definite outcomes in testing. Actually, in order to determine the most effective evaluation tool it is necessary to carry out the results analyses for the numeric indicators of the validity to be obtained.

The analysis of the relevant scientific and technical literature gave us an opportunity to highlight those common *types of validity*: *content validity*, *construct validity*, *concurrent validity*, *predictive validity*, *face validity* (Bachman, 2001, Fulcher, 2007). Let's view these types of validity briefly for their better understanding.

Content validity is a relevancy test and embraces the semantic sphere, in other words the test must correlate with the recommended language level thoroughly outlined in the syllabus (Common European Framework..., 2002).

Construct validity is a theoretical model of target skills to be measured. We consider the English professionally oriented competence being the test construct and the reading speech skills are the targets of such a construct (Hryshyna, 2011). The construct validity is the main criterion of the test validity indicating the reflection of the language skills within the test development, following the L. Bahman's research (Bachman, 2001).

Significantly, the standardization process of the test development according to the pre-set procedure of the strictly articulated construct guarantees the test validity. This type of validity is important, but it is not a mathematical property. It can be determined by a thorough analysis of test targets, evaluation criteria and empirical evidence (because there is no any valid test) are more or less reasonable interpretation of testing procedure itself.

Besides, we carried out a thorough analysis of the speech material and as a result, the accurate speech material selection for the testing was completed (Gryshyna, 2013). It is obligatory to refer to the requirements of the current syllabus outline taking into account all the pre-set speech targets alongside with the testees' interests. However, in addition to analysing and describing relevant materials, test content background and the construct validity positive outcomes, it's important to perform mathematically specified validity calculations in order to have the objective results introducing the test as an applicable and suitable tool for reading professionally oriented speech skills measuring.

The main characteristic of the *concurrent validity* is the ability to measure the validity of the pre-set test targets compared with the other test with the same construct. The results of one test are compared with the results of another test and a decision is taken with reference to the higher degree

of validity, i.e test construct correlation is a must. In particular, the specified validity shows such a correlation between the individual average score performance of the test task of a particular type and an average score of the student's progress in speech skills acquisition. The correlation, i.e the interaction density, is expressed by the correlation coefficient. Indicator of the concurrent validity, or the degree of real test task, is the correlation coefficient between their different results and the results of other test tasks or teachers' evaluation.

Predictive validity is the simulation of the effectiveness of the test, after the results of tests having being compared or correlated within a certain period of time.

Face validity ensures the test to be an appropriate means of measurement concerning the selected professionally oriented speech materials. Furthermore, the face validity is presented through the simulated content and test instructions of the professionally oriented communicative situations and the scientific stem of a test task that is accurately selected and written. Moreover, a certain number of items in the test task are a component of face validity and the indicator of the skills to be measured in different semantic ways.

With the reference to all the mentioned above, we consider the construct validity to be the most suitable quality characteristic of the communicative subtest and test task aiming at an objective measurement of the foreign speech reading competence.

Secondly, such quality as **reliability** aims at the stability of the test results. Being a reliable test means having the same results during the second performance. Reliability is one of the main characteristics of the test that is aiming at an objective measurement of the foreign speech reading competence (Kokkota, 1989). Such quality is based on the mathematical-statistical analysis of the test results.

Thirdly, the following quality of the test is **efficiency**. Mostly, it covers the following characteristics: 1) compact material; 2) rational test structure for subtests organisation; 3) time for test completion and time for evaluating. Efficiency provides the performance of accurate results within the pre-set goals of the test procedure during the repetitive completion of the test. Here, the foreign language competence is to be measured with minimum time on development, implementation and the time spent on the results processing. Also, test efficiency provides rational dislocation of test tasks in the subtest with the further test organization as a whole measurement tool. The right choice of test tasks in terms of the test format provides higher efficiency of the test performance (Kokkota, 1989). We assume the placement of tasks must be done according to the principle of the increasing complexity.

Thus, the efficiency stands as one of the main qualities of the subtest/test to objectively measure the foreign language competence in special reading of future engineers.

The fourth quality is considered to be **practicality** covering the following characteristics: 1) guidelines and test content accessibility; 2) testing procedure organization is relatively simple; 3) the possibility of testing procedure organization in universities must not be complicated concerning the technical facilities; 4) relatively simple processing of test results. Practicality provides simplicity of test development and the possibility of necessary resources accessibility.

The **item difficulty** (*facility*) is the fifth test quality being considered in our research. The complexity determination of a test becomes possible due to the percentage of the correctly answered items in the test. In this way, it is also possible to determine the quality of the proposed alternative distracters having analysed the percentage of the correct chosen distracter (Heaton, 1990, Huges, 1996). In our experimental study, we carried out the analysis of the test responses, which allowed us to determine the complexity and functionality of each test item.

The sixth important quality of the test is considered to be **authenticity**, i.e the maximum reflection of the foreign language content and the application of original speech material for the test development (Gryshyna, 2013, Petraschuk, 1997). Consequently, the effectiveness and efficiency of the test organization and procedure are caused by the pre-set qualities requirements. Among them, the most important requirement for the communicative test is the use of the authentic speech material. It goes with the main testing communicative/interactive principle in language teaching as well.

The realization of the interactive/communicative principle is necessary to be ensured simultaneously with the authenticity in testing development and administration.

Analyses features of the communicative approach of testing are being claimed as the main pronounced principles. As a result, definitions of the general methodology principles allow us to envisage and stipulate the main ways of implementation for testing of the communicative foreign language special reading competence.

The analysis provided the development of the test development for communicative testing and enabled an objective assessment and evaluation of the foreign language competence in special reading within the communicative approach framework. Professionally oriented communicative test tasks are the main component of the interactive testing methodology.

Test Content Organization Impact. With the reference to the qualities defined, the possibility of the specific ways how to create the interactive test tasks concerning all the requirements and test qualities has become evident. Actually, the test quality required as communicative and interactive characteristic causes the defining of the first one. Reasonably, it deals with the professionally oriented speech situations where a future engineer is going to use the English language as a primary tool for special goals while reading. Noteworthy, a brief summary of reading goals was given before. The professional situations were selected before the analyses of the engineers' needs in their academic study and research (in our study such situations were introduced as a specific system of typical professional situations we designed in the communicative testing Gryshyna, 2011).

According to the scientific methodology sources (Common European Framework..., 2002), Konisheva, 2004, Petraschuk, 1997), the future engineers often need to use English for specific purposes when taking part in the international conferences, at the library, using on-line library sources, while studying the professional literature for the writing articles or abstracts in English etc. We have carried out the modeling real-life domain situations in the test instructions. The modeling was performed as follows: 1) the creation of the professionally oriented situation according to the specified structure; 2) clear representation of the communication objectives in the test instructions caused by the future engineers' search for information while reading academic texts. Being an important factor influencing the correlation of the test task profile and test construct in terms of the communicative approach, the authentic material was selected carefully for reading test development. Indeed, the objective measurement of the reading competence is provided by careful presentation of the academic texts in test administration. We focused exclusively on the selection criteria in the previous studies.

Real-to-life Test Profile Features. The example of how to model the communicative situation lies in a test task specification written in terms of the speech targets of measurement – *to associate the scientific facts with the previously received data*; purposes of reading, kind and type of test task – *reading for details and the sentence completion*; the type of the communicative task in situation – *to search for information*; the measurement techniques – *objective*; and the situation itself – *the title and the number, according to the outlined system*.

Situation. You are a member of a scientific organization at your technical department. The second meeting is devoted to the problems of the metals extraction. You want to study the scientific paper to search for some new information.

Complete the following scientific statements according to the paper "Extraction of metals". The first has been done for you.

1. The field of extractive metallurgy encompasses many specialty sub-disciplines, each concerned with ...

- a) various physical and chemical processes*
- b) different field of material science*
- c) the most important chemical process*

Paper 1. Extraction of metals

[...] Processes for separating the metal from the impurities it is found with or other elements with which it is combined depend upon the chemical nature of the ore to be treated and upon the properties of the metal to be extracted. Gold and silver are often removed from the impurities associated with them by treatment with mercury, in which they are soluble. Another method for the separation of gold and silver is the so-called cyanide process. The Parker process, which is based on

silver being soluble in molten zinc while lead is not, is used to free silver from lead ores. Since almost all the metals are found combined with other elements in nature, chemical reactions are required to set them free. These chemical processes are classified as pyrometallurgy, electrometallurgy, and hydrometallurgy [...]

The given above example of the test task situation with all components are presented. Besides, we introduced in detail the process itself of the test task writing and the examples of the specification writing were given in the previous studies.

Modeling of the Communicative Situation:

Situation components:

1. Participants of interaction; place of interaction.

Test task construction example: *...a member of a scientific organization at your technical department.*

2. Communicative and cognitive speech task:

Test task construction example: search for information task:
to search for some new information

3. Sphere of Activity :

Test task construction example: academic:
to study the scientific paper

4. Kind of activity:

Test task construction example: *the preparation and gathering of the authentic information for the scientific and research study*

5. Output results:

Test task construction example: *the scientific and professional interpretation in Ukrainian*

6. Scientific text:

Test task construction example: *review paper "Extraction of metals"*

It is of primary importance to note that the accurate and simple pronunciation of the instructions is one of the main issues in the test design due to the real life profile domain it must have. Clearly and correctly formulated instructions cause the testees to solve the interactive cognitive tasks. Furthermore, the test instruction is a core of the test task offered in writing and in other words, definite organization and sequences of actions, the relative importance of parts of the task, time performance and the evaluation procedure (Hryshyna, 2011).

The instruction of the communicative test task includes the professional real life situation where student-future-to-be-engineer decides how to complete the task in the foreign language reading communication media. Moreover, such communicative instructions comprise an indirect measurement of the certain types of speech skills competency in special reading by future engineers. The speech skills in special reading determined are to be the construct of the communicative test in a whole.

Example:

Reading paper 2.

Situation. You are preparing for the seminar where the international guests are going to be present. You need to organize the following scientific texts in one article. To achieve this goal the arrangement of the parts is should be done according to the concept of the topic. Your teacher asked you to define the best sentences which convey the most important facts of the text.

Test criteria. The communicative testing technology requires a strict number of criteria to create a test task. The analyses provided a possibility to single out and group the following requirements as follows:

quality requirements (quality requirements appropriateness: reliability, validity, authenticity, interactivity, efficiency, item difficulty);

the content testing requirements (the test construct language/speech targets/purposes interaction, teaching domain and interaction of the types of test tasks, the proficiency level and program outline objectives interaction, the problem-solving profile, real life situation domain and

educational value, maximum communicative degree, the authentic interactive test construct and design, test task functionality id est. the compliance of the nature and content of the tasks with the needs and interests of testees, cognitive value, innovation features);

physiological requirements (compliance of test tasks with the testees' psychological age, sufficient motivation, clear and understandable instructions and test specifications that encourage the testees to use the foreign language for specific reading in order to obtain the information) (Heaton, 1990).

Summary and Perspectives. The research suggests the important features of the testing methodology comprising the development of special reading test construct, the design of a test, the test item construction. The test qualities and requirements based on outlined communicative testing principles provide the design of the test for the academic reading skills evaluation. Without a doubt, the outlined scheme of the situation modeling helps to understand and apply the techniques of the communicative reading test construction to design test. In terms of specific test qualities and requirements it will be possible to accurately elaborate the tools for objective evaluation of the students' reading competence.

References:

- Alderson, J. Ch. (1995). *Language Test Construction and Evaluation*. Cambridge : Cambridge University Press.
- Bachman, L. F. (2001). *Fundamental Considerations in Language Testing*. Oxford : Oxford University Press.
- Clapham, C. M. (1996). *The development of IELTS : A study of the effect of Background knowledge on reading comprehension* (Studies in Language Testing). Cambridge: Cambridge University Press.
- Common European Framework of Reference for Languages : Learning, Teaching, Assessment. Language examining and test development* (2002). [prepared under the direction of M. Milanovic (A.L.T.E.)]. Strasbourg : Language Policy Division.
- Fulcher, G & Dadidson, F. (2007). *Language Testing and Assessment*. London and New York: Routledge.
<http://dx.doi.org/10.4324/9780203449066>
- Heaton, J. B. (New ed) (1990). *Writing English Language Tests (Longman Handbooks for Language Teachers)*. New York : Longman.
- Hughes, A. (1996). *Testing for Language Teachers*. Cambridge: Cambridge University Press.
<http://dx.doi.org/10.1017/cbo9780511732980>
- Weir, C.J. (2005). *Language testing and validation: an evidence-based approach*. London : Palgrave Macmillan.
- Hryshyna, O. A. (2011). *Problema stvorenniya profesijno orientovanoj komunikativnoj situacii v chitanni anglijs'koyu movoyu majbutnimi inzhenerami* [The design of the communicative professionally oriented situations]. Nizhin: NDU im. M. Gogolya.
- Gryshyna, O. A. (2013). The selection of the scientific speech material for the communicative testing of the competence in reading of the future foundry engineers. *The Advanced Science Journal*, 9-12.
- Kokkota, V. A. (1989). *Lingvodidakticheskoe testirovanie : nauch.-teoret. posobie* [Linguistic testing: the theory manual]. Moskow, Russia: Vysshaya shkola.
- Konisheva, A. V. (2004). *Sovremennye metody obucheniya anglijskomu yazyku* [Modern English teaching technologies]. Minsk, Belarus: Tetra Sistemz.
- Petrashchuk, O. P. (1997). *Avtentichnist' yak oznaka komunikativnogo testuvannya* [Authenticity of the communicative testing]. Kyiv, Ukraine: Inozemni movi.

Received: 10.05.2016

Accepted: 25.05.2016