

Data Processing Technologies for Calculating Prognostic Validity of Educational Achievement Tests

Serhiy Rakov¹, Mariia Mazorchuk², Viktoriia Dobriak²

¹ Journal «Visnyk. Testuvannya i monitorynh v osviti», Bakulina, 11, of. 4-28,
Kharkiv, Ukraine

rakov_s@ukr.net

² National Aerospace University "Khal", Chkalova str. 17,
Kharkiv, Ukraine

mazorchuk.mary@gmail.com, viktoriya.dobryak@gmail.com

Abstract. Calculating test validity as an instrument of educational achievement level determination, is one of the main research tasks, especially on a state level. The system of HEI admission must be based on the principles of fairness and openness. This can only be achieved by conducting systematic admission test result analysis both in the context of organization processes and within the assessment of the system efficiency in general.

One of the main tasks in psychometry is the analysis of prognostic test validity. Calculating test validity cannot be performed once, it should be a permanent process, mainly because changes of external and internal factors influencing educational system cause changes in assessment mechanisms. Naturally, conducting permanent monitoring and analysis is impossible without using modern informational technology means and tools.

Existing data systems and technologies in Ukraine don't allow to conduct full analysis of External independent testing (EIT) quality. Data analysis is done with the use of modern methods and technologies; however, the analysis is not conducted regularly, it is often incomplete; and the process of data-gathering is complicated and not standardized. Results, which researchers receive on different stages of the analysis can have mistakes, and this can cause misjudgments in future.

In the paper the results of prognostic validity research conducted for EIT from 2008 to 2014 are represented. The analysis of problems with data-gathering, processing and interpretations is shown. The perspectives of the development of unified data system for assessment of EIT quality in Ukraine are studied.

Keywords: Validity, Assessment, Test, Education, Information Technologies, Data Processing, Data Analysis.

1 Introduction

Information technologies are the powerful instrument of educational system innovative development. Information technologies are widely used in e-learning, as well as in the open online courses. There are a lot of papers on using these technologies for computer

adaptive testing. Efficiency and availability of modern data processing systems and technologies give almost unlimited opportunities for big data analysis and for conducting research and in education. Information technologies gave the opportunity to conduct HEI admission exams in Ukraine in the form of standardized national testing. These exams are called External Independent Testing (EIT).

HEI admission exams must provide equal access to education and should be based on principals of transparency, openness, fairness and social acceptance. Admission exams in the form of testing are conducted in many counties, as testing technologies are the most effective for educational achievement assessment [1-3]. The EIT system was introduced in Ukraine in 2008. Since then it has been used as an instrument for equal access to high education. Enrollment conditions have changed several times since the launch of EIT. Until 2010 EIT had been the only criteria for students' admission, in 2011 the grade point average was introduced as one of the admission criteria. It is worth mentioning that the algorithm of grade point average use has changed several times.

How well does the admission system based on the EIT tests assure the quality of the students' selection? To answer this question researches have to analyze the quality of EIT test items, EIT tests in general, test administration and preparation, as well as the reliability of EIT grades, etc.. Much attention to the issues of authenticity and "clearness" of data is also paid in the other countries, where tests are used for HEI admission. For instance, a whole range of standards for test development and its quality analysis is developed in the USA [4-6].

Ukrainian Center for Education Quality Assessment carries out and publishes the results of such statistic researches annually. However, existing research results do not answer the question about EIT tests validity. That means that one cannot answer the question on how well EIT assures qualitative selection of students for HEI admission according to meritocratic principals. To answer this question it is essential to investigate not only the internal test characteristics (for example, reliability of the test can be evaluated based on single testing results conducted to a certain population of participants), but to conduct analysis of the EIT results in conjunction with students' educational achievements at the universities.

In such a way, it is necessary to research, how efficiently criteria selection of enrollees lets prognoses successfulness of studying in university. In terms of mathematical statistics such correspondence expresses by correlation between dimensions of properties of students selection criteria with properties of quality their education in university.

Here is a problem of how to provide collecting and analyzing the current results of student's education in universities, who enrolled by results of the EIT and evaluate the correlation between these values. It is almost impossible to provide relevant gathering the qualitative data and it's analyzing at the state level without using any modern information technologies. It is essential to develop the approaches, methods and methodology for gathering, saving, processing and analyzing of data aimed to get qualitative and reliable property at tests validity is essential.

The aim of this paper is to provide results of the EIT tests validity analyzing research with use of the existent technologies and approaches, and also to form the principals of collecting and processing of data, providing further qualitative analysis and implementing in practice.

2 Input Data for Research

Prognostic validity is the most important property of the EIT tests quality. A lot of works was dedicated to researching of validity beginning from the 50's of the previous century [7-9], when the theory of psychometrical analysis was only arising, ending and beginning of 20's century [10-13], where main principals and tends of validity research were formulated, and until now [14-16]. For instance, just to the problem of validity research measuring the conference in area of educational measuring of 2015 was dedicated, which was holding in Kansas, the USA [13]. The main points of the validity studying, are presented in reports [17-18].

As it was mentioned earlier, the prognostic validity is a result of correlation between test marks and results of education in university. Currently the criteria of enrolling the universities of Ukraine is estimating the sum of three components: grade point average (GPA), grade average of EIT (GAEIT), grade average at creative exams (GACE). In its turn GPA consists of marks at many subjects, GAEIT – at three subjects of EIT. Creative exams also can be up to three.

The criteria of successfulness is usually the grade average of the 1-st course exams (by results of winter and summer session - Grade average of session). This choice is explained by the fact that, the first term of education is mostly adaptive, what influences the marks, on the other hand, with time passing the influence of original factors increasingly levels out.

We must consider, that the amount of disciplines, studied on the 1 course, at which there hold exams comes up to 10 and more, and they also refer to the different categories: general, professional, compulsory and optional (by choice) education. Descriptions of the same modules can be different depending on the universities, educational programs, and academic hours of studying. Various educational and professional fields also differ by their specifics. Universities differ her by sizes (by amount of students: big, medium and small), by form of property, by location, etc. Students can study “in presentio”, “in absentio”, remotely; study on a state-fund education or on contract. Students differ by age, gender and etc.

All abovementioned factors can influence the results of studying. The amount of data to be analyzed is enormous. For emphasizing significant parameters, as a rule, factor analysis is held. That's why it is necessary not only to describe and classify data by given conditions for further research, but also to develop the analytical information system (AIS) “Quality of enrolling universities basing on EIT”, by means of which anyone can study the aspects of prognostic validity he is interested in online.

For these reasons, validity research focuses on the analysis of prognostic validity of separate criteria of enrolling the university and their combinations, where the university educational level is estimated by the average grade after the 1-st course.

3 Data Analysis Technologies for Validity Research

Validity research of EIT tests was held according to CRISP-DM methodology [19, 20], describing inter-industry standard process of data research (see Fig. 1).

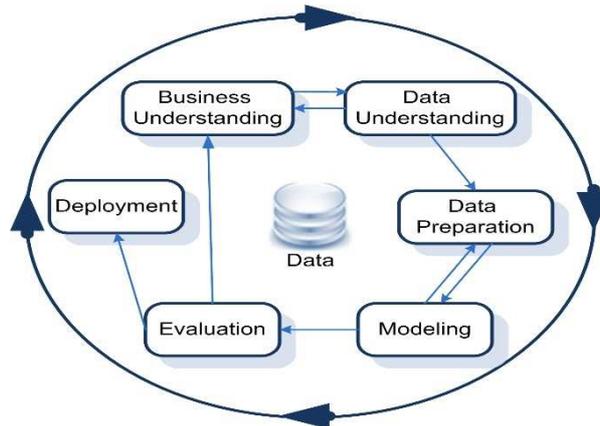


Fig. 1. Phases of the CRISP-DM Reference Model (Source: Shearer C., The CRISP-DM model: the new blueprint for data mining, J Data Warehousing (2000); 5:13 - 22.).

According to the stages of the process the first stage requires understanding the subject area. Here conceptual and structure analysis of the tests and EIT participants are held and the main aspects of the tests validity analysis are studied.

On the following stage primary data studying takes place. Here we determine types and scales of measuring key parameters, which reflect crosscuts of research, and also current results of education.

Preparing data for processing takes place next. This stage is the most difficult and continual, as the process of data gathering is decentralized, and even the data obtained with standard forms of data input is very “dirty”. The main problems are: different types and scales of measuring the same conditions; missed data; overshoot; inadequate meanings; sparse data; alternation of missed meanings by zeroes; mismatch amount of data fields in files; little sample by separate universities; absence of resulting parameters. That’s why there took place the processes of the exploratory analysis and data cleaning, decoding, transformation, filtration, calculation of result statistics and etc. Current stage takes more than 80% time from the whole process of data processing.

During the next stage the modeling is performed. Modeling is searching for conformity and building of models. The choice of correlation coefficient for validity analysis has been validated and the factor analysis has been held. The research completed with evaluating the quality of the received data.

The results of the validity research for 2008-2011 were published in [21]. This is the stage of implementing. Results can be used for further improvement of the tests quality, testing process and system of enrolling to the universities. Currently there results for 2012-2014 years are processed and the report is in work.

4 Numerical Results

This research was held by Union of principals of the universities in Ukraine cooperative with Institute of higher education Academy of pedagogical sciences of Ukraine during 2015-2016 years. The request-letters to provide the information about the first year students results in 2012/2013, 2013/2014, 2014/2015 educational years, information about EIT of these students, and also the results of entrance exams and creative contest in case, if they were held have been sent to 102 universities of Ukraine at the end of 2015. The information is provided in the Excel-table, which is applied.

Responses were received from 55 universities. There were represented the following information about students, enrolled these universities in 2012, 2013 and 2014 years:

1. Student's name (not always – some universities presented only name and patronymic, some didn't performed name at all, referring to the Personal Data Protection Act of Ukraine).
2. Field of education and professional field.
3. Grade point average about general second level of education.
4. Results of passing external independent testing.
5. Results of internal enrolling exams of universities (creative contests).
6. Winter session assessment (2012-2013, 2013-2014 и 2014-2015 years).
7. Summer session assessment (2012-2013, 2013-2014 и 2014-2015 years)

Some universities fed information in wrong form, so it became impossible to analyze their results.

On Fig. 2 performed description of one of the Excel-files with data results before processing.

1	Прізвище	Ім'я	По-батькові	Код ВНЗ	Код напрямку(спеціальності)	Середній бал з екзамена	Укр. мова і літер.	Математика	Історія	Іноземна мова	Фізика	Хімія	Біологія	Географія	Зарубіжні(світові) література	Всесвітня історія	Російська мова	Теоретич конкурс	Спеціальні	Укр. мова	Історія України	Іноземна мова2	
2	Аксев	Олександр	Олександрович	16.010103	164	127	124				134											4	2
3	Бабак	Катерина	Викторівна	16.010103	147	157	142				172											4	4
4	Борисова	Анна	Сергіївна	16.010103	171	127	124				142											4	2
5	Данда	Олена	Петрівна	16.010103	175	147	125				169											4	2
6	Ковач	Світлана	Сергіївна	16.010103	144	146	157				154											5	4
7	Лобода	Сергій	Анатолійович	16.010103	137	142	153				152											4	4
8	Лях	Ольга	Юрківна	16.010103	177	154	124				129											4	2
9	Макарич	Вадим	Костилівич	16.010103	150	139	125				154											3	2
10	Онищ	Олександр	Олександрович	16.010103	154	154	125				154											4	4
11	Радуть	Андрій	Сергіївич	16.010103	172	124	144				167											5	3
12	Савинов	Владислав	Олександрович	16.010103	170	167	167				145											3	4
13	Сименко	Ігор	Геннадійович	16.010103	171	125	157					139										3	4
14	Тимченко	Оксана	Ігорівна	16.010103	155	153	148					124										3	3
15	Олейнік	Антон	Олександрович	16.010103	162	141	166				151											5	3
16	Ангелов	Володимир	Віталійович	16.010104	160.0	152.5	155.5				159.0											4	4
17	Біланок	Дмитро	Володимирович	16.010104	177.2	149.0	160.5				151.5											4	3
18	Гавриш	Артем	Віталійович	16.010104	200.0	174.0	165.5				174.5											5	5
19	Кліменок	Віталій	Анатолійович	16.010104	163.0	134.0	170.5				178.0											2	ніз
20	Козубенко	Світлана	Олександрівна	16.010104	168.5	143.0	160.5					146.0										5	3
21	Корочен	Ірина	Валерійівна	16.010104	159.5	177.5	159.0				140.0											4	4

Fig. 2. Data results before processing (Source: Own work)

After verification there was provided a data base of research, that contained the information approximately about 30000 of students (information included the EIT results for 2012, 2013 and 2014 and results of first and second terms sessions of revealing educational years).

The data was processed with program instrumentation for statistics analysis SPSS, however many procedures of preprocessing were made in Excel and Python. To save the data the power of PC was enough.

Validity analysis was held on different levels of data aggregation in crosscut: countries, field of education (professional field by classification of 2015), universities (55, which have open source participated in research). There was also valued incremental prognostic validity for different student's categories (type of ownership of universities, their location, cutoff by EIT results, field of education, by EIT subjects). For example, Table 1 provides general parameters of entrance and education results by selection of students in whole by the country for 2014.

Table 1. General characteristics of parameters of student's entrance and education results in 2014

Parameter	Validity meanings	Minimum	Maximum	Mean	Std. Deviation
Grade point average	28727	0,00	60,00	47,55	7,06
Grade average EIT	29453	115,85	200,00	162,27	14,57
Grade average at creative exams	2181	110,0	200,00	165,86	18,83
Grade average of session	29516	0,00	5,0	3,56	0,60

Table 2 provides general results of validity values by all universities. As it is clearly from results, amount of participants is less, then in original selection. At that validity parameter of creative contest and session results are low. That caused first of all by multivalued data, that was not managed to be equated correctly to one scale. Secondly, we have a small sample in many university, which has had a negative impact on results. However, by separate universities, validity parameters were different and not always low, as within the limits of the university or field of education scale of measuring marks had no influence on correlation coefficient.

Table 2. Values of prognostic validity for 2014

Prognostic validity	Value (p<0,01)
Grade point average (GPA) - Grade average of session	0,444
Grade average EIT (GAEIT) - Grade average of session	0,454
Grade average at creative exams (GACE) - Grade average of session	0,295
GPA+GAEIT - Grade average of session	0,498
GPA+GAEIT+GACE - Grade average of session	0,550

Fig. 3 presents the common validity values for all years of analysis. As it is clear from results, the correlation of the grade average of EIT with results of session after 2011 is higher, then GPA (after 2011 there rules of enrollment were changed: to EIT marks were added GPA).

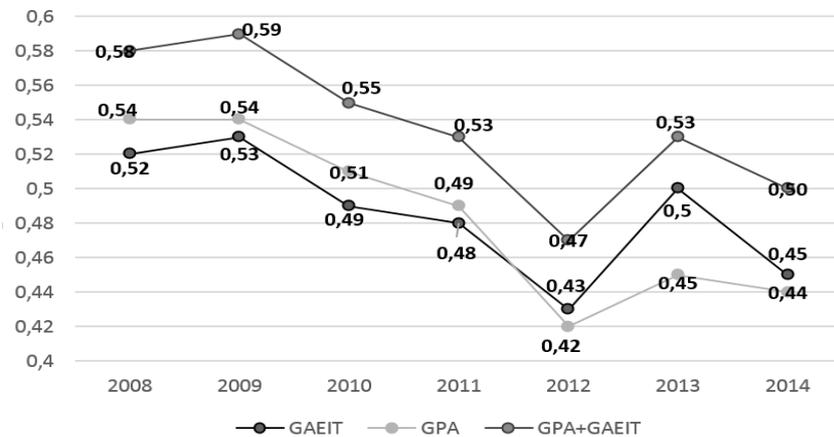


Fig. 3. Dynamics of predictive validity testing, GAEIT, GPA, GPA+GAEIT 2008–2014 (Source: Own work).

5 Conclusion

Evaluating test validity of educational achievements, used for enrolling the universities, requires permanent monitoring the student's educational results. Gathering this data from universities of Ukraine is problematic, as up to now there is no unified information technology of gathering and saving personified and personalized data of EIT participants and their results of education. There is no unified system of titles of subjects in universities and their classifications.

According to the results of validity researches of previous years, we can make a conclusion, that the existing approach doesn't provide reliable results. Data by educating of students has a lot of missed meanings, different scales of assessment, overshoot and inadequate meanings, what first and foremost equated to absence unified methodic and means of gathering of data. Many factors, that can be significant and influent on validity are not available to study. We cannot analyze the results of studying in dynamics. There are still problems with creating the representative selection for deeper research of tests and their influence on educational system.

Thus, the mentioned problems are the basis for further research and investigating new approaches, methods and technologies for tests results processing. To solve these problems it is essential to unite forces of different specialists in different areas of education, information technologies and data analysis.

References

1. Lane, S. Performance assessment: The state of the art. (SCOPE Student Performance Assessment Series). Stanford, CA: Stanford University, Stanford Center for Opportunity Policy in Education (2010).

2. Lane, S., Raymond, M.R., Haladyna, T.M.: Handbook of Test Development (1st ed.). London: Routledge (2011).
3. Haladyna, Thomas M.: Perils of Standardized Achievement Testing. *Educational Horizons*, vol. 85(1), pp. 30–43 (2006).
4. ETS standards for quality and fairness: Educational Testing Service, <https://www.ets.org/s/about/pdf/standards.pdf>, last accessed 2017/02/07.
5. Standards for Educational and Psychological Testing (AERA, APA, & NCME), <http://www.teststandards.org/>, last accessed 2017/02/07.
6. International Test Commission: International Guidelines for Test Use (*International Journal of Testing*, https://www.intestcom.org/files/ijt_testuse_guidelines.pdf, last accessed 2017/02/07).
7. Cureton, E. E.: Validity. In E. F. Lindquist. *Educational measurement*. Washington, DC: American Council on Education, pp. 621–694 (1951).
8. Cronbach, Lee J., Meehl, Paul E.: Construct validity in psychological tests. *Psychological Bulletin*, <http://psychclassics.yorku.ca/Cronbach/construct.htm> (1955), last accessed 2017/02/07.
9. Cronbach, L. J.: Test validation. *Educational measurement*. Washington, DC: American Council on Education, 2nd ed., pp. 443–507 (1971).
10. Moss, P. A.: Shifting Conceptions of Validity in Educational Measurement. *Implications for Performance Assessment (Review of Educational Research)*, vol. 62(3), pp. 229–258 (1992)
11. Messick, S.: Validity of psychological assessment. *American Psychologist*, 50, pp. 741–749 (1995).
12. Fair admissions to higher education: recommendations for good practice. Admissions to Higher Education Steering Group, <http://dera.ioe.ac.uk/5284/1/finalreport.pdf>, last accessed 2017/02/07.
13. Moss, P.A., Girard, B., Haniford, L.: Validity in educational assessment. *Review of Research in Education*, 30, pp. 109–162 (2006).
14. Sireci, Stephen G.: On Validity Theory and Test Validation. *Educational Researcher*, 36(8), pp. 477–481 (2007).
15. Moss, P.A.: Reconstructing validity. *Educational Researcher*, 36 (8), pp. 470–476 (2008).
16. The Three Most Important Considerations in Testing: Validity, Validity, Validity (41st Annual IAEA -2015 Conference), http://iaea2015.cete.us/sites/default/files/IAEA_conf_2015_flier.pdf, last accessed 2017/02/07.
17. Differential Validity and Prediction of the SAT: Research Report 2008-4, <https://research.collegeboard.org/sites/default/files/publications/2012/7/researchreport-2008-4-differential-validity-prediction-sat.pdf>, last accessed 2017/02/07.
18. Validity of the SAT for Predicting First-Year College Grade Point Average: Research Report, No. 2008-5, <https://research.collegeboard.org/sites/default/files/publications/2012/7/researchreport-2008-5-validity-sat-predicting-first-year-college-grade-point-average.pdf>, last accessed 2017/02/07.
19. Shearer, C.: The CRISP-DM model. The new blueprint for data mining. *Journal of Data Warehousing*, vol. 5 (4), pp.13–22 (2000).
20. Shafique, Um., Qaiser, H. A: Comparative Study of Data Mining Process Models (KDD, CRISP-DM and SEMMA). *International Journal of Innovation and Scientific Research*, 12(1), pp. 217–222 (2014).
21. Study of the quality of students competitive selection of higher educational institutions by results of external independent testing: analytical materials. Under the editorship of Countance V.V. and Rakov S.A., 160 p., Ukraine (2015).