Journal of Social Sciences *Vol. 1, no. 2 (2018), pp. 30 - 35*
Fascicle *Social Science* *ISSN 2587-3490*
Topic *Pedagogy and Psychology* *eISSN 2587-3504*

DOCIMOLOGY – A MODERN PERCEPTION WITH A CENTURY STAGE OF DEVELOPMENT OR THE OTHER SIDE OF EVALUATION

Ala Șișianu

T*echnical University of Moldova, 168 Ștefan cel Mare și Sfânt Bd , Chișinău, Republic of Moldova*
ala.sisianu@ia.utm.md

Received: September, 20, 2018
Accepted: October, 28, 2018

Abstract. In the article there was made an attempt to rediscover the core and significance of docimology as a science, in general, and of the docimological test, as basic method of knowledge evaluation, in particular, as well as to introduce the principles and conditions for gaining essential results while being evaluated, and to expose the circumstances that appear as obstacles in attaining the success when students try to pass an exam.

Keywords: *docimology, test, result, principle, validity, reliability*.

CZU 371.26

Introduction

The beginning of the 20th century represents the debut of the first scientific researches in the field of academic evaluation, initiated by the H. Pieron, who named this preoccupation docimology. Today, after approximately an entire century, the idea, importance and principles of the very subject are on the same level of interest.

The term "*docimology*" has Greek roots: *"dokime"* which means trial, test, and *"logos"* which means science; so docimology means the science of tests, exams. An exam is a form of social evaluation, by which one realizes a brief evaluation at the end of learning period (the BAC marks, the end of the high school, the university studies end with the university degree exam, the final exam after attending a course), and by graduating it, one can obtain a diploma, which allows the possessor to occupy a "social role" [1, p. 140].

Another side of evaluation

The psychologist Vasile Pavelcu considered the entire period of our existence as a succession of exams, which marks the ending of certain steps in the life of the individual. Due to their social importance, these exams are criticized severely, being reproached the absence of some proper evaluation instruments and the strong subjectivism in grading.

The researches which have been done have shown a multitude of factors that appear purposely or unintentional in evaluation, generating a low objectivity in estimating the results. These factors can be grouped in several categories, being reported to [2, pp. 23-27]:

1. Teacher: the "halo" effect, the "kind" effect, the generosity error, the Pygmalion effect, the "contamination" effect, the "contrast" effect, the examiner's personal equation, the error of central tendency, the logical error, the effect of Gauss curve, the teacher's personality factors;

2. The subject referred to: the papers in subject such as Physics, Math, or Chemistry can be evaluated more objectively than the papers in subjects such as Philosophy, Literature, etc.
3. Learner: personality particularities – temperament, mood, age, skills, abilities;
4. The social circumstances in which the evaluation is performed: the leaders, the mates, the parents' interventions on the examiner for a certain student, the tolerated deviation of cheating.
5. The authors experience allows us to stand that these negative factors are not enough known by the staff in general, but especially to ones in technical superior education. The phenomenon of the tolerate deviation of cheating is one of the most harmful and hard to analyze and to discourage.
6. In these conditions, the validity of the obtained information is doubtful. The quality of this information strictly depends on the objectivity of the evaluation process and on the quality of the used instruments.

The testing process will be objective if:

7. The process of applying the test will be objective - the same task given to all the students under the same condition;
8. The results will be objectively amazed by imposing a straight criterion of evaluation or a sample correct answer, subjectivity being in this way reduced to minimum;
9. The results will be objectively interpreted, meaning that the same performances are evaluated and marked in the same way by different examiners.

One of the main directions the contemporary docimology sustains concerning the objectivity of the learning evaluation is using the docimological tests, founded on the base of docimological principle. The docimological principles are "fundamental theses, general rules with descriptive and normative character in order to ensure their scientific consistency and efficiency" [2, p. 85].

The most relevant docimological principles, which lead evaluation activity, are:

1. The principle of evaluation objective character that refers to the structure and organization of evaluation, so that the learners' performances to be reflected and evaluated in a real and relevant manner, reducing as much as possible the influence of external factors;
2. The principle of evaluation interactive character expresses the fact that learning evaluation is inherently connected and determined both by the evaluation made by the teacher and by the pupil's self-evaluation activity;
3. The principle of the learner's performances contextualization: regards the fact that during the evaluation there have to be considered the performances and there have to be used such tasks that could reflect the reality, meaning to attach the pupil's capacity to adapt the knowledge to various situations.

The docimological test

As basic method of objective evaluation, the docimological test is “a set of questions with the help of which one can check and evaluate the knowledge and the capacities acquiring to operate with them, by reporting the answers to a sample appreciation scale, previously elaborated” [3, p. 401]. In the specialized literature, we will also come across other terms: pedagogical test, knowledge test, learning evaluation test, performance test or simple test to designate the instrument and method of evaluation that has a specific

element - the item, being characterized by a greater objectivity in evaluating the results. The quality of information offered by testing depends on two sets of attributes the test has to possess:

1. Psycho-pedagogical: the test must be appropriate to its specific purposes and comprehensive;
2. Statistical: which guarantees the perfection of the test as a measurement instrument; the most important being the accuracy and validity.

The accuracy of the test, also called constancy or exactness designates the trust we can have in the respective instrument, the degree of exactness of measurement.

The principal conditions a test must have in order to possess this quality refers to the stability of the results:

1. the scholar is being examined by different teachers, who do not know the previous results of the examinee;
2. the conditions in which the testing is made are modified. Because of this, a test applied in different objective and subjective conditions has to touch almost identical results. The most important objective conditions are: the position of the desk in the classroom and the examinee in the desk, the lightning in the classroom, the weather, the degree of the room ventilation, the moment the exams is being taken, the influences coming from the mates, and so on. Some of subjective conditions could be mentioned in this context as well. They include: degree of tiredness of the examinee, his/her previous experience, the social importance of the exam, the wish to pass the exam, the moral and the experience, the parents' pressure, the degree of nervousness;
3. the appearance of the items is changed: the grammatical form, the addressing manner, the replacement of one word by its synonym, the changing in the items order;
4. “time”: applied successively, the test must give the same results; if not, it is not an accurate one.

The accurate coefficient can be designed as follows [4, pp. 78-80]:

1. repetition, the applying of the test at a certain period replicating the test must not record measurement deviation, or, at least, the error must be precisely anticipated and measured; the constancy coefficient of the results will show the accurate degree of the test, by the correlation between the values obtained at two distinct and outdistanced application in time.
2. comparing the results obtained by its application with the result obtained at other equivalent tests. The qualitative and quantitative correlations show the equivalent coefficient. So, we will establish that tests are alike or are entirely different;
3. by halving, considering the even items score with odd items sore, the degree of correlation represents the homogeneity coefficient.

The reliability of the test depends on the difficulty of the items. When the test contains items with high difficulty, the individual resorts to guessing the correct answer. The greater the number of guessed answers, the more the scores distribution takes a binomial form. Therefore, we could stand that between the difficulty of the items and the quality of the test there is a reversely proportional rapport. If the test contains items with low difficulty the individuals seldom resort to guessing, and the distribution of the results is uniform. Therefore, the test cannot be useful to evaluate all the individuals. From the

theoretical point of view, the longer the test is (it contained a greater number of items), the more accuracy it has [5, p. 205].

The validity represents the most important quality of the test; shows whether the instrument measures what it proposes and how well makes that.

While establishing the validity of a test, two questions are offered:
1. does the test measure what it is meant to?
2. can it be used in taking the right decisions?

Concerning the purpose, validity can be:
1. of content: the test must refer to those contents referred to during the instruction
2. of criteria: involving the rapport to an external criterion

The first operation made after applying the test is to correct and spot the answers. There are two points of view in marking the answers of the tests' items [5, p. 231]:

1. The first point of view is founded on the concept of "errorless activity", used in preparing the specialists in different fields of techniques. According to this point of view, in the professional activity (programming, building plans, extraterrestrial flights) there are no options between minor and major errors. The computer program works or not. Therefore, while evaluating the specialist preparing, it is not considered the indice of items difficulty, but only the fact that they are solved right or wrong, omitted or partly solved it is not considered.
2. The second point of view is founded on the thesis that the marking of the item must reflect its index of difficulty (the relative frequency of the individuals who have answered correctly out of the all examinees who answered to that item). Thus, when marking, the teacher distributes the learners in class of results (e.g. the mark "7" will be given both to the students who have acquired 37 points and to the ones who have acquired 40). Therefore, when giving marks, the teacher levels the scores of the scholars pertaining to a class results.

After some years of application of these docimological principles and tests to the students of different specialties at Technical University of Moldova, I had the opportunity to study all those elements in practice. It has been demonstrated that a correct application of docimological principles can bring innovations at the level of whole coaching and final evaluating process. Because the main pursued goal is to eliminate the cheating, for this we must try to change the mentality, both the professors and students.

The contradiction consists in fact that nobody from the professors assumes the responsibility to improve his style of work, against the changing using the reason of previous successful generations of students. In this context, only the comparisons between the different generations of students are accepted.

Besides, from the exterior, it can be imposed to them what they have to do. The model of the perfect professor does not exist. The students benefit by the whole spectrum of the university values, both on the scientific and didactic plan, at the same degree as in the past. What is different and must be corrected is the inertness of the students. The students' involvement must be forced to rise to the level imposed by his or her future profession. The instrument destined to realize this correction is just the subject of our discussion: the university degree exam.

The university degree exam is used both in the various goals and in levels of engineering preparation. The justification of this kind of exams is important, but the fact

that such a brief exam can be useful for the final correction in university preparation. The final correction means:

1. covering the minimum knowledge required by practicing that profession, in the instruction process;
2. Assignment of a relevancy for this kind of exams by praising the deserving students and clear establishing the score of promoting.
3. These two principles assure the individual's motivation. It is well known that motivation (positive or negative) is one of the aspects of the success. The docimological criteria of the university degree exam have the role of stimulating the motivation mechanisms.

Way of work:

- Translation of the individual responsibilities towards the collective system of the university degree exam;
- Principle for scaling results of the evaluation: Gauss distribution situated between n% repel and m% maximum rating;
- Method of work: questionnaire focus on the engineering specialties, on extreme complexity:
- INFERIOR Extreme;
- SUPERIOR Extreme.

The indicated way of work means the fixing of the precise level in the area of the obligatory minimum engineering knowledge (physics phenomena, measures, calculus relations, etc.).
The accomplishment of the students' evaluating can become efficiently if the following aspects characteristic to the docimological tests are charged:

Specific aspects:

- Variable weight for each answer
- Vague algorithms for the results interpretation
  - Answers with positive bonus
  - Answers with penalty bonus, for the fault case
  - Variable and particularized tolerance
- Clear definition for the admissibility threshold
- Questionnaire with a supplementary number of items (300 – 350)
- Informatics system for generating the questionnaires
- Eluding the psychological obstacles:
  - free access for the self-evaluation

The variable weight of the answers is necessary because it is working both with very simple and very complicate knowledge.

The vague algorithm used in results analysis does not mean subjectivity and absence of precision. There is a possibility not to sanction a very good paper for few common faults generated by the tiredness, carelessness; but repeated wrong answer at common questions brings to the examinee an unsatisfying score.

To fight against cheating, the number of the items must be in excess. The number of the good answers being given will be also a relative criterion.
In order to avoid the inherent suspicious about the secret of elaborating the questionnaires, these are “gathered” in the morning of the exams day, by random generating.

The elimination of psychological obstacle can be realized by students' well-knowing of the evaluation process. In this way a self-instruction can be also realized, this being the final goal of the whole learning process.

Conclusion:

The use of the performing system of evaluation can contribute to the improvement of the entire educational process.

Strategic results:

- The graduation mark comes into prominence;
- The graduation exam becomes an incentive event.

The significance of the exams degree mark means not only a supplementary distance between the examinees, but also the possibility to eliminate the candidate risk to be in the situation of professional imposition, whether they will receive a diploma which they don't deserve.

Even the opinions about cheating are disputed; using this testing system, which realized a real competitor frame, is one among those that can change the mentalities.

Bibliographical References:

1. Strugă, C., Evaluarea şcolară, Ed. de Vest, Timişoara, 1999.
2. Stan, C., Autoevaluarea şi evaluarea didactică, Ed. Presa Universitară Clujeană, Cluj Napoca, 2001.
3. Nicola, I., Tratat de pedagogie şcolară, Ed. Aramis, Bucureşti, 2000.
4. Jinga, I. (coord.), Evalurea performanţelor şcolare, Ed. Aldin, Bucureşti, 1999.
5. Cabac, V., Evaluarea prin teste în învaţământ, Ed. Universităţii "Al. Russo" Bălţi, 1999.
6. Holban, I., , Testele de cunoştinţe, EDP, Bucureşti,1995.
7. Ionescu, M., Didactica modernă, Ed. Dacia, Cluj Napoca, 2001.
8. Moise, C, Evaluarea şcolară, note de curs.
9. Muster, D., Verificarea progresului şcolar prin teste docimologice, EDP, Bucureşti, 1970.
10. Pavelcu, V, Principii de docimologie, EDP, Bucureşti, 1968.
11. Bonciu, Stefan, Copiatul la examene ca devianţă tolerată, în vol. Câmpul universitar şi actorii săi, sub coord. Adrian Neculau, Editura POLIROM , Iaşi, 1997.
12. Niculescu, Rodica Mariana, Formarea formatorilor, Editura ALL EDUCATIONAL, Bucureşti, 2000.