

## THE DEVELOPMENT OF OPTIMIZATION METHODS FOR KNOWLEDGE BASE ENRICHMENT PROCESSES

*Olga Serdiuk*

*Department of Computer Systems and Networks,  
SIHE «Kryvyi Rih National University»  
Kryvyi Rih, Ukraine  
olgajs28@gmail.com*

### **Abstract**

The paper presents the concept of approach to the research and evaluation of the processes of intellectual activity associated with the enrichment of the knowledge base. A feature of the research of the process dynamics is the need of simultaneous consideration of such diverse factors as the complexity of information perception, the presence of the deviations of the response from the standard in the process of reproduction and accounting of the test time.

A significant influence on the methods of optimization of the knowledge base enrichment process is exerted by a considerable duration of the task learning process. This causes the use of the multifactor experimental design theory to accelerate the progress towards the optimum.

The research results can be used in the development of technologies for efficient knowledge assimilation, automation of skills, and also in the development of expert systems for diagnostics of the processes of intellectual activity.

**Keywords:** evaluation of intellectual activity, optimization of intellectual property, evaluation criteria of intellectual activity, evaluation of memory processes, evaluation index.

© *Olga Serdiuk*

### **1. Introduction**

The intellectual potential of society to a greater extent determines the capabilities of the nation with respect to the potential of using the opportunities of available natural resources. Therefore, one of the most important objects of research related to the development of personality are the processes of intellectual activity.

The increase of efficiency of any thought process rests on solving the problem associated with the technology of expansion of the knowledge base fundamentals.

The knowledge base enrichment processes are very resource-intensive. This fact significantly complicates the study since the design of at least one experiment requires several daily intervals, and the identification of the process is complicated by the need to consider many different factors.

However, it is the knowledge base enrichment processes that predetermine the success of other types of intellectual activity.

In this regard, the efficiency of the knowledge base enrichment processes is an important scientific and practical task.

The aim of the paper is to optimize the knowledge base enrichment processes using the methods of the multifactor experimental design theory.

### **2. Overview of the problem**

The issues of efficiency evaluation of the knowledge base enrichment processes are fairly complex in the estimation theory [1]. This is due to the fact that not all tasks may have a quantitative estimate of their complexity. That is why the information provided in test tasks must meet very specific requirements [2].

In addition, the entire process of knowledge assimilation should be divided into relatively equal amounts of information in order to allocate individual operations. And since a significant amount of knowledge is assimilated gradually, a comparison of the amount of assimilated knowledge with the standard at the intermediate stages of evaluation is required.

This, in turn, imposes severe restrictions on the criterion of evaluation of these processes [3].

One of the most relevant areas for the development of the test theory is studying the features of the memory system [4, 5]. However, some sources show that the processes of the complex

system of human memory are seen as a single continuous process [6]. Such an approach is not quite correct because human memory, as a complex system, represents the operation of many interconnected simple systems. The processes of these systems are associated with the enrichment of the knowledge base, the formation of behavioral responses, accumulation and extraction of information, and so on.

As a rule, "memory tests suggest learning of a set of artificial words (meaningless syllables) with checking their memorizing" [7].

Such studies are not complete from a scientific point of view because significant factors affecting the efficiency of the memorization process in actual practice, namely, the usefulness of information and interest of the subject in its assimilation are lost [8].

Also, the attempts to describe the principles of operation of the memory system can be noted [9-11]. Thus, the efforts of many researchers are aimed at developing the simulation models of the memorization processes [12].

However, the general trend of the known methods is an attempt to characterize the object of research either based on incomplete data, or by a multi-criteria approach [13, 14].

Hence, today there are no scientifically sound principles of effective enrichment of the knowledge base of the person, and the solution of this problem is an important scientific and practical task [15].

### 3. Materials and Methods

Many existing methods for the effective consolidation of information are designed for a universal approach [16], and all attempts to make the assimilation process effective and individual end in providing recommendations of a psychological nature [17].

The result of the learning process can be optimal only if it is possible to vary the magnitudes of the input actions (factors) in a certain range of values. In case of the possibility of varying two or more factors, the volume of investigations increases significantly, which greatly complicates achieving the optimum since the number of operations of searching for the influence factors can be equal to or even exceed the total number of operations of the experiment.

In this situation, the optimization process can be successfully completed using the multifactor experimental design theory [18].

**Materials and equipment:** a set of text dialogues in a foreign language, a timer and a spreadsheet to compute the evaluation index.

### 4. Experimental procedures

The research is aimed at evaluating and diagnosing the features of the development of the knowledge base enrichment (formation) process of the test subject.

The essence of the experiment was that the test subject was asked to memorize a series of 25 dialogues in a foreign language. The process of memorizing each dialogue lasted for seven days and consisted of three stages.

At the first stage (lasting one day), the test subject memorized a dialogue (about 1000 characters). In view of the limited capacity of short-term memory buffer [19], the full text of the dialogue was divided into several approximately equal fragments, of the size of 180-220 characters.

Before memorization, the test subject typed a piece of text by the touch method with ten fingers at maximum speed. As a result of the preliminary typing, the reference value of the evaluation index  $Q_E$  was determined [3].

$$Q = \frac{[\alpha(\kappa - 1) - fol]^2}{T^2 \alpha(\kappa \alpha - fol)}, \quad (1)$$

$Q$  – the evaluation index;

$\alpha$  – the volume of the test task;

$\kappa$  – the complexity of the task;

$fol$  – the degree of deviation of the solved task from the reference value.

On the basis of this value ( $Q_E = 1$ ), two threshold levels (0.8 or 0.5 of the maximum level), indicating the completion time of the cycle of repetitions for typing one of the fragments were determined.

Memorization was carried out by preliminary learning and further reproduction of fragments from memory, with fixing the result by the touch method with ten fingers. The duration of preliminary learning of the text fragment was included in the total duration of the operation.

The duration of each test operation was recorded.

Upon completion of the test operation, the result was checked against the reference, and the number of mistakes in the reproduced fragment was determined. Both distortion of a single word and the word order change were considered as mistakes. The results of the session to memorize the first fragment are given in **Table 1**.

**Table 1**

The data about reproduction of one of the fragments on the first day of the session of the experiment

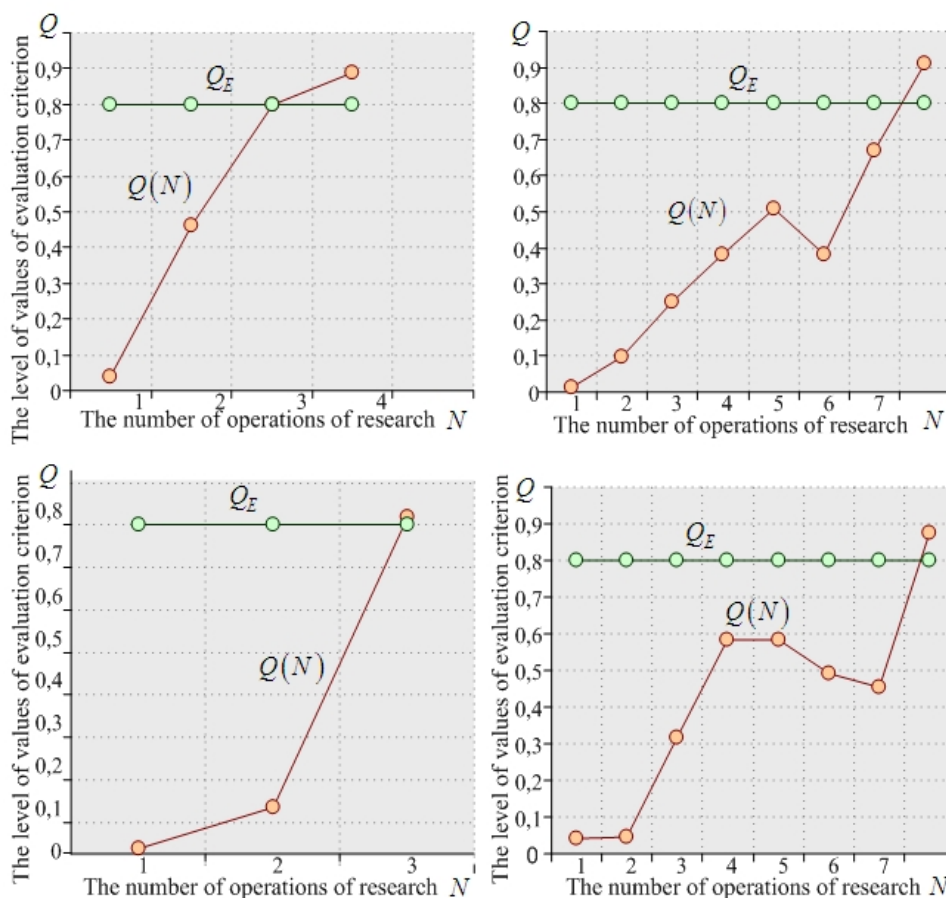
| Date       | N | T     | Mistake | $\alpha$ | $\alpha k$ | Q      | Level | v      |
|------------|---|-------|---------|----------|------------|--------|-------|--------|
| 05.12.2015 | 1 | 6,5   | 12      | 408      | 480        | 0,0036 | 0,8   | 36,92  |
| 05.12.2015 | 1 | 5     | 2       | 268      | 480        | 0,0805 | 0,8   | 48,00  |
| 05.12.2015 | 1 | 3,12  | 1       | 254      | 480        | 0,2479 | 0,8   | 76,92  |
| 05.12.2015 | 1 | 3     | 0       | 240      | 480        | 0,3200 | 0,8   | 80,00  |
| 05.12.2015 | 1 | 2,2   | 1       | 254      | 480        | 0,4986 | 0,8   | 109,09 |
| 05.12.2015 | 1 | 2,2   | 1       | 254      | 480        | 0,4986 | 0,8   | 109,09 |
| 05.12.2015 | 1 | 2,083 |         | 240      | 480        | 0,6638 | 0,8   | 115,22 |
| 05.12.2015 | 1 | 2,083 | 1       | 254      | 480        | 0,5561 | 0,8   | 115,22 |
| 05.12.2015 | 1 | 1,85  |         | 240      | 480        | 0,8415 | 0,8   | 129,73 |

At the second stage (the second day of tests), all text fragments were reproduced until the value of the evaluation criterion for the examined operation reached the level of 0.8. The results of the session to memorize the entire text fragment are shown in Table 2, Figure 1.

**Table 2**

The results of the reproduction of the whole text fragment on the second day of the session of the experiment

| Date       | N exp | T    | Mistake | $\alpha$ | $\alpha k$ | Q    | Level | Speed  |
|------------|-------|------|---------|----------|------------|------|-------|--------|
| 06.10.2015 | 1     | 8    | 1       | 254      | 480        | 0,04 | 0,8   | 30     |
| 06.10.2015 | 1     | 2,5  |         | 240      | 480        | 0,46 | 0,8   | 96     |
| 06.10.2015 | 1     | 1,9  |         | 240      | 480        | 0,8  | 0,8   | 126,32 |
| 06.10.2015 | 1     | 1,8  |         | 240      | 480        | 0,89 | 0,8   | 133,33 |
| 06.10.2015 | 2     | 7,44 | 7       | 304      | 440        | 0,01 | 0,8   | 32,258 |
| 06.10.2015 | 2     | 3,5  | 4       | 268      | 440        | 0,1  | 0,8   | 68,571 |
| 06.10.2015 | 2     | 2    | 5       | 280      | 440        | 0,25 | 0,8   | 120    |
| 06.10.2015 | 2     | 2,13 | 2       | 244      | 440        | 0,38 | 0,8   | 112,68 |
| 06.10.2015 | 2     | 2,18 |         | 220      | 440        | 0,51 | 0,8   | 110,09 |
| 06.10.2015 | 2     | 2,13 | 2       | 244      | 440        | 0,38 | 0,8   | 112,68 |
| 06.10.2015 | 2     | 1,9  |         | 220      | 440        | 0,67 | 0,8   | 126,32 |
| 06.10.2015 | 2     | 1,63 |         | 220      | 440        | 0,91 | 0,8   | 147,24 |
| 06.10.2015 | 3     | 6    | 1       | 193      | 368        | 0,04 | 0,8   | 40     |
| 06.10.2015 | 3     | 2,6  | 4       | 220      | 368        | 0,14 | 0,8   | 92,308 |
| 06.10.2015 | 3     | 1,44 |         | 184      | 368        | 0,82 | 0,8   | 166,67 |
| 06.10.2015 | 4     | 3,56 | 8       | 350      | 476        | 0,04 | 0,8   | 67,416 |
| 06.10.2015 | 4     | 3,9  | 7       | 336      | 476        | 0,05 | 0,8   | 61,538 |
| 06.10.2015 | 4     | 2,5  | 2       | 266      | 476        | 0,32 | 0,8   | 96     |
| 06.10.2015 | 4     | 2,2  |         | 238      | 476        | 0,59 | 0,8   | 109,09 |
| 06.10.2015 | 4     | 2,2  |         | 238      | 476        | 0,59 | 0,8   | 109,09 |
| 06.10.2015 | 4     | 2    | 2       | 266      | 476        | 0,49 | 0,8   | 120    |
| 06.10.2015 | 4     | 2,5  |         | 238      | 476        | 0,45 | 0,8   | 96     |
| 06.10.2015 | 4     | 1,8  |         | 238      | 476        | 0,87 | 0,8   | 133,33 |



**Fig. 1.** The dynamics of changes of evaluation criterion in the performance of the test task on the second day of the experiment:  $Q_E$  is the level of the reference value of the task quality

By the results of two-day testing using the results of the multifactor experimental design, the optimum parameters of influence factors (a threshold of the evaluation index and the number of repetitions of the series), the minimum memorization duration of the text fragment were determined.

At the third stage of testing, the process of improving the efficiency of skill formation was investigated. At this stage, reproduction of the full text was performed.

Since this stage is not related to the subject of research, materials of the third stage are not provided in the paper.

With the purpose of the structural optimization of the experiment. The multifactor experimental design theory was used for determining the course of the experiment and in data processing.

To improve the accuracy of identification of significant factors and the extent of influence on the knowledge base formation effectiveness, 25 series of experiments (sessions) was carried out.

## 5. Results

Having taken the object of our research as a cybernetic system "black box", there are the factors affecting the state of the object at the input of the system and the optimization parameter as the response that determines the behavior of the studied process of evaluation at the output. Factors that may affect the test result can be as follows:

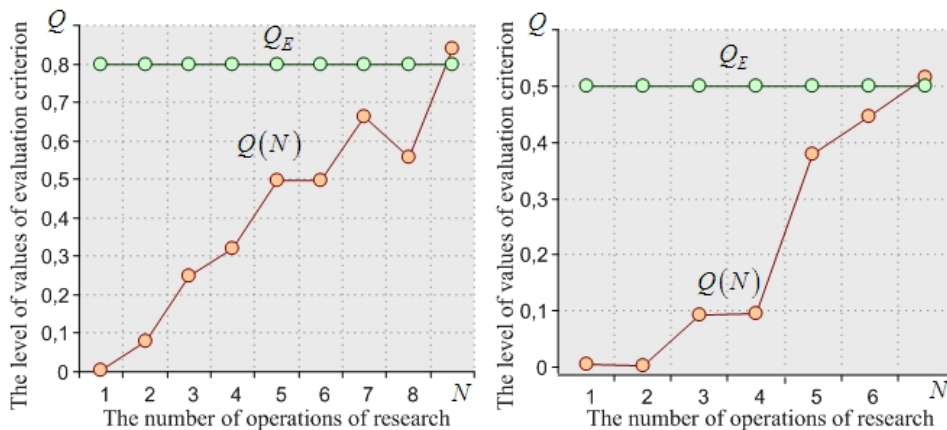
- the scope of the task  $\alpha$ ;
- the complexity of the task  $\kappa$ ;
- the duration of breaks between the repeated sessions  $t$ ;
- mistakes  $fol$ ;

- the number of times of fulfilling one task  $n$ ;  
- the level of the evaluation criterion which needs to be achieved in the reproduction of the text.

The challenge now is to identify significant factors for efficient memorization of the material. The criterion for the result achieved is the value of the evaluation index  $Q$ , the computation of which includes the parameters  $\alpha$ ,  $fol$ ,  $k$ ,  $T$ .

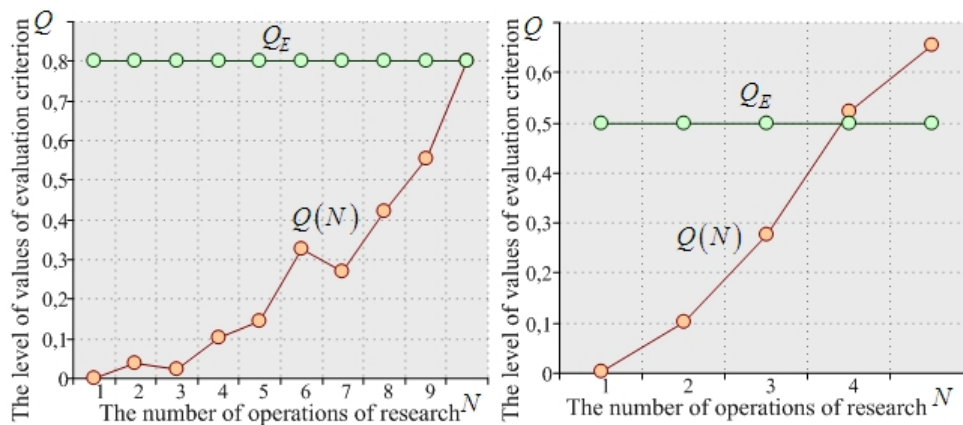
The research revealed the significant factors affecting the efficiency of achieving the result are as follows:

1. Bringing the reproduction of the memorized fragment to the lower or upper level of the given speed -  $X1$  (**Fig. 2**).



**Fig. 2.** The dynamics of changes of the evaluation criterion of the process of memorizing the first and second fragments on the first day of the session of experiments and bringing the quality of their reproduction to the reference level  $Q_E$

2. The number of repetitions of reproduction  $n$  of the memorized fragment is 1 or 2 times -  $X2$  (**Fig. 3**)



**Fig. 3.** The dynamics of changes of the evaluation criterion of the process of memorizing the third text fragment on the first day of the session of the experiment at the double repetition of reproduction and bringing the quality of their reproduction to the reference level

As the optimization parameter  $Y$ , we represent the time that was spent on memorization of the material  $T$  to the quality  $Q$  predetermined by the evaluation criterion.

The levels of the factors of the experiment are summarized in **Table 3**.

**Table 3**  
Determination of the levels of significant factors

| Factors        | Levels |      |     | Variability interval |
|----------------|--------|------|-----|----------------------|
|                | -1     | 0    | +1  |                      |
| X <sub>1</sub> | 0.5    | 0.65 | 0.8 | 0.15                 |
| X <sub>2</sub> | 1      | 1.5  | 2   | 0.5                  |

The matrix of factors and the results of computations for determining the influence coefficients are presented in **Table 4**.

**Table 4**  
The matrix of the levels of significant factors and computation of data for determining the influence coefficients

| N          | X <sub>1</sub> |            | X <sub>2</sub> |          | Y             | X <sub>1</sub> X <sub>2</sub> |            | X <sub>0</sub> Y | X <sub>1</sub> Y | X <sub>2</sub> Y | X <sub>1</sub> X <sub>2</sub> Y |
|------------|----------------|------------|----------------|----------|---------------|-------------------------------|------------|------------------|------------------|------------------|---------------------------------|
| 1          | 1              | 0,8        | 1              | 2        | 60,68         | 1                             | 1,6        | 60,68            | 60,68            | 60,68            | 60,68                           |
| 2          | 1              | 0,8        | -1             | 1        | 44            | -1                            | 0,8        | 44               | 44               | -44              | -44                             |
| 3          | -1             | 0,5        | 1              | 2        | 53,24         | -1                            | 1          | 53,24            | -53,24           | 53,24            | -53,24                          |
| 4          | -1             | 0,5        | -1             | 1        | 54,5          | 1                             | 0,5        | 54,5             | -54,5            | -54,5            | 54,5                            |
| <b>Sum</b> | <b>0</b>       | <b>2,6</b> | <b>0</b>       | <b>6</b> | <b>212,42</b> | <b>0</b>                      | <b>3,9</b> | <b>212,42</b>    | <b>-3,06</b>     | <b>15,42</b>     | <b>17,94</b>                    |

Computations of the coefficients are presented in **Table 5**.

**Table 5**  
The values of the coefficients of the degree of significant factors

| Influence coefficients | Value         |
|------------------------|---------------|
| b <sub>0</sub>         | 53,105        |
| b <sub>1</sub>         | <b>-0,765</b> |
| b <sub>2</sub>         | 3,855         |
| b <sub>12</sub>        | 4,485         |

Accordingly, the response function has the form  $y=53,105-0,765*X_{1+}+3,855*X_{2+}+4,485*X_{1}X_{2}$ . A negative value of the coefficient b<sub>1</sub> shows that with respect to the value of the threshold, the optimum point has been passed. Positive values of the coefficients b<sub>2</sub> and b<sub>12</sub> indicate the need for increasing the number of repetitions of the information reproduction towards the increase and mutual influence of factors.

## 6. Discussion

The technologies of execution and the evaluation of the results of research of the processes of intellectual activity are the most difficult with respect to the evaluation of the processes associated with short-term memory [20, 21].

One of the features of these studies is their duration. For this reason, to evaluate the tactics of their execution, the method of the multifactor experimental design was used.

Certainly, this method is especially useful in situations where the systems approach to the knowledge base enrichment, based on the assimilation of large amounts of information is applied. In this case, the method of the multifactor experimental design can be successfully used to search for the optimum parameters of operations of the studied process.

According to the author, the use of the index developed in [3], which takes into account all the factors affecting the evaluation of the process efficiency [22]: the complexity of the task, expert assessment of the result, deviation from the standard and duration of the studied operation has no alternative.

Creating a software product based on the use of materials of this research may contribute to the development of the fundamentally new technology of assimilation of new knowledge.

Surely, this can not cover the whole range of tasks of an educational process. However, these practices will be very useful to solve a well-defined range of problems.

It may be special tasks aimed at learning the linguistic structures of a foreign language.

With regard to increasing the efficiency of the knowledge base enrichment, the issues of motivation were not a subject of special studies, but were taken into account at the stage of the problem statement to get an adequate response from the respondent processes.

## 7. Conclusions

1. The use of the developed criterion allows to evaluate the dynamics of the knowledge base enrichment process due to simultaneous coverage of all the major parameters of the evaluation operation, subject to measurement.

2. The method of the multifactor experimental design provides the possibility to optimize the knowledge base enrichment process, which is particularly important for improving the efficiency of long-term memory processes.

3. Determination of significant factors and identification of the degree of influence on the efficiency of the studied process allows to individually find the optimum in the organization of the experiment and to make general rules for developing an effective expert technology for the identification of the processes of intellectual activity.

4. The research results can be used to create software tools, in particular when creating expert systems to evaluate the efficiency of the processes of intellectual activity.

---

## References

- [1] *Sikora, L. S., Tkachuk, R. L., Manishin, I. R., Antonik, M. S.* (2005). Logichna struktura procedur formuvannja intelektual'nih testiv dlja ocinki intelektual'nogo rivnja osobi. Modeljuvannja ta informacijni tehnologii: zb. nauk. pr., 35, 135–147.
- [2] *Atkinson, R.* (1980). *Chelovecheskaya pamyat i protsess obuchenija*. Moscow: Progress, 528.
- [3] *Serdiuk, O.* (2015). Criterion of identification of test steps in training and educational systems. *Metallurgical and Mining Industry*, 9, 605–608.
- [4] Bol'shaja jenciklopedija psihologicheskikh testov. (2006). *Jeksmo*, 416.
- [5] *Granovskaja, R. M.* (1974). *Vosprijatie i modeli pamjati*. Nauka, 362.
- [6] *Anastazi, A., Urbina, S.* (2003). *Psihologicheskoe testirovanie*. Saint Petersburg: Piter, 688.
- [7] *Barret, Dzh.* (2004). *Protestirujte sebja: sposobnosti, lichnost', motivacija, kar'era*. Saint Petersburg: Piter, 254.
- [8] *Kurgaev, A. F.* (2000). Vyzritel'nye vozmozhnosti metajazyka predstavlenija znaniy. *USIM. Inst. Kibernetiki im. , №3*, 73–84.
- [9] *Homskaja, E. D.* (2005). *Nejropsihologija*. Saint Petersburg, MGU, 496.
- [10] *Bannikov, I. I.* (1993). Nekotorye voprosy teorii intellektual'nyh processov. *Kibernetika i sistemnyj analiz*, 1, 171.
- [11] *Bondarev, V. N., Ade, F. G.* (2002). *Iskusstvennyj intellekt*. Sevastopol': NTU, 615.
- [12] *Gavrilova, T. A., Horoshevskij, V. F.* (2000). *Bazy znaniy intellektual'nyh sistem*. Saint Petersburg, 384.
- [13] *Gavrilova, T. A., Horoshevskij, V. F.* (2000). *Bazy znaniy intellektual'nyh sistem*. Saint Petersburg, 384.
- [14] *Gudvin, Dzh.* (2004). *Issledovanie v psihologii. Metody i planirovanie*. Saint Petersburg, 558.
- [15] *Durnjak, B. V., Sikora, L. S., Antonik, M. S., Tkachuk, R. L.* (2013). Avtomatizovani ljudino-mashinni sistemi upravlinnja integrovanimi ierarhichnimi organizacijnimi ta virobnichimi strukturami v umovah riziku i konfliktiv. *Ukrains'ka akademija druzarstva*, 514.
- [16] *Karter, F.* (2005). *Testy, ocenka lichnosti, sklonnostej i intellekta*. Saint Petersburg, 207.
- [17] *Sikora, L. S., Tkachuk, R. L., Tkachuk, G. V.* (2013). Kognitivni komponenti prijnattja rishen' intelektual'nim agentom v umovah, pov'jazanih iz rizikom. *Naukovi zapiski. Serija «Psihologija i pedagogika»*. Tematichnij vipusk «Aktual'ni problemi kognitivnoi psihologii». Vidavnictvo Nacional'nogo universitetu «Ostroz'ka akademija», 24, 172.
- [18] *Adler, Ju. P., Markova, E. V., Granovskij, Ju. V.* (1976). *Planirovanie jeksperimenta pri poiske optimum'nyh uslovij, izd. 2, pererabotannoe i dopolnennoe*. Nauka, 279.
- [19] *Druzhinin, V. N.* (1987). *Jeksperimental'naja psihologija*. Nauka, 256.
- [20] *Bannikov, I. I.* (1993). Nekotorye voprosy teorii intellektual'nyh processov. *Kibernetika i sistemnyj analiz*, 1, 171.
- [21] *Sikora, L. S., Tkachuk, R. L., Antonik, M. S., Minishin, I. R.* (2006). Informacijni koncepcii identifikacii mehanizmv intelektual'noi dijali'nosti v umovah riziku. *Chast. 1. Informacijno-komunikacijni tehnologii v suchasnij osviti: dosvid, problemi, perspektivi: zb. nauk. pr. LDU BZhD*, 179–185.
- [22] *Lutsenko, I.* (2015). Identification of target system operations. Development of global efficiency criterion of target operations. *Eastern–European Journal of Enterprise Technologies, Issue 2 (74)*, 35–40.