TESTING OF MENTAL ABILITIES

Vladimir D. Shadrikov,
Anna A. Taldykina
State University – Higher School of Economics
Moscow

This article reports on the results of development of human mental abilities measurement method based on the stated theoretical positions. Such positions are represented by the following: the thesis formulated by M.I. Sechenov that a thought is generated in the process of establishing a connection between the object and its properties; the definition of thinking given by S.L. Rubinstein who wrote that in the process of solving a problem the object manifests its new properties and qualities, it shows a new side, a new content is "bailed out" of it, as well as the thesis about mental development of abilities formulated by V.D. Shadrikov (2007).

Keywords: abilities, mental abilities, abilities testing, intelligence test, psychometric procedure, validity, reliability.

The issue of abilities testing is determined as having high practical and social value, however, it has to be stated that, in spite of numerous tests, often quite witty and provided with detailed instructions, experimental testing did not demonstrated any progress. The turning point was 1905 – when A. Binet together with T. Simon published the "metric scale of intelligence." The scale made it possible to measure and to determine mental development of schoolchildren by years, and even by months. This remarkable work opened new prospects and laid the foundation of the present-day testology. At the same time it should be noted that developing his method of mental abilities measurement Binet did not aim at detecting the structure of abilities and functions descriptive of a child's intelligence, but at something that would make it possible to determine the mental development level of one child in comparison with others.

"The main idea of such measurement was as follows, it was necessary to invent a significant number of tests that would have been quick and reliable at the same time and featuring increasing difficulty; to try these tests with a large amount of children from different are groups, to mea-

sure the results, to identify the test that are successful at a certain age, but cannot be (on average) performed by children at least 1 year younger; to compose, accordingly, the metric scale of mental abilities that allows to determine whether the intelligence level of a particular person is in accordance with their age or it is characterized by delay or advance, as well as how many months or years it corresponds to" (Binet, 1998, p. 49).

It is suffice to randomly give a list of tests developed by Binet and Simon: "3 months – *voluntary* eyesight; 3 years – ability to show one's nose, ear, mouth; to repeat two numbers, to specify persons and objects in a picture; to say one's surname; to repeat six syllables; 7 years – to show empty spaces in a picture; to count one's fingers; to copy a written phrase, to copy a rhomb, to repeat 5 numbers, etc." (Ibid.). As we can see from the list of tests, their performance requires, more likely, habits, than reveals mental abilities. Thereby, certainly, it is presumed that habit formation manifests cognitive abilities of a child.

The mentioned drawback of selecting specific tests for mental abilities testing is virtually typical for all other well-known and widely used mental abilities testing methods. Among the most well-known intelligence tests used by domestic psychologists we may mention the tests by D. Wechsler, R. Amthauer, J. Raven, Stanford – Binet, J. Cattell, H. Eysenck (1971). These tests have a number of significant drawbacks degrading their effective use. The major one among them is insufficient construct validity. The authors of the tests do not explain why they include these or those notions, logical relations, graphic material. They do not seek to prove if people of certain age and educational level must master the intellectual skills determined by the test, to know the chosen words, terms. Eventually the question "What is, after all, intelligence?" remains open.

Subject to the foregoing, for many decades one and the same question has hanged like a sword of Damocles over all intelligence testing methods: whether the detected constructs are real mental formations or this is only a form of classifying the subjects by test items (Kholodnaya, 1997). Analyzing current situation around the notion of *intelligence* in testology, Kholodnaya describes it as a critical one and determines its essence with two words: "Intelligence disappeared" (Kholodnaya, 1997, p. 47).

In this work, an attempt is made to develop a mental abilities (intelligence) measurement method based on clear theoretical grounds. The

latter are represented by a thesis formulated by M.I. Sechenov that a thought is generated in the process of establishing a connection between the object and its properties (Sechenov, 1952). Secondly, the characteristics of thinking given by S.L. Rubinstein, who wrote that in the process of thinking "the set problem with the whole variety of its objective properties and principles enters into yet new connections and owing to this manifests itself in new properties and qualities that are fixed in new notions; thus, yet new content is a kind of bailed out of the problem, it is a kind of shows its new side, it reveals new properties" (Rubinstein, 1958, p. 38). The third premise for designing a mental abilities (intelligence) test was our understanding of abilities implying that an ability may be understood only as a unity of three dimensions: an individual, an actor and a personality (Shadrikov, 2007). Abilities of human as an individual reflect also their natural (biological) essence. They developed to ensure human survival in natural conditions. Abilities of a human as an actor develop on the basis of natural abilities of an individual by complementing them with a system of intellectual operations and their usage mechanisms determining how and when intellectual operations shall be used. Abilities of a personality are abilities of an actor put under the control of a human's moral standards and conscience. They determine social success and qualitative originality of the social cognition and actions of a person. For the purposes of our research, intellectual operations of thinking shall be given special prominence. As S.L. Rubinstein wrote, "thinking correlates the data of senses and perception – matches, compares, discriminates, discovers relations, mediations, and through relations between immediately sensuously present properties of things and phenomena, reveals new, their not immediately sensuously present abstract properties; revealing interconnections and getting to know the reality through such interconnections, thinking gets deeper understanding of its essence" (Rubinstein, 1999, p. 309). Already in this quotation we may single out a number intellectual operations characteristic of thinking. These are: matching, comparison, discrimination, discovery of relations, connections. It should be underlined that all these operations are based on sense data of properties of things. Analysis of the core intellectual operations such as analysis and distraction (abstraction), analysis and synthesis, associations, comparison, generalization, systematization and classification, etc. shows that they are based on discrimination ability (this ability is often named perme-

ability). However the question may be asked: "What is discrimination ability conditioned by?"

And then we come to the conclusion that it is due to two factors: the ability of causing a subject of thought (by I.M. Sechenov) and subjective prior knowledge, because even William James noted that the ability to discern is subject to the general rule: "Only those mental elements, which we already know and who can play separately, can be distinguished from the total sensory experience" (Ezorsky, 1967, pp. 427-428). This provision W. James called the law of dissociation with accompanying changes in the elements. Allocation of object properties opens the way for their various connections. Each such connection has its own idea; i.e. selection of individual properties has conditions for the appearance of thought.

In the testers a lot of attention paid to the separation and study of fluid and crystallized intelligence. In their constructions, we follow the idea expressed by S.L. Rubinstein: "Intelligence is one" (Rubinstein, 1958). "The only correct concept in which ... can be turned on all the concrete material research, obviously, is as follows: the entire child's mental development at an early age is a single process, but within this unity there is a qualitatively different level." Therefore, diagnoses of mental abilities, we believe that unity. In the mental abilities manifested: the ability to isolate specific properties of things that are given testing situation, and the ability to operate on a subjective prior knowledge ascribed to the properties of things.

The study can be divided into two phases: the first phase it was necessary to find an object that meets the following criteria: 1) subject to a certain extent should be familiar to the subject and have a large number of intrinsic properties, and 2) stimuli should help differentiate the subjects on the number called signs. In other words, we needed was a subject, the interaction with which would generate a lot of substantive ideas.

The second phase took place all the necessary psychometric procedures associated with the development of test: verification of the main types of reliability evaluation of the main types of validity and its validity, the calculation of regulatory indicators (standardization).

After preliminary searches, as the desired object was selected CD-disk. It meets all the requirements given above: the subject is familiar to the subject, is quite simple and does not have high loading, the number of available characters reached 64. The subject is well differentiated by the number of subjects called signs. Thus, the theoretical definition construct was diagnosed operationalize, but she studied variables – dis-

played on the level recorded by observable and measurable manifestations and the facts.

Turning to the second phase of the research – testing psychometric test in the first place we turned to this indicator as validity as a measure of consistency of how the methodology and results of research to meet the requirements, is one of the most important characteristics. Of the varieties used in psychodiagnostics validity, we have chosen indicators Construct, criterion validity and statistical, as well as seeking to ensure the content validity of the developed techniques (Kondakov, Romanyuk, Sorokina, & Shishlyannikova, 1999). Construct validity test shows how the results can be considered as a measure of a theoretical construct, in this case, measures severity of mental abilities. In our case we can construct validity is proved theoretical assumptions underlying the developed test. The sense of validation for the construct is to highlight the basic concept – a construct, which is aimed at measuring the test, put in correspondence with a set of empirical Implicators and justify this correspondence. Rate criterial validity of the developed test we conducted using the method of contrasting groups (test significance of differences of the results between male and female sample).

The verification method of contrasting groups we conducted, using the results of retesting the total sample of subjects (N = 752), while male sample consisted of 384 subjects, and women – from 368 subjects. As a result of U-Mann-Whitney, we have found between male and female sample of reliable differences (at 5% or 1%) in the expression of the quality of diagnoses test. The values of the criterion presented in Table 1.

Table 1

Values of U-Mann-Whitney test to determine differences in expression between the characteristic male and female sample

	The properties of the CD-disk	The properties of the metal ball
The values of the U-Mann-Whitney	5824.500	5792.000
The level of statistical significance of differences	0.136	0.147

Consequently, we have developed a test does not differentiate between men and women in terms of mental abilities. Statistical validity of serving as a maximum likelihood test the validity, subject to the same

variances of items is also a high 91.4%, which is a confirmation of the entire above see Table 2.

Table 2

Statistical indicator validity

The developed test	Number of subjects, N	Statistical indicators,%
Validity	752	91.4
Variance	41	8.6

In addition, we developed methodology can be considered external validity (valid content), since, as has been noted previously, the development of methodology, we were guided by specific theoretical positions of I.M. Sechenov and S.L. Rubinstein. Thus, the final form of the developed method meets all the basic psychometric requirements.

The problem of measuring the reliability of the test specifies a more general problem of measurement in psychology – the question of the degree of confidence in the accuracy and stability of the test. Under the reliability, in the broadest sense, refers to characteristics of the extent to which the identified differences in the subjects on test results reflect actual differences in the measured properties (true variance), and in which they can be attributed to the action of random factors (variance of error). The coefficient of reliability of a specific test can be calculated as the ratio of "true" variance to "real" empirical (Kondakov et al., 1999, p. 112).

$$R_t = D_t / D_x$$

Where R_t – the coefficient of reliability, D_t – "true" variance, D_x – the real dispersion.

Of the options we have identified for the reliability of our test two types of reliability: retest and alternative, as well as the two most important parameters: reliability, as the stability and reliability of measurement, as accurate measurements. Retest method is to re-interview a sample of subjects the same test after a certain period of time when the same conditions. The time interval depends on the age, as well as events with the subject in life. A common assumption is that the test is "reliable" in a two-week interval, but there is no rationale in favor of this fact.

For the reliability index was adopted Pearson's correlation coefficient between the results of two tests.

$$r_{xy} = SP_x / \sqrt{SS_x * SS_y},$$

Where $SP_x = \Sigma(x_i - x)(y_i - y)$, $SS_x = \Sigma(x_i - x)^2$

We compared the results of primary and repeated testing for determining the properties of the object. The correlation coefficients obtained by comparing the results of the first and re-testing are presented in Table 3.

Table 3

Statistical analysis

	Correlation coefficient	**Number of subjects**	**The developed test (second series)**
The developed test (First series)	Pearson's correlation coefficient (r)	752	0.992

As can be seen from the table, the values of the correlation coefficient $r = 0.992$. These figures demonstrate very high reliability of the test in terms of sustainability of its results over time. We also used a reliable alternative forms, which is checked retesting the same sample tested in parallel form the dough through a minimum interval of time when the same conditions. Possibility of deception on the part of the test, its sophistication, the events in the interval between testing, do not exert much influence (as in retest method) on the reliability of the test. If you factor in training reduced when tested in parallel forms, the effect of the transfer of the principle tasks often takes place and should be considered in the development of parallel forms. As a parallel form, we take a new object – a metal ball, had a similar, described above, psychometric procedures and compared the figures.

For the reliability index is taken correlation coefficient between the test results of two parallel forms of the test. The correlation coefficient in our case, $r = 0.973$, such a high correlation coefficient and a large interval between the two tests indicate that different forms of the test are close to each other. The results are presented in Table 4.

Table 4

Statistical analysis

	Correlation coefficient	Number of subjects	Parallel forms of the test
The developed test (CD-disk)	Pearson's correlation coefficient (r)	752	0.973

The obtained reliability coefficients were significant, which meets the requirements of psychometrics and indicates sufficient reliability of our method.

Also, one of the procedures, psychometric testing was rationing. The nature of psychological measurement is that the interpretation of the primary result obtained specific subjects while performing a certain test is meaningless without the possibility of comparing this result with the rules of execution of this test in the standardization sample, or other standard indicators. Norma is a quantitative measure of assessment test results, determining the severity of psychological properties as the object of measurement, which shows what the place of the individual is in a reference group with certain characteristics.

Figure 1 shows a histogram of the distribution of responses of the subjects in the course of our technique using as stimuli – CD-disk.

Analysis of the figure, demonstrating the features of the distribution of responses of the subjects, and enables to compare the empirical distribution with a theoretical normal distribution, strengthen our assumptions about the normality of the sample and the distribution of responses. For your convenience, we carried out the procedure for valuation and set of Scale assessment, the relevant scales of traditional methods (150 points). Following the procedure of developing standards (comparing the empirical distribution with the normal application of the "three-sigma rule," Translation in Z-evaluation, etc.) was developed by rating scale – the implementation of universal norms of the test.

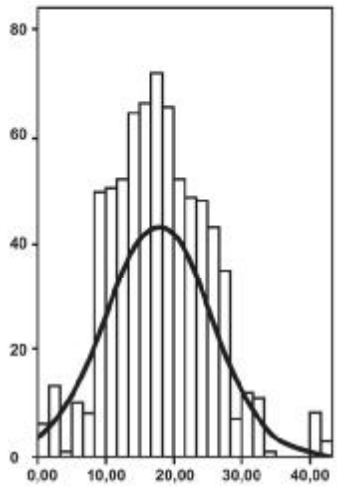

Figure 1. Distribution of answers of the subjects, obtained in the analysis of stimuli – CD-disk

It is worth noting that we also conducted a correlation study with existing intelligence tests, to establish communication. Between scores on two tests calculated the correlation coefficient. A positive correlation indicates that the developed test measures roughly the

same psychological quality that reference method. It is important to note that it does not require too high a correlation coefficient, because if it turns out that the new and reference tests are almost identical in content, and the technique does not have the advantage of brevity and ease of use, it means a duplication benchmark. However, in our case, the high correlation coefficient would indicate only the benefits but not the drawbacks.

For comparisons and to identify correlations, we used D. Wechsler test, and intelligence tests by H. Eysenck and J. Cattell, as the most frequently used and validated methodology. To calculate the necessary correlation coefficients used Spearman rank correlation coefficient. As a result, we have defined a measure of the relationship between results of tests by H. Eysenck, J. Cattell, D. Wechsler, and ours.

Spearman correlation coefficients reflecting the relationship between test results of our test and tests by H. Eysenck, J. Cattell, D. Wechsler, respectively, are equal to 0.959, 0.899, 0.886, indicating that hit all four of the coefficients in the region of 1% regulation of statistical significance ($p < 0.01$), therefore, we developed the method meets the requirements of psychometrics to the parameters validity. Analyzed the correlation coefficients are presented in Table 5.

Table 5

Statistical analysis

	Correlation coefficient	Eysenck test	Cattell test	Wechsler test
The developed test (CD-disk)	Spearman's correlation coefficient (r)	0.959	0.899	0.886

Results of our statistical procedures permit the conclusion according to our method of traditional psychometric requirements for validity of the developed test. From the combination of psychometric procedures performed, we can say that the designed test measuring mental abilities has high validity and reliability. Note that the normalization of the developed test carried out only for a sample of subjects aged 18-25 years. As the subjects were 752 men university students of different types (technical, human, natural science). Below, the developed test diagnosis of mental abilities.

Test measurement of mental abilities

Advice and guidance to the user

Time of test run is strictly limited to 15 minutes, conducting the test; you must clearly follow a temporary measure. Stimuli subjects were presented at the beginning of the test and remain available to them during the entire time of the assignment. Before the main testing procedure, subjects were invited to perform adaptive task for 3 minutes. Upon completion of the implementation of the adaptation task, you must answer all questions on the test and eliminate ambiguity.

Adaptive assignment

Answer Sheet

Full Name_____

Age_____ Date_____

Adaptive assignment:

Pencil
Properties:
1) long
2) wooden
3)
…

Main task

Instruction for subject is: "You will be represented by an object whose properties you need to write for 15 minutes. In the course of all the work you can perform any manipulation with the object. The properties are described by a noun or an adjective."

Main task:

mini CD-disk
Properties:
1)
2)
3)
4) …

Key and result processing

Necessary to calculate all the mentioned subjects of property (raw scores), the relevant comparative list of features CD-disk (Table 6) and in accordance with the scale (Table 7) – to determine the level of mental abilities (Table 8). If the list of test words are presented does not match, then they are ignored or taken as zero.

Table 6

Comparative list of the properties of CD-disk

Comparative list of the properties	
1. Capacious	2. Rigid
3. Cheap	4. Holed
5. Closed	6. Recording
7. Dry	8. Hard
9. Even	10. Light
11. Fast heated / fast chilled	12. Breakable
13. Flat	14. Striped
15. For computer	16. Tiny
17. Fragile-strong	18. Clear
19. Grey	20. System
21. High-quality	22. Compact

23. Invented	24. Straight
25. Lightweight	26. Flying
27. Manufactured	28. Informative
29. Melting	30. Plastic
31. Multiple-use	32. Artificial
33. Non-transparent	34. Processed
35. Operating	36. Rainbow-like
37. Readable	38. Rough
39. Refringent	40. Small
41. Round	42. Rotating
43. Self-coloured / multi-coloured	44. Coloured
45. Sharp	46. Reflecting
47. Slippery	48. Modern
49. Smooth	50. Double-sided
51. Sound-recording	52. Mirror-like
53. Sparkling	54. Writing
55. Tasteless	56. Shining
57. Thin	58. Cold
59. Translucent	60. Silverish
61. Water resistant	62. Water proof
63. Wide	64. Bright

Table 7

Scale of results

Brutto points (properties)	Mental abilities level (points)
0-5	50 points
6-11	60 points
12-18	70 points
18-23	80 points
24-29	90 points
29-35	100 points
36-41	110 points
42-47	120 points
48-53	130 points
54-59	140 points
60-64	150 points

Table 8

Value scale

Points	Level of mental abilities
50 – 70 points	low level of mental abilities
80 – 100 points	average level of mental abilities
110 – 150 points	high level of mental abilities

Thus, in this paper presents an approach of designing a test of mental abilities on the basis of specific theoretical positions. This test enriches the arsenal of diagnostic tools measuring abilities and intellect; he had operatives in conducting and has high construct validity, which allows you to make meaningful judgments about the mental faculties of man. Proposed test may be of interest to professionals working in the field of designing tests abilities, as well as for practitioners concerned with psycho-diagnostics.

In the future it is planned to standardize the test values for different are groups. Preliminary experiments showed that the test can work with the age of 10 years and over.

References

Binet, A. (1998). *Izmerenie umstvennyh sposobnostej* [Measuring Mental Abilities]. Saint Petersburg: Soûz.

Eysenck, H.J. (1971). *Race, Intelligence and Education*. London: Temple Smith.

Ezorsky, G. (1967). Pragmatic Theory of Truth. In P. Edwards (Ed.), *Encyclopedia of Philosophy. Vol. 6.* New York, London: Macmillan.

Kholodnaya, M.A. (1997). *Psihologiâ intellekta: paradoksy issledovaniâ* [Psychology of Intelligence: The Paradoxes of the Study]. Tomsk: TGU; Moscow: Bars.

Kondakov, I.M., Romanyuk, E.I., Sorokina, O.L., & Shishlyannikova, L.M. (1999). *Razrabotka testovyh zadanij. Psihometrika* [Development of Test Items. Psychometrics]. Moscow.

Rubinstein, S.L. (1958). *O myšlenii i putâh ego issledovaniâ* [On Thinking and Ways of its Study]. Moscow: AN SSSR.

Rubinstein, S.L. (1999). *Osnovy obŝej psihologii* [Fundamentals of General Psychology]. Saint Petersburg: Piter.

Sechenov, I.M. (1952). *Izbrannye proizvedeniâ. T. 1. Fiziologiâ i psihologiâ* [Selected Works. Vol. 1. Physiology and Psychology]. Moscow: AN SSSR.

Shadrikov, V.D. (2007). *Mentaľnoe razvitie čeloveka* [The Mental Development of Man]. Moscow: Aspekt Press.