COMPARATIVE ANALYSIS OF RUSSIAN AND FOREIGN SYSTEMS FOR THE NEUROPSYCHOLOGICAL DIAGNOSIS OF CHILDREN FROM THE STANDPOINT OF THE PSYCHOMETRIC APPROACH AND ITS LIMITATIONS WHEN USED IN CLINICAL PRACTICE

Alena V. Astaeva,
Mikhail A. Berebin
South Ural State University
Chelyabinsk, Russia

Modern approaches to the psychological diagnostics of children are based on need analysis and the systematization of existing trends in domestic and foreign children's neuropsychological diagnoses. This article highlights the strengths and weaknesses of each approach, as well as identifying possible areas of integration. One of the most acute problems is the development of experimental psychological methods to determine the quantitative and expressed characteristics of the mental development of children by flexibly combining qualitative and quantitative approaches, with a view to putting into practice foreign neuroscience principles and standards of evidence. An analysis of contemporary publications on neuropsychological diagnosis reveals the need to consider the standardization of neuropsychological research in the context of current approaches, requirements, and psychodiagnostic criteria. In the domestic neuropsychological literature, these issues are need attention: standardized procedures for presenting stimuli are not, as a rule, described; basic psychometric assessment of the results of the tests is lacking; no investigation of their clinical and psychometric validity is carried out. An analysis was made of the nature of the psychometric approach in foreign child neuropsychology, which relies on mathematical procedures of processing qualitative (ordinal) data converted into quantitative indicators. We examined separately the mathematical software for clinical trials (based on the principles of «evidence-based medicine»), which relies on the «abnormal» nature of the distribution of clinical phenomena.

Keywords: neuropsychological diagnostics, psychometric approach, norm, qualitative and quantitative assessment of results.

A systemic approach to mental functions requires definition of the specific content of the formation during ontogeny of the mental functions of children. However, any manifestation of mental activity should be clear grounds for the qualification of «normal» or «individual» or «pathological» types of higher mental functions.

Methodological issues in diagnostic work with children are often treated in the domestic and foreign literature (Shereshevsky, 2007; Vasserman & Shereshevskij, 2006).

The traditional approach to neuropsychological evaluation involves the allocation of a neuropsychological factor (a pathological functional system with primary defects and secondary developed symptoms) as an explanation for the observed disorders, followed by their design in the form of a specific neuropsychological syndrome.

Russian neuropsychologists indicate the need for a systematic structure of the neuropsychological syndrome with three components: first, the primary injured level of the functional system; second, the causal link with the related complexes of the mental health disturbances; third, the mental manifestations of compensation for these complexes. In fact, we are talking about restructuring the previously defective functional system, manifesting the formation of its new management, including the damaged elements, and developing, as a result of their compensation, mental formations (Korsakova, Mikadze, & Balashova, 2001). Often, the identification of all components of a functional system is possible only by conducting special studies (sensitized and functional neuropsychological tests, neuroimaging, and so forth).

In domestic infant neuropsychological assessment, several batteries of neuropsychological techniques have been created, mainly since the mid-1990s, but until recently there has been little systematic review of their composition and the characteristics of their application. However, some monographs discuss some fairly obvious and most urgent problems of neuropsychological status (for example, Ahutina & Pylaeva, 2003; Glozman, Potanina & Soboleva, 2006; Korsakova, Mikadze, & Balashova, 2001; Simernitskaya, 1991; Tsvetkova, 2001). In general, these techniques and batteries are characterized by their focus on the use of qualitative assessments and quantitative data breaches for reporting diagnostic results. In addition, as a rule, they do not use standardized procedures for the presentation of stimuli; they do not use basic psychometric approaches for assessing results; and they provide no psychometric and clinical validity for the methods and batteries in general. This situation attests to a lack, in domestic neuropsychological diagnosis, of psychometric software for experimental psychological research. As a consequence, the results of experimental psychological research into clinical practice cannot be assessed in accordance with the criteria for the standardization of norms for psychodiagnostic methods (Fletcher, Fletcher, & Vagner, 1998).

In foreign neuropsychology different approaches are seen for the qualitative and quantitative evaluation and standardization of research results (Apperly, Samson, & Humphreys, 2005; Haladyna & Downing, 2006; Joy, 2001; Rabin, 2005). Test batteries in neuropsychology in English, as a rule, have a standardized procedure for the presentation of stimuli and for their calculations of psychometric norms and evaluation characteristics; these procedures are determined by the characteristics of their validity and reliability on the basis of their fulfillment of compulsory norms for clinical research.

However, in foreign child neuropsychology, techniques based on the results of modern interdisciplinary studies of the ontogeny of higher mental functions in children are a relative novelty. As a consequence, diagnostic systems have insufficient theoretical and methodological validity of the rich content of the projects, and the results are interpreted in the discrete categories of «normal» or «pathological» without specific consideration of the ontogeny of higher mental functions in view of the variance in and individual nature of their formation (Furr, 2008; Groth-Marnat, 2003; Strauss, 2009).

At home and abroad neuropsychology has described in detail the methods and techniques of the existing diagnostic approaches and psychodiagnostic batteries, the history of their creation, their adaptation and development, and their possible applications. Determining the difference in these systems is fundamentally distinguishing the bases of their assessment and qualification of the research results.

Thus, domestic neuropsychology is characterized by:

- a reliance on qualitative descriptions
- attempts to introduce quantitative scales for calibrating assessments of disturbances in accordance with their degree of severity
- a relatively arbitrary sequence of research

In foreign neuropsychology we mainly observe:

- standardization of the neuropsychological instruments themselves in classical form (the determination of indicators of representativeness, reliability, validity)
- attempts to calculate specific techniques for obtaining methodological norms for different population and nosological samples based on the statistics of a normal distribution
- the presence of relatively rigid algorithms for research used in conjunction with a battery of both unstandardized (classical neuropsychological) and standardized (test) methods (Berebin & Astaeva, 2008)

This kind of comparative analysis of the methodological, systematic, and classificatory bases of the development and qualification of the methods and techniques of psychodiagnostics attests to the presence of at least two major problems in modern neuropsychological assessment.

First, it is necessary to develop adequate psychometric techniques that preserve the informative value of qualitative and quantitative diagnostic data, as well as their combinations.

Second, it is necessary to introduce into practice adequate psychometric procedures for standardization and regulation, taking into account the specific aspects of clinical research and the particular features of the ontogenesis of higher mental functions in normal, pathological, and individual versions of development.

On the one hand, the solution to these problems must be based on modern concepts of the theory of psychological measurement, mathematical and statistical methods for processing their results, and classical psychodiagnostic requirements for psychometric standards. On the other hand, a solution that adequately quantifies the results of neuropsychological studies should relate to clinical needs by adopting medical-evaluation categories: "normal," "pathological," "nosological," "pre-illness," "individual variation."

Most Russian classical techniques of qualitative neurological research are designed to analyze syndromes of disturbances through the identification of the qualitative characteristics of performance tests, through the impaired release of a factor, and, possibly, through the definition of a local focus. In doing so, the symptoms (the qualitative characteristics of the sample run) are usually measured in these ways:

- On a scale of intensity or severity: "no," "weakly," 'strongly" (Homskaja, Efimova, Budyka, & Enikolopova, 1997);
- On a four-point rating scale: "the absence of errors or errors inherent in healthy subjects," "mild disorders," "average degree of error," "gross disorder," accompanied by a sufficiently detailed explanation system (Vasserman, Dorofeeva, & Meerson, 1997);
- On a four-point scale with different interval assessments (0-5-10-15 or 0.5-1.0-1.0 points); the intervals are a result of a study that reflects the distinctive features and the extent of the severity and number of allowable errors as a percentage of the expressed ratio of the sums of all accumulated points to the greatest possible result, and the interpretation of such results is carried out in relation to the norms of the boundaries of one of the four bands of the success of the subjects (Fotekova, 2003);
- On the scale dimension 0–1 or 0–4 points; this method comes from the calculation of the final outcome of the procedures as described by Fotekova (2003), taking into account the number of errors, the volume of tasks performed correctly, and the time spent on them (Ahutina & Inshakova, 2008)
- On a scale from 0 to 3 points (with the «inner» interval estimate of 0.5 points), taking into account a system of "fines" – adding points for faulty execution or exceeding the time limit; the final grade is the ratio of the total score for all violations to the number of neuropsychological tests (Glozman, 1999)

In general, the above systems are presented as more sophisticated versions of the famous psychometrics qualitative methodological approach "presence–no sign" (Vasserman & Shchelkova, 2004), which fixes the result of detection of "point features" – "no violation" (0 points) – "a violation" (1 point). At the same time, there is significant deviation from the requirements and principles of the theory of measurement principles and psychometric data representation.

Thus, in Fotekova's proposed system, scale ranges of norms are not correlated with any of the known statistical scales of psychometric standards. These scales show evidence of criterion standards: the lower boundary has been successfully established at 50% or 55% (for children in elementary school and middle school, respectively), but the step of ranges in the next gradation scale of success is uneven: 15%, 15%, and

20% for younger students and three times 15% for older students. In this case in fact the relative (specific, percentage) presentation of the results provides no statistical basis for the psychometric transformation of results. And, in general, this kind of system does not adequately translate the point results or the percentage of results even in percentile values of the scales – for example, the values of the ranges of subjects' success do not correspond to the values of the step percentiles and *z*-assessments, calculated by the classical procedure of transferring raw scores into *z*-assessments through the calculation of percentile ranks, and do not correspond even to the standard percentage representation of the ranges of success (Sidorenko, 2002): "very bad" (10% of results), "bad" (20%), "average" (40%), "good" (20%), and "excellent" (10%).

In the proposed Glozman ratings system, the performance tests are the basis for assigning scores that are taxonomically distinct. For example, there is a 0.5-point increase for the estimated latent period of execution, but 1.5 points for the presence of 1–2 errors with their self-correction, 2 points for inertia and perseveration, 2.5 points for failure to test for interfering effects. In addition, the success of the performance on some of the tests correlates with the time standards; when they are exceeded, "penalty" points accrue in direct proportion (for example, exceeding the norm by 25% gains 1 additional point). For other researchers, in order to ensure the certainty and consistency of their assessment of violations, a tabular form for conformity to clinically significant disturbances or to a particular score is proposed.

In some foreign publications also, fairly detailed systems for regulating results are presented; they are essentially qualitative neuropsychological studies of the interval (including psychometric) scales (Furr, 2008; Strauss, 2009). Note that in the Russian literature on the problem of clinical and psychological research, regulations of this kind are virtually absent.

Thus, in a review Strauss (2009) presents an analysis of the foreign neuropsychological batteries and methods used to diagnose disorders and characteristics of higher mental functions in children of different age groups. For example, a neuropsychological battery, NEPSY-II, designed to determine the neuropsychological status of children 3 to 16 years old, is described in accordance with all the accepted requirements for a standardized psychological battery: indicators of validity, test and retest reliability, given age-performance standards for each of the study populations. There are separately described algorithms for the evaluation of the results: figures, localized in the range of fewer than 2 percentiles (P_1–P_2), showed a very low level of expression of the variable, with values in the range of P_3–P_{10} – on the reduced level, as, for example, in the range of P_{26}–$P_{75,}$ recorded as "sufficient," and in the results for the values P_{75} and above, recorded as "quite good."

For interpreting the results of one of the most common neuropsychological batteries from abroad, the Halstead-Reitan (Hebben, 2010), the total performance of the test is determined or an overall assessment of neuropsychological defects is made using decimal values of the scale dimension of performance from 0 to 1 point. The boundaries of the evaluation scale are in increments of 0.2 points primarily: 0.0–0.2 points corresponds to the normal output, and 0.8–1.0 corresponds to gross violations of performance. The increase in the step range "inadequate performance" up to 0.3 points (0.5–0.7 points) approximates a scale constructed by taking into account the laws of the normal distribution of steps. The overall index of neuropsychological impairment is calculated by totaling the values of 42 variables, calculated on a scale of 0 to 4. Established criteria standards evaluate the level of the defect: values of 0–25 points are determined to be the "norm"; for example, 68 or more points indicate serious or flagrant violations. The heart of this system is also based on qualitative assessments of the actual performance of each sample. Either the number of errors ("no mistakes," 0 points; and "3 errors and more," 3 points, etc.) or the evaluation of specific features of the execution of one test or another (from "no violations," 0 points, to "expressed disturbances," 3 points) is taken into account.

Some foreign neuropsychological tests (for example, the Boston Naming Test and the Wisconsin Card Sorting Test) do not contain standards tested for the normality of the distribution of the results of the sample studied, as they were obtained from small samples. Moreover, according to the developers of these tests, an increase in the number of subjects leads to the automatic allocation of specific groups (subsamples) – for example, subjects with too high (too low) estimates, discrete groups of populations with different characteristics of the effects of treatment or rehabilitation. Instead of rules based on indicators of dispersion, these tests use the values of fashion and the media. Accordingly, the population data are not symmetric with respect to the arithmetic mean and the mean (as is observed in the normal distribution), and, in regard to other indicators, they are not used in obtaining psychometric standards – in particular, the performance mode (Strauss, 2009). However, for these methods all the necessary psychometric characteristics (reliability, validity, measurement error, the error of testing the accuracy of the interpretation of the phenomena) are determined, and, as well, there are requirements for the qualifications, jurisdiction, and competence of the psychologist.

This system of qualitative and quantitative evaluation of the results of neuropsychological studies has both positive and negative sides. The system of evaluation, "the presence or absence of signs" (nonverbally presented in metrical form by discrete values 0 or 1), allows one to make a quantitative estimate by using qualitatively different manifestations of disorders of mental functions in accordance with their degree of severity – for example, by application of the procedure of "quantification" (Wasserman & Shchelkova, 2004). In psychometric terms, this problem can be solved by establishing compliance of the qualitative evidence for altered mental functions and their degree of severity with a particular scale of measurement (Berebin & Astaeva, 2008). In these systems the measure of the severity of violations is assessed on a scale with values from 1 to 3 points, which requires precise qualification of the metric nature of such a scale of measurement. In a study Ryazanova (2010) has shown that the measure of the severity of violations has, first, the attributes of the qualitative nominative because it contains the verbalization of different levels of qualitative changes of mental activity (such as "weak violation"). Second, it has attributes of the qualitative ordinal (rank) scale inasmuch as it allows one to increase the measured attribute (assign it a rank) in proportion to the increase in its intensity ("ranking in ascending order": "weak"–"moderate"–"strong"). In addition, the described system of qualitative results of a study is actually a dichotomous scale of rank, as a discrete value – "violation is" ("1") – is subsequently converted into a continuum of values of continuous data; again, it is subjectively discretely divided, for example, at three intervals – "poorly ..." ("1.1"), "moderately ..." ("1.2"), "strongly marked disturbance" ("1.3"). In the process the actual size of these intervals is, at least mathematically, not exactly defined, or they are obviously unequal to each other. In this case, a more precise definition for carrying out this procedure is not the "quantification" of results ("split apart" from the Latin *quantum* – an indivisible part of any quantity) but their representation in an orderly, ordinary form.

In general, it should be noted that the translation of the verbalized characteristics of mental disorders into a point scale is, in this case, an example of "increasing the capacity of the scale" (from nominative to rank), but this increase does not go beyond quality measurement because the nominative and ordinal scales measure the qualitative, not quantitative, properties of the object. For quantitative measurements to be obtained, at the least, scale intervals are needed (Suppes & Zines, 1967). This conclusion allows further analysis of the psychometric status of neuropsychological studies assuming adequate mathematical skills for measuring the characteristics of the clinical and neuropsychological status of the subjects.

We noted above that foreign child neuropsychology in general is characterized by a psychometric approach based on mathematical procedures for processing quantitative data. In their most general form, these procedures are defined mathematically as an adequate conversion of the summed scores (in absolute values or above the "raw" or "arithmetic-mean-of-sample" scores) into the value on a percentile scale and of the construction on its base curve of the distribution of research results and the subsequent determination of values of the standardized z-scale (actually, getting the research results converted into the values of the z-interval scale). In particular, the most common way to construct such an interval scale is either by grouping the point values on the principle of the equality of cumulative frequency, thus producing a percentile scale – for example, the method of successive intervals – or by grouping intervals according to the principle of the equality of cumulative frequency (Sidorenko, 2002). The latter method is seen as an adequate basis for scales only in the case of the normal distribution of the investigated trait. In psychometrics it is believed that percentile scales also behave like ordinal scales because they reflect the allocation of subjects according to their success without regard to how or how much the success of one subject is more (or less) comparable to the success of another subject. Further transformation (the so-called normalization of the interval) is executed by moving from a percentile scale to the sigma scale using the inverse of the integral (the ordinate distribution percentile is the abscissa of the transition to a normal distribution). In turn, the sigma scale is the basis for z-assessments based on a calculation of the classical formula

$Z = (xi - X) / \sigma$ (Strauss, 2009). Building a defined amount of data in this way (by increasing power measurements) is carried out either on the basis of the empirical distribution or on the basis of an arbitrary model of the theoretical distribution (usually only on the model of the normal distribution).

In addition to the purely statistical problems and limitations in the standardization of neuropsychological techniques and in obtaining their psychometric norms, it should be emphasized that methodological problems exist in extrapolating the static approach to the research area of clinical phenomena. And the main question is whether the distribution measured in the clinical psychodiagnostics of the phenomena is a normal distribution (the Laplassa–Gaussian distribution).

The foreign literature on the mathematical support of clinical research (including literature on "evidence-based medicine," "clinical epidemiology") rather convincingly presents the view that, first, mathematical, statistical, or any other theorems that allow the prediction of the shape of the distribution of clinical measurements do not exist, and, second, the researchers make a priori assumptions about the normality of the distribution of clinical measurements, guided primarily by convenience and the ease of calculation of averages and standard deviations (Fletcher, Fletcher, & Vagner, 1998), so that the similarity of the real distributions to the distributions of the normal curve often accidentally or actually is not generally subject to "normal" law.

Indeed, clinical trials are usually performed separately on "normal" and "pathological" samples. The distribution of the most discussed phenomena in medicine (the variables) cannot be divided into "normal" and "abnormal" because clinical symptoms are mainly distributed in discrete-dichotomous form. Thus, manifestations of health problems or violations of mental functions in "normal" and "pathological" subjects may not show a distinct break in the continuum between "health" and "disease" because, for example, they can occur in those who are healthy and those who are sick. In addition, the graphs of the distribution of the "healthy" manifestations of mental functions and the "broken" versions of their development cannot always be represented by two obvious peaks (the excesses of the distribution curve), one of which would represent a normal distribution in the "healthy" group and the other in the "pathological" group. Highly problematical from a mathematical and clinical presentation of the position is the distribution of the variety of results from neuropsychological tests into a single bimodal or normal curve in which one peak (the "tail distribution") would be consistent with the result in the "normal" group and the other peak would fit the result in the "pathological" group (Fletcher, Fletcher, & Vagner, 1998).

There are several explanations in clinical epidemiology for this situation.

First, the division of subjects into "healthy" and "sick" categories is done on the basis solely of measurement results that are the consistently changing (increasing or decreasing) values of the exponent, attesting to a disturbance of health; such a division is impossible because of the specifics of the operation of etiopathogenetic mechanisms. Indeed, an increase (decrease) in any indicator does not always suggest a pathological process, and, conversely, an existing pathological process is not always accompanied by an increase (decrease) in the values of any parameter. In the case of mental health problems, these trends are even more blurred. As a consequence, the distribution of the measurement results obtained does not make it statistically justified to divide the examined sample of subjects into "normal" and "pathological" subsamples (Fletcher, Fletcher, & Vagner, 1998).

Second, the subjects in the "normal" and "pathological" groups in fact belong to two different populations: healthy populations and patient populations. In the general population, these populations are combined in a ratio that is usually unknown at the beginning of the study (if there are no results from "cohort," population-based, epidemiological, cross-cultural, and any other studies, including foreign studies of sometimes tens of thousands of subjects). As a consequence, it is problematical to create a research sample that represents the ratio of healthy and diseased people in the population. In this case, almost any sample is not representative or, at best, is a "biased sample." The distribution in clinical studies of a phenomenon in such a mixed and (or) biased sample of "healthy people" and "patients" may not always be described as a normal distribution: Among different "patients," one and the same index can take on different values, overlapping (larger or smaller) values of this parameter in "healthy" people. It is assumed that the proportion of "healthy" people in the population is much greater than the proportion of "patients." Therefore, the distribution curves of the values of the variables in the group of "patients" often "absorb" much of the curve describing their distribution among the «healthy.» As a consequence, the distribution of measurement results in such a biased sample does not allow one to get the standard indices for an unambiguous separation of "patients" and "healthy" people (Fletcher, Fletcher, & Vagner, 1998).

The foregoing discussion provides a reasonable basis for concluding that clinical (in particular, clinical-psychological) phenomena are likely to describe distributions that are different from normal. Therefore, the traditional method of interval normalization of results (z-receipt estimates for the sigma-scale representation of the resulting transformation of the percentile distribution of the values of clinical and psychological symptoms, measured using qualitative ordinal scales) may give nonrigorous or even erroneous results.

First, these results are connected to relatively irregular measurement by the qualitative ordinal scale, yielding opportunities for the interval scale to transform the results obtained by psychometric indices. The most adequate form of mathematical representation of this type of data is a representation of violations – measured data in the form of "a violation is" (1, 2, 3, etc., points) – in the values of the percentile scale. Further data conversion performed by constructing a psychometric scale interval to increase the level of measurement should be performed only if there is sufficient evidence that the sequences presented in the values of the variables have a normal distribution. Above it was clearly shown, in our opinion, that the normal distribution is characteristic only of the quantitative properties (variables); such a characteristic allows operations such as arithmetic operations of addition (Suppes & Zines, 1967). For high-quality sequence data, such arithmetic operations are not always adequately reflected in the nature of the results.

Therefore, at the present stage of development of ideas about the features of psychometric support for clinical neuropsychological research it seems most appropriate to us to conclude that the practice of obtaining only the percentile of normalized results needs to be restricted because this practice allows the researcher to not have to think about the shape of the distribution. The results of research in the form of psychometric standards for high-quality sequence data require substantial justification in each case of experimental diagnostics.

Second, nonrigorous and erroneous results are connected to the lack of substantiated evidence to support the distribution of normal and clinical phenomena in the universe and in the populations of "normal" and "pathological." As a consequence, any "experimental" and "control" samples generated for research purposes are, in fact, nonrepresentative, "biased" samples. In this case, the construction of traditional psychometric scales (based on proven criteria according to the empirical distribution of the normal distribution model) for distributions other than normal ones seems to us mathematically inadequate. At the same time for distributions of any standards based in assessment measures, obtaining the deviation of the results obtained from the average (in the form of deviations from the values of the ranges of $M \pm \sigma$ ($\pm 2\sigma$, $\pm 3\sigma$) and in the form of z-values) is even more inadequate. However, in domestic and foreign literature on psychological diagnostics and psychometrics (Fletcher, Fletcher, & Vagner, 1998; Shmelev, 2002; Suppes & Zines, 1967) there are extensions of the recommendation of this approach to obtaining rules without taking into account the description above of the features of the distribution examined in clinical psychology phenomena.

As proof, we use the citation in the fundamental U.S. text on the problem of the standardization and use of neuropsychological techniques, the *Compendium of Neuropsychological Tests: Administration, Norms, and Commentary* (3rd edition); the position is that, in the absence of a normal distribution of the obtained results, a standardized assessment may not reflect the actual population figures. In particular, it is pointed out that percentile data obtained on the basis of the conversion of grouped frequencies should not be used for traditional psychometric indicators of tests: "When distributions are skewed, the mean and median are not identical because the mean will not be at the midpoint rank and z-scores will not accurately translate into sample percentile rank values" (Strauss, 2009, p. 8).

It is important that the foreign literature, as a rule, be taken into account. We have described above restrictions on the use of the psychometric approach and the principal unresolved issue of the "normal / not normal" distribution of pathological phenomena in populations of "healthy" people and "patients." Therefore, in publications, there should be clearly formulated recommendations about the use of neuropsychological batteries and tests for only those specific populations (for example, different age children with hyperactivity, patients with varying degrees of severity of head injury, psychiatric patients with specific diseases) for which the test norms are calculated (Vanderslice-Barr, Lynch, & McCaffrey, 2008). In the first place, focusing on this particular kind of rule makes the presentation of percentile results statistically and logically justified because of the apparent homogeneity of the sample, and, in the second place, such a rule allows one to statistically justify the use of a minimum of criterion standards in their traditional form – for example, the results of “unsuccessful” subjects or from the value of P_2 (analogue values – 2 σ in the interval quantitative *z*-estimates on the quality percentile scale) or from the value of P_{16} (analog – 1 *z*-value). If there is no possibility of obtaining adequate statistically sound clinical and epidemiological psychometric standards for a population as a whole, use of criterion norms is a clear “step forward.”

In particular, this step creates the preconditions for a transition from a diagnostic paradigm of “the presence or absence of the characteristic,” as discussed above, to a paradigm of “finding a point on the axis of the continuum” that establishes the location of an individual test according to the extent of its expression of the tested characteristic. This paradigm, as described in the classifications of Vasserman and Shchelkova (2004), already applies to the test (measuring) form of diagnosis.

Thus, when creating neuropsychological methods for studying children’s mental development, it is necessary to develop qualitative and quantitative (including psychometric) performance criteria for each sample and the range of possible violations of the test, differentiated depending on the level of the formation of test functions, as well as ways to help children in the survey process, which reveals a zone of proximal development of the function.

In each study, the research specialist should, first, rely on such criteria and, second, clearly categorize the obtained experimental data in accordance with their (metric) measurement of clinical, epidemiological, and psychodiagnostic (including psychometric) characteristics.

References

Ahutina, T.V., & Inshakova, O.B. (2008). *Nejropsihologicheskaja diagnostika, obsledovanie pis’ma i chtenija mladshih shkol’nikov* [Neuropsychological assessment and examination of the writing and the reading of children of the elementary school]. Moscow: Creative center by V. Skachov.

Ahutina, T.V., & Pylaeva N.M. (2003). *Diagnostika razvitija zritel’no-verbal’nyh funkcij* [Assessment visual and verbal functions]. Moscow: Academy.

Apperly, I.A., Samson, D., & Humphreys, G.W. (2005). Domain-specificity and theory of mind: Evaluating neuropsychological evidence. *Trends in Cognitive Sciences*, *9*(12), 572-577.

Berebin, M.A., & Astaeva, A.V. (2008). K voprosu o kachestvennom i psihometricheskom podhodah v sovremennoj nejropsihologicheskoj diagnostike [On the qualitative and psychometric approaches in contemporary neuropsychological assessment]. *Vestnik SUSU* [Bulletin of South Ural State University] (Chelyabinsk), *32,* 19-28.

Homskaja E.D., Efimova I.V., Budyka E.V., & Enikolopova E.V. (1997). *Nejropsihologija individual'nyh razlichij* [Neuropsychology of the individual differences]. Moscow: Russian Pedagogical Society.

Fletcher R., Fletcher S., & Vagner Je. (1998). *Klinicheskaja jepidemiologija* [Clinical epidemiology]. Moscow: Medya Sfera.

Fotekova, T.A. (2003). *Sostojanie verbal'nyh i neverbal'nyh funkcij pri obwem nedorazvitii rechi i zaderzhke psihicheskogo razvitija* [State of verbal and nonverbal functions in the general underdevelopment of speech and mental retardation]. Moscow: MSU.

Furr, R.M. (2008). *Psychometrics.* Thousand Oaks, CA: Saga.

Glozman, J.M. (1999). *Kolichestvennaja ocenka dannyh nejropsihologicheskogo obsledovanija* [The Quantitative evaluation of neuropsychological assessment data]. Moscow: Center Lechebnoy Pedagogyky.

Glozman Zh.M., Potanina A.Ju., & Soboleva A.E. (2006). *Nejropsihologicheskaja diagnostika v doshkol'nom vozraste* [Neuropsychological assessment of children in preschool years]. St. Petersburg: Piter.

Groth-Marnat, G. (2003). *Handbook of psychological assessment* (4th ed.). New York: Wiley.

Haladyna, T.M., & Downing, S.M. (2006). *Handbook of test development*. New York: Taylor & Francis.

Hebben, N. (2010). *Essentials of neuropsychological assessment*. Hoboken, NJ: Wiley.

Joy, S. (2001). Quantifying qualitative features of block design performance among healthy older adults. *Archives of Clinical Neuropsychology, Feb. 16*(2), 157-170.

Korsakova, N.K., Mikadze Ju.V., Balashova E.Ju. (2001). *Neuspevajuwie deti: nejropsihologicheskaja diagnostika trudnostej v obuchenii mladshih shkol'nikov* [Children with learning difficulties]. Moscow: Rospedagentstvo.

Rabin, L.A. (2005). Assessment practices of clinical neuropsychologists in the United States and Canada. *Archives of Clinical Neuropsychology, Jan. 20*(1), 33-65.

Ryazanova, A.Y. (2010). *Differencial'naja psihodiagnostika osnovnyh variantov organicheskogo rasstrojstva lichnosti pri voenno-vrachebnoj jekspertize podrostkov* [Differential psychodiagnostic main variants of organic disorder of person at the military-medical examination of adolescents]. Chelyabinsk: SUSU.

Shereshevsky, G. (2007). *Transkul'tural'nyj analiz razvitija detskoj nejropsihologicheskoj diagnostiki*: Disserattsiya kand. psihol. nauk [Transcultural analysis of development assessment in children neuropsychology: PhD dissertation]. St. Petersburg: SPSU.

Shmelev, A.G. (2002). *Psihodiagnostika lichnostnyh chert* [Psychodiagnostic of personal traits]. St. Petersburg: Rech'.

Sidorenko, E.V. (2002). *Metody matematicheskoj obrabotki v psihologii* [Methods of mathematical processing in psychology]. St.-Petersburg: Rech'.

Simernitskaya, E.G. (1991). *Nejropsihologicheskaja metodika jekspress-diagnostiki "Lurija-90"* [Neuropsychological method of the express assessment of "Luria-90"]. Moscow: Znanie.

Strauss, E.A. (2009). *Compendium of neuropsychological tests: Administration, norms, and commentary* (3rd ed.). New York: Oxford University Press.

Suppes, P., & Zines, R. (1967). Osnovy teorii izmerenij [A foundation of the measurement theory]. In *Psihologicheskie izmerenija* [Psychology of measurement] (pp. 9-110). Moscow: Mir.

Tsvetkova, L.S. (Ed.), (2001). *Aktual'nye problemy nejropsihologiji detskogo vozrasta* [A actual problems of the children neuropsychology]. Moscow: Moscow Institute of Psychology and social.

Vanderslice-Barr, J.L., Lynch, J.K., & McCaffrey, R.J. (2008). Screening for neuropsychological impairment in children using Reitan and Wolfson's preliminary neuropsychological test battery. *Archives of Clinical Neuropsychology, 23*(3), 243-249.

Vasserman, L.I., Dorofeeva S.A., & Meerson, Ja.A. (1997). *Metody nejropsihologicheskoj diagnostiki* [Methods of neuropsychological diagnosis]. St. Petersburg: Stroylespechat.

Vasserman L.I., & Shereshevskij G. (2006). Sovremennye batarei dlja detskoj nejropsihologicheskoj diagnostiki, ispol'zuemye v anglojazychnyh stranah (obzor). Soobshchenie 1 [Modern batteries for children's neuropsychological diagnostics used in the English-speaking countries (Review). Post 1]. In *Obozrenie psihiatrii i medicinskoj psihologii im. V. M. Behtereva* [Review of Psychiatry and Medical Psychology after B.M. Bekhterev], *3*, 13-15.

Vasserman, L.I., & Shchelkova O.J. (2004). *Medicinskaja psihodiagnostika Teorija, praktika i obuchenie* [Medical psychodiagnostic. Theory, practice and education]. St. Petersburg: Akademija.