Problems and prospects of complex psychological phenomena measurement

S Kostromina¹ and D Gnedykh¹,²
¹ Department of psychology, Saint-Petersburg University, 7/9 Universitetskaya nab., St. Petersburg, 199034, Russia
² Author to whom any correspondence should be addressed

E-mail: d.gnedyh@spbu.ru

Abstract. The article is dedicated to the key problems of complex psychological phenomena measurement: the difficulties of measurement systems correspondence in the evaluation of the psychological phenomenon by qualitative and quantitative measurement methods; the difficulties in the comparability of the measurement results in psychological phenomena studying. The problems that arise in the methodology of psychological science and in the organization of interdisciplinary research are discussed. In conclusion, the direction of solving the described problems aimed at joining efforts of psychologists and metrologists is proposed.

1. Introduction
One of the key methodological problems of psychology is the objectification of the studied psychological phenomena. The solution of this problem mainly determines the identity of psychology to Humanities or Natural sciences. The controversy of research methods used in psychology did not stop throughout the XX century which determined the formation of two main areas: psychometric and ideographic with their merits and limitations. The evidence of certain psychological patterns was based on the mathematical statistics methods (STATISTIKA, SPSS, R software package), computer processing of obtained data, complex multivariate methods of statistical analysis: regression, factor and discriminant analysis, structural equation modeling.

The development of psychophysiological and neurophysiological methods of nervous system and human brain investigation mainly contributed to the strengthening of biological approach and natural-scientific directions in psychological research. The question of linearity of the functional brain and mental activity remains open [1, 2]. The experience of the past decades accompanied by the study of complex psychological phenomenon using neurophysiological or genetic methods (e.g., the study of patriotism, or emotive qualities – kindness, aggression, etc.) showed the limitations of direct interpretation and the impossibility of transparent transfer of the received data to psychological phenomenon.

Thus, the main contradiction of the modern methodology of measurements in psychology is the impossibility of the integral assessment of psychological phenomena through the lens of linear functions and direct interpretations. Call for transdisciplinarity [3] and search for a new measurement paradigm different from traditional psychometric rise to a crescendo.
This call is based on the facts that, firstly, phenomena, which are comprehensive in its manifestations and cannot be ‘measured’ using a questionnaire or devices (for instance, consciousness, personality, psychological time, decision-making, psychology of choice, life scenarios, etc.), are increasingly in the focus of psychological research. Most of them can hardly be defined. Secondly, their evaluation requires not only quantitative but also qualitative methods. The need of development of multidimensional design of psychological research is determined by: the intersection of qualitative description of complex psychological phenomenon with those that are objectively and quantitatively recorded; the duality of psychological phenomena genesis, therefore it is difficult to identify the biological and social factors contribution. Third, most psychological research studies any mental phenomenon as a static condition, based on the idea that a personality or any other psychological phenomenon is a phenomenon with stable structure. Hence, in psychology there is the dominance of structuralism, traits theory and system approach, as well as the corresponding measurement technology: concept operationalization, indicators identification, selection of measurement tools and scale for each indicator; their calculation, comparison, and searching for the structure (correlation analysis), classes (clustering), types (discriminant analysis) or cause-and-effect relationships (regression analysis or structural equation modeling) between indicators.

Meanwhile, like any living system, the mental system is characterized by the processes of growth, relative stability, degradation and, sometimes, irreversibility. From the point of view of dynamic approach, sequence and variability are two dimensions of the mental processes nature. That means sequential movement in certain areas of mental development, but in different ways within and between areas.

The lack of methodology to conduct such studies leads to: 1- spontaneous combination of different measuring and evaluation instruments; 2 - discrepancy between studied psychological phenomenon content and measure procedure. As a result, there are unsustained interpretations and conclusions. An additional problem is the rapidly changing environment, in which people live, and mobility of mental structure. Therefore, researchers are faced with difficulties of replication of studied psychological phenomenon for new samples and temporal spans.

2. Measurement of complex psychological phenomena: problems and prospects

Based on the difficulties and contradictions mentioned above, we identified two main problems that arise when choosing of measuring methods and means in psychological research.

1) Difficulties of measuring systems correspondence in estimating of a phenomenon by means of qualitative and quantitative assessment methods.

2) Difficulties in comparability of the results of psychological phenomenon measurements in split-level interdisciplinary analysis.

The first problem is related to the fact that the complexity of a psychological phenomenon is the reason why researchers use several tools for its measurement and evaluation. This difficulty arises not only at the stage of their choice (the question of technique validity), but also at the stage of correspondence of the obtained data. Despite the use of standard methods of statistical analysis, the researchers need to develop a unified system of correspondence of disparate results and their interpretation. The difficulty is that in psychological science there are still no unified principles for the identification of such correspondence, or universal algorithm applicable to the interpretation of data based on the different measurement variables.

Exemplify the problem. The study by Zolotareva [4] aimed to establish similarities and differences in the understanding and description of shame and guilt among generations within a family. Shame and guilt are social emotions associated with the feeling of a mismatch between the act and the normative regulation (shame), or with a deep repentance activating the desire of a person to correct the situation and to change himself (guilt). The emotional context necessitates the use of such measuring means that allow to assess the degree of the emotional expression at the level of understanding (cognitive component), experience (sensory component) and action (behavioral component). In this regard, the most suitable methods are the psychodiagnostic test and the collection of information with
a focus not on statistical measurements, but on understanding, explaining and interpreting information about the subject of the study. That means the combination of quantitative and qualitative methods.

However, if the quantitative measurement methodology is based on the answer to the question ‘How much?’, in qualitative methods not numbers are used, but words that allow to answer to the questions ‘What?’, ‘How?’, ‘Why?’ discovering the processes in the human mind. Such methods include interviews, narratives, focus-group method, discursive and ethnographic research strategies. As Melnikova and Khoroshilov [5] note, there are fundamental principles that allow to suggest the conceptual unity of qualitative research which are fundamentally different from quantitative ones: (1) contextual sensitivity; (2) understanding; (3) interpretative reconstruction; (4) reflexivity.

Contextual sensitivity implies the measurement of context specificity (historical, political, social, etc.), the separation of micro- and macro-social context, the identification of context and its ontological status [6]. The principle of understanding in qualitative methodology is implemented as a process of categorization and typology of the obtained data, and refers to the overcontextual generalization of repetitive semantic patterns/units (content analysis). Interpretative reconstruction of qualitative research involves the definition of meanings from the point of view of the subject (empathic-descriptive level of interpretation) and in depth understanding of the phenomenon based on existing theoretical traditions and prospects (critical-hermeneutic level of analysis). Reflexivity allows to evaluate critically the data analysis.

Thus, the methodology of qualitative research exists in the categories of meaning, word and context. It intersects with the quantitative thesaurus only in the frequency of use of certain words/meanings, and in the percentage expression of dominant judgments that determine different interpretations.

Existing integration models of different types of indicator variables do not always cope with the task of studying complex psychological phenomena. For example, the Rasch measurement model allows to analyze the qualitative data by quantitative methods [7]. However, the values of indicator variables of psychological phenomenon, which being subjected to linear transformation, are often behavioral characteristics. In this regard, the difficulty is that not all psychological phenomena (e.g., personality) have unambiguous "behavioral correlates", based on which it would be possible to predict their development and variability. Thus, this model can be more successfully applied in management and education (to assess the quality of education, achievement tests) [8-11] than in psychology.

The results of qualitative methods (interviews and descriptions of the situation) in the mentioned study was data on the frequency of answers matching in relation to the understanding of shame and guilt or the experience of these emotions in different situations (Likert scale from 0 to 10). When comparing the responses of representatives of different generations within one family, that kind of assessment allowed to conclude about the degree of proximity of their perception of the emotional context and the meaning of shame and guilt. The descriptions of situations, in which subjects experienced shame or guilt, allowed to define the types of situations and calculate the frequency of experience in each of them. Thus, there was the wide phenomenology of the obtained ontological data but their measurement model was limited.

In the described study there were following measuring steps. Data for each subject was represented in the set of numbers: the occurrence frequency of the answers to questions; the number of situations of shame or guilt in a particular area; the percentage of matches; the percentage of the differences in responses between the generations. Each set of numbers served as a model of the subject.

These models can be understood in different ways depending on the problem statements – as a linear space, a multidimensional (one–dimensional) random variable, etc. However, in psychology these models most often are limited to the calculation of frequencies, as well as the percentage of certain thematic blocks/answers. At the same time, even these manipulations do not guarantee the research validity. The attribution of numbers to subjects should make it possible to use these numbers in accordance with the requirements of mathematics. For a psychologist, this turns into a problem because not every numerical relation has the corresponding empiricism. And if some numerical
relation does not correspond to the empirical object properties, its practical use becomes problematic. Especially in respect to scales of low types - nominal, ordinal, etc. [12].

In order to overcome the limitations of qualitative research, the results of interviews or narratives are supplemented by psychodiagnostic tests related to quantitative measurements. In the case of the shame and guilt study, two validated questionnaires ‘Guilty Inventory’ and ‘Test of Self-Conscious Affect’ were used. Both techniques use the classical psychometric approach with the identification of scales allowing to measure the distinct manifestation of shame and guilt key indicators. Thus, the research can be considered as the complementary study of a complex psychological phenomenon. However, as a result, these two kinds of measurement (qualitative and quantitative) exist as two parallels.

The qualitative strategy provided an adequate transition to the system of features describing the observed phenomena that lead to the opportunity for the identification of variables characterizing the processes of understanding and experiencing the shame and guilt. The quantitative parameters reflected the information about the degree of manifestation of shame and guilt presented in the form of the measuring scale. Is it possible to apply a new level of mathematical and statistical data processing? For instance, obtaining of the ontological facts of the second level – the calculation of correlations between parameters of the quantitative and qualitative parts of the research? Is it possible to combine them to calculate the regression model? A reasonable measurement methodology should be developed for this intersection, allowing to combine quantitative and qualitative estimating methods into a unique model [13, 14]. Otherwise, these methods will continue to exist as complementary but non-overlapping measurement strategies.

Another problem of the methodology of complex psychological phenomena measurement is interdisciplinary studies that require comparison of variables at different levels.

As an illustration of this problem, the study of neurophysiological features of cognitive control (CC) in the process of learning activity, for students with different levels of academic success is presented [15]. The main objective of neurophysiological studies of mental phenomena is the search and explanation of physiological mechanisms as the basis of human mental activity [1, 2, 16].

Cognitive control is the multi-level psychological phenomenon [17] with many components and functions [18], involved in regulation of behavior, thoughts and emotions [17]. Cognitive control provides the adjustment of cognitive processes (attention, memory, thinking) to the task solution depending on certain conditions.

The analysis of the methods, tools and instruments used to measure cognitive control allows to identify the main approaches:
- neurophysiological methods (e.g., electroencephalography through power (mV), waves frequency (Hz) and latency (ms) measurement);
- collecting behavioral data (recording of reaction time (ms); analysis of tasks results, e.g., number of correct / incorrect answers/decisions);
- psychodiagnostic empirical methods (psychodiagnostic tests and questionnaires) based on the participant’s decisions about his own behavioristic characteristics (motivation, ability to planning, constancy to purpose, etc.).

Thus, for a comprehensive study, many indicators that allow to record cognitive control’s occurrence and process, and its relationship with other phenomena should be taken into account.

In the mentioned study CC was studied at neurophysiological (by means of electroencephalograph), behavioral (recording of reaction time and number of incorrect answers, data on academic success) and personal (determination of personal characteristics using psychodiagnostic techniques) levels. Mathematical and statistical analysis was the following: nonparametric methods for comparison of independent samples, correlation analysis, MANOVA.

Despite of possibility to compare the data received on the basis of different scales and presented in different quantitative units of measurement, the full correspondence of the results, obtained at different levels of analysis of the same phenomenon, was not achieved. Registration of neurophysiological activity of the brain while subjects performed the task was aimed at identifying the characteristics of
CC through the measurement of physical quantities such as time (ms) and power (mV) of electric waves. Behavioral data related to the manifestation of cognitive control mechanisms reflected the time of decision–making (MS), and its correlation with the correct answer (using the nominal scale). Conclusions about students’ personal characteristics were based on the analysis of scoring system data. The results of academic success (also presented in the scoring system) allowed to divide students into two groups (successful and not successful) and to compare them with each other.

By means of correlation analysis the relationships between certain indicators of cognitive control at the behavioral and personal levels were revealed. For instance, the more time a subject spends on experimental tasks, the more mistakes he makes; the more emotionally stable the student is, the better system of self-management is developed; the higher the student's need to communicate with people, the worse he passes the first exam period, etc.

However, it was difficult to correlate the results of neurophysiological part of the study with other data. In this case, it is not correct to establish the relationship between the power amplitude (mV) of evoked potentials (EP) (as the neurophysiological characteristics of cognitive control) and extroversion (as a personal characteristic associated with CC) not only from the point of view of comparability of these units of measurement, but also from the point of view of logic of interpretation and understanding of mental phenomena.

Regarding to behavioral data, such comparability seems more possible. For example, P300 component (EP at 300 ms) is recorded in the situation of identification, memorization, involvement and decision-making [19] and can by suggested as one of the neurophysiological indicators of cognitive control. Thus, the correlation can be as follow: the higher the P300 amplitude, the less mistakes in the control task are made by the participant of the study. In other words, the more a person is involved in the task, the more successfully he solves it.

Thus, the comparability of measurement results should be reasonable not only from the point of view of logic, but, first of all, from the point of view of Metrology [20]. Whereas these approaches to the measurement of cognitive control are based on different units of measurement, the results of behavioral measurements (at the level of personal characteristics) might not be comparable with neurophysiological ones. However, commonsense reasoning ‘the more a person is involved in the task, the more successfully he solves it’, which was made by the researcher, may seem quite logical. Nonetheless, there was no objective confirmation of this statement based on the assurance of uniformity of measurements of involvement and success (due to the impossibility to measure success in mV in comparison with involvement (which is associated with the component P300)).

Consequently, when measuring complex psychological phenomena, comparability of the results obtained at different levels of analysis becomes difficult. In this regard, there are also difficulties in correlating different results at the stage of their interpretation. Would be methodologically correct their generalization in the absence of mathematical and statistical basis? If the results, based on different methods of cognitive control measurements, are that students with high academic success have a high P300 amplitude (mV) and their self-management ability correlates with extroversion, can we conclude that the magnitude of P300 amplitude depends on students’ desire to build social ties (or vice versa)? Obviously, until the mathematically and statistically reasonable confirmation will not obtained, such conclusion is not legitimate.

In our opinion, the key decision of these problems is the search for a new approach to the definition and development of ‘universal’ units of psychological phenomena measurement and measurement methods that can connect different levels of mathematical and psychological analysis of complex psychological phenomena.

3. Conclusions
Thus, despite achievements in the application of new statistical methods in psychology, ‘zone of proximal development’ of psychometrics is the development of an integrated methodology of measurement and evaluation of complex psychological phenomenon, establishing evidence-based principles for combining qualitative and quantitative research methods, psychological, sociological
and neurophysiological techniques, as well as their interrelated mathematical and statistical processing. Psychological phenomena that exist at different levels, have different forms of manifestations, in certain contexts change their meanings, and characterized as self-developing systems are increasingly in the focus of psychologists. Nowadays, psychology is in need of a measurement methodology that allows to consider the processes of growth, irreversibility, variability, fluctuation and the potential of self-change of the mental system. That is, the methodology should base not on structuralism, but on procedurality (the idea of measuring complex multidimensional nonequilibrium systems).

Acknowledgments

Supported by RF Government grant contract No.14.W03.31.0010.

References

[1] Roik A O and Ivanickij G A 2011 Neurophysiological model of cognitive space J. Higher Nerv. Act. 61(6) 688–696 (in Russian)
[2] Rose N 2016 Reading the human brain: how the mind became legible Body Soc. 22(2) 140-177
[3] Gusel'tseva M S 2015 Psychology and new methodologies: epistemology of complex Psych. Res. 8(42) 11 (in Russian)
[4] Zolotareva A S 2017 Understanding of the Phenomena of Shame and Guilt by Representatives of Different Generations within One Family (SPbU) p 109 (in Russian)
[5] Melnikova T A and Khoroshilov D A 2013 Methodological principles of qualitative research in psychology Bulletin of Moscow University Ep. 14 Psychology 3 4–14 (in Russian)
[6] Hammersley M 2008 Context and Contextuality The Sage Encyclopedia of Qualitative Res. Methods Ed L M Given (Thousand Oaks, CA: Sage) vol 1 pp 122-123
[7] Mari L and Wilson M 2014 An introduction to the Rasch measurement approach for metrologists Measurement 51 315-327
[8] Ingebo G S 1997 Probability in the Measure of Achievement (Chicago: MESA Press) p 148
[9] Maslak A, Moiseev S and Nasonova T 2018 Application of the Rasch method of evaluating latent variables in management and administration Society. Integration. Education: Proc. Int. Scientific Conf. VI 330-340
[10] Maslak A, Anisimova T, Rybkin A and Titenko E 2015 Measuring and monitoring of education infrastructure of southern federal district of the Russian Federation Proc.—Social and Behavioral Sciences 214 1062-69
[11] Pozdnyakov S A 2007 Study of Latent Variables Measurement Accuracy in Education (Slavynsk-na-Kubani: Izd. centr SGPI) p 118 (in Russian)
[12] Tolstova YU N and Maslennikov E V 2000 Qualitative and quantitative strategies: empirical research as a dimension in the broad sense of the word Sociol. Res. 10 101-108 (in Russian)
[13] Fisher W P Jr and Stenner A J 2011 Integrating qualitative and quantitative research approaches via the phenomenological method Int. J. of Multiple Research Approaches 5(1) 89-103
[14] Laba L YA 2004 Approaches to integration of qualitative and quantitative methods Sociological research 2 124-129 (in Russian)
[15] Kostromina S N, Mkrtchian N A, Kurmakaeva D M and Gnedyk D S 2017 The interrelationship between cognitive control and academic success of first-year students: An interdisciplinary study Psych. in Russia: State of the Art 10(4) 60-75
[16] Ivanickij A M 2012 ‘Reading the brain’: achievements, prospects and ethical problems J. Higher Nerv. Act. 62(2) 133–142 (in Russian)
[17] Posner M I and Snyder C R R 1975 Attention and cognitive control Information Processing and Cognition Ed R L Solso (Hillsdale N Y: The Loyola Symposium. Hillsdale N Y) pp 55–85
[18] Miyake A, Friedman N P, Emerson M J, Witzki A H, Howarter A and Wager T D 2000 The unity and diversity of executive functions and their contributions to complex ‘frontal lobe’ tasks: a latent variable analysis Cognitive Psychol. 41 49–100
[19] Gnezditsky V V and Shamshinova A M 2001 Experience of Application of Evoked Potentials in Clinical Practice (Moscow: AOZT ‘Antidor’) p 480 (in Russian)

[20] Biserova V A, Demidova N V and Jakoreva A S 2007 Metrology, Standardization and Certification (Moscow: EKSMO) p 10 (in Russian)