Development of a test to determine the level of intellectual development of 5-6-year-old children

Authors:

Ilaha Gulaliyeva, Moscow State University, Baku Branch; *gulalieva_ila@mail.ru*

Mahammad Adigozalov, State Examination Center of the Azerbaijan Republic; *maqo1483@hotmail.com*

Key words: standardization, intelligence test, preschool age.

Abstract

Most of methodological literature for practical psychologists in the market often includes tools which have not been properly tested. On the other hand some specialists and schools apply foreign norms and standards which may not culturally appropriate for our society to diagnose children.

We carried out theoretical and methodological analysis of existing intelligence tests, and tried to describe psychological problems of children that arise when entering school. The study revealed the lack of a methodology in Azerbaijan that fully complies with all standards of approved methods for identifying the level of intellectual development of preschool age children.

This article reflects the study of the mental characteristics of preschool age children and the development of a new standardized toolkit for diagnosing the level of their development.

We used generally accepted structure of the mental development of children of preschool age and developed methodology for the comprehensive diagnosis of the level of intellectual development consisting of verbal and non-verbal blocks, as well as separate questionnaires for parents and teachers to obtain comprehensive data on socio-psychological factors. Verbal block includes assessment of general awareness, abstraction ability; non-verbal block includes assessment of arithmetic ability and spatial recognition. The paper describes the process of development and the results of the initial testing of the methodology. The primary version of the test has been standardized, tested for validity and reliability. The reliability of the test was determined by repeated testing after 3 months, the validity was checked by comparing the diagnostic results with the Raven test. Predictive validity was determined by comparing peer review with test results, as well as the success of passing the exam for admission to school.

Chi-square methods and Spearman's rank correlation coefficient were used to determine the significance of differences. The results of the subtests diagnosing 5 characteristics showed a high degree of compliance ($p < 0$, () 1).

Introduction

The education in Azerbaijan is undergoing reforms that require modern educational institutions to utilize effective methods for monitoring and evaluating student's knowledge.

The test tools used in the educational field cause concern among parents and teachers that implies a thorough analysis of the quality control of standardization as reliable tools for assessing educational achievements. The issues of improving the control of test methods need extensive research.

Mainly intelligence tests are applied for diagnosing the levels of student's knowledge and development. The diagnosis of preschooler's level of intelligence is difficult due to age-related characteristics as well as a small number of reliable test methods.

Azerbaijan started using tests inn educational system at the end of the 20^{th} century for assessment of knowledge on various subjects, for selection of students to universities, but psychological tests were put in practice much later.

An acute shortage of high-quality psychological toolkit led only to the borrowing and using western materials in the Azerbaijan sample often without going through all the stages of the adaptation.

This point raises another issue: the need for design and justification of standards in the field for training of specialists in various educational areas, the further development of the national theory of pedagogical measurements.

In this way the object of our activity is the development of a testing toolkit for diagnosis of intelligence level of 5-6-year-old children.

Tasks of activity:

1. theoretical analysis of intelligence tests for 5-6-year-old children,
2. design of a new tool for diagnosing intelligence by analogy with Wexler,
3. standardization of the test on a sample of primary school children in Azerbaijan.

Child's basic mental qualities develop by 5 years. The analysis of these qualities allows us to compose ideas about the formation of the children's mental state and outline ways/directions of development or correction.

Today major large multiscale intelligence tests are developed abroad that's explained by the increasing popularity of this trend in the science of Western countries (Sternberg, 2002; Binet, 1916; Raven, 1936; Wexler, 1939; Eysenck, 1995). Domestic development of tools for measuring the intelligence mostly limited to the adaptation of foreign techniques.

As a result of the analysis of the literature, two main approaches to the classification of intelligence tests can be distinguished: psychometric (Galton, Spearman, Galton, Eysenck) and system theories (Gardner, Sternberg).

There are two basic methods in the study of the child diagnosis: test and clinical. Test examination allows you to diagnose objectively, excluding the influence of certain factors. Therefore there are certain requirements for the research procedure, prescribed instructions and stages of the examination.

Psychological diagnosis in our country is in the stage of initial formation that is development of baselines and methods for implementation of these baselines.

We expect that this work will contribute to the science and we will be able to obtain significant results. In this regard, our main task involves the development of a comprehensive questionnaire that could be applied for practical purposes. We are interested not only in the final score, the number of correct answers, but also the nature of the mistakes made by the child, the analysis of his behavior during diagnosis and the collection of data from parents and teachers.

All these processes permit us to get a more detailed picture of the mental development of the child.

Research Methodology:

Having collected information on existing methods for diagnosing the intelligence, as a sample we took Wechsler Intelligence Scale for Children (WISC) - the test developed by David Wechsler for testing children and adolescents (from 6 to 16 years).

The test design process included the following steps:

1. Analysis of various tests for children and selection of the best working scales (according to the literature).
2. Selection of subtests and creation tasks for each subtest by the level of complexity - easy, medium and complicated levels.
3. Initial verification of the toolkit was carried out using mathematical processing methods on a sample of 80 children.
4. Removing items from the original version of the test that don’t work and have low reliability.
5. Development of new tasks and the final version of the methodology.
6. Data collection for experimental verification of the toolkit and in parallel on the Raven test for validation. There were examined 20 children during the testing.

The designed test includes 4 subtests divided into 3 levels of complexity (easy, medium and complicated). There are 98 tasks in total.

The verbal block consists of 2 subtests: "the level of conversance", "the ability of generalization and abstraction". The conversance is aimed at revealing the breadth of cognitive interests as well as intellectual functions: memory and attention. The analysis showed that this section is predictive with regard to verbal and general intelligence. Ability to generalize and abstract - aims to identify the ability of logical generalization, the presence of the ability to abstract and the ability to qualify, organize information and develop conceptual thinking.

There are also 2 subtests in the non-verbal block: "the spatial thinking" and "the arithmetic unit". The spatial subtest implies the ability to correlate the special and the general, the dealing according to the standard which includes heuristic components of thinking. The arithmetic subtest determines the presence of basic mathematical knowledge, properties of attention, visual-motor coordination, ability to construct the consistent chain and to solute the logical decision tasks.

Results:

There were 53 test items in the original embodiment and the empirical data were collected to verify tasks on 5-6 years old children groups in kindergartens, schools in Baku.

The children's answers were converted to points: 1 - corresponds to the correct answer, 0 - incorrect answer. We calculated the average scores for each subtest in order to distribute the answers according to the level of complexity. We separated tasks by using averages that were easily solved by children, and vice versa, a large number of mistakes were made. The raw scores have been converted into standardized scores. The empirical mean was calculated M = 0.57, the standard deviation σ = 0.13 thus the normal range for the distribution over the average values was from 0.44 to 0.70. At the next stage, we conducted an analysis to check the significance of the obtained empirical average and how the deviation of the empirical average is statistically significant from the theoretical one.

Table №1. One-Sample Kolmogorov-Smirnov Test.

	Test Value = 0.5					
	t	df	Sig. (2-tailed)	Mean Difference	95% Confidence Interval of the Difference	
					Lower	Upper
VAR00092	3.049	53	.004	.06074	.0208	.1007

The level of significance is below 0.05 which indicates the deviation of the obtained the empirical mean is significantly different from the theoretical one. Thus, the sample is shifted to the right, which indicates the presence of easily solved problems.

Testing for normality

Table №2. One-Sample Kolmogorov-Smirnov Test.

	1 subtest	2 subtest	3 subtest	Total
N	80	80	80	80
Kolmogorov-Smirnov Z	,758	,952	1,185	,865
Asymp. Sig. (2-tailed)	,614	,325	,120	,442

This table specifies the level of significance of Kolmogorov-Smirnov Test that used to test the normality of research's distribution. If the level of significance is greater than 0.05 then we accept the 0 hypothesis - a representative sample. The table shows us that the distribution corresponds to the principle of normality.

Testing the consistency of items and the reliability of the scale.

Table № 3. Alpha Crombach's values for 3 Subtests.

	1 subtest	2 subtest	3 subtest

Alpha Crombach	,622	,864	,870
Number of items	20	21	28

These subtests have a high level of reliability, items are well coordinated with each other. The parameter's value is close to 1 which indicates a high congruence of items among themselves. The statistics of the correlation of the item with the total score for subtests allowed us to identify tasks that, after a meaningful analysis, can be shifted or a decision on deletion made.

The next step during the statistical analysis of results was conducting the sphericity test that is carried out in order to make a decision on factor analysis. As we can see from Table №4 the level of significance of the Bartlett Sphericity Criterion is 0 which indicates the possibility of using factor analysis. The correlation matrix of primary variables differs from the spherical one so we can conduct factor analysis.

Table № 4 . KMO and Bartlett's Test

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		,574
Bartlett's Test of Sphericity	Approx. Chi-Square	345,875
	df	210
	Sig.	,000

Factor analysis will be carried out after increasing the sample.

Test tasks were changed and transformed based on mathematical analysis. Unreliable tasks were deleted and new ones with different levels of complexity were added. The 4th subtest was completely redone as it showed low reliability.

The 2nd stage of test designing included the adding process of new tasks, 98 tasks were created instead of the previous 68. Correlation analysis was carried out between the test indicators and expert assessment ($p \leq 0,05$, $R=0,243$) and also test Raven ($p \leq 0,05$, $R=0,32$) for validity verification. The process is at the data collection stage.

Conclusions:

Standardized testing has gone a long historical way of formation and has become an effective tool for assessing students' knowledge, skills and abilities. Its effectiveness depends on a number of factors - the formulation of the goal, the selection of the control content, the choice of methods, criteria and assessment scale, principles and technology for the designing of tasks and test, the phased realization of testing, proper statistical processing of the results, etc.

This article described the process of the initial designing of test tools for diagnosing the level of intelligence of children of 5-6 years old in Baku. Initial approbation was carried out on a sample of 80 people and showed high reliability of tasks and test items, at the same time, the predominance of light types of tasks over complex ones was revealed. This state posed the following task for us - the adjustment of previous tasks and the creation of new ones more accurately to be able to differentiate the preparedness of children. At the second stage 98 tasks were created, the approbation process is at the data collection stage, therefore, statistical data are not presented in this article yet. The prospect of further research is to continue

approbation process of test items and using them for diagnosing the level of development of children at preschool age.

Литература:

- *Айзенк Г. Дж.* Коэффициент интеллекта. — Киев: Гранд, 1994. — 112 с. — 50 000 экз. — ISBN 5-7707-3600-8.