Anton Agapov, READ Program Expert

Using SAM to Measure Computer Literacy

This work is devoted to the development of a diagnostic tool aimed at assessing information and communication literacy, and based on the SAM level model. The relevance of the work is due to the high pace of development of modern information technologies, their penetration into the educational process and the everyday life of adolescents. In response to this, part of educational policy in different countries has become the introduction of information and technology literacy as an expected outcome of school education. This, in turn, has led to the emergence of various monitoring studies related to ICT literacy.

The developed instrument, in addition to its testological characteristics, must have technological simplicity in development and interpretation in order to expand the possibilities for its use in the CIS countries. The authors chose a practice-oriented approach to developing tasks, which will allow assessing ICT literacy in a situation as close as possible to reality, and facilitating the interpretation of the results.

The main stages of the work were the analysis of the main modern theoretical models for measuring ICT literacy; analysis of the age characteristics of the target audience and the requirements of the Federal State Educational Standard in the field of ICT literacy; development, examination and testing of the test and its web version.

There are many approaches to defining and building the structure of IT literacy in the world. In PISA, it is "the interest, attitude and ability of people to use digital technologies and communication tools appropriately to access, manage, integrate and evaluate information, gain new knowledge and communicate with others to participate effectively in society." In ETS, it is "the use of digital technologies, communications and/or networks to access, manage, integrate, evaluate and create information to function in a knowledge society."

Computer literacy structures are similar from source to source. Activity-based structures typically include the following components:

- Access - knowing and knowing how to collect and/or retrieve information.
- Management - application of an existing organizational or classification scheme.
- Integration - interpretation and presentation of information. It includes generalization, comparison and contrast.
- Evaluation - making judgments about the quality, relevance, usefulness or effectiveness of information.
- Creation - generating information by adapting, applying, designing, inventing or creating information.

An alternative approach can be called structuring IT literacy according to content elements:

1. Use of computer technology

- Working with files – open, read (including unknown type), edit
- Installing and uninstalling programs, data backup
- Using application programs (photo editor)

2. Working with information

- Using search

- Relevance check
- Analysis, synthesis, visualization of data

3. Communication on the Internet

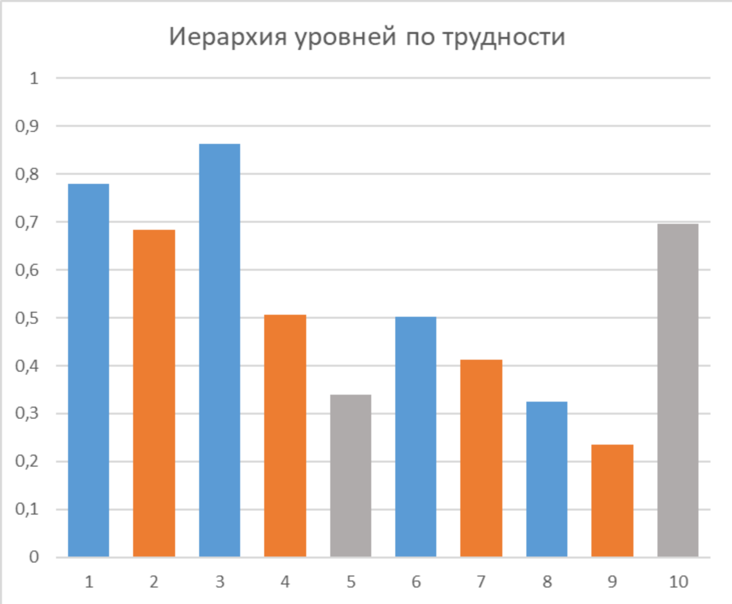
- Sending letters, using instant messengers

This structure was chosen for further development of tasks. Our goal was to create a practice-oriented tool that would allow us to observe the manifestation of literacy in the process of working with computer systems, and at the same time retain our affiliation with SAM theory. To do this, it was necessary to correlate the levels of mastery of subject tools within IT literacy with the formats of the tasks.

Level	I	II
Skills	Organization of works	Integration, Interpretation
Use of computer		
Work with files (to open, to read, to edit)	To open file of unknown format (doc, txt), to enter data into the answer field	To download unknown file, to find a way to open it, to enter its data to relevant field
Use of applications	Download the graphic file, enter the text that is written on it in the response field	Download the picture, recognize the text on it using the editor, enter it in the answer field. Download the audio file, play it backwards / slow down / speed up, enter the listened text in the answer field
Work with information		
Use of search engines	Find specific objective information (dates, numerical data, last names)	Find information by combined query (search for contemporaries)
Check for relevance	True/False choice	True/False on combine requests
Analysis, generalization, visualization	Select the diagram that best describes the data	Calculate total/average values for the found indicators
Internet communication		
Emails and messengers	Sending emails, entering replies	Restore the sequence of events from letters (analyze dates and chain)

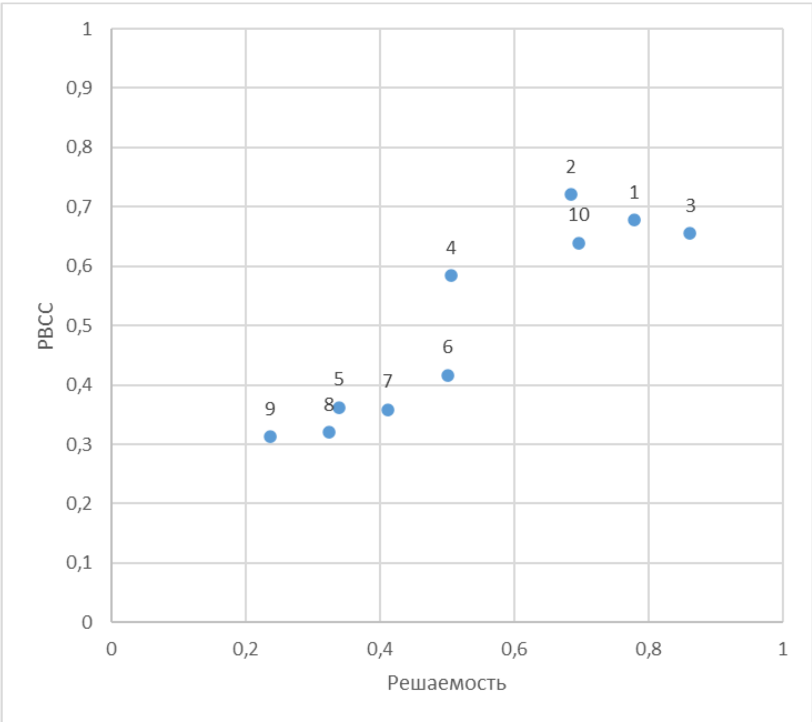
The given task formulations were implemented in an online format. The files were processed on the subject's device, which provided additional contextual information. 955 students in grades 8-10 took part in testing the tool. Below we present the test results.

The tasks demonstrate a hierarchy of difficulty in accordance with the theoretical level, which indicates compliance with the theoretical model. The difficulty of tasks at each level is presented in the diagram. Items 5 and 10 do not have a “pair”, that is, an item assessing the same skill at a different level. Otherwise, level 2 tasks are more difficult than level 1 tasks on the same content element.



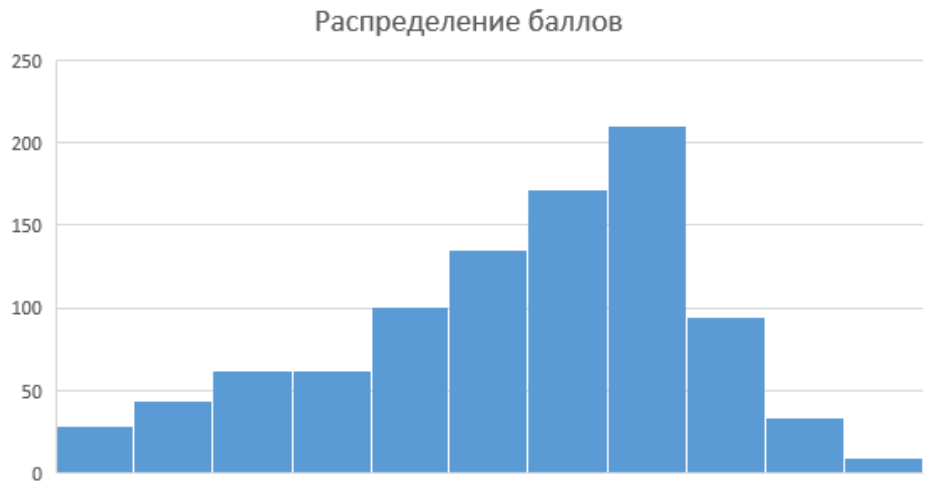
The table below shows the main characteristics of the test items: solvability and the point-biserial correlation coefficient, used as an indicator of the discriminative ability of the items.

	1	2	3	4	5
solvability	0,779	0,684	0,863	0,507	0,339
PBCC	0,678	0,721	0,655	0,584	0,362
	6	7	8	9	10
solvability	0,502	0,412	0,324	0,236	0,697
PBCC	0,417	0,359	0,320	0,313	0,639



Overall, the tasks demonstrate a good level of discrimination and solvability. There are no extremely difficult or extremely easy tasks. Discriminativity is interpretable across the entire set of tasks and lies in the range (0.3; 0.75).

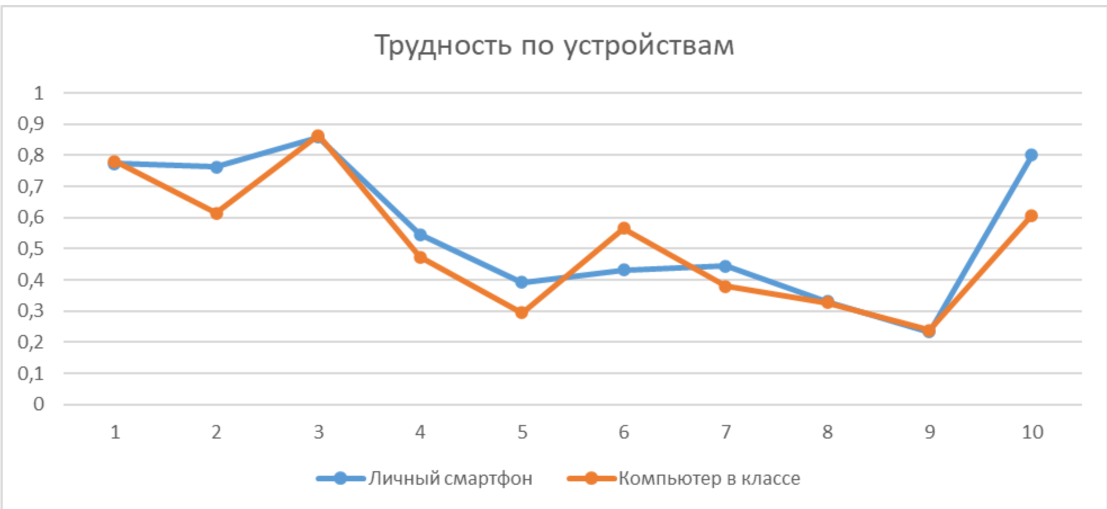
The following chart shows the distribution of test scores:



As can be seen from the diagram, the scores are distributed quasi-normally with an upward shift. That is, the test turned out to be easy for a sample of subjects, but, nevertheless, the distribution indicates the one-dimensionality of the test.

The following is a breakdown of the devices that subjects used while taking the test and the corresponding differences in difficulty.

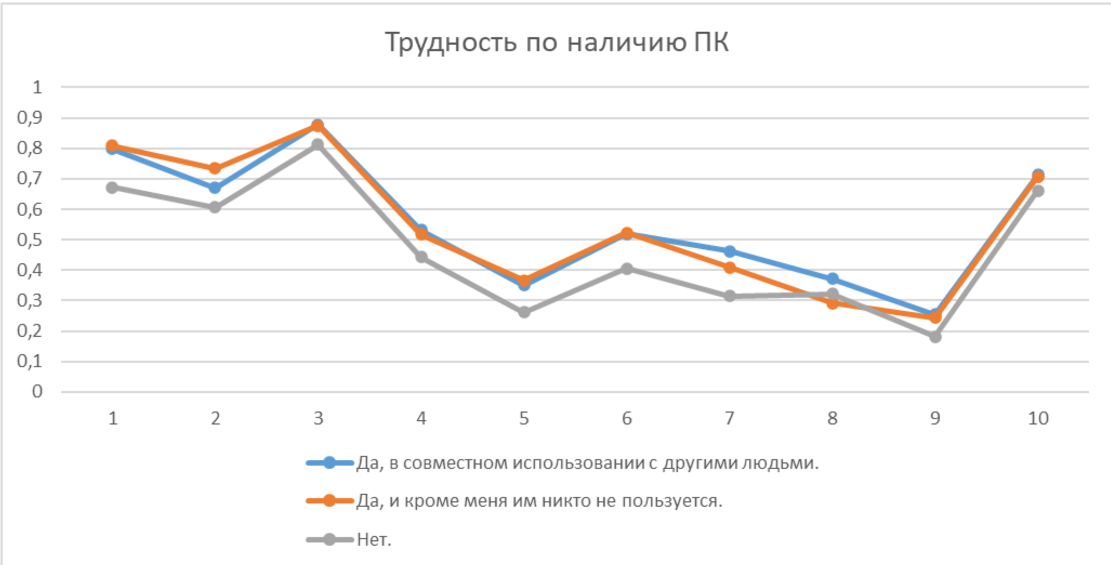
	1	2	3	4	5	6	7	8	9	10
Personal smartphone	0,773	0,762	0,858	0,546	0,391	0,431	0,445	0,330	0,234	0,800
Computer in a class	0,780	0,614	0,864	0,474	0,295	0,565	0,380	0,327	0,238	0,608



As can be seen from the table and graph, most tasks function the same on all types of devices. However, tasks 2, 6 and 10 show significant differences in difficulty depending on the device, and require additional analysis.

The following provides evidence of differences in item difficulty as a function of various contextual data. The table and graph below show the difficulty of the tasks depending on whether the test subject has a personal computer.

	1	2	3	4	5	6	7	8	9	10
Yes, when shared with other people.	0,798	0,671	0,878	0,531	0,350	0,519	0,463	0,371	0,255	0,712
Yes, and no one uses it except me.	0,809	0,733	0,874	0,517	0,366	0,522	0,409	0,292	0,244	0,708
Нет.	0,673	0,606	0,812	0,442	0,261	0,406	0,315	0,321	0,182	0,661

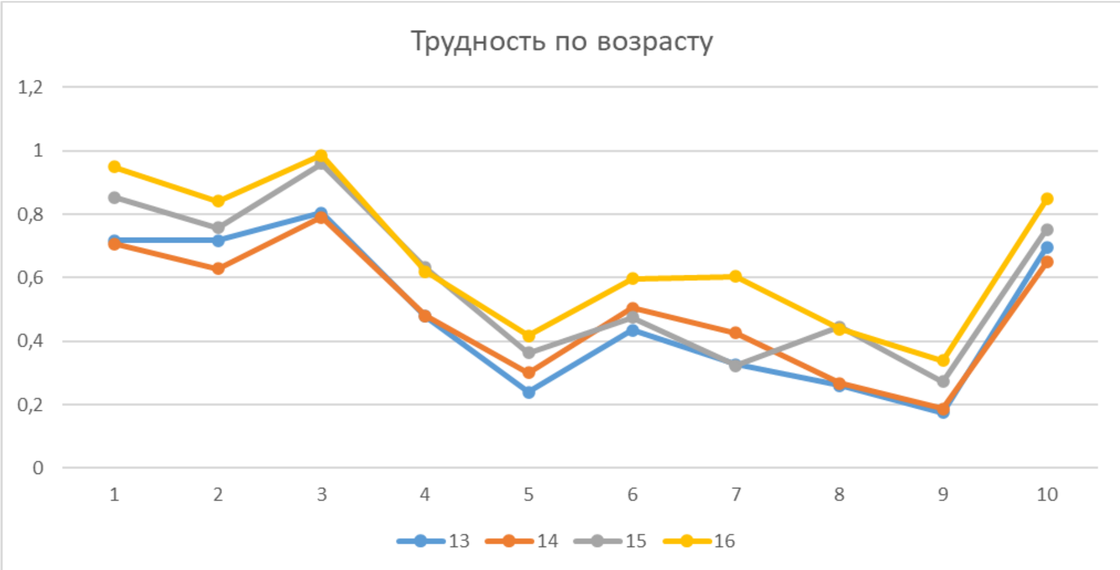


As can be seen from the data, subjects who do not have a personal computer cope with tasks in general worse than those who do have one. While the factor of computer sharing has a lesser effect on task completion.

The impact of having your own computer on success in completing tasks is understandable. Those with a personal computer have more experience interacting with files and data (differences in performing non-PC tasks are not as obvious), while subjects who do not have a computer are less likely to work with files on their own.

Below are data on the performance of tasks by subjects of different ages.

	1	2	3	4	5	6	7	8	9	10
13	0,717	0,717	0,804	0,478	0,239	0,435	0,326	0,261	0,174	0,696
14	0,707	0,628	0,791	0,480	0,300	0,504	0,426	0,267	0,187	0,650
15	0,854	0,758	0,960	0,631	0,364	0,475	0,323	0,444	0,273	0,753
16	0,950	0,842	0,986	0,619	0,417	0,597	0,604	0,439	0,338	0,849



The data shows that older participants (15-16 years old) cope better with tasks than younger ones (13-14). Apparently, this is associated not with age, but with the class of study. However, such a distribution can serve as an indicator of the validity of the test (the test measures a skill that develops over time).

The difference in results among subjects of different ages is quite understandable. Older subjects take a computer science course at school and gain experience interacting with PCs in other contexts - project work, and others. Therefore, it is logical to assume that more experienced subjects will cope better with the tasks.

Cronbach's alpha method was used to assess reliability.

This method compares the spread of each item with the overall spread of the entire scale. If the spread of test results is less than the spread of results for each individual question, then each individual question is aimed at exploring the same common ground. They produce a meaning that can be considered true. If such a value cannot be developed, that is, a random scatter is obtained when answering questions, the test is not reliable and the Cronbach alpha coefficient will be equal to 0. If all questions measure the same attribute, then the test is reliable and the Cronbach alpha coefficient in this case will be equal to 1.

Cronbach's alpha will generally increase as the intercorrelations of variables increase, and is therefore considered a marker of internal consistency in assessing the validity of test scores.

Since maximum intercorrelations between variables across all items are present when the same thing is being measured, Cronbach's alpha indirectly indicates the extent to which all items are measuring the same thing. Thus, alpha is most appropriate to use when all items are aimed at measuring the same phenomenon, property, phenomenon.

However, it should be noted that a high coefficient value indicates the presence of a common basis for a set of questions, but does not indicate that there is a single factor behind them - the one-dimensionality of the scale should be confirmed by additional methods. When a heterogeneous structure is measured, Cronbach's alpha will often be low. Thus, alpha is not suitable for assessing the reliability of intentionally heterogeneous instruments (for example, for the original MMPI, in which case it makes sense to conduct separate measures for each scale).

Professionally designed tests are expected to have an internal consistency of at least 0.70. The alpha coefficient can also be used to solve other types of problems. Thus, it can be used to measure the degree of agreement between experts assessing a particular object, the stability of data during repeated measurements, etc. The Cronbach alpha coefficient for the given test is 0.6, which does not allow us to make an unambiguous conclusion about the reliability of the test, although this value may be explained by the heterogeneity of tasks in terms of the skills used, and their small number.

The developed tool is a combination of a flexible and easy-to-use theoretical framework, easy-to-produce tasks, developed in a practice-oriented approach and allowing for a qualitative assessment of students' ICT literacy.