SSAC Computerized Item Banking and Test Development Model

Maleyka Abbaszade
Chairperson
State Students Admission Commission of the Republic of Azerbaijan
office@tqdk.gov.az

Tural Mustafayev
Head of the Technological Control and Information Security Sector
State Students Admission Commission of the Republic of Azerbaijan
turalm@yahoo.com

Rahman Mehraliyev
Chief consultant at the Item Bank Department
State Students Admission Commission of the Republic of Azerbaijan
rehman_mehrali@mail.ru

Abstract

An item bank is a repository of assessment items and associated software. Item banking software enables items to be stored, added, edited, selected and retrieved. In order to support the selection and retrieval of items, each item needs to be stored with metadata called item parameters.

The State Students Admission Commission (SSAC) of the Republic of Azerbaijan has developed the in-house Computerized Item Banking and Test Development system named "Test Builder". This paper details the item banking and test development approach adopted at SSAC along with the features the "Test Builder" provides.

Key words

Computerized item bank, automated test development and administration, educational and psychological measurement

Automated Item Banking

For many years tests have been administered for various reasons, such as educational testing or assessing particular skills. Results of these tests are used, for instance, as a factor in determining whether an examinee should be admitted to study at an educational institution. Such tests are typically administered to a large number of examinees. Due to some concerns, including administration costs and security, automation of the entire testing process becomes an issue of great importance. Automated item banking and test development make a significant contribution to this process. [1]

Test constructed from item bank is a very powerful tool in educational and psychological measurement. An item bank is a large collection of test items all measuring the same domain of knowledge, stored in computer storage together with their parameters. Known item parameters allow the test constructor to have strict control of them and make it possible to search for items according to a number of criteria.

The goal of the testing is to construct a test efficiently for the purpose of measuring a skill, ability, etc. Therefore, each test is constructed to conform to a test specification, which defines the rules and/or constraints for selecting the items. In constructing a test, items are selected from item bank, so that the combination of selected items satisfies the test specification.

A test is typically divided into sections of questions. The test specification generally defines the number of items to be presented in the test, the number of test sections, the number of items in each section, and the time set for taking the test. The test specification

also defines criteria for item selection. These are based on item characteristics, which include, for example, item content, e.g., mathematical questions relating to arithmetic, algebra, or geometry and parameters of items, difficulty of the selected items as an example. [2]

The approach to design of an automated item banking and test development system can be envisioned as an integration of three phases. Phase I involves the development of an item bank that contains each item's content and parameters, and has full editing capabilities. The parameters include, but are not limited to, item statistics, such as item difficulty, and statistics of the test where the item appeared. Phase II provides the capability to retrieve items from the item bank in order to construct tests with prespecified characteristics. Item parameters are used for this purpose. Finally, in Phase III, the actual test booklets are published with specific format. [3], [4]

In order to provide these three phases the State Students Admission Commission (SSAC) of the Azerbaijan Republic has developed the in-house Computerized Item Banking and Test Development system named "Test Builder". This system consists of two parts. The main part of the system is the item bank which has been conceptualized as a database. The second part of the system is the computer program designed to efficiently perform various tasks using this database. Program features include user-friendly interface, ease of editing or adding new items to the bank (the powerful text processor MS Word that is integrated into program's user interface is used for item entry and editing; this eliminates characteristic of many similar programs' limitations related to handling special characters and non-text data), retrieval of statistical reports on the state of the item bank (both for electronic use and printing), creation of various types of exams, intelligent item selection based on predefined request, and compilation and auto-formatting of test booklets.

Classification of item parameters and item format

Since 1994, SSAC has been developing item bank for administering entrance examinations to Bachelor's level of higher schools. In 2005, SSAC launched a new project of development of the item bank intended for use at entrance examinations to Master's level of higher schools. Today SSAC is running the item bank consisting of over 200 thousand items and this bank is continuously growing.

SSAC item bank is subject-based. The organizing structures of the bank have been designed and developed by discipline committees of teachers. Paying particular attention to what is now being taught throughout the Republic, they have developed comprehensive course outlines. Test items have been collected from a large number of teachers. Each item in SSAC item bank has 6 main parameters, namely the subject, the topic within that subject (content parameters), the degree of difficulty, the cognitive level called "relevancy" (skill parameters), the "course", indicating types of examinations an item is to be used in, and indicator of similarity called "analogousness". As already mentioned, items are grouped by subjects and each subject domain is divided into a number of topics. For each item subject experts determine the most corresponding topic. They also assign the degree of difficulty and cognitive level to each item. Subject experts use 3 difficulty levels (easy, moderate and difficult) to characterize item difficulty. Further, their estimations are checked and revised when an item is used in examination. Subject experts use 7 cognitive levels to distinguish the different thought processes required to respond. The cognitive level reflects the manner in which knowledge is being assessed, rather than the technical difficulty of the content. Items coded to cognitive level I (knowledge of terminology) elicit knowledge of terminology. Items coded to cognitive level II (knowledge of facts) elicit knowledge of specific facts, methodology, principles, theories and structures, and ability to memorize, repeat and identify. Items coded to cognitive level III (reasoning based on generalization) elicit the ability to generalize and differentiate specific facts, principles, theories and structures, and the skills such as comparison and classification. Items coded to cognitive level IV (interpretative

reasoning) elicit the ability to interpret events or facts rather than naming them, often involving cause/effect relationships. Items coded to cognitive level V (predictive reasoning) elicit ability to predict consequences. Such items describe a specific or hypothetical situation, and require forecasting consequences when certain period of time passes or when some factors determining a situation are changed. Items coded to cognitive level VI (modeling or application) elicit ability to use information, methods, concepts and theories in new situations, to solve problems that have single or best answers by applying acquired knowledge, and techniques and rules in a different way. Items coded to cognitive level VII (calculation) elicit ability to solve problems using mathematical principles and formulas. In this sense calculation can be regarded as one of the forms of application.

Each item has also an indicator of similarity assigned to indicate whether an item has an analog or parallel, i.e. an item or items with content, difficulty and cognitive level similar to its. Analogous or parallel items are used to build parallel (equivalent) test forms.

At present SSAC administers paper and pencil tests mainly using traditional multiple-choice items. Items are written by teachers and reviewed comprehensively by the experts in the field. Each item consists of a stem containing the introductory question or partial statement that the examinee must answer or complete, four or five alternatives, called distractors, marked by the letters A, B, C, D, and E. One of the distractors, the key, is the correct, or best, response to that stem. Other information stored on items includes unique item identifier, exams in which it has been used (generally, a used item is not included in the tests administered in the next two years), names of the item writer and subject experts revising the item, date created, and other notes concerning the item. These notes may include, for example, solution of a problem presented in the item.

Automated test development

SSAC has developed the in-house Computerized Item Banking and Test Development system named "Test Builder" which consists of several integrated procedures designed to handle an entire spectrum of test development process.

Adding, storing and editing test items.

Items are added and stored in database as MS Word files in XML type. <ADD NEW ITEM> command opens a program window with MS Word window integrated into it (Fig. 1). This command also provides a special template where user must complete appropriate fields. This template is strictly structured which prevents a user from corrupting item structure accidentally. User enters each element of an item (namely, item stem and answers, or alternatives), and selects appropriate subject, topic, difficulty level, relevancy, course, author, revising experts, school grade, which graduates are expected to solve the problem presented in item, from respective lists. If the item is added as an "analogue" to already existing item, a user must enter analogousness ID of that item in respective field. Otherwise, program will assign a new analogousness ID to added item. Finally, an item is added to database with unique item ID generated by program.

There are several alternatives for adding new item to item bank. Items typed in certain formats can be added to item bank automatically. Besides, it is possible to type new item using an already existing item as a template. This is particularly useful when adding analogues: user only edits item stem and answers, because parameters of analogous items must be identical. These alternative methods add extra flexibility to item adding process and help to save time and resources.
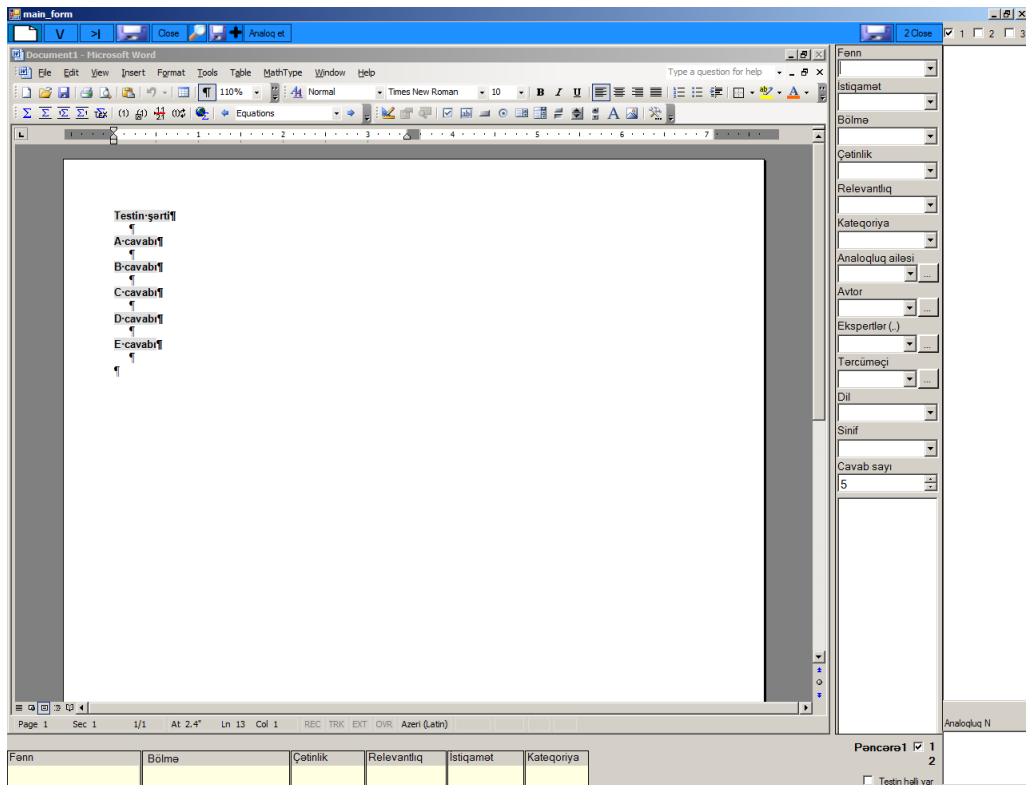
Fig. 1

Each item can be edited using similar program interface as in adding new item. Items are edited using MS Word text processor, which allows making virtually any change to the text, using graphics, tables, mathematical formulas via OLE technology, utilizing all of the powerful features provided by MS Word. It is also possible to edit several items at once, which is particularly useful while editing items selected for exam. A user opens all items to be edited at once, not one by one. This feature utilizes MS Word subdocument technology and allows for saving edited items one by one, in groups or all at once. (Fig. 2)

Fig. 2

Retrieval of statistical report ("map of the bank") on the state of item bank.

This feature can also be envisioned as a part of search capabilities "Test Builder" provides. Advanced search capabilities of the program include parameter-based and content-based search which allows for locating items based on any combination of item parameters as well as item content. For example, it's possible to find all items from mathematics subject domain that have difficulty level equal to 3 (difficult) and contain the word "derivative" in their correct answer. Statistical report on the state of item bank – a three dimensional electronic table containing number of items by difficulty level and relevancy within each subject topic – is necessary for test development. Based on this report, the software developed at Statistical Research and Scientific Modeling department of SSAC generates a special-formatted request, which, in its turn, is used for intelligent item selection. Statistical report can be retrieved for electronic use as well as for printing. Statistical report for electronic use is a very handy tool for getting detailed information on items with the same search parameters. Double clicking on the cell at the intersection of topic of "X", difficulty level of "Y" and relevancy of "Z" opens a new window where it is possible to view everything on respective items, including items themselves. (Fig. 3)

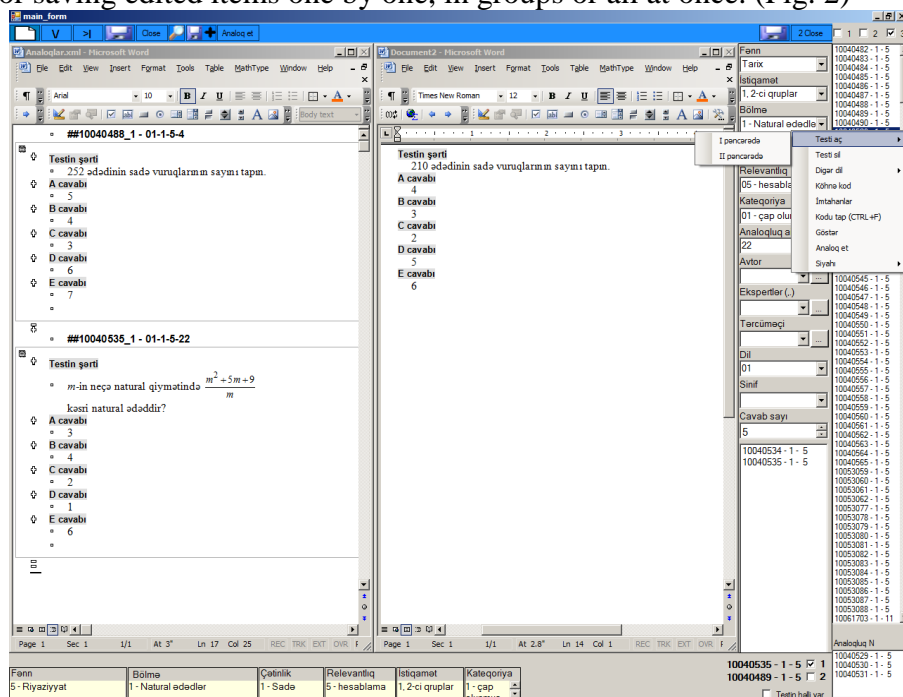| N | Bölmənin adı | cəm | Mürakkəblik dərəcəsi | | | Mürakkəblik dərəcəsinə görə relevantliq | | | | | | | | | | | | | | | | | | | | | |
			sadə	orta	çətin	sadə 1	2	3	4	5	6	7	orta 1	2	3	4	5	6	7	çətin 1	2	3	4	5	6	7
1	Natural ədədlər	273	143	105	25					63		80					41	2	62					11	8	6
2	Adi və onluq kəsrlər	142	78	52	12					46		32					24		28					8		4
3	Faiz, nisbət, Tənasüb	299	87	190	22					79		8					127	4	59					12	3	7
4	Həqiqi ədədlər	162	84	57	21					54		30					16	4	37					6	3	12
5	Tam cəbri ifadələr	150	62	72	16					38		24					24	2	46					6	2	8
6	Çoxhədlinin vuruqlara ayrılması	165	83	60	22					41		42					34		26					4	2	16
7	Cəbri kəsrlər	278	145	106	27					83		62					40		66					11		16
8	Kvadrat köklər. Həqiqi üstlü qüvvət	851	427	375	49					367		60					260		115					17	2	30
9	Birməchullu tənliklər və məsələlər	879	289	486	104					191		98					303	6	177					56	10	38
10	Tənliklər sistemi	368	114	192	62					83		31					117	2	73					30	4	28
11	Bərabərsizliklər	594	245	275	74					192		53					185	5	85					34	8	32
12	Silsilələr	580	233	261	86					191		42					151	4	106					31	4	51
13	Triqonometrik funksiyalar	813	447	308	58					261		186					161	2	145					22	4	32
14	Triqonometrik funksiyalar üçün toplama t...	615	222	320	73					141		81					228	4	88					54	2	17
15	Triqonometrik tənliklər və bərabərsizliklər	548	164	311	73					85		79					202		109					31	6	36
16	Funksiyalar və qrafiklər	866	382	409	75					197		185				2	219	4	184					23	12	40
17	Törəmə və tətbiqləri	554	170	294	90					93		77					168		126					36	16	38
18	Üstlü və loqarifmik funksiyalar	613	228	306	79					141		87					182	2	122					55	6	18
19	Üstlü, loqarifmik tənliklər və bərabərsizliklər	1081	253	613	215					189		64					394	4	215					130	6	79
20	İbtidai funksiya və inteqral	372	120	193	59					50		70					114	2	77					23	2	34
21	Həndəsənin əsas anlayışları	229	141	84	4					105	2	34					54		30						4	
22	Üçbucaqlar	693	295	330	68					195		100					183	5	142					32	8	28
23	Çevrə və dairə	567	224	275	68					128		96					183		92					35		33
24	Dördbucaqlılar	633	261	320	52					180		81					204	10	106					16	2	34
25	Fiqurların sahəsi	723	291	321	111					194		97					218	2	101					61	6	44
26	Hərəkət. Oxşarlıq	339	170	143	26					86		84					87		56					6		20
27	Vektorlar. Koordinatlar metodu	810	442	308	60					296		146					181	8	119					25	2	33
28	Fəzada düz xətlər və müstəvilər	323	161	120	42					106		55					92		28					23	2	17
29	Çoxüzlülər, onların səthi və həcmi	580	192	258	130					108		84					207		51					82	11	37
30	Fırlanma cisimləri	391	150	175	66					79		71					125		50					24		42
31	Ardıcıllıqlar	124	59	47	18					57		2					33		14					8	2	8
32	Funksiyanın limiti	262	90	118	54					86		4					70	2	46					37		17
33	Məntiqi testlər	19	0	19	0														19							

Fig. 3

Adding and editing exams.

SSAC provides a wide range of exams. "Test Builder" provides user-friendly tools to create any kind of exam in a few simple steps. <CREATE NEW EXAM> command opens appropriate program window where a user is asked to select exam type, set exam provision date, define exam name and specify Global Random Number (GRN) that will be used in processes requiring randomization. In next step a user is asked to select requests for item selection for each subject that will be used on exam.

It is possible to have several requests for one subject. In this case, it is recommended to give meaningful names to requests to avoid confusion. In next step user is required to set the following parameters for each subject:

1. Number of items.
2. Number of initial test forms. It is a common practice to select more than needed items from bank in order to give subject experts who make final revisions to items to be used

on exam a little freedom of choice. For example, if two distinct final test forms will be administered on exam, number of initial test forms is usually 4 or 6.

3. Number of final distinct (original) test forms. This number is always less than or equal to the number of initial distinct test forms
4. Total number of final test forms. Some final test forms may consist of same items and differ only in both positions of items and positions of distractors within items.
5. Subject-level random number for random selection of items from bank (SRNI).
6. Subject-level random number for building final test forms from initial ones by randomly replacing items and distractors within items (SRNF).
7. Minimal number of occurrence of any of letters A, B, C, D or E as correct answer (MINN).
8. Maximal number of occurrence of any of letters A, B, C, D or E as correct answer (MAXN).
9. Maximal length of series of any of letters A, B, C, D or E as correct answer (MAXL).
10. "Block structure" (BS), reflecting blocks of items to be handled independently while building final test forms. For example, if items at position from 16 to 20 on initial test form are connected in some way, they form a separate block. This block should occupy the same position range on final test form, i.e. they should be replaced only within block.

As stated before, a used item is not included in the tests administered in the next few years. For each exam created a list of so-called "excluded exams" should be specified. In final step user creates this list, thus defining, which items will not be considered in item selection process. After this final step user saves parameters of new created exam to database. It is possible to edit exam parameters at any time.

Intelligent item selection.

Intelligent item selection means selection of items randomly within specified categories considering user-defined options. Categories are specified in special-formatted request file. First, a group of subject experts prepares a blueprint of the objectives across a given content area and defines the number of items representing each difficulty and cognitive level at the test, as well as subject topics or topic groups that should be represented in the test. Generally, for each subject the number of topics is greater than the number of items that represent given subject in the test. For this reason experts assemble close topics into groups so that the test covers the whole curriculum. Using data provided by subject experts and map of the bank, the software developed at Statistical Research and Scientific Modeling department of SSAC generates a special-formatted request. This request contains parameters of each item to be included in the test. Based on the request, program randomly selects items from the bank, using certain combination of GRN and SRNI. Selection algorithm considers complex criteria defined in request file as well as user-defined options. Criteria defined in request file may include a list of topics, difficulty levels, cognitive levels, block and analogousness information and whether they should be handled arbitrarily or in certain order. User-defined options are mainly related to the use of analogous items. Besides, user may choose whether or not to include solution of the problem presented in item, list of exams where item has been used, etc. in assembled test form.

"Test Builder" is capable of simultaneously developing parallel test forms, i.e. test forms containing items that are similar in terms of content and parameters by making use of analogous items. These test forms are called "initial test forms". As a common rule, number of initial test forms is always greater than that of final original test forms.

After assembling initial test forms, they are reviewed by the experts. "Test Builder" includes two procedures for editing drawn items. One of them, "Item search", is used to find

an item by specifying the number of form and number of item within that form. The second one, "Item Replacement", is used to replace a drawn item with parallel item on another form. After the final expertise program assembles a specified number of test forms called "derivative forms" or final forms from the items of selected initial forms, randomly changing both the position of items and the position of alternatives within each item in the initial form, using a certain combination of GRN and SRNF. This contributes to preventing cribbing during an examination while providing identical items on the test. This procedure also takes into account MINN, MAXN, MAXL and BS values which provides for control over the frequency of occurrence of each of the alternatives as a key.

Compilation and auto-formatting of test booklets.

After the assembly of final forms a procedure called "Booklet Compiler" compiles test booklets from final forms. This procedure also auto-formats test booklets and makes them ready for use at the examination. (Fig. 4)
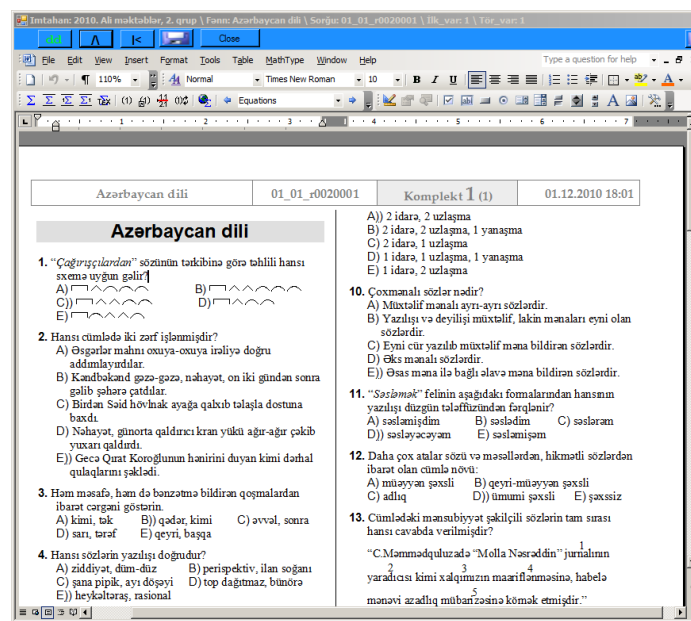

Fig. 4

Future plans

At present, SSAC administers only paper and pencil tests. But in the nearest future it is going to provide computer-based tests for admission to master's level of higher schools. Computer-based testing has several advantages over traditional paper and pencil testing. The biggest advantage of using computers to deliver tests is the level of control over the conditions of the testing session it gives to instructors. For example, the amount of time used for the test and the type of feedback (either on individual questions or on the entire test) can be varied. Such timing information can assist test designers in eliminating bad or poorly worded questions and in comparing the knowledge level of test-taking candidates.

A second set of advantages includes the variety of multimedia learning objects that can be incorporated into test questions and the fact that most question types can be scored instantly. Using computer based testing also enables the use of adaptive tests. Adaptive tests are tailored to the individual ability of the test-taker and are able to provide somewhat better estimates of ability with shorter tests. [5] SSAC staff is working on implementation of a computer-based testing in admission examinations in a few years to come.

REFERENCES

[1] Annie W. Ward, Mildred Murray-Ward, "Guidelines for Development of Item Banks," ITEMS Instructional Topics in educational Measurement, *Module 17*, pp. 263-268, 1994.

[2] David J. Weiss, "Item Banking, Test Development, and Test Delivery", The APA Handbook in Testing and Assessment

[3] Vale, C. D. "Computerized item banking". In S. M. Downing & T. M. Haladyna (Eds.), *Handbook of test development.* Mahwah NJ: Erlbaum., 2006

[4] Assessment Systems Corporation (2010b). *User's manual for the FastTEST 2.0 Item Banking and Test Development System.* St. Paul MN: Author. Available from http://www.assess.com.

[5] Wainer, H. (Ed.). *Computerized adaptive testing: A primer* (2nd ed.). Hillsdale NJ: Erlbaum., 2000