

Direction to Prepare Item Bank for the Purpose of On-Line Evaluation

Dr. Prashant M. Dolia

Assistant Professor,

Department of Computer Science & Applications, Bhavnagar University, Bhavnagar Gujarat, India

Email: prashant_dolia@rediffmail.com

Abstract—Straight forward meaning of On-Line Evaluation is that, few questions are extracted from the question bank and user is asked to “mark” the response. Normally, questions are assumed to be of “objective type”. In this paper, this traditional idea is extended and abstract concept of “Item Bank” is introduced. The researchers have developed and tested Item Bank software earlier. The guidelines discussed in the paper are the outcome of this real life experience. The scope of the Item bank would not be confined only the Evaluation of competency of respondents but it covers assessment of the person for fresh recruitments/ Assignments or for the Reward or to Accredite the achievement of the employee performance with the use of client server architecture.

Index Terms—Item Type – True-false, Completion, Interpretation, Objective, Multi Choice, Item Bank, On-Line Evaluation

I. Introduction

A. Item Bank Concept

Item bank is a large collection of items organized and catalogued like the books in library. The idea is that the user selects items as required making up particular Evaluation process. Item bank is an important part of computer-based Evaluation and collection of items used to conduct the Evaluation process. Item bank is a database focused on group of items and a set of design specifications to conduct the Evaluation process with the use of client server architecture. Item bank is an advanced system, developed for the windows an

environment; it allows managing an unlimited number of items, for storing and administrating them as well as for the automatic creation of very diverse kinds of Evaluation process. Items are described with their difficulty in the Item bank. In order to develop test or subtest, one does not need to go through the laborious process of developing large set of items for evaluation. Runder [1]

B. Item Concept

Creating new items are very time-consuming endeavor. Administrator need not to compose the items. They also must determine each item’s difficulty so that the Evaluation process should be neither too hard nor too easy. The evaluator designs items in very diverse formats; from the very well known multiple choice format to more complex items as alternative answer (true/false), alternative answers with reason, multi choice answers, complete the blank and group answers. The developer has the responsibility to write a quality set of items for such category. Items that are written poorly are misleading the candidates and declining the quality of item banking. If topic contains too many difficult or too many very easy items, candidate may not response difficult items correctly or response all of easy items correctly. Over time, items become irrelevant, new items must be added and items that turned out to work poorly, despite to the developer’s best efforts to write only the best items, must be replaced. It must be stated that the design of the item banking software includes the strictest evaluation necessities. The design of the system contains a series of advantages as for it allows covering more than 80% of

the general necessities of the evaluators of different knowledge areas. According to Davis[2], the construction of solid and reliable evaluation requires considerations of quantitative information regarding difficulty of a item and discrimination power of each item.

C. Evaluation Concept

The Education Commission of India (1964-66) has understood that “Evaluation is continuous process and integral part of total system of education which is intimately related to educational objectives.” The techniques of evaluation are means of collecting evidence about the Student’s (candidates) development in desirable directions. These techniques are valid, reliable and practicable. In essence, Daniel[3] suggested the educational evaluation should concern itself with the identification of individual differences in what candidates have learned. Concept of evaluation was to evolve from this general idea has been reflected in educational practices ranging from “Homogeneous grouping” through individualized direction.

II. Advantages of Item Bank

- Using Item Response Theory method, such as Rasch model items are placed on a common scale per items type. It is very similar to IRT1 in that only considers one parameter (difficulty). In Rasch measurement, extreme data (e. g items which are consistently well or poorly responded) are discarded and model is fitted to the remaining data. As per Linacre [4,5] the scale indicates the relative difficulty of the items and they are placed into item banks without extensive testing.

- Item bank permits you to add and withdraw items as per the requirement.

- Item banking provides substantial saving of time and energy as compared to conventional test development environment.

- It is necessary to provide user-friendly interface between candidate and administrator of the on-line evaluation process.

- It is very useful to increase the productivity, efficiency, flexibility and evaluator’s control. Productivity means large number of quality base Item sets can be generated through item banking efficiently.

- Item banking software is a unique and has multiple advantages regarding item administration based on program, which handles databases.

III. Features of Item Bank

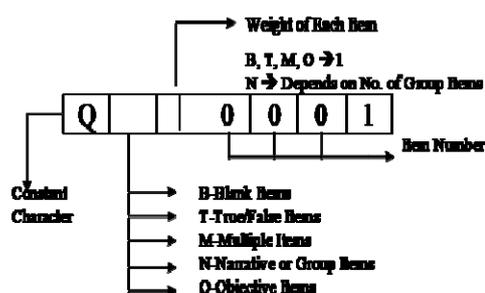
- Item banking software has a module for developing On-Line Evaluation in order to implement computer-based Evaluation (CBE).

- It provides wide range of different kinds of items.

- It allows the input of ‘group’ items that are formed by a general reading and items are associated to it. These items are very common in analysis of texts, verbal reasoning, fill in the blanks etc.

- It allows storing more than one correct response for an item.

- Each item is identified by a item number and its format is as follows:



(Figure :1 Format of the Item Number)

Third column indicate the weight of the each item. For item type B, T, M and O the weight is 1 and for item type N the weight is based on number of sub items i.e. if item type N containing 3 items then weight is 3.

- It includes set of options for database handling such as search, sort etc. Searching is based on item type or item number.

IV. Item Types

There are five different item types like True/False items, Completion items, Interpretation items, Objective items and Multi choice answer items are considered. In the following sections the advantages of each item and guideline to prepare the items are also mention. These guidelines are helpful to item creator for standardized Item bank.

A. True-False Items

Spinoza stated, “*Who would distinguish the true from the false must have an adequate idea of what is true and false.*” True false test items enjoyed considerable popularity in evaluation process in the early days. True false items tend to be regarded as one of the least satisfactory forms for evaluation. There are several reasons for its disfavor.

1. Ordinary true false items are typically less reliable and chances of errors, due to the guessing on some items are probably part of reason.

2. They are frequently judged to be trivial. If such sentences are not considered carefully, evaluation process based on them may be filled up with trivia.

3. They are directly based on sentences from textbooks, it is danger that they encourage and reward sheer verbal memory.

4. If sentences are not worded with enough care or not included with enough background information then candidates cannot judge true or false answer.

5. The true false item does not provide any explicit alternative in relation to which the relative truth or falsity of item could be judged.

6. The true false items have been criticized as educationally harmful on the ground that they expose the candidate to make error.

There are several things, which justified, favor and overcome above problems with the use of true false items with the following measures:

1. The knowledge is knowledge of propositions. One essential purpose of evaluation is to test the truth or

falsity of hypothetical propositions. It provides a simple, direct and fundamental test of student’s knowledge.

2. Efficiency of true false items is lost as a result of ambiguity in the items or as result of candidate guessing. In proposed evaluation process, candidate has to specify that why the items is false to obtain the efficiency.

3. Care must be taken to select important propositions as a basis for the items to use originality in phrasing them so as to make it difficult for a candidate to respond correctly on a basis of superficial verbal memory and to word them precisely so as to avoid intrinsic ambiguity.

The following are the guidelines to develop the True false items:

a. As per the Robert [2], each true statement that is included in a test should permit the test constructor to say of it: “*The truth of this statement is not obvious to everyone, but I can provide evidence to show that it is essentially true.*” Similarly as it has been said by Robert [6], each false statement that is included in a test should permit the test constructor to say of it: “*This statement reads like a plausibly true statement, I can provide evidence to show that it is essentially false.*” If the item pass the test, they are likely to be good true false test items.

b. Lee & Alexander [7,8] found that false statements tend to discriminate somewhat more sharply between students of high and low achievement than do the true false statements. This may be due to what is called an “acquiescent response set”. In absence of firm knowledge, a student seems more likely to accept than to question a declarative statement whose truth or falsity he must judge. If the false statements tend to be higher in discrimination, it would be advantageous to include higher proportion of them, as many as 60%.

c. Make true statements true enough and the false statements false enough so that candidate would unanimously agree on the answer.

d. Chose and state Propositions so that their truth or falsity is not immediately obvious to every reasonable person.

e. Word the statements so that sheer memory of words, empty phrases or meaningless sentences will permit the correct answer.

f. According Ina [9], guard against irrelevant clues, such as specific determiners, which would permit test wise but unprepared examinee to respond correctly. Specific determiners refer to sweeping terms like 'all', 'always', 'none', 'never', 'impossible', 'inevitable' etc. Statements including such items are likely to be false. On the other hand, statements including such qualifications as 'usually', 'sometimes', 'often' etc. are likely to be true.

g. Good true false items express single, not multiple ideas.

B. Completion Item

The Completion items are frequently classed with selection type items and all are referred to as objective type items without providing any guidance. The Completion items are statements with one or more missing words that must be provided by the candidate. The essential features of these kinds of items are that the response consists of only one or two words. These words must be supplied by the candidate rather than selected from a list of available alternatives. When the candidate is expected to have memorized certain terms, terminology, symbol etc. to fill up the blank. Completion or supply type items are most valid and most effective form of test item to use. The advantages of the Completion items are as shown below:

1. They are quite discriminative and testing knowledge in the subject.
2. These items are reliable and valid
3. Factual replies have to be given in their answer.
4. They are easy to construct and does not require a list of plausible distracters.
5. It provides little opportunity for guessing; either the candidate supply the response or he cannot. He has no chance to guess from among provided alternatives.

The following are the guidelines to develop the Completion items:

a. Do not insert more than one or two blanks in any completion items

b. Wording of each item is so precise there is only one desired response.

c. Do not use statements copied from a textbook or other lesson source.

d. Do not ask long, ambiguous or indirect items.

e. Try to have items that have only one correct response. This is often impossible and led to major problem in the use of completion items. All the possibilities are included in the correct answer.

C. Interpretation Items (Narrative Items)

The new interpretation items are illustrated by interpretations of data items. In this type of item, data are presented to the candidate, who responds to the data on the basis of what he knows about the subject. These data comes from any source like verbal description, program segment, table, result of experiment etc. The only criteria to be applied to the data are that the data be new to the candidate. It must be something that he has never seen before, something, which has not been used in class.

Items are written after the new, relevant and practical data have been selected. Narrative items are purely based on situation mentioned by the examiner. The items either true false, objective type or completion type. Here, student has a chance to enter into evaluative process and to learn about the own thinking process.

The following are the criteria to be met in constructing these items:

- a.** Newness to the candidate
- b.** Similarity to life like situation
- c.** Significance of problem to daily living
- d.** Appropriate vocabulary level

D. Objective Items

Objective items are currently the most highly regarded and widely used form of item. They are adaptable to the measurement of most important educational outcomes-knowledge, understanding and judgment, ability to solve the problems and to make predictions. The form of the objective item, with the stem asking or implying a direct

question, provides a realistic, naturally appropriate setting for candidate achievement. Main goal of the objective item is to differentiate candidates of higher and lower achievement.

The typical objective item consists of an introductory statement and several statements listed beneath it, which complete the stem. One of the statements is the correct or best answer and the others, which are incorrect, are called distracters or decoys. The following are the reasons:

1. This type of items is adaptable to many different types of situations.
2. This type of items measure candidate's capacity for interpretation and discrimination.
3. They are less indirectness and artifice compare to the other items.
4. It is less susceptible to chance errors resulting from guessing by instructors and students.

The following are the guidelines to develop the Objective items:

- a. There are five alternatives and last one is 'None of These'.
- b. Objective items should be based on sound, significant ideas, which could be expressed as independently with meaningful propositions.
- c. The wording of objective items should not follow familiar textbooks phraseology so closely that verbal memory without comprehension will provide an adequate basis for response.
- d. The distracters in an objective item should be definitely incorrect but plausibly attractive to the uninformed.
- e. Difficulty would usually arise from irrelevant ambiguity not from the presentation.
- f. A candidate who selects the correct response to an objective item be eliminating the incorrect responses, which demonstrate useful achievement.
- g. All of the responses to an objective item should be parallel in point of view, grammatical structure and general appearance.

h. The item writer can make the correct response easier to select by making the alternatives more heterogeneous and harder by making them more homogeneous. An item so easy that all examinees answer it correctly, or so difficult that none answers it correctly, yields no information about relative levels of achievement. Do not consistently avoid any one of the response positions as the place for the correct answer. e. g never putting the correct response in the third position.

i. The response 'None of These' is used for discriminating good student with others and also reduces the guessing problem.

E. Multi Choice Item

In this type of item, data presented to the candidate and he/she has to make a decision or to draw a conclusion that which ones are most consistent with the fact in the light of his knowledge. Usually five conclusions are given. Out of five all are true or any numbers of conclusions are true. Candidate has to classified the fact and find out the most relevant truth from the given fact. In other words, we are investigating the candidate's thinking, both from the viewpoint of accuracy, from the number and types of error made. It might also be added that, for more reliable scores, a test should be made up of a series of these application items, when we are attempting to appraise this objective of how individuals carry on critical thinking.

Smith [10] stated, individuals are constantly being confronted with so-called 'proofs'- why they should do or believe various things. How do our students react to these 'proofs'? Analysis showed that these behaviors included disposition to analyze proofs critically, the recognition of basic assumption upon which a conclusions depends. The examinee goes through these and marks those, which explains or supports the conclusion. The following are the guidelines for preparing multiple choice items:

- a. This type of test items should adaptable to many different types of situations.

- b. This type of test items must measure candidate's capacity for interpretation and discrimination.
- c. Conclusions should be purely based on the fact.
- d. Proper prepositions must be used in the wording.

V. Database Design Tool

Personal or PC base database management systems, such as Microsoft's MS-Access and visual FoxPro, are usually stored on a user's desktop computer system or a client computer. These database packages are developed primarily for single user applications. When such package is used for a multi-user or a shared access environment, the database applications are stored on a file server or a server and files are transmitted to the client computers over the networks. The following are the disadvantages of single user application refer from Nilesh [11]:

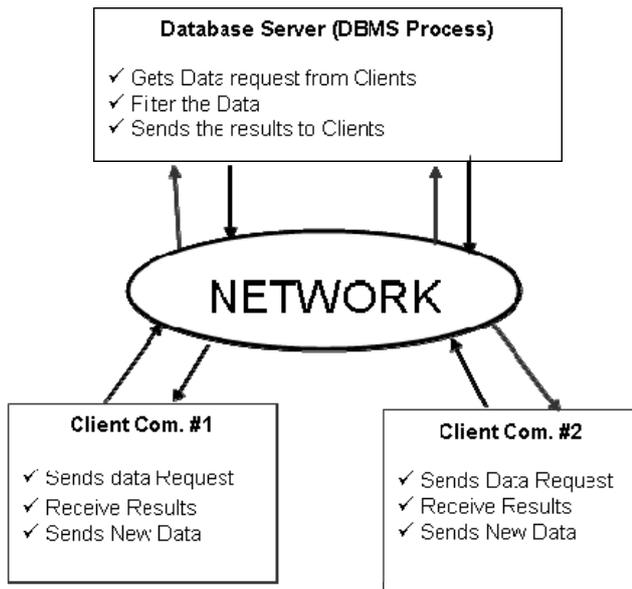
1. If the client requires a small piece of data from the server's large database, the server has to transmit the entire database to the client over the network. Client computer hardware has to handle heavy demand and the network has to handle heavy traffic in both directions.
2. The personal database system assumes that no two transactions will happen at the same time on the one table, which is known as optimistic locking. In optimistic locking, the tables are not locked by the database system.
3. When a client is performing record insertions, deletions or updates, those records are locked by that client and are not available to the other clients. Now, if the client with all the record locks fails due to software or hardware malfunction, or due to power outage, the locked record stay locked. The transactions in progress at the time of failure are lost. The database can corrupt and needs to be repaired. In order to repair the database, all users have to log off during the repair, which can take few minutes to a few hours. If the database is not repairable, data can be restored from the last backup, but the transaction since last backup are lost and have to be reentered.

4. Personal databases do not have file based transaction logging. Transactions are logged in the client's memory. If the clients fail in the middle, some transactions are written to the database and some are not. The transaction log is lost because it is not stored in a file.

Solution of above personal database problem is Client/Server database such as oracle, run the DBMS as a process on the server and run client database application on each client. The client application sends a request for data over network to the server. When server receives the client request, the DBMS retrieves data from the database, performs required processing on the data and sends only the requested data back to the client over the network.

The following are the justifications for Client/Server database like Oracle:

1. The client computer does not run the entire DBMS, only the client application that requests data from the server. The client does not store the entire database on its local drive but receives only the requested data from the server. Data processing is performed on the server's side. The demand at the client's end is minimal. The clients request data from server and only requested data is sent back via the network, which results in less network traffic.
2. The DBMS takes care of the locking automatically and it involves no extra effort on application programmer's part. In client/server system, such as Oracle, when an agent reads a table, it is locked totally or partly by the DBMS. The second agent cannot read the part of the table. e. g answer table of the one candidate cannot read by the other candidates.
3. In case of a client failure, the Client/Server database is not affected. The other clients are not affected either. The failed client's transactions in progress are lost. If the server fails, a *Central Transaction Log*, which keeps a log of all current database changes, allows the Database Administrator (DBA) to complete or rollback unfinished transactions. Client/Server database like Oracle has fast and powerful recovery utilities.



(Figure 2 A Client/Server Database in Multi-User Environment)

VI. Conclusion

Theory is much further advanced than practice when it comes to item bank but it seems certain that these new techniques will make a major impact on the educational scene within the next few years. They offer important advantages over traditional types of standardized tests. The standardized item bank and Computer based evaluation process measure the competency of respondents for the purpose of Fresh Recruitment/ Assignment or for the Reward or to accredit the achievement of an employee performance with the use of client server Architecture.

References

- [1] Runder Lawrence (1998), "Item banking , Practical Assessment" Research and Evaluation 6(4)
- [2] Davis Federick, "Item Analysis in Relation to Educational and Psychological Testing", Psychological bulletin (1972), p. 97-121
- [3] Daniel Strach (1916), "Standard Tests as Aids in the classification and promotion of pupils", Fifteenth Yearbook of the National Society for the study of Education, Part-II, p. 143
- [4] Linacre J.M. Bruce Choppin (1996), " Rasch Measurement Transactions" P. 394
- [5] J.M.Linacre(Ed.) "Rasch Measurement Transactions", part-2, Chicago:MESA Press. www.rasch.org/memo49.htm
- [6] Robert Ebel, "How to use true false tests", Chapter-5, Measuring Educational Measurement, Michigan State University. P. 145
- [7] Lee J. Cronbach, "Studies of Acquiescence as factor in true false test", Journal of Educational psychology, 401-415 Robert L. Ebel (1960), "Some Tests of competence in Educational Measurement", The 17th yearbook of the
- [8] Alexander G Wesman, "The usefulness of Correctly Spelled Word in Spelling Test", Journal of Educational Psychology, P. 242-246 National Council on Measurements used in Education (East Lansing, Michigan State University) P. 93-104
- [9] Ina H. Brinkmeier and G.M.Ruch, "Specific Determiners in True-false Statements", Journal of Educational Research, XXII P. 110-118
- [10] Smith E.R. & Tyler R.W, "Apprising and Recording Student Progress", Progressive Education Association Evaluation in Eight years Study. New York: Harper
- [11] Nilesh shah, "Database System Using Oracle", Client Server Databases, P. 35-39, PHI

Dr. Prashant M. Dolia is Assistant Professor at Department of Computer Science & Applications, Bhavnagar University, Bhavnagar, Gujarat, India. His educational Qualification is B.Sc (Physics), M.C.A. and Ph.D. (Computer Science & Application) from the Bhavnagar University. He is having **Twelve** years of teaching experience at Post Graduate level. He has published **NINE** research papers in International journals and **ONE** research paper in National journal. He has also published **SIX** books and **ONE** Chapter book in subject of Computer Science & Applications. His area of research includes Web Intelligence, Data mining & Warehousing, Wireless technology, Linux platform and Kernel Programming.