

2013

Issues in Contemporary Evaluation

UNIVERSITY OF IBADAN LIBRARY

Edited by

Olajide Olorunnisola
Adams Otuoze U. Onuka
Oyebamiji Babalola
Aderemi I. Alarape
S. A. Babarinde

Published by
The Postgraduate School,
University of Ibadan,
Ibadan, Nigeria.

E-mail: postgrad@mail.ui.edu.ng

Website: <http://www.postgraduateschool.ui.edu.ng>

First Published 2013

No part of this publication may be reproduced, stored in a retrieval system or transmitted by any form or by any means, electronic, mechanical, photocopying, recording, or otherwise without the expressed written consent of the copyright owner.

All rights reserved

ISBN 978- 978-50948-7-9

Printed by Sapphire Prints
08034892535. 08070709894

Chapter 14

Item Banking

Adams O. U. Onuka

Introduction

Measurement has been defined as the process of assigning numerical values to entities or objects. Thorndike (1997) posits that whatever exists in some quantum can only be known if its quantity and quality is thoroughly deciphered. Psychometricians believe that measurement is of two kinds: namely physical measurement and psychological measurement. These two forms of measurement are governed by rules for assigning numbers to objects.

Test development, administration and conclusion are onerous tasks, involving drafting pool of worthy test items/questions, verifying their validity leading to selecting the real test and banking them for future utilisation, makes them different from the teacher made questions which often are processed and as such their validity is thus questionable. However, in our clime, the various examining bodies at the secondary education level undertake rigorous testing by which we mean all the processes involved in examining from the beginning to the end. The testing activity encompasses composing new test items and determining each item's difficulty in order to ensure that the test will neither be too hard nor too simple.

An item bank is the repository of test items that belong to a particular testing programme containing all information relating to those same items. It is often a relatively large collection of easily accessible test items. Item banking is alternatively referred to as 'question banks', 'item pools', 'item collections', 'item reservoir' and 'test item libraries'. Item banking, therefore, is the process of creating and maintaining a repository of test items from which examination questions can easily be drawn when it is necessary. In most cases of testing and assessing, the items in the bank are of the objective type, but any format can be stored in the item bank. In evaluation, we believe that the most objective way of assessing candidates/students is by administration as it provides the

equitable pedestal for all the testees to prove their worth. Test items could then be pulled from the bank and assigned to test forms for publication either as a paper-and-pencil test or an *e-assessment*. One good feature of item banking is that anyone knowledgeable in the said area can administer as well as score the candidates on the test. An item bank will normally include the text of each item, and also extensive information regarding test development and psychometric characteristics of the items and the keys in the case of objective test or marking schemes, each a product of thorough moderation by a conference of examiners. Such information will likely include:

- Item author
- Date written
- Item status (e.g., new, pilot, active, retired)
- Correct answer/key
- Item format
- Classical test theory statistics
- Item response theory statistics
- User-defined fields

Item banking describes the process of maintaining all test items or assessment tasks in a common pool. Usually an item bank contains information about the content and statistical performance (if available) of each test item or test question. Furthermore, in addition to finding tools for compiling a customised question bank such as each department/faculty can have, taking cognizance of the area of learning, it is quite possible to come across existing online question banks, which can save one a great deal of time, but must be contextualised. Item banking software does facilitate easy construction of tests that could meet certain pre-defined qualitative and quantitative constraints in terms of the content and item difficulty balance of included assessment items. A large number of desktop item banks exist and can be used to quickly assemble paper-and-pencil tests to be printed. These are popular for classroom assessment. Nowadays, many item banks provide web-based facilities for test item authoring, test assembly, and even online test delivery.

Need for regular revision of tests to avoid administering the same test year after year cannot be overemphasized. It is noteworthy that creating

new tests can be a very time consuming endeavour, yet it is a necessary 'evil' we must regularly undertake. Not only do test writers need to compose the test items, they also must determine each item's difficulty level, discrimination index, reliability and validity, and if it is to be administered on a large number of testees, standardised it [i.e. producing the test psychometric properties] in order to ensure that a test is usable uni-dimensional etc. Item banks are files of various suitable test items that are "coded by subject area, instructional level, instructional objective measured, and various pertinent item characteristics (e.g., item difficulty and discriminating power)" included (Gronlund, 1998, p. 130).

Disadvantages and limitations of item banking

Item banking and item response theory are not cure-alls for measurement problems. Persistence and good judgment must remain vital aspects in any test construction and test usage exercise. Every possible effort must always be made to include only quality items in the item bank. The same care and effort must be given to test item writing. Items purchased or 'imported' from external sources must be evaluated carefully to match your curriculum as well as for technical quality.

Item banking involves equating various tests and items. It is entirely possible, mathematically, to equate tests which cover entirely different subject matters in a course, which of course is undesirable because it will lead to distortion. At the practical level, this means that it is also possible to equate items are subtly assessed, but with significantly different skills. In order to avoid this undesirable situation, the item review process must also include a careful evaluation of the skill assessed by each item and tests must be carefully formulated.

The intent of compiling a test using latent trait theory is to be able to make a prediction of the composite test characteristics. While the prediction is often surprisingly accurate, it must be validated. Tests developed using latent trait theory should still be field tested.

In the United States of America, while some education districts have implemented very successful item banks and Rasch calibrated testing programmes without knowing anything about IRT, good practice calls for a staff that is comfortable with and is knowledgeable of what the

are doing. A district undertaking an item banking project should have full understanding of the practical as well as the mathematical/theoretical aspects of item banking.

An item bank really consists of multiple collections of items with fairly unidimensional content area, such as mathematical computations or vocabulary. Collections of items would usually span several levels. In developing the bank, many tests must be calibrated, linked (or equated), and organized. This requires a great deal of work in terms of preparation and planning and in terms of computer time and expertise. Once the item bank is established, however, test development time, effort and cost are reduced.

Planning an item bank

The most crucial step in developing an item bank as in test development itself is planning. This involves the preparation of individuals, the identification of what one has to start with as an item bank with [content], and the identification of what you hope to accomplish with an item bank [objective/goal]. Everyone on the staff of the organisation preparing an item bank should have enough familiarity with Rasch measurement principles and item banking to be able to knowledgeably discuss and explain the project. You can formally train your staff by using in-house personnel, bringing in a travelling workshop, or having people attend a pre-session at a workshop or conference.

You should have senior level personnel available to answer technical questions that might arise. You should also have computer experts that are capable of doing the following tasks: 1.) modifying computer programmes, 2.) establishing a data base system, and 3.) capable of running packaged programmes.

If you intend to do any item bank exchanges or purchases, you should have someone on your staff who knows what is available. You need personnel capable of critically evaluating test items for technical quality, curriculum match, unidimensionality, and potential bias, thereby, accurately calibrating test items and to establish requisite scales, items need to be presented to examinees with a wide range of ability.

Linking the various grades and levels within a content area requires common anchor items. Such anchor items must be administered along with the real items within a given grade/level. The form and anchor items are calibrated together. The anchor item parameter values based on calibration with one form are compared with the anchor item parameter values based on calibration with another form. The difference in parameter values is then used to link the forms. You are required to also identify for which content areas you have administered overlapping subtests and the number of students responding to the set of items. You may find out that you will need to gather additional item response data to link grades and levels. The data processing staff should examine literature and programmes on item banking to determine what programmes must be developed and what programmes can be modified.

Do make every effort to identify the projected testing needs for the next five years. This would involve identification of which subtests you will need to revise, what additional areas you may need to assess, and how objectives might be differently stressed.

Item Bank Start-up Activities

The start-up activities would most likely involve administrative activities and the data processing staff. Each test would have to be calibrated and equated to a parallel form and adjacent grade levels. The data processing staffs have to adapt existing computer programmes to the local system to develop a database system. They would then calibrate each test, equate the tests, and store the equated item parameters and their descriptors in a database system. With a large number of tests and test items, it becomes a prominent undertaking.

The administrative staff have to coordinate item banking activities to insure that the data requirements are met. During a planning process, a chart can be constructed to identify the tests and anchor items that have been and therefore, need to be administered to a relevant sample. Working from these charts, testing coordinators will need to organize the administration of tests and subtests needed to calibrate and equate all the items going into an item bank. This process involves compiling test booklets, making testing arrangements, collecting response sheets,

and preparing the data arising there from for data processing. The frequency of students taking multiple subtests from different levels and forms would determine whether this too can become a major project.

Options in setting up an item bank

Different test users have different/varying needs and concerns, which culminate from their varying circumstances and or the purpose for which tests are desired to achieve. A postgraduate teacher who, for example, is interested in only working with a small item bank to produce classroom tests, has a situation that is very different from that of the co-ordinator of a professional licensing bureau that manages and produces several dozens of examinations in the course of a year. Departments or Faculty bodies are concerned with problems such as sharing of the bank while maintaining the security of the items and often with prepared alternate forms of examinations. They are also interested in interfacing the scoring and analysis of information with the item entry process. The postgraduate teacher may be more concerned with programmes that would assist in the item writing process and that could operate on any available computer system.

The following are the three basic options for establishing an item bank:

Purchase of a Bank of Items, along with the software with which to operate it:-

The advantage of this option is that the labour of developing items and the software is eliminated, though one must learn to master the use of such software appropriately. Also, some of the item banks provide items that have been carefully developed and evaluated, and whose programmes are generally easy to operate. The major disadvantage is that much of the savings in developing items is neutralised by the necessity to check the items against the objectives or test blueprint, so that they conform to the blueprint. Also, many of the items may turn out to be flawed and have to be replaced. Furthermore, the constraints of the software limit what can be done about customising the item bank. One may be limited to a single item type, for example, and editing or replacing items may be very difficult. The purchase of an item bank saves only a part of the labour, that of preparing items. It is still essential that the coverage of an assessment be carefully specified which can be done through a curriculum guide, a list of objectives, or a test blueprint.

Purchase or Leasing of Item Banking Software With Local Development of Items:

The advantage of this option is that it offers the possibility of maximal flexibility, requiring only a minimal computer know-how. The disadvantage is that selection of software packages requires a careful analysis of the users' needs and of the features of each software package, thus, the need to master the manipulation of such software effectively. It also requires that the items, item entry procedures, and formatting be adapted to fit the capability of selected programmes and it's would be operators, and these capabilities vary widely from one programme/person to another.

Local Development of Software Programmes and Items:

The advantage of this option is that the programmes would be tailored right from the outset to provide the specific characteristics the test developer wants, so it is the obvious choice for a test developer who has the requisite computer know-how. The obvious disadvantage is that development of item banking software requires a level of computer sophistication beyond that of most test developers. However, if the user has a good word processing programme, and if time permits each test to be hand-formatted with the items hand selected, much of this disadvantage is lost. If, however, the developer chooses to collaborate with a person with requisite computer know-how, this challenge would be easily overcome

Steps in Item Banking

- **Item Entry:** There are several processes by which items may be entered into an item bank. They can be entered by keying in on a computer keyboard; they can be purchased already entered into an item bank that runs on the programme; they can be entered by reading a disk or tape that has been created using a word processing programme and they can be entered from a hard copy by use of a document reader; or the item banking programme may provide algorithms that create items.
- **Editing or adding new items to the bank:** Many of the considerations for item entry also apply here. However, if changes in the item bank are made frequently, the ease of entering and editing items becomes of even greater importance. The bank must of a necessity be amenable to easy editing so that

new additions can be made with ease, while items that have become obsolete can easily be deleted as well.

- **Classification/coding:** Regardless of how the items are initially entered, there must be a provision capable of making retrieval of the items in the bank possible in order to make revisions or create a test. It is essential that the items be coded in some fashion, and that certain information could be readily available. It is proper classification or coding of items in the bank that will engender this process. The primary reason for coding is to simplify the task of item retrieval so that tests can be assembled to fit the desired coverage of the assessment. The codes must be specified in advance so that they could be entered into the files at the time the items are entered. Once a coding scheme is set up, there should be no modification in the coding scheme unless the entire bank is recoded.
- **Formatting and printing tests; preparing an item map:** Some programmes were designed to present items for computer administration, whereas others produce printed forms. Some programmes will produce tests in either form, and some of the programmes for computer administration permit tailored testing; that is, items are selected in accordance with the examinee's performance on previous items.
- **Scoring and Item analysis:** When tests are administered by computer, the scoring is usually, but not always, handled immediately. Some programmes calculate a scaled score and produce an immediate report. Other programmes accumulate examinees' scores based on a given administration of tests. When printed examinations are administered, the scoring and analysis are usually separate processes.

Administering an Item Bank

An item bank will allow one to withdraw items as may be needed to develop new or special tests and subtests. There are principally two activities involved in running an item bank - making deposits and withdrawing items to develop a test. However, there are viable options for making deposits to the item bank. The "large deposit" option involves collapsing an existing item bank with one's own. If the existing item bank has been IRT calibrated, then it is necessary to administer a

subset of items (per content area) from the new bank along with items already in one's item bank. Remember, each item bank uses its own anchor items and allows one to equate the scales. This part involves testing with a relatively small group of students. The anchor items from the new item bank can be appended to the current group. Coordination would be similar to what one does in starting your own item bank.

The major task involved in using items from another item bank is a thorough, careful review of the items. All potential entries must be evaluated for technical quality, curriculum match, and potential biases. This process would involve some test development experts, curriculum/instructional staff, and coordination between the two.

After an item review, the available non-calibrated test items at one's reach could be treated like items developed by your staff. "Small deposits" could be made by calibrating and equating a few items at a time. One very efficient approach to collecting the relevant data is to append subtests of new items to original groups. The items within the original group would serve as anchor items for the new subtest(s) of items. In this way, you can constantly be adding to your item bank.

When fully developed, an item bank would be positioned to provide the advantages discussed above. If one has to develop new subtests from the bank, you would develop a blueprint/table of specifications to outline what or how you want your new subtest to be like. Curriculum specialists and test development experts would then go to the item bank and identify which items in the bank appear appropriate in terms of content and in terms of their relative difficulty. If they find an insufficient number of items, then one can make arrangements to add new items to the bank.

If the bank contains a sufficient number of items of the appropriate nature, such items can then be grouped to form a new subtest. Without pilot testing, the characteristics of this new subtest can be predicted, and with reasonable accuracy, you will know how much skill an examinee needs to obtain any given total raw score on the new subtest. The prediction should be validated by administering the subtest to two parallel sets of students, one having received appropriate instruction

and the other set not having received such instruction. This can also be accomplished by appending items to the existing forms. This validation would need a sample as large as you used in field testing the original group of items.

An item bank provides a scale of relative difficulty of tasks that covers multiple grade levels and skills within the particular content areas under consideration. As a service to the instructional/curriculum staff, one can provide information on the relative difficulty of different tasks within and across grade levels. For example, one can identify which sets of fraction problems in Mathematics, 200 level undergraduates find as difficult as certain decimal problems also in Mathematics; or one may identify which construction skills taught in 400 level Engineering class can be mastered by students in that level. It could also be used to help organize special programmes differently for both gifted and remedial students respectively.

Item bank is prepared for:

Item bank is useful to anyone whose duties involve preparing or conducting assessments or evaluation on a regular basis. In education, these people include classroom teachers, university instructors, directors of system-wide testing programmes, directors of state testing programmes and evaluators of special projects. Personnel in licensing and certification agencies and test publishers will also find the information in an item bank not only valuable but overtly useful. The needs of the users may definitely vary very greatly, and there are many options available, for individuals to find the kind of help that is right for them.

Advantages of Item Banking

The primary advantage of item banking is in test development. Using an item response theory method, such as the Rasch model, items from multiple tests could be placed on a common scale, one scale per a subject matter. The scale indicates the relative difficulty of the items. Items can be placed on the scale, i.e. into the item bank, without extensive testing. New sub-tests and tests, with predictable characteristics, can be developed by drawing items from the bank. For example, if one is interested in developing a new sub-test to cover complex numbers at one hundred or two hundred in the Department of

Mathematics, s/he can go to the item bank, identify items related to their objectives and then predict the characteristics of a sub-test composed of those items. The effect of including or excluding particular items can also be predicted.

Another advantage of an item bank is that it permits one to "deposit" additional items in the bank, to be withdrawn as the need arises. The size of the testing programme determines whether or not one or two practical approaches is/are adopted for depositing items in the item bank. One can make "large deposits" by merging one's item bank with the one of another unit's item bank. "Small deposits" can be made by adding a few locally developed items at a time. The large deposit option will involve purchasing from or trading items with another relevant/congruent department and then equating their scale to one's. The small deposit option involves piloting a fewer number of items with examinees in similar levels. This can easily be accomplished by adding a supplementary page containing experimental items to be administered along with booklets from the university.

Item banking provides substantial savings of time and energy over conventional test development option. In traditional test development, items can only be described relative to the other items within the test and to whom they were given. That is, item characteristics are grouped and test specific. With item banking, items are described in terms of their relative difficulty across grade levels. Developing a new test or sub-test, does not require going through the laborious process of developing a large set of items for piloting and evaluating. Instead, one just draws from the bank. Further, drawing from the bank allows one to make fairly accurate predictions concerning composite test characteristics.

One additional advantage of item banking is that it helps to establish a language for discussing curriculum goals and objectives. The items usually describe individual tasks students are capable or incapable of doing. The location of the items on a calibrated scale allows one to identify the relative difficulty of particular tasks. This provides a way to discuss possible learning hierarchies and ways to better structure the curriculum.

Conclusion

A careful review of the foregoing discourse leads to the obvious conclusion that the importance of item banking in the testing or examining process need not be overstressed. Item banking, in essence, ensures not only security of test items but also engenders the availability of test items for varied use at all points in time. It equally makes it easy for the adaptation of test items by varied users for various purposes. Item banking makes room for easy updating of items to make them meet the test of time and avoid obsolescence. It assures test quality because various psychometric processes in ensuring their validity and that they are of high standard. It invariably provides for multi-disciplinary collaboration among experts in testing/evaluation, curriculum development and information and communication technology among others. Item banking allows an expert who may not have designed a test to mark the test in the absence of the designer, thus avoiding delays that may have occurred if the situation were otherwise. This system of test preparation and administration as well as scoring and grading that are provided when developing the bank is recommended to the university for adoption.

References and Bibliography

- Gronlund, N.E. (1998). *Assessment of Student Achievement. Sixth Edition*. Needham Heights, MA: Allyn and Bacon.
- Lord, F.M. (1980). *Applications of item response theory to practical testing problems*. Hillsdale, N.J. L. Erlbaum Associates.
- Mengel, Bill E.; Schorr, Larry L. (1992) *Developing Item Bank Based Achievement Tests and Curriculum-Based Measures: Lessons Learned Enroute*. (ERIC Document Reproduction Number ED344915).
- Thondike, R.M (1997) *Measurement and Evaluation in Psychology and Education, 6th Edition*, Prentice-Hall, Inc.
- Ward, A.W.; Murray-Ward, M. (1994). Guidelines for the development of item banks. An NCME instructional module. *Educational Measurement: Issues and Practice*, 13(1), 34-39.
- Wright, B.D.; Stone, M.H. (1979). *Best Test Design. Rasch Measurement*. Chicago, IL: MESA Press.