

Principles of Item and Task Bank Construction

Sripen Srestasathiern and Fred Davidson

- ▼ [1. Introduction](#)
- ▼ [2. Item Banks and Task Banks](#)
- ▼ [3. A Practical Example](#)
- ▼ [4. Concluding Remarks : The Well-reasoned Task Bank](#)
- ▼ [5. Footnote](#)

Abstract

An item or task bank is a computer database in which the records are actual items/tasks and associated information (e.g., item performance statistics, usage history, etc.). Such database has great utility both for testing and teaching. We present several samples of task banks and discuss our experiences in creating a bank for some new tests used in the Department of Languages and Social Science at King Mongkut 's Institute of Technology North Bangkok. We demonstrate the use of off-the-shelf software for the department's item bank, in this case Access 97 (part of Office 97 from Microsoft, inc.) Our experiences in the creation of this new operational test item bank have yielded for us a series of principles - general statements of task banking which should apply to many settings, for example,

The Adventure Principle

Try to view your database construction as an adventure or an exploration.

Be brave. Trial and error is a norm, not the exception. As you travel through your adventure, you may need to record what you do.

1. Introduction

In this paper, we will describe test item banks , which are a special case of task banks. We will illustrate several types of task banks. We also detail our discoveries while creating a task bank on computer for a university in Thailand. That computer task bank will make use of commonly available off-the-shelf software , and we will propose that it is not necessary to purchase specialized item banking computer packages. As we discuss these topics, we will generalize several principles about task banks to help language test development at your setting.

2. Item Banks and Task Banks

Educational measurement companies , schools , teaching administrative regions (e. g. school districts) , and classroom teachers often have a need to organize and track their test items. Over time , test questions accumulates as do data about them : was the question easy? Was it difficult? Which curricular goals does it match? How frequently has is been used? An item bank can help to answer questions such as these.

An item bank is a database of test items. There is a rich and long-standing literature on the creation and maintenance of item banks. For example, Gronlund advocated (some years ago) a card file with one card per test item , where the card is to contain the text of the item (stem and choices), the "course , content , outcome" , and statistical analysis of the item (Gronlund, 1965 , fig. 11.2 on p. 210 ; see also pp. 215-216). Gronlund did not invent the cardfile for test items - an earlier reference appears in Davis (1951) . A paper by Groth, Allen and Grayson (1971) was an early description of item banks on computer. And an early book - length treatment of the topic is found in Wood and Skurnik (1969). The idea of an item bank is well-established in scholarly literature.

The basic unit of an item bank is a record corresponding to (typically) one test item. The record can contain the text of the item itself, associated usage history , performance statistics , and related information. For example , we might consider an English grammar item such as the following :

Mary gave _____ apples to her sister.

- a. any
- b. that
- c. the
- d. one

The test item about Mary and her apple-loving sister might appear as an item bank record such as shown in Figure 1.

Figure 1: Sample Item Bank Record

```
* _____ * [Begin record]
  Stem : Mary gave _____ apples to her sister.
  Choice1 : any
  Choice2 : that
  Choice3 : the
  Choice4 : one
  P1 : .50
  D1 : .26
  T1 : 26OCT1996
  P2 : .55
  D2 : .30
  T2 : 15MAY1997
* _____ * [End record]
```

The information to the left of the colons (' Stem ', ' Choice1 ' etc.) is known as the field , field name , or variable. The information to the right of each colon is called the entry or value , which is hypothetical in this example. In Figure 1 , the item has been administered twice , on Oct 26, 1997 and on May 15 , 1997. Results were calculated at each administration. At the first usage, fifty percent of the examinees got the item correct (P1=.50) and the item - to - total correlation was .26¹. The item was slightly easier on May 15 (P2 = .55) and correlated a bit better with the total scores achieved by test - takers on that date (D2 = .30).

Prior to widespread use of computers in education , item banks were often kept on index cards. Today , it is more convenient to key in data about each item and to create a computer item database. The creation , entry, debugging and manipulation of an item bank entail many of the same considerations that arise in creation of any computer database. One such consideration is the design of the item bank record itself. There are several possible ways one might organize an item bank record - these alternatives we illustrate in Figures 2 through 4.

Figure 2 : Sample Item Bank Record: Alternate Design

```
* _____ *
  Stem : Mary gave _____ apples to her sister.
  Choice1 : any
  Choice2 : that
  Choice3 : the
  Choice4 : one
  USE1 : 26OCT1996 , .50, .26
  USE2 : 15MAY1997, .55, .30
* _____ *
```

Figure 3 : Sample Item Bank Record : Alternate Design

```
* _____ *
  Stem : Mary gave _____ apples to her sister.
  Choice1 : any
  Choice2 : that
  Choice3 : the
  Choice4 : one
  P - values : .50 , .55
```

D - values : .26 , .30
 Dates : 26OTC1996 , 15MAY1997
 * _____ *

Figure 4 : Sample Item Bank Record : Alternate Design

* _____ *

Stem : Mary gave _____ apples to her sister.
 Choice1 : any
 Choice2 : that
 Choice3 : the
 Choice4 : one
 Key : 3
 P - values : .50 , .55
 D - vaules : .26 , .30
 Dates : 26OTC1996 , 15MAY1997
 * _____ *

Figures 2 and 3 display the p-values and d-values differently. Figure 4 is a variant of Figure 3 in which the 'key' (correct answer) is identified in a separate field. An alternative is to enter the key in the same field in each record , and then when building a test from the item bank to randomize the order of the choices at a later stage. These decisions reflect choices that confront the designer of the item bank : where to put information, how to arrange it, whether or not to include certain pieces of information (e. g. the key) - these are important decisions that must be made.

Such decisions lead us to our first principle of item bank construction:

Principle 1 : The Ultimate Use Principle

Item bank design is determined by the ultimate use of the item records, including both the product of the data bank and the creative steps required to obtain that product.

For example , the choice of where to place the p-values and d-values may depend on a simple consideration : will the item bank be used for automated test building of any sort? Such machine-constructed tests imply that a computer program will analyze the item bank and then choose items by some user-specified criterion ("give me item with a p-value of .40 to .60 and a d-value of at least .20"). The product is then sent to a destination like a word processing file or the computer screen if the test is to be administered by computer. The precise location of the item statistics, whether each item statistic is consulted or only the most recent result , and so forth are all considerations which require knowledge of that end product and of the programming needed to achieve that end product.

Item banks need not include only item statistics ; indeed , they might not be used for automated test building at all . Principle 1 does not imply that statistical information is the only criterion by which test items are selected nor does it imply that such information should be necessarily retained. It might be that design of the item bank is motivated by other issues. For example , we work in the area of second/foreign language testing. For many years now , our discipline has sought to break free of the straightjacket of highly discrete test items which focus largely on the grammar of the language being tested². Such tests are superb for situations where a normal curve score distribution is desired , because they tap an important core skill of language ability. However , modern language teaching paradigms encourage attention to a wider array of skills³. Our wider mandate has implications for an item bank design. Our second principle follows on from the first and has certain social and practical implications.

Principle 2 : The Item/Task Principle

An item bank is a special case of a "task" bank. Task banks are not wedded to any particular

teaching philosophy. They should answer to changes in theory , parctice , curriculum , social preferences and other flexible features of any educational systems.

All test items are actually educational tasks. We "task" a student with a multiple-choice question. We also "task" a student when we require a term-paper, a vault in gymnastics , a certain movement or gesture in the blocking of the school play, a particular tone on the oboe during the

Mahler symphony , a safety practice during Driver ' s Education , or a heartfelt piece of advice in a confidential guidance counseling session after a conflict with another student. All these are activities that happen frequently in schools. Educators carry repertoires of tasks which they use with students as part of teaching and testing. Any set of tasks is amenable to database banking. An item is just a particular case of a task.

For example , a science teacher might create an item/task bank of records about assignments which have been particularly successful in an introductory biology lab. Instead of field names like 'stem' 'choice' , and 'key', this task bank might include field names like : 'objective' , 'materials', and 'time required'. It might also include a usage history of the biology tasks. A particular record -for instance dissection of a frog - might show how often the teacher has actually had classes perform the dissections, which classes , and even which students. This bank is both a lesson plan record and a tool to help construct tests , for the teacher can consult the bank to see what was done before preparing the final exam.

Principle 1 applies in this case, because the teacher designs the bank with its ultimate goal in mind. Principle 2 can be seen in comparing this science example with the grammar item about Mary and her sister. The grammar setting is a situation where the item bank creators probably want to generate a test. The science setting may help to create a test , but its purpose is richer and more complex (we could expect the science database entries to be longer and more detailed). The difference between the grammar and science settings illustrates Principle 2. Both settings have a task bank , but the designs are quite different.

It may be argued that Principle 2 is implied by Principle 1. If the use of a task bank is to facilitate the teaching and testing of a science class, then it will be designed one way. If its use is to produce a grammar test , then it will be designed another way. The use determines the design.

We disagree with this contention. Item banking originated and has evolved greatly in the educational testing industry. We perceive that its innovations are slow to spread to other facets of teaching , and by separating Principle 2 from Principle 1, we advocate that standard practices of item bank design at educational measurement institutes could and should have great usefulness to classroom teachers , to school building administrators , and even to students.

3. A Practical Example

If a school administrator were to contact a testing company and enquire of their item banking practices , one negative result would ensue. The administrator would be shocked at the cost to create , maintain and use a task bank. Testing companies spend large sums for personnel , hardware, and software to maintain their test record banks. There are some commercial item banking software packages , but these are somewhat expensive , because great effort went into tailoring the software specifically for test databases.

An alternative is to use an off-the-shelf, lower-cost computer database program to create a task bank. We have selected to do so here with Microsoft Access 97, part of Microsoft Office 97, under the Windows 95 operating system. The entire Office 97 software package (Access and several other programs, including a word processor and spreadsheet) are much more reasonably priced than specialized item banking software.

We would like to illustrate the use of Access 97 in the creation of a task bank for a real testing situation. As we do so , we will evolve additional principles which concern the use of off-shelf database software for task banking.

Our testing situation is the assessment of English as a Foreign Language (EFL) for university and government personnel staff improvement in Thailand. Thai government agencies and government-supported universities have a need to assess the EFL ability of candidates who apply for government scholarships. Frequently , these candidates seek to further their study abroad in an English-speaking country, and often such study is intended to result in a Master 's or Doctoral degree. Each year, the Thai government grants quite a number of such staff-upgrade awards. For a number of years , this group took an EFL test prepared and implemented by the government 's Office of the Civil Service Board. In 1995 , this testing was decentralized and distributed to various Thai government and university departments in which upgrading of human resources was a priority. The particular testing situation of interest here is the EFL test under development at the Department of Languages and Social Science at the King Mongkut 's Institute of Technology North Bangkok.

The Department has organized this effort by creating a Testing Unit , which is charged not

only with the scholarship testing but also with creating an EFL test that meets other purposes. The Department also wants a test that can be used as local EFL certification for university personnel and other interested parties - people who simply want to document their study of English. It also wants the test to serve as admissions evidence for its own Master 's degree in English for Business and Industry. This is to be a multi-purpose test answering several constituencies and mandates.

Results for the test will be effective for two years. As is typical in large-scale foreign language testing , there is a concern that over time language ability might degrade , and since candidates might therefore re-take the test , multiple test forms are needed. Multiple forms would also improve test security by guarding against the loss of an entire test booklet (an alternate form could then be used). Hence , the Department has decided to create a testing item bank. The main purpose of the bank will be to store test items , to consult pilot statistics and expert analysis about the items , and thence to create equivalent forms of the test. We have selected sample non-operational reading items from this context to illustrate the topics of our paper. The item banking at the Department is just beginning , and so our illustrations are primarily of initial design and creative decisions. As the bank grows and evolves over the coming years, the bank will grow and evolve as well, and it is likely that the decisions we make now might change, a point in keeping with Principle 2 above.

We discussed the goals of this project at length and agreed that a database 'form' or 'template' was needed. After some experimentation, we learned that Access 97 calls these 'views' or 'form views' and the process to establish a view (to arrange it on the screen) is called a 'query'. Regardless of the name, the function of a view is rather simple. It allows the database designer to structure the computer screen so users can enter and view data in consistent manner. Let us assume that we want to work with the Department 's reading items , some of which have associated graphic images and all of which have associated reading passages. Figures 5 and 6 illustrate our data view.

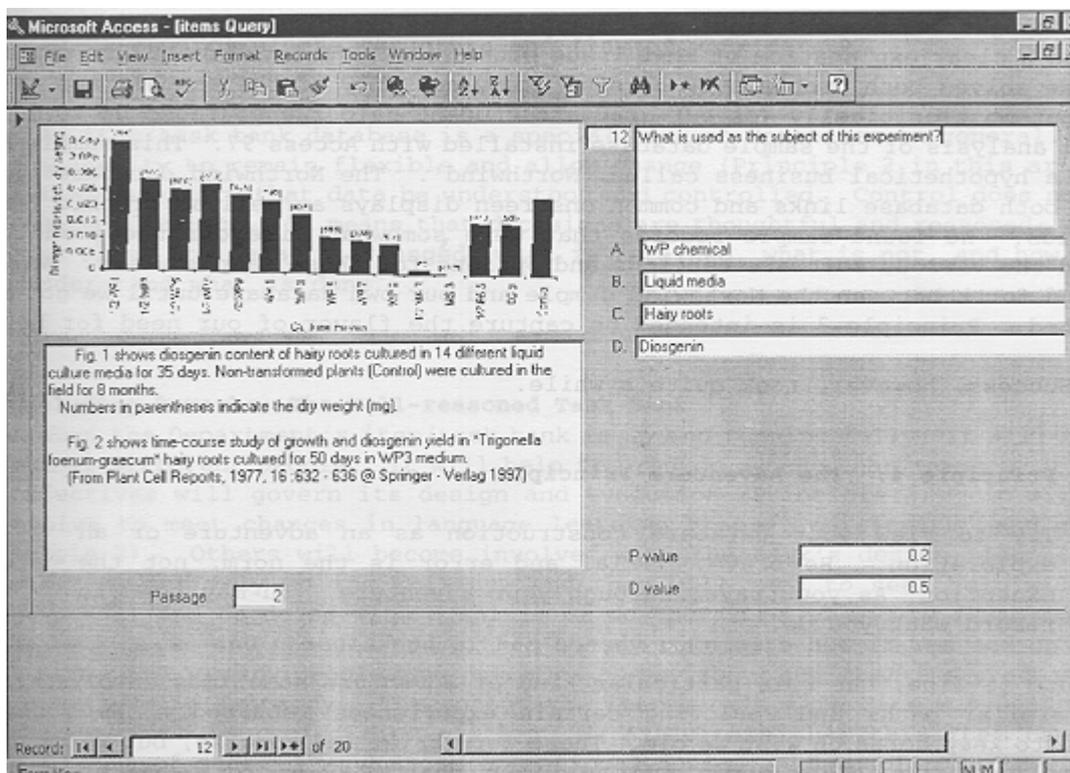
In Figures 5 and 6 , Principle 1 is still at operation, because the database is designed to show an item at the right and the associated reading passage at the lower left. If there is an associated graphic, then it is displayed at the upper left. If we were creating a grammar test-such as the "Mary" item above-we would have a different view. It would be much simpler, because it would not have an associated passage and graphic image.

Figure 5 : Access 97 Data View : reading item with passage

The screenshot shows the Microsoft Access 97 interface in Data View. The window title is "Microsoft Access - [Items Query]". The menu bar includes File, Edit, View, Insert, Format, Records, Tools, Window, and Help. The toolbar contains various icons for navigation and editing. The main area is divided into three sections:

- Top Right:** A question box labeled "5" containing the text "What is the purpose of the writer?". Below it are four multiple-choice options:
 - A. To teach
 - B. To inform
 - C. To explain
 - D. To instruct
- Bottom Right:** Two input fields for "P value" and "D value", both containing the number "0.5".
- Bottom Center:** A "Passage:" label followed by a text box containing the number "1".
- Bottom Left:** A record navigation bar showing "Record: 1 of 20" and a "Form View" button.

Figure 6 : Access 97 Data View : Reading Item with Passage and Graphic



We do not want to teach Access 97 in this article. Rather, we wish to reflect on our experiences with this database and from that reflection, induce several additional principles. We want to generalize from our work to create the data view shown in Figures 5 and 6. Our first generalization concerns help. We were not able to design the database without help, and our experiences in seeking help suggest our next principle.

Principle 3 : The Help Principle

Make use of all help available to you : printed manuals , online help, online sample files , and consultants.

For example , we had quite a bit of difficulty with displaying both the items and the passages on the same view form (onscreen at the same time). This is because we actually created two databases : one for items and one for passages. We did so to conserve resources. One reading passage might be used by as many as ten items, so if we had copied the passage onto each item record, we would have wasted space.⁴ We had correctly linked the two databases - this we knew from a different Access 97 screen that displayed the link as a line onscreen. However we could not get the screen to show us each item and its passage together.

In addition , we wanted certain passage records to show a graphic as well as a reading passage text itself. We scanned in the graphic image but could not figure out how to define a field in our passage text itself. We scanned in the graphic image but could not figure out how to define a field in our passage database for graphics. We saw how to define a field for characters and numbers (e. g. the hypothetical p-values and d-values) and we thought we knew the right way to approach the graphics problem , but we could not find a clear explanation of that issue either in the printed or online manuals.

We solved both problems by careful study not only of the manuals and online help , but most critically (as our graduate student helper encouraged us to do) by is for a hypothetical business called 'Northwind'. The Northwind database illustrates both database links and common onscreen displays as well as graphic images as fields . We found sample records that were somewhat like our needs and then studied the various software settings and options for those sample views. We went back and forth between the Northwind sample and our own database until we got what we wanted. Principle 3 is intended to capture the flavor of our need for help.

Success , however , took quite a while.

Principle 4 : The Adventure Principle

Try to view your database construction as an adventure or an exploration. Be brave. Trial and error is the norm , not the exception. As you travel through your adventure, you may need to record what you do.

Adventure is fine, but this particular kind of adventure sometimes involves traveling similar paths and re-living certain experiences repeatedly. We found it helpful to keep notes on what we did. These were brief and concise, but we returned to them and edit them again and again so that they would serve as a later reference. We managed our adventure with our note-taking.

Figure 7 : Excerpted notes (these pertain to the creation of the form/view itself ; i . e . , they apply to the point

in time of designing the actual screen shown in Figures 5 and 6)

Designing the Screen Form

(1) Open the itembank

(2) Press the **Forms** tab

(3) Click **New**

(4) Click **Forms Wizard**

(5) Choose the **new query** in the wizard window by selecting **New Query** in the

Tables/Queries list box. (To find this ,

click the up or down arrow at the corner of the window.)

(6) Click the fields that you want to appear on the screen. If you want all of the fields, click

>>. Then click **Next**

(7) Click **By Item**, then **Next** and **Next** again. The **Style** window will appear. Choose

Standard and then press **Next**.

(8) Name it and press **Finish**.

What is the benefit of notes like these? Perhaps we are (in effect) writing our own Access 97 manual for use by other staff members in the Department. Perhaps we are recording our notes for a later date when, after entering operational data for several months, we wish to return and re-design the view form - that is, perhaps our notes guard against our own forgetfulness. Perhaps our learning style is such that we only feel we have achieved a goal when we can write it down in our own preferred note style.

Perhaps these notes are just a desire for control :

Principle 5 : The Data Control Principle (quoted from Principle 2.1, Davidson, 1996)

Take control of the structure and flow of your data.

An item/task bank database is a special case of databases in general. While it is necessary to remain flexible and allow change (Principle 2 in this article), it is also necessary that data be understood and controlled. Control does not mean bull-headed fixation. It means that decisions are thoughtful and reasoned : We need to think about how data is arranged , what is retained , what is not , and how others will understand what is done.

4. Concluding Remarks : The Well-reasoned Task Bank

When the Department 's item/task bank is up and running fully , it will be used by several different people. It will help to produce a test that meets many needs. Its objectives will govern its design and evolution (Principle 1). It will grow and evolve to meet changes in language learning theory, curriculum , and mandate (Principle 2). Others will become involved with the bank ' s design, learning and contributing actively to its utility, and they will need to seek help willingly (Principle 3) and to do so with a spirit of co-operative adventure (Principle 4). If the bank grows and changes and either nobody remembers how it was set up in the first place or too many people have too much to say about it , then its product - the test - will be in jeopardy. Throughout the item bank ' s life , its mutability must be balanced with control (Principle 5). We anticipate that this item bank will reach a point of stable , productive operational design , and we plan to re-visit it in a year or two and report additional principles which are yet to surface.

Footnote

¹The p-value is an index of item difficulty. The higher the p-value, the easier the item. The d-value is an index of item discrimination and is usually computed by applying a correlation formula between the right-wrong item scores and the total score. D-values are best judged relative to other items in the test and testing situation. Testing situations which desire a normal , bell-shaped curve often seek many items with a p-value of about .5 and as high a d-value as possible. Modern item statistics also make

use of item analysis from Item Response Theory (IRT) , which gives additional useful statistics about each item's performance.

²This prevalent status quo is precisely why we chose a grammar item earlier in this paper - it is by far the most familiar type of language test question.

³For some seminal literature this wide array of language skills, see Canale and Swain , 1980 ; Bachman , 1990 : Chapter 4.

⁴Our items and passages are, together , known as 'relational' database, in which the records of one database are linked to the records of another. In this case , it is a 'heirarchical relational' database , because one database (the passages) has far fewer records than the other database (the item) .

References

- Bachman , L. F. 1990. Fundamental Considerations in Language Testing. Oxford , UK: Oxford University Press.
- Canale , M. and M. Swain. 1980. Theoretical bases of communicative approaches to second language teaching and testing.
Applied Linqustics 1:1 , pp. 1-47.
- Davidson , F. 1996. Principles of Statistical Data Handling. Thousand Oaks , CA : Sage.
- Davis , F. B. 1951. Item Selection Techniques. In E. F. Lindquist (Ed). Educational Measurement. Washington , DC :
American Council on Education , pp. 266-328
- Gronlund , N. 1965 . Measurement and Evaluation in Teaching. New York : MacMillan.
- Groth , W. P. , D. W. Allen , and A. Grayson. 1971. Computer programs for test objective and item banking.
Educational and Psychological Measurement 31:1, pp. 245-250.
- Wood, R. and L. S. Skurnik. 1996. Item Banking: A Method for Producing School-based Examination and Nationally Comparable Grades. Upton Park, UK: National Foundation for Educational Research in England and in Wales.

	<p>NAME : Associate Professor Sripen Srestasathien</p> <p>EDUCATION : M.Ed (English Language & Literature)</p> <p>POSITION : Associate Professor in Department of Languages and social Sciene King Mongkut ' s Institute of Technology North Bangkok</p>
	<p>NAME : Associate Professor Dr. Fred Davidson</p> <p>EDUCATION : Ph. D. (Applied Linguistics)</p> <p>POSITION : Associate Professor of English as an International Language University of Illinois at Urbana- Champaign ,</p>



This document was last modified on