DOCUMENT RESUME

ED 462 435 TM 033 699

AUTHOR	Anzaldua, Ric M.
TITLE	Item Banks: What, Where, Why and How.
PUB DATE	2002-02-00
NOTE	33p.; Paper presented at the Annual Meeting of the Southwest Educational Research Association (25th, Austin, TX, February 14-16, 2002).
PUB TYPE	Reports - Descriptive (141) -- Speeches/Meeting Papers (150)
EDRS PRICE	MF01/PC02 Plus Postage.
DESCRIPTORS	*Item Banks; Item Response Theory; Test Construction; *Test Items; Validity
IDENTIFIERS	Calibration; Linkage

ABSTRACT

This paper discusses item banks calibrated to indicate levels of difficulty to assist in test development. The item bank topics discussed are: (1) purpose; (2) development issues; (3) advantages and disadvantages; and (4) practical issues. The most common issues are content validity, reliability, concerns with software purchase and programming, classification and analysis of items, maintenance and upkeep, and good item writing practice. It is important to remember that item response theory remains an integral part of interpreting ability measurements for the forms generated from the item bank. The paper cites examples of studies in which form linking and sample size have become issues. The calculations involved in linking forms are simple and straightforward, but when there are hundreds of items across a multitude of forms administered to thousands of examinees, linking forms becomes a massive undertaking. It is beneficial to take advantage of existing item bank software to develop an item bank. An example of linking forms is presented using a heuristic data set. (Contains 3 tables, 3 figures, and 18 references.) (SLD)

Item Banks:

What, where, why and how

Ric M. Anzaldua

University of North Texas

Paper presented at the annual meeting of the Southwest Educational Research Association, February 14-16, Austin, TX

Abstract

This paper is a presentation of item banks calibrated on levels of difficulty that assist in test development. The item bank topics discussed will be (a) purpose, (b) development issues, (c) their advantages and disadvantages, (d) and practical issues. This paper will cite examples of studies where form linking and sample size have become issue. Finally, an example of linking of forms will be presented using a heuristic data set.

Introduction

An item bank is a collection of test items, calibrated on levels of difficulty that assists in test development. These items are organized and catalogued "like books in a library" for use (Testing the Questions, October 2001). Such a bank will allow for retrieval of items for use as a test form that will determine a student's ability in one content area. The resulting score is a measurement of ability that can be placed in relation to other students' abilities.

Ability is determined regardless of similarity of items used across forms administered to students. This is possible because Item Response Theory (IRT) is the basis for item banks. Crocker & Algina (1986) say,

> With item response theory the test developer assumes that the responses to the items on a test can be accounted for by latent traits that are fewer in number than the test items. Indeed, most applications of the theory assume that a single latent trait accounts for the responses to items on a test. At the "heart" of the theory is a mathematical model of how examinees at different ability levels for the trait should respond to an item. (p. 339)

Under IRT, item banks are created to measure one ability, called unidimensionality. This unideminsionality means that all individuals are being measured on the same ability.

Item difficulties must also be invariant. In IRT, items are represented graphically by an item characteristic curve. This curve should exhibit the same characteristics, that is have the same parameter estimates, despite the characteristics of any group of examinees to which that item was administered. It is invariance that allows for the item difficulty to be interpreted as level of ability measured among individuals. For more information on the ICC and its parameter estimates, see Crocker & Algina (1986), Embretson & Reise (2000) and Hambleton (1983).

The bank itself is more than just a collection of items. Whether on note cards or in a computer program, the items are stored, indexed and retrieved for use. A good item bank is well organized with items indexed (Roid, 1989) and will allow for addition of items and for their editing and updating. It will also produce diagnostics that will determine usefulness and exposure of each item.

This paper is a presentation of item banks. The item bank topics discussed will be (a) purpose, (b) development issues, (c) their advantages and disadvantages, (d) and practical issues. This paper will also cite examples of studies where form linking and sample size have become issue. Finally, an example of linking of forms when building an item bank will be presented using a heuristic data set.

Purpose

Item banks were originally developed for the purpose of assisting school teachers and university instructors in developing classroom tests (Item Banks, October 2001). The intent was to help teachers tailor tests to immediate objectives. Doing so meant that each teacher would have a pool of items from which to draw. This pool was intentionally developed so that any subset of items (e.g., a test form) would be properly calibrated with the core of items in such a manner that determination of ability is possible. (Item Banks, October 2001).

Suppose two instructors wish to test their students' ability on a particular content area. Both have the option of accessing the item bank. Each one creates a form from items drawn from the item bank. Unless the instructors are working together to develop their test forms, it is unlikely they will each select the same items for determining their students' abilities. Further, each instructor must determine whether or not their students have reached the desired level of ability in that content area. It would be desirable to have the items set up in such a way that any group drawn from the item bank would deliver a true measure of the students' abilities. For this reason it is necessary that the forms are comparable, that is, they are on the same scale. Note that comparability would hold across time. This is relevant when an instructor wants to know how their current students' abilities compare to the ability of prior students.

Computers

Developing an item bank has been greatly enhanced thanks to computerization. Prior to the use of computers and development of software programs, item banks were developed on cards, kept in files and managed manually (Lord, 1962, Roid 1989).

Word processors have allowed for quicker and easier test form development. They have increased efficiency of assembly and printing of multiple forms (Roid, 1989). The advent of personal computers has impacted development and maintenance of item banks. Computers have increased the capacity to archive and index large numbers of items as well as assisted in indexing, updating, editing and retrieving items with greater ease (Roid, 1989). Additionally, they allow for designing test forms based on specifications.

Computers also deliver test form and item diagnostics. Diagnostics are especially important during test bank development, building and testing. Computer programs, such as SPSS (Version 10.0), Rascal (Assessment Systems Corporation, 1995) and Bilog (Mislev & Block), produce information on items to test for item difficulty and reliability. Rascal and Bilog also provide information to test for unidimensionality. Unidimensionality is a necessary condition for use of IRT in item banks. This is the idea that the items on a form, and thereby in the bank, are measuring one and only one ability. This is key in measuring one's ability using interpretations from item difficulty. Reliability, item difficulty and unidimensionality must all be considered in IRT and thus when developing an item bank.

Yet, simply measuring one's ability is insufficient. It must be comparable to others whose abilities have been measured using the same item bank. Wood & Skurnik, (1969) say, "the strength of an item characteristic curve method of item analysis is appropriate in item baking because it offers a logical framework for describing precisely how an item functions in relation to a global ability group". To accomplish this, IRT is necessary so that an item characteristic curve can be calculated. Discussion of ICC's and IRT is beyond the scope of this paper. To see a detailed explanation of this, see Crocker & ALgina (1986), Embretson & Reise (2000) and Hambleton (1983).

Computers have also allowed for the development and use of tests administered on computers themselves, specifically, computer adaptive testing (CAT). No other area is more dependent on rapid composition of test forms (Roid, 1989), item indexing, editing, and retrieval, as well as the delivery of the ability measurement using IRT and item diagnostics. For more information on CAT, see Mills and Stocking (1996).

It is important to note that since computers have begun to be used in test development using item banks, the Rasch model (IRT) has not been changed since its introduction (Wood & Skurnik, 1969). This model for test development and item calibration, then, is integral in item banks and testing situations that utilize these banks.

Item Bank Development

Prior to computerization there were several issues to address when developing an item bank. First, the items in an item bank must serve a purpose; they must measure some "fragment of learning" (Wood & Skurnik, 1969). To accomplish this first the interested parties would prepare a blueprint of the objectives across a given content area. Then only items developed or incorporated form other sources that measure the content area are entered into the item bank. these should range in difficulty level and in sufficient numbers to reflect the blueprint.

Secondly, the items must be organized. Each item must be stored and catalogued. They must also be determined to be representative of the broad range of difficulty levels to which the test will be administered. This will be assessed during the third step, the pretest. At that point, should the desired ability level not be achieved, the item can be edited, its level of difficulty adjusted, and then retested. In many cases, the item may be removed for not fitting into the blueprint, not contributing to reliability, or simply being a badly written item. Item banks only work if each item has a proper estimate of difficulty (Lord, 1962).

Further, it is desirable to have more items than needed. One recommendation is there be at least three times as many items as needed. This is necessary due to "wastage" (Wood & Skurnik, 1969) during the pretest.

Third, when developing an item bank the items must be calibrated. The items are placed on different forms and administered to a sample population. Item analyses are performed to determine item difficulty, form reliability, and unidimensionality. Once these are satisfactory, the item is entered into the item bank as part of the form it was on. This form is calibrated among common items with other forms.

In one case, Lord (1962), calibrated items by separating them into distinct test forms and then testing individuals representing the general population. The problem with this method was none of the items were on more than one form so the forms could not be linked. The result was that none of the items were calibrated with items not already on their form. This meant comparability became an issue; there is no correlation drawn between items from the different subtests. Comparability, then, was subjective and left the judgment of the instructors.

Also, in the case above, there were too few individuals tested on each form. The level of difficulty for each item, then, was to easily influenced by characteristics of each group tested. This violated the assumption of invariance of items which is important in IRT.

Ideally, during a pretest each item should be maximized for use (Lord 1962), that is, it should appear on as many forms as possible without the pretest process becoming unnecessarily unmanageable. Each item should be tested for unidimensionality. This can be accomplished through the use of computer softwares mentioned above. Items not faring well during pretest can be further examined. They may be changed or "fixed" so that they achieve unidimensionality or they may be removed altogether.

Good item bank development includes the inclusion of each item on different forms which are tested on different groups. This will assist in ensuring item invariance. Comparability will be achieved as each item appears on different forms thus allowing for their item difficulties to be calibrated across forms. Nunnaly (1968) recommends at least ten individuals be used to test each item and in no case should less than five individuals per item be used. Also, the range of difficulty of items on each pretest form should vary. Any student getting all answers correct or all answers incorrect will be removed from the item analysis since they contribute no information. A good practice is to use smarter people to calibrate the more difficult items and less smart people to calibrate the easier items (Wood, 1969, appendix 2). In any event, try to ensure that each individual contributes to the pretest by getting at least one right and one wrong. The calculation of item statistics will remove any observation that does not contribute information.

Calibration across forms is done through form linking. When linking is performed, any subset of items used for a test form will produce a measurement of ability for each student such that it can be compared to others who have been tested from this item bank. An illustration maximizing item use during a pretest and an example of test linking will be presented later.

The three issues addressed above were important during test development prior to computerization. Currently, however, there are several options that help simplify the development and management of item banks. Computer programs that manage item banks and item banks themselves are available for purchase. One has the option of (a) purchasing both the bank and the software to manage it, (b) purchasing the bank but not the software to manage it, or (c) purchase just the software to manage an item bank. Option (b) would require a skilled programmer to develop the software to manage the item bank. Option (c) would require development or import of existing items.

When considering a purchase as in option (a) there are additional considerations. Where items are already developed it must be ascertained that they measure the contents and weights as according to the blueprint (Ward & Murray, 1994), as developed during the first step of item bank development. Some items may be flawed or they may all be of a limited type, say all multiple choice (Ward & Murray, 1994). Further, editing, importation and replacing may be difficult. Other important considerations are flexibility of item types, required computer programming ability, local tailoring and form development. In any event, when considering these purchasing options keep in mind that extensive programming is beyond the skill of most test developers. It is then always appropriate for each researcher to calibrate item statistics specific to their own sample (Vacha-Haase, Kogan & Thompson, 2000)

Advantages and Disadvantages

Presently, item banks have many advantages in test development. Some of them, as mentioned previously, are local control of test development, allow for tailoring of specific objectives, and allow for editing and updating of items. These items are strictly tied together using IRT which means a student's measured ability is on the same scale as any other student who has been tested using the same item bank. Other advantages are flexibility, minimization of item exposure, time and energy savings.

Flexibility of item banks allow for the test to be created as a short or long form and still provide a statistically valid measure if the ability of the student. Minimization of item exposure means large numbers of items used in short forms can allow for repeated use without overexposure of an item. Any subset of items will provide the same ability measurement without the need for using the same items over and over.

Time savings are accomplished once the item bank is established. This is true of energy savings. Be aware, however, that a considerable amount of both is needed during the development and testing of an item bank. Once the item bank has been established, the time and energy needed to maintain the bank is minimal. It is a good idea to make this an ongoing process. Items and forms should be reviewed regularly for exposure and reliability.

Some of the disadvantages are cost, time and energy expended up front during development. Depending on the option taken when purchasing item banks, a required level of skill at software programming may also contribute to costs, time and energy expenses. Some item banks limit the number of items that can be stored. Others banks may limit the type of items that it contains. Some item banks are not maintained thus their items become overexposed. In this case, the bank has been in service too long. An item bank must be kept as a "living" bank with continual editing, updating and refining (Ward & Murray, 1994).

Software that is used for developing an item bank but does not have stored items will require linking of existing forms or items that have been entered into the bank. Therefore, a trained measurement expert would have to be consulted to evaluate item and form statistical issues, including reliability and calibration.

Aside from statistical issues, good item writing is necessary when developing an item bank. Remember, an item bank is only as good as the material put into it. A study by Hansen & Dexter (1997) found there were significant violations of good item writing practices in developed item banks.

Linking of Forms

Simply put, linking is calibration of items achieved by having items common to existing forms. Linking of forms is a necessary step when building an item bank from individual items, from existing forms, from a combination of existing items and existing forms or when adding items or forms to an existing item bank. It is linking that allows for placing items and different forms on the same scale of measurement.

When forms exist and are to be linked to each other, it is necessary that there be items common to both forms during a pretest. If no items exist, simply add some from each form to the other form. In the case that items are being added to an item bank, first create a form using the new items and existing items in the bank. Figure 1 is an example diagram of four existing forms being linked through common items. In this example, each of the existing forms measure a different range of ability from low ability (easiest) to high ability (hardest). The linking forms are composed of items taken from the two forms being linked.

Insert Figure 1 About Here

The next step is to pretest this form to get statistics on reliability and item difficulty as well as to assess unidimensionality. Once these have been deemed satisfactory, the item difficulties are then calibrated to the item bank by the common items. These common items provide the link to the item bank which will place the new items on the same scale as the existing items. Once in place, the newly added items now reflect a measure of ability.

As stated before, maximizing item use during pretest is important. Figure 2 depicts items and sample groupings that represent the maximization of item use. Notice that each form has five items. Items 2 through 5 from form A appear on form B along with item 6. Likewise, items 3 through 6 appear on form C along with item 7. This is repeated for all items. the last four items would be on a form with item 1. The last three items would be on a form with items 1 and 2. In this manner, each item will appear on five different forms.

Insert Figure 2 About Here

Suppose your recommended sample size is 10 students for each form item. Then, for the example depicted in Figure 2, each item is tested on 50 individuals rather than on the 10 that take any one form. Pretest item difficulties are estimating the true item difficulty. Therefore, the larger the sample the better the item difficulty estimate. Using a linking item across forms provides many more observations.

Another example of maximizing item use on subjects is a study (Wood & Skurnik, 1969) where 2,000 subjects were being tested over 500 items. In this example, the number of items is too volumnous to follow the example above. It was decided that each individual would be tested on 30 items. This means there would be

2,000 persons x 30 items = 60,000

item responses. Allowing for an even distribution of each of the items, then there are a possible

60,000 responses/500 items = 120 responses/item.

That is there were 120 persons tested on each item. Given that there were only 30 questions per form, then each item should appear on

120 responses / 30 items per form = 4 forms.

So rather than settle for the 30 responses per item, this study maximized item use and generated 120 responses per item by placing each item on four different forms.

Finally, consider each item in the above example a test form rather than individual items. This maximization of item use across forms represents a balanced incomplete block design using test forms as units. Statistically, that is sound and almost as good as a balanced complete block design. A balanced complete block design would mean giving every item to every individual during testing. That could prove logistically difficult and daunting. Given this fact, balanced incomplete block designs are more than adequate.

A heuristic data set was developed to provide an example of form linking. The data set contains 30 items and 720 observations, or responses to each item. Item 1 through 18 were on form A and items 13 through 30 were on form B. There were 6 items linking the two forms, items 13 through 18. Figure 3 is a diagram of this form linking.

Item difficulties are calculated for each of the two forms and a linking is performed. The first step is to place the item difficulties for the linking items side by side (D_A and D_B for forms A and B, respectively) and calculate their item differences, D (Table 1). Calculating the mean of their differences yields -0.47. This amount will be used to adjust the item difficulties of form B as an estimate of their item difficulties in relation to form A.

Insert Table 1 About Here

Table 2 shows the calculations that are made to center and scale the item difficulties so that they are calibrated between forms. Columns D_A and D_B contain item difficulties for all items on their respective forms. After centering, the mean of the item difficulties for each form is 0.00. The third column D_B + D is the centering of the scores in form B to the center of form A. Notice the mean for form B is now -.047. This is because the mean of the differences was added to each item difficulty in form B.

Insert Table 2 About Here

The third step is to use the information we have on the item difficulties for the linking items to make a better estimate of their item difficulty. Column $(D_A + (D_B+D_A))/2$ is the average of the linking items' difficulties from form A and the re-centered item difficulties from form B. We now carry over the item difficulty estimates into column D_C, referred to as the common item difficulty. For items 1 through 12 we simply carry over the item difficulties from column D_A to column D_C. For items 19 through 30 we carry over the re-centered item difficulties, $D_B + D$, to column D_C. For the averaged estimate of the linking items, carry column $(D_A + D_B + D_A)/2$ over to column D_C.

Notice now that the mean item difficulty for D_C is -0.16. Once again we re-center the mean to 0.00. To do this we subtract the mean D_C from each item difficulty in column D_C. Here, D_C.-Mean D_C. = D_C.-0.16. Now the two forms are calibrated and have mean item difficulty of 0.

Verifying that this method properly calibrates the items from the two forms, the item difficulties calculated during the forms linking are placed in ascending order alongside the item difficulties from the entire 720 observations (Table 3.). All items align in the same order and very nearly the same item difficulty except for items 25, 8 and 26. Each of these items had item difficulties calculated during the linking slightly different from those calculated when all observations are included.

Insert Table 3 About Here

This reflects the heuristic data. This data was created using a much smaller data set of about 300 responses where the lower 60 responses were copied to give a total of 360 responses. Then the lower 120, which is the last 60 observations used twice, were copied an additional three times to add 360 responses for a total of 720. this means the responses for Form A represented 300 different individuals with 60 of them being repeated. The responses for Form B represent only 60 individuals, each being repeated 6 times. This is reflected in the differences between the calibrated item difficulties and the total data set item difficulties.

A better developed heuristic data set would have provided closer estimates for item 26, 1.59 versus 1.25 from entire set, and for item 8, 1.07 versus 1.16. Even so, the estimates for item difficulties were close because there were at least 20 observations per response. This twice the number of observations recommended by Nunnaly (1968). Therefore, the observation reuse performed in creating Form B still allowed for the item difficulty estimates to be very close to those calculated using the entire data set.

Summary

The calculations involved in linking forms are simple and straightforward. But when there are hundreds of items across a multitude of forms that were administered on thousands of examinees, linking forms will be a massive undertaking. It is beneficial therefore to take advantage of existing item bank softwares to develop an item bank.

As discussed here, there are many issues to address when developing an item bank. The most common issues are content validity, reliability, concerns of software purchases and programming, classification and analysis of items, maintenance and upkeep, as well as good item writing practices. Throughout the process, it is important to remember that item response theory remains integral to the ability to interpret ability measurements for forms generated from the item bank.

Currently, thanks to the increased use of computers in testing there are other item bank capabilities being addressed. One is the ability to create forms using algorithms that respond to requirements entered. Another is the ability to import and export graphs and to have performance and audio-visual item capabilities. For a discussion of these see Ward & Murray (1994).

Sources

Assessment Systems Corporation. (1995). Rascal (Version 3.50e) [Computer software]. St. Paul, MN: Author

Crocker, L., & Algina, J. (1986). *Introduction to classical and modern test theory.* Fort Worth, TX: Holt, Rinehart and Winston.

Embretson, S.E., & Reise, S. (2000). *Item response theory for psychologists.* Mahwah, NJ: Erlbaum.

Hambleton, R.K. (Ed.). (1983). *Applications of item response theory.* Vancouver: Educatonal Research Institute of British Columbia.

Hansen, J.D. & Dexter, L. (1997). Quality multiple-choice test questions: Item-writing guidelines and an analysis of auditing testbanks. *Journal of Education for Business, 73, 3*, 94-7.

Item Banks: What, Why, How. Retrieved October 29, 2001 from http://www.rasch.org/memo43.htm

Lord, F.M. (1962). Estimating norms by item sampling. In R. Wood & and L.S. Skurnik, (1969), *Item sampling: A method for producing school-based examinations and nationally comparable grades*, Appendix 3. Slough, Wales: National Foundation for Educational Research in England and Wales

Lord, F.M. (1980). *Application of Item Response Theory to Practical Testing Problems.* Hillsdale, NJ: L. Erlbaum Associates.

Mills, C.N. & Stocking, M.L. (1996). Practical issues in large-scale computerized adaptive testing. *Applied Measurement in Education, 9, 4*, 287-304.

Mislev, R. & Block, R.D. (n.d.) Bilog (Version 3) [Computer Software] Lincolnwood, IL: Scientific Software International.

Nunnaly, J.C. (1968) *Psychometric Theory.* New York: MacGraw-Hill.

Roid, G.H. (1989). Item writing and item banking by microcomputer: An update. *Educational Measurement: Issues and Practice, 8, 3*, p17-20.

Rudner, L. (1998). Item Banking [Electronic Version]. *Practical Assessment, Research & Evaluation, 6,4.*

Testing the Questions - The Rasch Model and Item Banking. Retrieved October 29, 2001 from http://www.rasch.org/memo49.htm

Vacha-Haase, T., Kogan, L.R., & Thompson, B. (2000). Sample compositions and variabilities in published studies versus those in test manuals: Validity of score reliability inductions. *Educational And Psychological Measurement, 60, 4*, 509-522.

Ward, A.W, & Murray-Ward, M. (1994). Guidelines for the Development of Items Banks. *Educational Measurement: Issues and Practice, 13, 1*, 34-9.

Wood, R. & Skurnik, L.S. (1969). *Item Banking: A method for producing school-based examinations and nationally comparable grades.* Slough, Wales: National Foundation for Educational Research in England and Wales

Wright, B.D. & Stone, M.H. (1979). *Best Test Design.* Rasch Measurement. Chicago, IL: MESA Press.

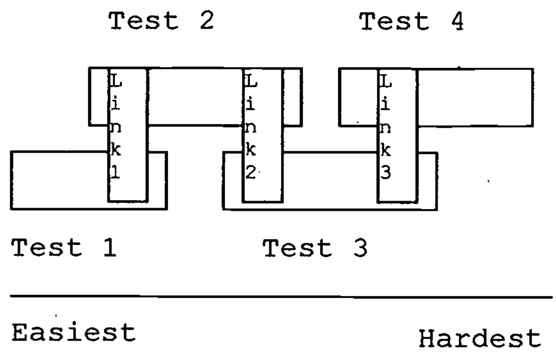

Figure 1. Linking four tests, that measure different ability levels, through common items.

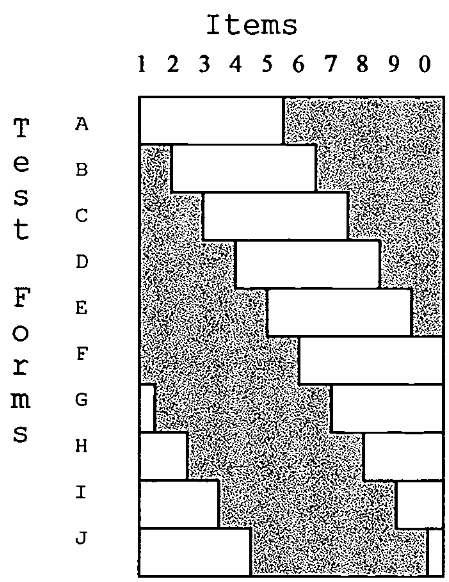

Figure 2. Maximizing use of items on pretest forms. Each item appears on 5 forms.

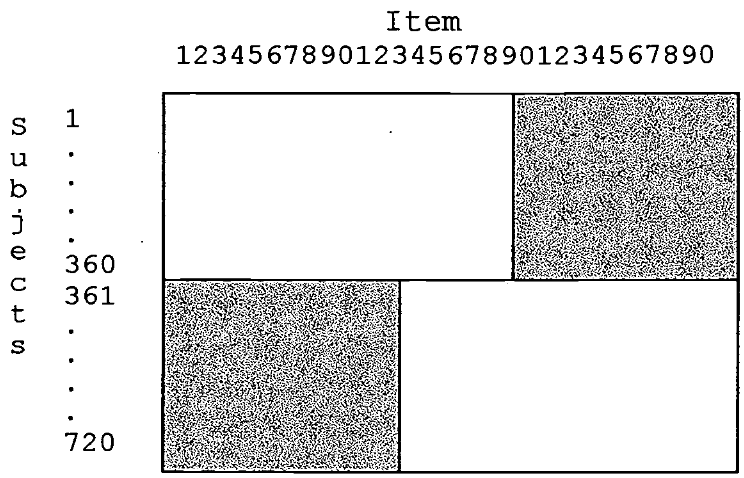

Figure 3. Example of linking of two forms, each with 360 observations and 18 items. Six items link the two forms.

Table 1.

Item difficulties of six linking items on two pretest forms.

Item	D_A	D_B	$D = D_A - D_B$
13	-0.39	-0.54	-0.39
14	-0.70	-0.27	-0.43
15	-0.01	-0.55	-0.57
16	-0.83	1.66	-0.83
17	-1.48	-1.18	-0.30
18	-1.48	-1.18	-0.30
Mean			-0.47
Std Dev			0.21

Table 2.

Linking of two forms, each with eighteen items and having six items in common.

Item	D_A	D_B	D_B + D	$(D_A+D_B+D)/2$	D_C	D_C-Mean D_C
1	0.19				0.19	0.32
2	1.41				1.41	1.57
3	2.08				2.08	2.23
4	-1.82				-1.82	-1.66
5	0.16				0.16	0.32
6	0.25				0.25	0.41
7	1.24				1.24	1.40
8	0.91				0.91	1.07
9	1.16				1.16	1.31
10	1.16				1.19	1.31
11	-1.32				-1.32	-1.17
12	-1.64				-1.64	-1.48
13	-0.93	-0.54	-1.01	-0.97	-0.97	-0.81
14	-0.70	-0.27	-0.74	-0.72	-0.72	-0.56
15	-0.01	0.55	0.09	0.04	0.04	0.19
16	0.83	1.66	1.19	1.01	1.01	1.17
17	-1.48	-1.18	-1.65	-1.56	-1.56	-1.41
18	-1.48	-1.18	-1.65	-1.56	-1.56	-1.41
19		-0.14	-0.61		-0.61	-0.46

Table 2. (Continued)

Linking of two forms, each with eighteen items and having six items in common.

Item	D_A	D_B	D_B + D	$(D_A+D_B+D)/2$	D_C	D_C-Mean D_C
20		1.32	0.85		0.85	1.01
21		-0.69	-1.15		-1.15	-1.00
22		-1.00	-1.47		-1.47	-1.32
23		-0.54	-1.01		-1.01	-0.85
24		-0.02	-0.49		-0.49	-0.33
25		0.77	0.30		0.30	0.46
26		1.91	1.44		1.44	1.59
27		-0.84	-1.31		-1.31	-1.15
28		-0.84	-1.31		-1.31	-1.15
29		-0.40	-0.87		-0.87	-0.72
30		1.43	0.96		0.96	1.12
Mean	0.00	0.00	-0.47	-0.63	-0.16	0.00
Std Dev	1.27	1.02	1.02	1.00	1.18	1.18

Table 3.

Comparison of item difficulties.

Common Link		Entire Data Set	
Item	Difficulty	Item	Difficulty
4	-1.66	4	-1.75
12	-1.48	12	-1.56
18	-1.41	18	-1.39
17	-1.41	17	-1.39
22	-1.32	11	-1.23
11	-1.17	22	-1.23
28	-1.15	28	-1.08
27	-1.15	27	-1.08
21	-1.00	21	-0.94
23	-0.85	13	-0.81
13	-0.81	23	-0.81
29	-0.72	29	-0.69
14	-0.56	14	-0.57
19	-0.46	19	-0.45
24	-0.33	24	-0.34
15	0.19	15	0.16
1	0.32	25	0.35
5	0.32	1	0.35
6	0.41	5	0.35
25	0.46	6	0.44

Table 3. (Continued)

Comparison of item difficulties.

Common Link		Entire Data Set	
Item	Difficulty	Item	Difficulty
20	1.01	20	0.80
8	1.07	30	0.89
30	1.12	16	1.07
16	1.17	8	1.16
9	1.31	26	1.25
10	1.31	9	1.43
7	1.40	10	1.43
2	1.57	7	1.52
26	1.59	2	1.70
3	2.23	3	2.42

U.S. Department of Education
Office of Educational Research and Improvement (OERI)
National Library of Education (NLE)
Educational Resources Information Center (ERIC)

ERIC®

TM033699

REPRODUCTION RELEASE

(Specific Document)

I. DOCUMENT IDENTIFICATION:

Title: Item banks: What, where, why and how	
Author(s): Ric M. Anzaldua	
Corporate Source:	Publication Date:

II. REPRODUCTION RELEASE:

In order to disseminate as widely as possible timely and significant materials of interest to the educational community, documents announced in the monthly abstract journal of the ERIC system, *Resources in Education* (RIE), are usually made available to users in microfiche, reproduced paper copy, and electronic media, and sold through the ERIC Document Reproduction Service (EDRS). Credit is given to the source of each document, and, if reproduction release is granted, one of the following notices is affixed to the document.

If permission is granted to reproduce and disseminate the identified document, please CHECK ONE of the following three options and sign at the bottom of the page.

The sample sticker shown below will be affixed to all Level 1 documents	The sample sticker shown below will be affixed to all Level 2A documents	The sample sticker shown below will be affixed to all Level 2B documents
PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY Sample TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) 1	PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL IN MICROFICHE, AND IN ELECTRONIC MEDIA FOR ERIC COLLECTION SUBSCRIBERS ONLY, HAS BEEN GRANTED BY Sample TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) 2A	PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL IN MICROFICHE ONLY HAS BEEN GRANTED BY Sample TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) 2B
Level 1 ☑	Level 2A ☐	Level 2B ☐
Check here for Level 1 release, permitting reproduction and dissemination in microfiche or other ERIC archival media (e.g., electronic) *and* paper copy.	Check here for Level 2A release, permitting reproduction and dissemination in microfiche and in electronic media for ERIC archival collection subscribers only	Check here for Level 2B release, permitting reproduction and dissemination in microfiche only

Documents will be processed as indicated provided reproduction quality permits.
If permission to reproduce is granted, but no box is checked, documents will be processed at Level 1.

I hereby grant to the Educational Resources Information Center (ERIC) nonexclusive permission to reproduce and disseminate this document as indicated above. Reproduction from the ERIC microfiche or electronic media by persons other than ERIC employees and its system contractors requires permission from the copyright holder. Exception is made for non-profit reproduction by libraries and other service agencies to satisfy information needs of educators in response to discrete inquiries.

Sign here,→ please

Signature: Ricardo Anzaldua	Printed Name/Position/Title: Ric M. Anzaldua, UNT	
Organization/Address: University of North Texas	Telephone: 940-369-7894	FAX:
	E-Mail Address: ranzaldua@unt.edu	Date: 02/16/02

(over)

III. DOCUMENT AVAILABILITY INFORMATION (FROM NON-ERIC SOURCE):

If permission to reproduce is not granted to ERIC, or, if you wish ERIC to cite the availability of the document from another source, please provide the following information regarding the availability of the document. (ERIC will not announce a document unless it is publicly available, and a dependable source can be specified. Contributors should also be aware that ERIC selection criteria are significantly more stringent for documents that cannot be made available through EDRS.)

Publisher/Distributor:
Address:
Price:

IV. REFERRAL OF ERIC TO COPYRIGHT/REPRODUCTION RIGHTS HOLDER:

If the right to grant this reproduction release is held by someone other than the addressee, please provide the appropriate name and address:

Name:
Address:

V. WHERE TO SEND THIS FORM:

Send this form to the following ERIC Clearinghouse:

However, if solicited by the ERIC Facility, or if making an unsolicited contribution to ERIC, return this form (and the document being contributed) to:

ERIC Processing and Reference Facility
4483-A Forbes Boulevard
Lanham, Maryland 20706

Telephone: 301-552-4200
Toll Free: 800-799-3742
FAX: 301-552-4700
e-mail: ericfac@inet.ed.gov
WWW: http://ericfac.piccard.csc.com

EFF-088 (Rev. 2/2000)