

DOCUMENT RESUME

ED 140 007

CE 011 106

AUTHOR Misanchuk, Earl R.
 TITLE Multiple Matrix Sampling Applied to Needs Analysis: An Introduction and Assessment of Efficacy.
 PUB DATE Apr 77
 NOTE 20p.; Paper presented at the Adult Education Research Conference (Minneapolis, Minnesota, April 1977)
 EDRS PRICE MF-\$0.83 HC-\$1.67 Plus Postage.
 DESCRIPTORS Computer Oriented Programs; *Measurement Techniques; *Needs Assessment; Questionnaires; *Sampling; *Statistical Surveys
 IDENTIFIERS *Multiple Matrix Sampling

ABSTRACT

A common approach to needs analysis involves the development of a self-report questionnaire administered to members of the intended audience. Such an approach poses a problem however: Concern for data reliability demands polling as many respondents as possible, while concern for cost-effectiveness dictates minimizing the size of the sample. The solution involves a sub-group of the population responding to all items in the questionnaire, and parameters estimated from the sample data. Yet this approach may not be the most effective and efficient one. Multiple matrix sampling (MMS), a technique whereby both the population of respondents and the universe of items are jointly sampled, may be the answer. Basically, this technique involves the random assignment of some items to a random sample of respondents, with other random samples of respondents receiving different random items. A computer estimates the parameters from the sample data gathered. The study reported herein begins with the pool of data generated by a self-report questionnaire needs analysis and proceeds to estimate those same parameters through the use of MMS. These estimates are then compared to the known values and to corresponding estimates from traditional respondent-sampling techniques. The paper concludes with a summary of the advantages and disadvantages of MMS. (Author/BL)

 * Documents acquired by ERIC include many informal unpublished *
 * materials not available from other sources. ERIC makes every effort *
 * to obtain the best copy available. Nevertheless, items of marginal *
 * reproducibility are often encountered and this affects the quality *
 * of the microfiche and hardcopy reproductions ERIC makes available *
 * via the ERIC Document Reproduction Service (EDRS). EDRS is not *
 * responsible for the quality of the original document. Reproductions *
 * supplied by EDRS are the best that can be made from the original. *

#14 6442

ED140007

MULTIPLE MATRIX SAMPLING APPLIED
TO NEEDS ANALYSIS: AN INTRODUCTION
AND ASSESSMENT OF EFFICACY

Earl R. Misanchuk
Extension Division
University of Saskatchewan

U S DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION
THIS DOCUMENT HAS BEEN REPRO-
DUCED EXACTLY AS RECEIVED FROM
THE PERSON OR ORGANIZATION ORIGIN-
ATING IT. POINTS OF VIEW OR OPINIONS
STATED DO NOT NECESSARILY REPRESENT
OFFICIAL NATIONAL INSTITUTE OF
EDUCATION POSITION OR POLICY

Paper presented at the Adult Education Research
Conference, Minneapolis, April, 1977.

CE 011 106

A typical approach to identifying the training needs of a specified group, especially in cases where a particular set of skills or certain knowledge is required (as for job-related training), is to perform an analysis of the tasks that comprise the job role, then to somehow measure whether or not the individual is capable of performing each task. Any discrepancy between the job requirements and the individual's performance level is, ipso facto, the need for training.

This paper provides a description of how such a task analysis was used recently as a base for a needs analysis. The needs analysis is then used to demonstrate a potential application of multiple matrix sampling, and to discuss the efficacy of the procedure.

Multiple matrix sampling is, as the name implies, a sampling strategy with related statistical procedures which is particularly useful to needs analysts who employ a metric approach. Its worth as a tool for the measurement of needs lies in its capacity for increased precision with decreased expenditure of time and energy (and, therefore, dollars) in comparison to other commonly-used sampling strategies.

When an attempt is made to measure a variable in a number of people, typically a number of parallel, or alternative, measures must be used. Psychometry simply has not progressed far enough to allow very precise measurements of psychological variables using only one measurement. Usually, this limitation means that in order to get one answer, several questions must be asked. For example, to answer the question "How competent is a given individual as a manager/administrator?"--a question, incidentally, which must

be answered under the needs identification model described earlier in order to determine whether the individual has any need for training--it may be necessary to ask a number of questions relating to ability to define staff roles, develop policy, interpret and implement procedures appropriate to pertinent policy or legislation, regulate admissions, develop and foster communication patterns, supervise employees, delegate responsibility, and so on. In fact, to be technically correct, it would probably be necessary to ask several sub-questions to answer each of the questions just identified.

A typical means to answer a question like the example above is to administer a questionnaire or test of some kind to the individual whose needs are to be assessed. (The exact form of the data collection technique is less important than the fact that the data so collected can be adequately represented by a score which is the sum of the individual's responses to a set of questions.)

Unfortunately, within the real-world constraints of time and money, while it is often possible to ask each person about his or her needs as an individual, it is seldom possible to act upon those needs for each individual, as an individual. In other words, when the needs of 50 or 100 manager/administrators are to be assessed, they are typically all asked the same questions, then their answers pooled to obtain a profile of the so-called "average" manager/administrator. The learning experiences are then designed on the basis of the "average" needs thus derived.

It is in just these situations, where the focus is on a group of people rather than on an individual, that multiple matrix sampling displays its power.

Sampling Strategies

Usually the way in which sampling is done to gather information depends on how many potential respondents there are. If there were only 30 people in the population of concern, it is probable that an attempt would be made to gather information from each and every one of them. This strategy is called census data-gathering. On the other hand, if there were 3,000 people in the population of concern, randomly sampling say, 100 or 300 or 1,000 people (depending on the return rate we expected) might be done in order to generalize their responses to the population. This strategy is called respondent sampling. Somewhat less comprehensible in the context of needs identification (albeit quite comprehensible in the context of developing standardized test items) is the notion of item sampling, which involves randomly sampling items from the total item pool (or universe) and administering them to all respondents in the population.

Multiple matrix sampling can be conceptualized as a combination of respondent sampling and item sampling. The application of multiple matrix sampling simply involves administering random samples of items (which constitute sub-tests) to random samples of respondents, usually in such a way that every item in the universe is responded to by some respondents, and every respondent responds to some items. Thus (usually) all sub-tests are responded to, and all respondents respond, but no respondent responds to all sub-tests, nor do any sub-tests get responded to by all respondents.

Incidentally, it is worth noting that neither the sub-tests comprising the "total" test (or item universe) nor the respondent samples comprising

the population need be equal in size to one another. Although it is frequently convenient to have $\underline{N} \div \underline{t}$ respondents (where \underline{N} = the population and \underline{t} = the number of sub-tests) per sample, and to administer $\underline{K} \div \underline{t}$ items (where \underline{K} = the number of items in the universe) to each sample, there is no requirement by the mathematics of multiple matrix sampling that either the item sample size or the respondent sample size be constant.

Advantages and Limitations of Multiple Matrix Sampling

Sufficient work has now been done with multiple matrix sampling in both theoretical and practical applications that certain generalizations can be derived (Shoemaker, 1973). Although by far the largest part of the research to date has been done with data from ability or achievement testing situations, where responses can be coded as correct or incorrect (1 or 0), at least a couple of studies (Loadman, 1972; Pugh, 1971) have used attitudinal or opinion scales as well. When considering the advantages and limitations of multiple matrix sampling outlined below, it should be kept in mind that it would be wise to replicate a number of the achievement/ability testing studies with needs analysis data before generalizing into that domain.

Nevertheless, there are indications that multiple matrix sampling in needs identification would have a number of advantages over other sampling strategies. Among these are:

- (1) the increased feasibility of using a large question pool.

Designers of survey questionnaires, either for needs identification or for other applications, are constantly and acutely aware of the length of the questionnaire. How often are important questions unasked for fear of creating

an inordinately long questionnaire? Multiple matrix sampling may help solve such problems.

(2) reduced standard error of estimate. In comparison to the respondent sampling schemes that are typically applied in needs identification surveys, multiple matrix sampling provides a greater degree of precision, as reflected by a reduced standard error of estimate (when the number of observations is kept constant). This outcome has been demonstrated both theoretically (Lord & Novick, 1968) and practically (e.g., six studies cited by Shoemaker, 1973).

(3) shorter questionnaire completion time. There is a social cost associated with the completion of a questionnaire. Each of a certain number of people must expend a certain amount of time and energy in order to generate the information sought by the researcher. A researcher might tend to think of the length of a questionnaire only in terms of how it influences the rate of return of the questionnaire, but taking a wholistic point of view, it seems almost unethical to waste respondents' time by asking for information that might be obtained more efficiently and more precisely through a modification in methodology.

(4) likely more palatable to respondents. Return rate and ethical considerations aside, it could be argued that researchers owe it to themselves and to future researchers to make the respondents' jobs as palatable as possible. If respondents were to stop providing researchers with data, needs identification would become an even more difficult task than it is now.

Multiple matrix sampling also has some limitations which need to be recognized and taken into account. Among these are:

(1) the procedure applies to groups, not to individuals. This paper began by pointing out that sometimes an individual-oriented sampling scheme is used when all that are wanted are group data. However, there are times when there is a legitimate need for data on individuals, and in those cases, of course, multiple matrix sampling could not be applied.

(2) logistics and computations. The application of a multiple matrix sampling scheme requires little more planning than do more conventional schemes. To be more specific, the item sampling must be rationalized and performed, and alternate forms of the questionnaire must be printed and differentially handled for mailing.

The computations involved in correcting for the fact that some of the respondents answer some of the sub-test are quite lengthy and complex. Fortunately, a computer program is readily available (Shoemaker, 1973) to perform the computations.

(3) the context effect may be overlooked. Multiple matrix sampling assumes that the respondent will answer a question in the same way on a short questionnaire as he or she would answer the same question within the context of a longer questionnaire. Studies by Burton and Remer (1972), Hill (1975), Huck and Bowers (1972), Shoemaker (1970b), Sirotnik (1970), and Sirotnik and Wellington (1974) have indicated that context effects are minimal. However, Novak (1974) and Feldt and Forsyth (1974) both found some context effects. The latter study is especially interesting since it did not find a context effect for an English examination, but did find one for a mathematics examination, indicating a possible context effect by subject-matter interaction. This finding leads to the speculation that different kinds of needs analysis might differentially show evidence of context effects. Further study in this

area of context effects, especially in needs analyses, seems warranted.

(4) assumption of random sampling. Multiple matrix sampling requires that both the items and the respondents be sampled randomly, without replacement. In the application of multiple matrix sampling to need identification, there will be times when this assumption can be met to a greater or lesser degree. Apparently, the validity of application of the procedure is proportional to the degree to which the assumption can be met. Here again, empirical evidence on the robustness of the procedure would be welcome.

The first and last limitation, you will note, refer not so much to the procedure of multiple matrix sampling itself as to the potential for its application in specific situations.

Procedures Involved in Multiple Matrix Sampling

The procedures involved in multiple matrix sampling are quite straight-forward. In practice, sub-tests are created from the item pool by randomly sampling without replacement to generate the number of sub-tests desired, making certain that all items are represented in at least one sub-test. (Technically, there is no requirement that all items be represented, but it makes little sense in the needs identification context to identify an item as worthy of an answer, then not attempt to get an answer for it.) The items may be selected under a stratified random sampling scheme, if such a scheme is appropriate to the situation. Each sub-test is then administered to a randomly-selected (without replacement) subset of the respondent population. The respondent sample, too, may be generated using a stratified approach. A number of rules of thumb are available (Shoemaker, 1973) to guide

the researcher in choosing the most appropriate sampling plan to minimize the standard error of estimate under both normal distribution and skewed distribution conditions.

The application of the multiple matrix sampling procedure will be illustrated through reference to the needs analysis mentioned earlier. Understanding why multiple matrix sampling would be advantageous in this context requires a brief description of the circumstances surrounding the needs analysis.

An Example of a Needs Identification

The population of concern was managers/administrators of social housing in the Province of Saskatchewan--a group of some 260 people. A task analysis of the job role of the social housing manager was done by consulting with a number of managers and their superiors to identify the various tasks and sub-tasks whose performance was required for the satisfactory execution of the manager's job. As well, various job descriptions and training program syllabi from elsewhere were studied. The task analysis broke the managers' functions into seven broad categories, namely (1) board/authority relations, (2) family/relative relations, (3) tenant/resident relations, (4) staff relations, (5) maintenance and security, (6) managerial/administrative knowledge and skills, and (7) finance/office routine. Within each area, a number of task and sub-task descriptions, developed by the project coordinator in conjunction with the manager-consultants, sought to circumscribe the knowledge and skills required by the manager's role. In other words, the resulting task descriptions were such that if, when taken together, they

accurately described what a particular individual were capable of doing, then it was highly likely that that individual would be an effective and efficient manager. If, on the other hand, the individual under consideration could not do most or all of the tasks described, the individual would likely perform poorly as a social housing manager.

Conversion of the task analysis into a diagnostic questionnaire was accomplished by presenting each task and sub-task description individually and asking each respondent to identify his or her competence at the task. In other words,, the respondent was asked to react to each description by checking a box to indicate whether he or she felt (a) sufficiently competent in the task to not need training, (b) sufficiently lacking competence to desire some training, or (c) sufficiently lacking competence to definitely require training, with respect to that task.

In addition to finding out how competent each individual perceived himself or herself to be with respect to each task, the coordinator of the project felt it necessary to also determine how important each task was perceived to be with respect to the manager/administrator's role. Therefore, the questionnaire also asked each respondent to indicate to what extent each of the tasks and sub-tasks described was important in terms of his or her job role. Specifically, the respondents were asked to indicate whether each task description was (a) very important, (b) important, (c) not important, or (d) not applicable to his or her job.

The task analysis identified 204 tasks and sub-tasks associated with the role of the social housing manager. Each task and sub-task description

required two responses, one in terms of the respondent's self-perceived competence vis-à-vis the task, and the other in terms of the importance of the task vis-à-vis the respondent's job. When these items were added to the demographic data and data regarding the preferred format of learning experiences that the project coordinator wanted to collect, the result was a very lengthy questionnaire.

This researcher's lack of sufficient familiarity with multiple matrix sampling at the time of implementation of the needs analysis prohibited the application of the technique. Had greater familiarity obtained, it might have been possible to overcome some of the political and temporal exigencies that prevented the application of multiple matrix sampling. As it was, the census approach was selected, and the lengthy questionnaires were sent to all members of the client population.

Despite some evidence by Burton and Remer (1972), and Champion and Sear (1969), that the return rate of questionnaires is not adversely affected by questionnaire length, it is difficult to believe that some differential self-selection on the part of respondents does not occur in response to such extreme questionnaire length, or, at least, that within a systems view, a social cost of some kind is not exacted from the respondent population. It cannot be known whether or not a shorter questionnaire would have yielded a higher return rate, but it is fairly certain that with the application of a shorter questionnaire, a certain number of person-hours' worth of time could have been made available for other activities.

Nonetheless, the realities were such that the lengthy questionnaire

was used. It yielded a 54% return rate, and the data were used to determine the training needs of social housing managers. Responses to the competence rating were converted to numerical terms (1 = definitely need training, 2 = could be improved, 3 = do not need training), as were the responses indicating the importance of the task to the respondent's job (1 = very important, 2 = important, 3 = not important, 4 = not applicable). Using these item scores, subscale scores were computed for each of the seven major areas identified by the task analysis (i.e., board/authority relations, family/relative relations, tenant/resident relations, staff relations, maintenance and security, managerial/administrative knowledge and skills, and finance/office routine) by summing the responses to the items comprising the subscale.

The Efficacy of Multiple Matrix Sampling--A Preliminary Investigation

The abundance of available research describing the efficacy of multiple matrix sampling in applications to testing situations (e.g., Cook & Stufflebeam, 1967; Hill, 1975; Novak, 1974; Plumlee, 1964; Pugh, 1971; Shoemaker, 1970a; Sirotnik, 1970; Sirotnik & Wellington, 1974), together with the peculiar problems associated with the lengthy questionnaire in the current needs identification study prompted some preliminary investigation into the technique. Specifically, an attempt was made to use the data generated by the census-sampled needs identification in a post-mortem sense to provide an example of and to help gauge the efficacy of the multiple matrix sampling technique. The results were not especially encouraging. However, they do point to several aspects of research on the method that are needed before

widespread application of the multiple matrix sampling technique to needs identification can be recommended.

In post-mortem sampling, one begins with an extant data base, for which parameters are known. One then samples from that data base, pretending that the parameters are unknown, and estimates the parameters from the sample, be it respondent sample or matrix sample. Finally, one compares the estimates of the parameters to their known values.

Using the returned questionnaires as a data base, the post-mortem analysis compared the known parameters for two of the seven subscales with the estimates for those parameters using (a) a 10% respondent sample, (b) a 20% respondent sample, and (c) a multiple matrix sample set.

For each of the 10% and 20% respondent-sampling schemes, four random samples were obtained from the data base, and subscale scores were computed. For the multiple matrix sampling scheme, each of four randomly sampled (without replacement) respondent sets were "administered" a sub-test of randomly selected (without replacement) items (i.e., the responses to the appropriate items were culled from the data base for each sub-set of respondents), and processed with the multiple matrix sampling computer program provided by Shoemaker (1973).

Contrary to expectations, and to previous findings of a number of studies (e.g., Cook & Stufflebeam, 1967; Hill, 1975; Novak, 1974; Plumlee, 1964; Pugh, 1971; Shoemaker, 1970a; Sirotnik, 1970; Sirotnik & Wellington, 1974), in every case the estimate determined by the multiple matrix sampling scheme was more deviant from the known parameter than the corresponding

estimate determined by the respondent sampling scheme (see Table 1).

 Insert Table 1 here.

A couple of possible reasons for the relatively low concordance of the sampling schemes suggest themselves. First, the data base which was used for the post-mortem sampling was relatively small--139 respondents. Although examples of applications of multiple matrix sampling are found in the literature for relatively small respondent groups (e.g., three kindergårten classes used by Shoemaker and Okada, 1970) the studies in which statistics derived through multiple matrix sampling are compared to those derived via other sampling schemes typically range upward from several hundred (e.g., Novak, 1974; Pugh, 1971; Shoemaker, 1972; Sirotnik & Wellington, 1974). Shoemaker (1973) notes that although some of the original equations underpinning the multiple matrix sampling technique were limited to dichotomously-scored items administered to reasonably large (i.e., 500) numbers of respondents, the equations used in his program are ". . . more generally applicable and permit multiple matrix sampling being used with a wider variety of sampling plans [p. 34]." On the other hand, Jaeger (1974) seems to imply that the technique ought only to be used in situations involving more than 300 units of analysis. Perhaps there is a lower practical limit with respect to the number of respondents for the valid application of multiple matrix sampling. Some empirical clarification on this point would be most welcome.

Second, the data base used may have been idiosyncratic. Shoemaker (1973) notes that as the reliability of the distribution of test scores

increases, it becomes increasingly difficult to estimate parameters. The solution he recommends for this problem is obtaining a relatively large number of observations. Thus, this second possible explanation for the results observed leads in the same direction as the first. Since it would usually be expected that needs identification data would be highly reliable (if the task analyses underlying them were well done), this limitation of the multiple matrix sampling procedure might well interfere with its application to all but large-scale needs identification studies.

Implications

The single example described in this paper should not be confused with a rigorous study comparing multiple matrix sampling to other sampling schemes. This study does not undo what has been done before; it does not even cast doubt on earlier studies. Rather, it points out an area of concern for the legitimate application of multiple matrix sampling to needs identification--that of the lower limit for respondent sample size in the context of typically high reliability data.

The technique of multiple matrix sampling appears to hold much promise for needs identification, but prior to its wide-spread application, some questions must be pursued. This paper ends, as do so many others, with a plea for further study--this time on the application of multiple matrix sampling to needs identification.

REFERENCES

- Burton, N., & Remer, R. (1972) Item sampling in questionnaires: Rate of return and context effects. University of Colorado Laboratory of Educational Research Paper No. 57, February, 1972.
- Champion, D. J., & Sear, A. M. (1969) Questionnaire response rate: A methodological analysis. Social Forces, 47, 335-339.
- Cook, D. L., & Stufflebeam, D. L. (1967) Estimating test norms from variable size item and examinee samples. Educational and Psychological Measurement, 27, 601-610.
- Feldt, L. S., & Forsyth, R. A. (1974) An examination of the context effect in item sampling. Journal of Educational Measurement, 11, 73-82.
- Hill, R. K. (1975) Minimizing context effect when using multiple matrix sampling. Paper presented at the Annual Meeting of the National Council on Measurement in Education, Washington, D. C., Mar. 31 - Apr. 2, 1975. (ERIC Document Reproduction Service No. ED 110468)
- Huck, S. W., & Bowers, N. D. (1972) Item difficulty level and sequence effects in multiple-choice achievement tests. Journal of Educational Measurement, 9, 105-111.
- Jaeger, R. M. (1974) Some new developments and discoveries for evaluative analysis. Paper presented at the Annual Meeting of the American Educational Research Association, Chicago, April, 1974. (ERIC Document Reproduction Service No. ED 093904)
- Loadman, W. E. (1972) An inquiry concerning the use of item sampling as a method to reduce testing time. Indiana University, Educational Research and Evaluation Laboratory, Report No. 5.
- Lord, F. M., & Novick, M. R. (1968) Statistical theories of mental test scores. Reading, Mass.: Addison-Wesley.
- Novak, C. D. (1974) An empirical investigation of multiple matrix sampling in an elementary school setting. Paper presented at the Annual Meeting of

~~the~~ American Educational Research Association, Chicago, April, 1974. (ERIC Document Reproduction Service No. ED 093997)

Plumlee, L. B. (1964) Estimating means and standard deviations from partial data--an empirical check on Lord's item sampling technique. Educational and Psychological Measurement, 24, 623-630.

Pugh, R. C. (1971) Empirical evidence on the application of Lord's sampling technique to Likert items. The Journal of Experimental Education, 39, 54-56.

Shoemaker, D. M. (1970a) Allocation of items and examinees in estimating a norm distribution by item sampling. Journal of Educational Measurement, 7, 123-128.

Shoemaker, D. M. (1970b) Test statistics as a function of item arrangement. The Journal of Experimental Education, 39(1), 85-88.

Shoemaker, D. M. (1972) A note on allocating items to subtests in multiple matrix sampling. Los Alamitos, Calif.: Southwest Regional Laboratory for Educational Research and Development. Report No. SWRL-TM-3-72-05, June, 1972. (ERIC Document Reproduction Service No. ED 110519)

Shoemaker, D. M. (1973) Principles and procedures of multiple matrix sampling. Cambridge, Mass.: Ballinger.

Shoemaker, D. M., & Okada, M. The communication skills program and concomitant spelling proficiency. Los Alamitos, Calif.: Southwest Regional Laboratory for Educational Research and Development Technical Note.

Sirotnik, K. (1970) An investigation of the context effect in matrix sampling. Journal of Educational Measurement, 7, 199-207.

Sirotnik, K., & Wellington, R. (1974) Scrambling content in achievement testing: An application of multiple matrix sampling in experimental design. Journal of Educational Measurement, 11, 179-188.

Table 1

Efficiency of Estimating Two Subscale Means Using Various Sampling Strategies

Sampling Scheme	Subscale A	Subscale B
Census	37.295	82.561
20% Respondent Sample		
#1	32.655	86.862
#2	36.821	62.571
#3	38.821	82.821
#4	41.536	77.643
Average	<u>37.458</u>	<u>77.532</u>
10% Respondent Sample		
#1	38.429	79.643
#2	35.929	79.286
#3	38.000	83.286
#4	40.714	70.786
Average	<u>38.268</u>	<u>78.250</u>
Multiple Matrix Sampling ¹		
#1	29.405	105.360
#2	28.238	105.671
#3	28.500	107.954
#4	28.973	107.826
Average	<u>28.779</u>	<u>106.703</u>

¹ For subscale A, five subtests of 10 items and one of 9 items were "administered" to five groups of 23 respondents and 1 group of 24 respondents, respectively; for Subscale B, four subtests of 4 items were "administered" to three groups of 35 and one group of 34 respondents, respectively.

ADDITIONAL REFERENCES

Bunda, M. A. (1973) An investigation of an extension of item sampling which yields individual scores. Journal of Educational Measurement, 10, 117-130.

Forsyth, R. A. (1976) Estimating means via multiple matrix sampling: A note on the effects of selected data base characteristics. Educational and Psychological Measurement, 36, 275-282.

Shoemaker, D. M. (1973) A note on allocating items to subtests in multiple matrix sampling and approximating standard errors of estimate with the jack-knife. Journal of Educational Measurement, 10, 211-219.

Shoemaker, D. M., & Knapp, T. R. (1974) A note on terminology and notation in matrix sampling. Journal of Educational Measurement, 11, 59-61.

Sirotnik, K. (1973) Some notes on the estimation formulas in matrix sampling. Los Angeles: The Instructional Objectives Exchange. Technical Paper No. 9.

Sirotnik, K. (1974) Introduction to matrix sampling for the practitioner. In W. J. Popham (Ed.), Evaluation in education: Current applications. Berkeley, Calif.: McCutchan.