

CHALLENGES IN THE DESIGN AND IMPLEMENTATION OF THE ADULT EDUCATION LITERACY STUDY

Thomas Krenzke and Leyla Mohadjer, Westat;
Thomas Krenzke, Westat, 1650 Research Boulevard, Rockville, Maryland 20850

Key Words: Sample Design, Sampling Frame, Adult Education, Literacy Assessment

1. Introduction

Survey designs and sampling techniques are built on probability theory with the aim of producing reliable statistics for the reference population. At the same time the designs need to be practically feasible, and should satisfy the survey objectives at minimum cost. In particular, Kish (1965) states that “the statistician must sacrifice some theoretically available efficiency for the sake of practicability”. Achieving these guiding principles of sample design can be quite challenging in surveys where logistical complexities create situations that are far from optimal with respect to the survey objectives. In this paper, we illustrate the challenges in sample design and selection of a large-scale multi-stage sample.

The Adult Education Literacy (AEL) study, sponsored by the Office of Vocational and Adult Education (OVAE), and the National Center for Education Statistics (NCES), allows, for the first time, a reliable comparison of those enrolled in adult education programs with the general population. In addition, the 2003 AEL study is the first nationally representative sample that included a large enough sample of persons enrolled in adult education programs to permit a separate analysis of their literacy skills. Assessments in AEL were conducted in Spanish and English to compare literacy outcomes in both languages for Spanish speakers. The literacy assessment was individually administered to the adult education learners.

The eligible adult education programs for AEL consisted of courses in general adult education, including Adult Basic Education (ABE), Adult Secondary Education (ASE), and English as a Second Language (ESL). Courses tailored to a particular job are ineligible. The courses are generally administered in Local Education Agencies (LEAs), community colleges, community-based organizations, correctional facilities, and other types of locations, such as four-year colleges and libraries.

In the first chapter of Cochran (1977), he discusses the principal steps in a sample survey. We discuss the challenges of maintaining a probability design during some principal steps of the AEL sampling and data collection process, while facing

several deterrents to a theoretically optimal design. Principal steps in the AEL survey design are incorporated in the paper’s outline as follows:

1. The importance of conducting a pretest (discussed in Section 2);
2. The population to be sampled (discussed in Section 3);
3. The sample design (Section 4);
4. The frame (Section 5); and
5. The selection of the sample (Section 6).

Lastly, in Section 7, we include a short discussion about data release in the presence of a highly skewed population of adult education programs.

2. The Pretest

The 2003 AEL study was preceded by the 2002 Level 1 Study, which also served as a pretest to AEL. Level 1 was a small-scale study that consisted of a household and program component and was geared toward building analytical literacy scoring models for persons in the lowest end of the literacy scale. It involved over 900 assessed learners from 30 programs in five states with large adult education enrollment. Although a probability design was not required, the Level 1 study was a great opportunity for a trial run at testing a probability-based sample design. The pre-test of sampling procedures was invaluable for establishing sample design improvements for the larger AEL study. The AEL study would not have been as successful without the diligent attempts made at random sampling techniques in the Level 1 study. The benefits of the pretest are discussed throughout the remainder of this paper.

3. Target Population and Sample Size Requirements

The eligible programs were part of the National Reporting System (NRS) and/or those federally funded under Title II (i.e., Adult Education and Family Literacy, Title II of the Workforce Investment Act of 1988, D.L. 105-220). Learners were eligible for the AEL study if they were enrolled in ESL, ABE, and ASE courses at the time of sampling (spring of 2003) and had at least 12 hours of instructions since July 1, 2002. Learners enrolled in Beginning Literacy ESL course were ineligible for the assessment since the

English-based assessment was not designed for those who could not speak English at all.

There were 5,000 learner assessments required for the AEL sample, including 2,000 English assessments administered to Non-Hispanic adults, 2,000 English assessments administered to Hispanic adults, and 1,000 Spanish assessments administered to Hispanic adults (i.e., one-third of the Hispanic sample were to receive a Spanish assessment).

4. Sample Design

The sample design for the learner component was a multi-stage probability sample of adult education learners within the 50 states and the District of Columbia. The sampling stages included the selection of a sample of programs, the selection of sites within each program, the selection of classes or groups of classes within each site, and the selection of adult learners from each class. The following sections discuss the issues faced in the sample design, and how the sampling requirements were successfully reached.

4.1 Highly Transient Population

4.1.1 Level 1 Experience

The adult education population is highly transient. Within the ESL population, many learners are refugees and other immigrants that are trying to get settled in the U.S. by learning the English language in order to increase their prospects of landing jobs. They tend to move more frequently. The effect of the transient population was experienced during attempts at sampling learners in Level 1.

The sample design for the program component of Level 1 included three stages of sampling including selection of programs, sites, and learners within the selected sites. We chose a three-stage design for Level 1 after we confirmed that sites can provide listings of learners in all their classes. By selecting learners within sites, we improved the precision of the outcome statistics by reducing the clustering of students within classes. The learner lists were gathered in January/February of 2002 and learners were selected in February in time for the training of the Exercise Administrators (EAs)¹. After training, the EAs scheduled the interviews using the pre-sampled list of learners during the first week of March. The time gathering the lists of learners and administering the assessments proved to be too long for a study involving

adult learners as many learners moved or withdrew from class during this time period. The resulting response rate from the probability sample component of Level 1 was only 42.5 percent, with a high level of attrition coming from learners who withdrew (or were unlocatable) at the time of interview.

4.1.2 Sampling in Waves

Having experienced the challenges of sampling such a transient population in Level 1, we put a great deal on emphasis on developing procedures to improve the design of the AEL study. The main challenge was to establish a design and a selection process that arrive at the most complete (and unduplicated) list of learners and to select and administer the assessment with the least amount of attrition due to withdrawal. During Level 1, it became clear that course listings were fairly stable by the time learners registered in December or January. To reduce the time elapsed between the listing of learners and the selection of learners, two classes or two groups of classes were selected, and each class was randomly assigned to one of two waves. The first wave of sampling learners occurred right after the training of EAs in March, and the second wave occurred after a second training of EAs in April. The EAs were trained to construct the listing of learners and to select the sample. The scheduling, interviews and assessments began shortly after the sample was selected, thus minimizing the time between list construction and assessments. The scheduling of sampling activities in Level 1 and AEL is summarized in Table 1.

Table 1. Schedule of sampling tasks for Level 1 and AEL

Task	Level 1 dates	AEL dates
Sampling frame of programs	Summer	Summer
Program selection	September	September
Sampling frame of sites	October	October
Site selection	November	November
Sampling frame of classes	Not Applicable	December
Class selection	Not Applicable	February
List of learners	February	Wave 1 in March, Wave 2 in April
Sample learners	Spring	Wave 1 in March, Wave 2 in April

Sampling in waves was very successful as the completion rate for learners climbed from 42.5 percent

¹ EAs were program staff that administered the assessments.

in Level 1 to 76.8 percent in AEL, as shown in Table 2. The increase in completion rate at the site level was due to increased staffing and intensified follow-up procedures.

Table 2. Completion rates for Level 1 and AEL

Task	Level 1	AEL
Programs	92.7%	87.0%
Sites	68.0 ¹	84.7
Classes	--	99.7
Learners	42.5	76.8

¹ Includes original sites only

With sampling in waves, if in structured semester settings, sample bias could be introduced, for instance, through selecting learners in the Northeast from a beginning ESL class in early spring, while learners in the West from a beginning ESL class in late spring after the learners had been in class a month or two. However, in the adult education setting, the majority of classes are offered through open enrollment, and thus the learners are free to enter the course at any time. Another reason that sample bias from wave sampling is negligible is that new courses can start at any time in the springtime, such that learners sampled in Wave 2 (April) could be sampled toward the beginning of a new class session, or toward the end of a class session.

4.2 Predicting Response Rates

As discussed in Section 4.1, the Level 1 experience helped to make two significant improvements in the design: 1) establish a four-stage sample, and 2) select the sample of learners in waves. Although we knew the improved design will increase the response rates, we did not know how much of an increase we should expect. It is critical to make an accurate (as close to accurate as possible, of course) prediction of response rate when designing and selecting samples. Inflated predictions of response rates will result in a shortfall in the outcome sample, and the planned analyses will then need to be curtailed. On the

other hand, if response rates are underestimated, there will be a surplus, and the cost per additional assessment could overrun the budget. The predicted response rates are also important for scheduling, and for staffing plans.

Using all relevant information available to us at the time of the sample design, we predicted a response rate of 60 percent among the learners for AEL. As seen in Table 2, we experienced a far greater response rate than expected, which yielded conflicting results. We exceeded the AEL total sample size requirement by 20 percent and met all sample size goals for Hispanics. The larger sample size, of course, created a richer and more powerful data set for the analyst. At the same time the larger than expected sample size increased costs of assessing and scoring beyond what was planned for AEL.

4.3 Oversampling Hispanic Learners

The design needed to include an oversampling of Hispanics, given the disproportionate number of Hispanics required in the sample (60%), compared to the proportion of Hispanics in the general adult education population (about 42%). Developing the oversampling design became a major challenge since there were limited race/ethnicity data available at each sampling stage. To arrive at the oversampling rates, we used the experience from the Level 1 study that was conducted in five states with large enrollments. We assumed that 80 percent of the learners in ESL classes are Hispanic, and 20 percent in ABE/ASE classes are Hispanic. Considering the sample size requirements and the ethnicity distribution assumptions, we designed the sample so that two-thirds of the sample came from ESL courses. To accomplish this, programs were randomly assigned to two sample types: ESL and ABE/ASE, with about two-thirds of the programs designated for ESL. The upper portion of the chart in Figure 1 illustrates the assignment of programs to the ESL and ABE/ASE sample types. The ESL sample contained a sample of ESL learners only, while the ABE/ASE sample contained a sample of ABE/ASE learners only. This design, resulted in achieving over 3,000 Hispanic adults in the sample (including the 20% increase in sample yield, as discussed earlier).

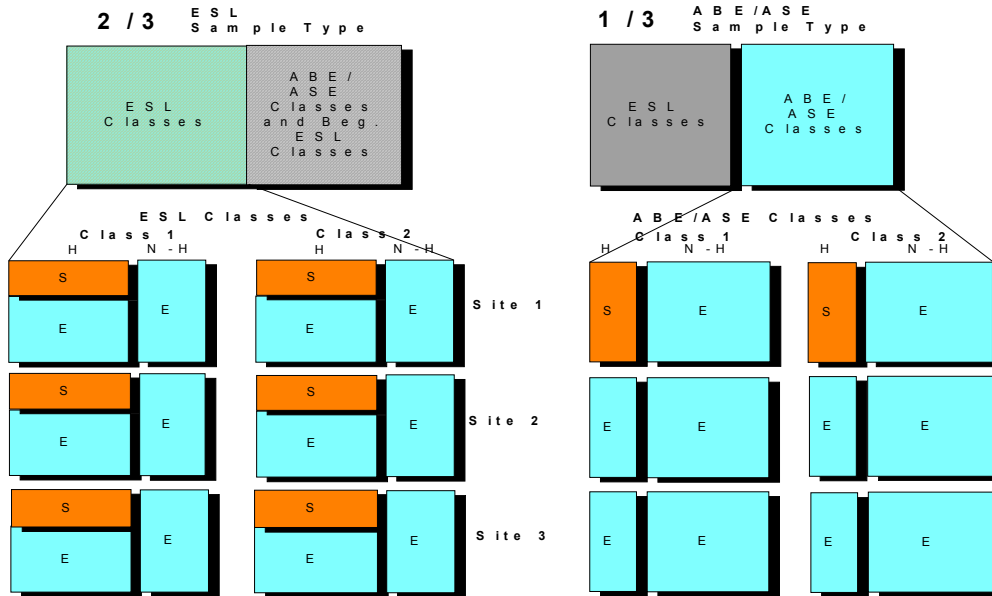


Figure 1. Illustration of Oversampling and Spanish and English Assessments Procedures

4.4 Handling Learners in Multiple Classes

The sample type assignment approach not only became the basis for the oversampling design for Hispanics, but also handled multiple chances of selection for learners in more than one class type, where class type was either ESL or ABE/ASE. In addition, for learners in more than one ESL class, or more than one ABE/ASE class, we minimized the duplicate chance of selection by keeping the learners in Wave 1 but removing the duplicates on the Wave 2 list of learners.

4.5 Administering Spanish and English Assessments

It was necessary to implement a process that randomly assigned Spanish and English assessments to Hispanic adults. In addition, it was necessary to limit the number of bilingual interviewers required for data collection. Since the ABE/ASE population contained only about 20 percent Hispanic adults, we clustered the Spanish and English assessments for Hispanics for ABE/ASE sites. That is one-third of the ABE/ASE sites were randomly selected to administer the Spanish assessment only to Hispanic learners. The clustering approach reduced the number of required bilingual interviewers. In ESL sites, to reduce the clustering of Spanish assessments, one of every three Hispanic learners was given the Spanish assessment. This design, helped to meet the goal of 1,000 representative

Hispanic adults assessed in Spanish (including the 20% overage in sample size).

The design aspects of administering the Spanish and English assessments are illustrated in the bottom portion of Figure 1. The illustration depicts an example of three sites and two classes within each site. Within each class, the figure shows the assignment of Spanish (S) and English (E) assessments for Hispanics (H) and for NonHispanics (N-H).

Operationally, the successful handling of materials under the requirement of random sampling was a great challenge. Equal distributions among the twelve individual English assessment booklet types, and among the four Spanish assessment booklet types were to be attained. Exercise Administrators were trained on selecting the appropriate booklet for the appropriate ethnicity of the selected person in a manner that retained the randomness in the booklet assignment. In addition, Spanish and English booklets had to be shipped in efficient numbers, that is, without knowing precisely the number of Hispanic adults, the number of Spanish and English booklets was assigned to each class so that a shortfall of booklets did not occur, and so that excess booklets would not become a resource problem. Lastly, it was necessary to maintain the distributions of the booklets across demographic and geographic variables, as well as programs types through the random assignment.

5. Sampling Frames

The survey frame is the data source used for stratifying, sampling and locating the selected units. It is therefore, a vital component of any survey, and is instrumental in establishing a solid ground on which the entire survey is constructed. Frame creation is a more critical step for multi-stage samples, such as AEL, since frames are required for each stage of sampling.

Statistics for the adult education system have recently become more available through the efforts of the OVAE office, which has implemented a reporting strategy known as the National Reporting System (NRS). The system aggregates reports of learner demographics, course completion rates, etc from geographic location of the course, to the program level, then to the state level. This NRS effort has begun to help give the population a structure to its data systems. This has helped to a certain extent, to pave the way for the AEL study, especially in constructing the sampling frames. The NRS data was instrumental in creating accurate frames for various stages of the AEL sample.

5.1 Sources and Scope

Four sets of sampling frames were needed for the AEL study. To construct the sampling frames for AEL, it was necessary to be in contact with 1) the state adult education departments (all 50 states and the District of Columbia), 2) the selected adult education programs, 3) sites within the selected programs, and 4) the classes (teachers) within selected sites. Figure 2 provides a chart of the sampling frames and their sources and scope.

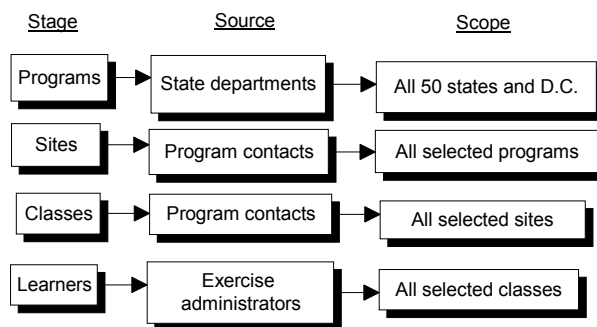


Figure 2. Sampling frame sources and scope

5.2 Sampling Frame Issues

The construction of the frames required collection of complete, up-to-date, and unduplicated data. It was also necessary to ensure that the data was consistent with data from other sources. For example, to

ensure that the program enrollment reflected unduplicated totals², the enrollment data were checked against NRS aggregate total enrollments. In the same manner, when site listings were received from the program contact, the data underwent several quality checks to maintain consistencies between the data provided by the state (about the program), and the aggregated site data. All non-negligible discrepancies were followed-up with the programs. Similar consistency checks were conducted on the class lists and the learner lists.

6. Sample Selection and Operational Constraints

Balancing the statistical efficiency and costs related to practical and operational problems in conducting a survey is always a challenging task. Operational constraints heavily influence the final design and selection of surveys, and result in deviations from an efficient statistical design.

6.1 Small Classes and Other Class Sampling Issues

Several frame and selection issues arose when compiling the class frame. Most of the classes were open enrollment, that is, the learners could attend or enroll at any time after the start of the session. Other classes varied in session length. To keep these types of classes in the sampling population, the target population was defined as learners enrolled at the time of sample selection. In addition, small classes (with less than 15 students) and one-on-one tutoring groups were grouped until there were 15 or more learners in a group. A group of classes was referred to as a pseudoclass. The formation of pseudoclasses clustered the learners, thereby reducing data collection costs significantly.

6.2 Small Sites

Since it was not cost-efficient to send EAs to small sites, we carefully determined a cutoff size that would allow us to exclude small sites from the study. In several instances, classes were administered in satellite locations; in churches for example. The cutoff threshold was set to reduce the high cost of reaching only a few learners in these sites, and at the same time reduce the impact of undercoverage resulting from excluding the satellite sites. The exclusion of satellite sites resulted in an undercoverage rate of about three percent that was

² That is, learners counted only once even though they may have been enrolled in more than one class.

corrected through adjustments to control totals during the weighting process.

6.3 Large Programs

A major operational constraint for AEL was the limit imposed on the number of units to be selected at each stage of sampling. Within each program, no more than three sites were to be selected. Within each site, no more than two pseudo classes were to be selected, and lastly, within each site, no more than 40 learners were to be selected.

In the AEL population there are a small number of very large programs that contribute a high percentage of the population of adult education students. These large programs deserved more sampled sites, classes, and learners than the constraints allowed. The operational constraints made the design sub-optimal, since it caused variation in the resulting selection probabilities, which increased the variances relating to the resulting estimates. Although procedures were taken to limit the variation in the weights, the coefficient of variation of the trimmed weights exceeded 100 percent.

6.4 Laboratories

Through the Level 1 study we learned the importance of selecting the appropriate time intervals for administering the assessments. The adult programs include education laboratories, which are offered to all adult education students. Laboratories are open to registered students, who could work during assigned hours in a week. To sample learners, the week was divided into hourly intervals and two intervals were randomly selected within the week. The learner sampling list was comprised of all persons that attended the selected time interval. A random sample was selected from the list.

7. Data Release

NCES and OVAE are obligated by law to reduce the risk of disclosing the identity of survey respondents

to data intruders through the release of survey data. However, the distribution of the adult education programs is highly skewed, with a very small number of outlier programs dominating the size distributions in terms of budget, enrollment, staffing, and number of sites. These outlier programs are highly visible, and create a high risk of data disclosure if their data is included in a public use file for the learners³. Currently, NCES and OVAE have decided to create a public-use file of the learner data with some minimal program data. In addition, a restricted use microdata file and a data tool for public use will be available for data users.

8. Summary

With the AEL study being the first nationally representative sample of literacy assessments within adult education programs, we discussed several challenges and issues in its sample design, sample frame construction, sample selection under operational constraints, and data release. Of note was the great learning experience from the Level 1 study, and the improvements made to the AEL sample design. The study design and handling of issues was highly successful in meeting the sample requirements, although the data collection was completed with a 20 percent surplus of assessments incurring additional costs to the survey.

9. Acknowledgements

The authors want to acknowledge the efforts of Sylvia Dohrmann, Laura Alvarez-Rojas, Eugene Brown who worked on various facets of the project that are mentioned in this paper.

10. References

- Cochran, W. G. (1977). *Sampling Techniques*. 3rd edition. John Wiley & Sons, New York.
- Kish, L (1965). *Survey Sampling*. John Wiley & Sons, New York.

³ Prior to the literacy assessments, a random sample of adult education programs was selected and asked to complete a program questionnaire.