

Part B: Collection of Information Involving Statistical Methods

The U.S. Department of Labor (DOL) has contracted with MDRC to conduct a rigorous evaluation of the 2011 YouthBuild program funded by DOL and the Corporation for National and Community Service (CNCS). The evaluation includes an implementation component, an impact component and a cost-effectiveness component. This data collection request seeks clearance for data to be collected during site visits to YouthBuild sites participating in the implementation component of the evaluation. In the future, DOL will submit a separate OMB-PRA clearance request for a longitudinal series of participant surveys that will be administered as part of the impact component of the evaluation. It is understood that OMB clearance of the site visit data collection protocols does not constitute clearance of the future participant follow-up surveys.

The implementation component of the evaluation will explore the design and operations of the YouthBuild programs, the perceptions and experiences of the participating youth, and the local context in which each program operates. The implementation and impact study are linked in the following ways. First, the findings from the implementation study will describe the program we are evaluating. What is YouthBuild, as it actually operates? How does the program vary across the participating sites, in terms of key components, emphases, services offered, and youth served? Second, the findings will be used to assess how impacts vary with program features. It is likely that there will be variation across the participating programs in strength and fidelity to the YouthBuild model. It will thus be important to identify whether this is true and, if so, to estimate impacts for high versus low fidelity programs. In addition, there is interest in assessing how impacts vary with various program features, such as the length of mental toughness orientation and the focus on post-secondary education. Data from the implementation

analysis will be used to estimate, using a hierarchical linear model, how impacts vary with a selected number of program features. The interviews will take place after sites have conducted random assignment (from summer 2012 to spring 2013), in order to assess the programs while they are serving youth who are in the study.

B. Sampling for Site Visit Data Collection Instruments

Data collection during site visits will include: interviews with program staff; focus groups or individual interviews with YouthBuild participants; observation and completion of the Classroom Observation Checklist by evaluation team members; and completion of a Cost Data Collection Worksheet.

1. Site Selection for the Site Visit Data Collection Instruments

DOL awarded grants to 74 YouthBuild programs in May 2011. According to program proposals, these 74 DOL grantees planned to serve a total of 3,304 youth this year. When determining which of the grantees to include in the site visits, three programs were dropped from the selection universe because discussions with site staff indicated that youth assigned to the control group were highly likely to receive services that were very similar to YouthBuild. These three programs planned to serve a total of 133 youth. The other 71 programs planned to serve a total of 3,171. Given that the excluded sites account for only 4.1 percent of expected enrollment among DOL-funded sites, our ability to extrapolate the study findings to the universe of DOL-funded sites is not compromised.

Sixty programs from the universe of 71 programs were selected for the impact component of the evaluation, using probability-proportional-to-size sampling. Each program had a probability of selection that was equal to its expected enrollment over a given program year. This method gives each YouthBuild slot (or youth served) an equal chance of being selected for the

evaluation, meaning that the resulting sample of youth who enroll in the study should be representative of youth served by these programs. Once a program is selected for the evaluation, all youth who enroll at that program between August 2011 and December 2012 will be included in the research sample. As mentioned earlier, a later submission will discuss sampling and analysis for the youth follow-up surveys.

In deciding on the total number of DOL-funded programs to include in the impact component of the evaluation, we attempted to balance the objectives of 1) maximizing the representativeness of the sample and the statistical power of the impact analysis and 2) ensuring high quality implementation of program enrollment and random assignment procedures. On the one hand, maximizing the representativeness of the sample and the sample size would call for including all grantees in the study. However, substantial resources are required to work with a study site to: a) secure staff buy-in and cooperation with the study; b) train staff on intake and random assignment procedures; and c) monitor random assignment. Given a fixed budget, the team determined that 60 DOL-funded programs should be selected for the evaluation. The quality of the enrollment process (and, potentially, the integrity of random assignment) would suffer if more than 60 sites were included. Given the expected number of youth served per program, 60 programs was deemed adequate to generate a sample size that would provide reasonable minimum detectable effects on key outcomes of interest.

There are 40 programs that comprise the universe of programs that receive CNCS but not DOL funding in 2011, and 23 of these programs received grants of \$95,000 or more from CNCS. These 23 programs were selected for the impact component of the evaluation.

As mentioned earlier, a key use of these data will be to describe the operations and features of the programs that are included in the evaluation, in order to interpret later impact

findings. However, there may be instances when the team will seek to generalize the findings to the broader population of YouthBuild sites. Because the 23 CNCS programs were not drawn randomly from the universe of CNCS only-funded programs, no inferences will be made from these 23 CNCS programs to the larger group of 40 CNCS-only funded programs. However, given that the DOL-funded sites in the study were selected randomly from the full universe of DOL-funded sites, findings from the process study can be generalized to the full universe of DOL-funded YouthBuild programs. While much of the process analysis will involve qualitative data, any efforts to use quantitative data to generalize to the larger sample of DOL-funded sites will weight the data accordingly, to account for larger sites' greater probability of being selected for the study. The weight given to each site will equal the inverse probability of selection into the study. With N grantees in the selection universe, and m_i equal to planned enrollment at program i , a given program's weight is equal to:

$$w_i = 1/p_i, \text{ where } p_i = m_i / \sum_{i=1}^N m_i.$$

MDRC uses the statistical software SAS, which offer a range of procedures to analyze data with various survey designs. PROC SURVEYMEANS, for example, uses the Taylor series expansion for variance estimation.

2. Procedures for the Collection of Information

The protocols are semi-structured scripts for site visitors to conduct their data collection while on site. The loosely structured scripts allow flexibility to tailor the composition of questions for each person as appropriate. These data, including information on the cost of provided YouthBuild services, will be collected via interviews with a range of individuals at each YouthBuild program site, including YouthBuild program administrators, recruitment and

intake staff, and case managers/counselor. Data from the focus groups will be obtained by engaging five-to-six YouthBuild participants at each site in a group discussion about the program. Cost data for program operations outlined on the Cost Data Collection Form will be collected for the cost-effectiveness analysis. Finally, site visitors from the evaluation team will conduct structured observations of YouthBuild education and training activities, using the Classroom Observation Checklist.

3. Methods to Maximize Response Rates and Data Reliability

Both DOL and CNCS grantees are required to participate in the evaluation as a requirement of their grant awards. Thus, we expect that all sites selected for the impact component of the evaluation will agree to participate and will agree to host the evaluation team during the proposed site visits, providing opportunities to the evaluation team to collect the requested information. The evaluation team will work with site staff to plan an efficient, yet productive, site visit that is most convenient for the program. Interviews with staff and others with connections to YouthBuild will last approximately one hour only and will be scheduled to accommodate the best interest of the individuals. Focus group discussions with YouthBuild participants will last one hour.

4. Tests of Procedures or Methods

In addition to having used similar scripts for prior research projects and having found them to be reliable and an efficient way in which to collect the necessary information, the evaluation team conducted a pilot site visit to one YouthBuild grantee. This visit was conducted by a senior member of the research team, and was used to test the protocols for their reliability and clarity and to measure the time required to collect the information. The data from this site visit will be included in the final analysis.

5. Consultants on statistical aspects of the design

There are no consultants on the statistical aspects of the design for the site protocols.

Contact information for key individuals responsible for collecting and analyzing the data:

Andrew Wiegand, Task Leader for the Implementation Component
Social Policy Research Associates
1330 Broadway, Suite 1426
Oakland, CA 94612
510-763-1499
Andrew.Wiegand@spra.com

Cynthia Miller, Project Director and Task Leader for the Impact Component
MDRC
16 East 34th Street
New York, NY 10016
Cynthia.miller@mdrc.org
212-532-3200