The National Guard Youth ChalleNGe Job ChalleNGe Evaluation

OFFICE OF MANAGEMENT AND BUDGET SUPPORTING STATEMENT PART B

The Employment and Training Administration (ETA) of the U.S. Department of Labor (DOL) is funding three National Guard Youth ChalleNGe programs to expand the program's target population to include court-involved youth and add a five-month residential occupational training component called Job ChalleNGe.

The goal of Youth ChalleNGe is to build confidence and maturity, teach practical life skills, and help youth obtain a high school diploma or GED. The program's numerous activities all address its eight core pillars: leadership/followership, responsible citizenship, service to community, life-coping skills, physical fitness, health and hygiene, job skills, and academic excellence. It has a quasi-military aspect in which participants, known as "Cadets," live in barracks-style housing in a disciplined environment for about 20 weeks—the residential phase. Cadets wear their hair short and dress in military uniforms. Upon completing the residential phase of the program, participants receive one year of structured mentoring designed to help them successfully transition back to their communities.

The addition of the Job ChalleNGe component to the existing Youth ChalleNGe model has the potential to bolster the program's effectiveness by incorporating occupational training. Job ChalleNGe will expand the residential time by five months for a subset of randomly selected Cadets who are interested in staying, and will offer the following activities: (1) occupational skills training, (2) individualized career and academic counseling, (3) work-based learning opportunities, and (4) leadership development activities. In addition, the program will engage employers to ensure Cadets' skills address employers' needs.

The National Guard Youth ChalleNGe Job ChalleNGe Evaluation, sponsored by DOL's Chief Evaluation Office (CEO), will use (1) a set of interviews with staff and youth to learn how these program enhancements are implemented, and (2) a randomized controlled trial (RCT) to measure the effectiveness of Job ChalleNGe. The RCT will compare Youth ChalleNGe graduates who attend Job ChalleNGe to graduates who were on track for Job ChalleNGe but were not randomly selected to attend. Expanding eligibility for Youth ChalleNGe and Job ChalleNGe to court-involved youth, who for the most part are not currently eligible to participate in Youth ChalleNGe, could make a difference in the lives of those youth who can be the hardest to serve. The evaluation will take place at sites awarded Job ChalleNGe grants in 2015: Fort Stewart, Georgia; Battle Creek, Michigan; and Aiken, South Carolina.

The CEO has contracted with Mathematica Policy Research, in conjunction with its subcontractors MDRC and Social Policy Research Associates (SPR), to conduct this evaluation. With this package, clearance is requested for four data collection instruments related to the impact and implementation studies to be conducted as part of the evaluation:

1. Baseline information form (BIF) for youth

2. Site visit master staff protocol

3. Site visit employer protocol

4. Site visit youth focus group protocol

The site visits interview protocols will include semi-structured interviews with grantee administration and staff, partners, and employers. The site visits protocols and youth focus group protocol will be conducted in all three program sites. No statistical methods will be used in the implementation analysis and discussions of the results will be carefully phrased to make clear that no generalization is intended.

An addendum to this package, to be submitted at a later date, will request clearance for the follow-up data collection from study participants, including communication tools (like advance letters and email text for non-response follow-up) and the survey instrument. We are submitting the full package for the study in two parts because the study schedule requires random assignment to take place and the implementation study to begin before the follow-up instrument and related tools are developed and tested.

A. Justification

1.   Respondent universe and sampling methods

Baseline data collection for the Impact Study

The impact evaluation will assess the impact of the Job ChalleNGe program on youth outcomes. The universe and sample will be drawn from the three Youth ChalleNGe grantee sites. As a condition of receiving the grant, grantees are required to participate in the evaluation.

The evaluation design is a random assignment design with a random assignment process conducted for eligible Youth ChalleNGe participants who are interested in participating in Job ChalleNGe.

The universe and sample of youth will comprise youth in the three grantee sites who are enrolled in Youth ChalleNGe, are expected to successfully complete Youth ChalleNGe (as of a cutoff period determined in conjunction with each grantee), are deemed by the grantee to be suitable for Job ChalleNGe, and are interested in participating in Job ChalleNGe. The evaluation is designed to have up to 1,620 youth undergo random assignment for the Job ChalleNGe program. Job ChalleNGe will also enroll youth in cohorts; prior to the start of each cohort, the evaluation team will assess with each grantee if there are sufficient applications in the incoming Job ChalleNGe cohort to allow for random assignment to take place.

For the random assignment procedures at each grantee, it is expected that about two-thirds of youth will be assigned to the treatment group and one-third will be assigned to the control group (Table B.1). For Job ChalleNGe, the expected sample size is 1,620 total, with 1,080 enrolled in Job ChalleNGe and 540 placed in the control group. Random assignment will be stratified within each grantee organization to ensure a balanced sample of court-involved and non-court-involved youth in the treatment and control groups. Based on the expected sample sizes, the treatment group would comprise 540 court-involved youth and 540 non-court-involved

youth, whereas the control group would comprise 270 court-involved youth and 270 non-court-involved youth. Sample sizes are further broken out by grantee site (Table B.2).

Table B.1. Estimated sample sizes

Analysis sample	Treatment	Control	Total
Job ChalleNGe			
Court-involved	540	270	810
Non-court-involved	540	270	810
Total	1,080	540	1,620

Table B.2. Estimated sample sizes by grantee site

Analysis sample	Georgia		Michigan		South Carolina	
	Treatment	Control	Treatment	Control	Treatment	Control
Job ChalleNGe						
Court-involved	180	90	180	90	180	90
Non-court-involved	180	90	180	90	180	90
Total	360	180	360	180	360	180

We expect high response rates (at least 95 percent) for both the consent form and the baseline information form, based on prior experience conducting an evaluation of YouthBuild, a similar program, for which we obtained a comparable response rate.

Site Visits for Implementation Study

For the implementation study, in-depth site visits will be conducted to each of the three sites to look more closely at program operations, challenges, and successes that can help support the impact study. The visits to the three grantee sites will involve interviews with up to 113 staff, and 3 employers during the first visit, and interviews with up to 104 staff, and 6 employers during the second visit. The visits will also involve focus groups with up to 42 cadets during each of the two visits. We will not use any statistical methods in the selection of staff, employers, or youth to interview or in the analysis of the interview data.

2. Procedures for the collection of information

Youth (and their parents/guardians, when needed) will be asked to provide consent for participation in the study prior to the collection of study data or random assignment. When possible, youth will be asked to provide consent and complete a form requesting baseline data early in the Youth ChalleNGe program. The consent form administered at that time also will be used to secure consent from the youth (and their parents or guardians, when needed) for the Job ChalleNGe random assignment and related impact study. In this way, consent will need to be collected only once, rather than twice.

As noted above, random assignment processes will be used to select up to 1,620 youth for the Job ChalleNGe impact analysis (1,080 treatment, 540 control).

a. Analysis methods for impact estimation

The central feature of the analysis to estimate the impacts of access to the Job ChalleNGe program on youth outcomes is the random assignment of program-eligible youth to a treatment group that will be eligible to participate in the Job ChalleNGe program, or a control group that will not be eligible for Job ChalleNGe but could receive the standard Youth ChalleNGe follow-up services. Experimental statistical methods will be used to yield unbiased estimates of the impacts of the Job ChalleNGe programs by comparing the mean outcomes of the treatment and control group members over time (see the section below on "Estimating impacts for the full sample" for a fuller discussion). Outcomes will be measured as either binary (0/1) variables (for example, whether or not the youth received a high school diploma) or continuous variables (for example, earnings). Impacts will be estimated not only for the full sample, but also for policy-relevant subgroups, such as court-involved youth. The analysis will be conducted using the SAS and Stata software programs.

Assessing baseline equivalence. If a random assignment design is conducted properly, there should be no systematic observable or unobservable differences between the treatment and control groups except for the services offered after random assignment. To assess whether randomization was conducted properly, statistical t-tests will be conducted to assess mean differences in the baseline measures of treatment and control groups using data from the BIFs. Because baseline data will be collected prior to random assignment, there should no differences in data quality or response between the treatment and control groups. In addition, t-tests will be conducted on each baseline measure in isolation, and a joint F-test will be used to assess the joint significance of the baseline differences.

Estimating impacts for the full sample. With a random assignment design, simple differences in the mean values of outcomes between students assigned to the treatment and control groups will yield unbiased impact estimates of program effects, and the associated t-tests can be used to assess statistical significance.

This study will calculate the impact of access to the Job ChalleNGe program for youth who have completed the Youth ChalleNGe residential phase and are interested in participating in Job ChalleNGe. The strength of the experimental design is that, when implemented properly, random assignment ensures that the treatment and control groups are similar to each other at baseline on both observed and unobserved characteristics. Any statistically significant post-baseline differences between the two groups can be interpreted as effects of the program. Our proposed approach for Job ChalleNGe impact estimation combines the strength of random assignment design with statistical modeling to improve the efficiency of the estimated impacts.

The primary analytical method will compare average outcomes for treatment group members and control group members pooled across the sites. Regression adjustments will increase the power of the statistical tests. For impacts on continuous outcomes, such as earnings during a certain time period, the following regression model will be estimated:

(1) $Y_{ij} = a + \beta X_{ij} + \delta P_{ij} + \gamma S_j + \varepsilon_{ij}$

where Y_{ij} is the outcome measure for sample member i at grantee site j, P_{ij} is an indicator for participation in the Job ChalleNGe program, X_{ij} is a set of background characteristics for sample member i at grantee site j, S_j is a grantee site fixed effect, ε_{ij} is a random error term, and α, β, δ, and γ are parameters to be estimated. The key coefficient of interest is δ, which represents the impact of the program on the outcome.

The Job ChalleNGe impact analysis will be limited to youth who participated in Youth ChalleNGe and expressed interest in participating in Job ChalleNGe. The key coefficient of interest is δ, which represents the impact of the Job ChalleNGe program on the outcome. The outcomes will be constructed from both survey and administrative data for the Job ChalleNGe impact analysis. (As noted above, clearance for the follow-up survey instrument will be requested in a separate clearance package.) In addition to the outcomes constructed from administrative data listed above, additional outcomes from the survey will include measures of (1) education success (such vocational certificates or credentials), (2) employment success (such as work-readiness skills, work maturity, responsibility and flexibility, weeks of employment, fringe benefits, and occupation type), and (3) delinquency and criminal justice involvement (such as drug use, delinquent behaviors, time spent in juvenile detention or incarceration, and probation or parole status).

Estimating impacts for subgroups. To understand how impacts of Job ChalleNGe programs may vary by youth characteristics, subgroup impacts will be calculated using a "split-sample" approach in which the full sample is divided into two subgroups. Impacts for subgroups will be estimated using a straightforward modification to equation (1), where the model includes terms formed by interacting subgroup indicators with the treatment status indicator variable and using F-tests to assess whether differences in impacts across subgroup levels are statistically significant. During the study design phase, the evaluation team will work with CEO to determine the subgroups for which a subgroup analysis should be conducted. Subgroup analysis will be conducted for court-involved youth versus non-court-involved youth, and potentially for groups defined by age and race/ethnicity, given that the National Job Corps study found different impacts for subgroups defined by these characteristics (Schochet et al. 2008). Other potentially important subgroups could be defined by the youth's level of economic disadvantage and the youth's program.

The analyses described above will provide estimates of the impact of having access to Job ChalleNGe services. Commonly referred to as "intent-to-treat" estimates, these estimates are based on all treatment group members, whether or not they participate in the program (as well as all control group members, whether or not they would participate if given the opportunity to do so). However, because some treatment group members will not participate, the impacts for only those participants who enroll in program services will also be investigated. Using an assumption that there are no impacts on treatment group members who do not participate, a "treatment on the treated" impact will be estimated per participant by dividing the intent-to-treat impact estimate by the proportion of treatment group members who participate (Bloom 1984). In the unlikely event that control group members receive program services (known as crossovers), they will be accounted for analytically.

b. Precision calculations of the impact estimates

We assume that our estimates will generalize to a super population of all eligible youth who may have attended Job ChalleNGe at the three grantee sites, and therefore calculate standard errors for the impact estimates.

Based on study sample sizes of up to 1,620 youth for Job ChalleNGe across all three grantees combined, the evaluation will have sufficient statistical power to detect meaningful impacts on key study outcomes that are similar to those found in impact studies of similar types of interventions. The Job ChalleNGe study is expected to detect a significant impact on involvement in a productive activity of 7.0 percentage points, and an impact on weekly earnings of $45. With the intense vocational focus and additional residential time on the Job ChalleNGe intervention, the MDIs seem reasonable.

Table B.3. Minimum detectable impacts on key outcomes

Outcome	Minimum detectable impacts
Involved in productive activity (percentage point)	
Full sample	0.070
Court-involved youth	0.099
Current weekly earnings (dollars)	
Full sample	45
Court-involved youth	64

Note: Calculations assume that the control group is half the size of the treatment group and court-involved youth represent half of both groups. Data for outcomes are assumed to be from survey data for which we assume a response rate of 80 percent. We assume a rate of involvement in productive activity (work or education/training) of 66.4 percent, and a standard deviation of weekly earnings of $308. We assume that covariates in the regression model will explain 20 percent of the variation in the outcome measures.

c. Analysis methods for implementation study

The implementation analysis will not utilize statistical techniques or specialized sampling procedures.

3. Methods to maximize response rates and deal with nonresponse

For this study, we are requesting approval for the Baseline Information Form (BIF) to be completed as sample members go through an intake process, and for protocols to be used during visits to study sites. No monetary or nonmonetary incentives will be provided to respondents. Below, we discuss the methods to maximize response rates and data reliability, first for the BIF and then for the implementation study protocols. Based on prior experience, we expect to achieve at least a 95 percent response rate for consent and for the BIF; therefore, nonresponse bias should not be a significant issue for analysis of the baseline data. (As noted earlier, a future OMB package will request clearance for a follow-up survey instrument for the evaluation, and will include a discussion of analysis and weighting to address survey nonresponse.)

a. **Intake documents: Baseline information forms**

Response rates. The baseline information form will be administered, after consent is obtained, to all eligible youth during the sample intake period. The project team will work closely with program staff to identify ways to integrate the BIFs into their normal program procedures in order to minimize respondents' burden and increase response rates.

To maximize response for the consent/assent forms and BIFs, the study team will use methods to ensure that the study is clearly explained to both study participants and staff, and that the forms are easy to understand and complete and will be used. These methods have been successful in many other random assignment studies. Care has been taken in these forms to explain the study accurately and simply to potential participants. In addition, the forms will be available in Spanish to accommodate Spanish-speaking students and parents/guardians. Grantee staff will be thoroughly trained to address study participants' questions about the forms and to check that the forms have been filled out properly. Grantee staff will also be provided with a site-specific operational procedures manual prepared by the research team, contact information for members of the research team, and detailed information about the study. Based on experience with similar data collection efforts, the evaluation team expects that nearly 100 percent of eligible program applicants will participate in the study and will complete the BIF. Given this high expected response rate, nonresponse bias should not be a significant issue for analysis of the baseline data.

Data reliability. All forms required at intake are unique to the current evaluation and will be used across all three Youth ChalleNGe grantees, ensuring consistency in the use of the forms and in the collected data. The forms have been extensively reviewed by evaluation staff and staff at DOL and thoroughly tested in a pre-test. Building on our successful model from the YouthBuild evaluation, a paper-and-pencil hard-copy form will be used to collect baseline information on all youth who will participate in the evaluation. Program staff will hand out the form for youth to fill out. Program staff will send the forms to the evaluation team using a secure mailing system such as FedEx. These data will be double entered into an electronic database by specially trained data entry clerks on the evaluation team staff. Paper-and-pencil administration avoids costly programming of a web-based system and does not require the program staff to enter the data.

b. **Site visit data**

The data will be collected through semi-structured interviews and focus groups held at grantees sites. Experienced researchers will conduct two-day site visits.

4. **Tests of procedures or methods**

Most of the items in the BIF are either identical or similar to questions used in previous studies (including other DOL studies such as the Evaluation of the Impact on the YouthBuild Program and the Evaluation of Youth Career Connect, as well as the previous National Guard Youth ChalleNGe Study) or national surveys. As such, these items have been thoroughly tested on large samples. Additionally, the consent form and BIF have been pretested with nine youth in another National Guard Youth ChalleNGe site. After the pre-test participants completed the form, members of the evaluation team conducted a debriefing with the participants using a standard debriefing protocol to determine whether any words or questions are difficult to

understand or answer. Slight changes were made in the instruments and consent forms at that time to respond to the feedback.

The implementation study protocols were developed largely from protocols that were piloted in previous studies, including studies of YouthBuild, grants for Youth Offenders, the Los Angeles Reconnections Career Academy, RExO, and Youth Career Connect, as well as the WIA Gold Standard Evaluation. Protocols from these studies helped us identify key topic areas, a list of potential questions and a universe of answers for questions with more well-defined answers, such as those concerning the use of particular curriculum, types of partnerships, and so forth. Because the Youth ChalleNGe and Job ChalleNGe programs differ from existing programs, we also conducted a close review of existing documents for these two programs, including the successful grantee proposals, to distill questions about service design options.

5. Individuals consulted on statistical methods

Consultations on the statistical methods used in this study have been used to ensure the technical soundness of the study. The following individuals were consulted on statistical aspects of the design, and will also be primarily responsible for actually collecting and analyzing the data for the agency:

Mathematica Policy Research

- Ms. Jeanne Bellotti   (609) 275-2243
- Dr. Jillian Berk   (202) 264-3449
- Dr. Eric Isenberg   (312) 994-1009
- Dr. Karen Needels   (609) 750-4043

Additionally, the following individuals were consulted on the statistical methods discussed in this submission to OMB:

MDRC

- Mr. Dan Bloom   (212) 340-8611
- Ms. Megan Millenky   (212) 340-8670

Inquiries regarding the statistical aspects of the study's planned analysis should be directed to:

- Dr. Jillian Berk   (202) 264-3449
- Dr. Molly Irwin   (202) 693-5091

REFERENCES

Bloom, H. "Accounting for No-Shows in Experimental Evaluation Designs." *Evaluation Review*, vol. 8, 1984.

Millenky, M., D. Bloom, S.M. Ravett, and J. Broadus. "Staying on Course: Three-Year Results of the National Guard Youth ChalleNGe Evaluation." New York: MDRC, 2011.

Schochet, P., J. Burghardt, and S. McConnell. "Does Job Corps Work? Impact Findings from the National Job Corps Study." *American Economic Review*, vol. 98, no. 5, 2008, pp. 1864–1886.