Evaluation of Employment Coaching for TANF and Related Populations

OMB Information Collection Request
New Collection

Supporting Statement

Part B

July 2017

Revised January 2018

Submitted By:
Office of Planning, Research and Evaluation
Administration for Children and Families
U.S. Department of Health and Human Services

4th Floor, Mary E. Switzer Building
330 C Street, SW
Washington, D.C. 20201

Project Officers:
Hilary Forster
Victoria Kabak

B1. Respondent Universe and Sampling Methods

The Office of Planning, Research and Evaluation (OPRE) within the Administration for Children and Families (ACF) at the U.S. Department of Health and Human Services (HHS) seeks approval for data collection activities conducted for the *Evaluation of Employment Coaching for TANF and Related Populations*. The objective of this evaluation is to provide information on coaching interventions implemented by Temporary Assistance for Needy Families (TANF) agencies and other employment programs. The evaluation will describe up to three coaching interventions and assess their effectiveness in helping people obtain and retain jobs, advance in their careers, move toward self-sufficiency, and improve their overall well-being. The evaluation will include both an experimental impact study and an implementation study.

Programs selected for the evaluation will include a robust coaching component and have the capacity to conduct a rigorous impact evaluation, among other criteria. Each program is expected to recruit 2,000 eligible people, for a total of 6,000 participants across all three programs. After participants consent to participate in the study (see Attachment A), half will be randomly assigned to the treatment group and will be offered coaching services; the other half will be randomly assigned to the control group and will not be offered these coaching services.

This information collection request (ICR) covers data collection activities for both an impact and an implementation study. Data collection activities for the impact study include: (1) baseline data collection and (2) two follow-up surveys (the first of which is covered in this ICR). Data collection activities for the implementation study include: (1) semi-structured staff interviews; (2) a staff survey; (3) in-depth participant interviews; (4) staff reports of participant service receipt; and (5) video recordings of coaching sessions. A subsequent ICR will request approval for the second follow-up survey of the impact study.

Impact Study

This ICR seeks clearance for two instruments associated with the following data collection efforts for the impact study:

1. **Baseline data collection (Attachment B).** Baseline data will be collected from all study participants (approximately 2,000 in each program, or 6,000 across all three programs). Under most circumstances, program staff will administer the baseline data collection. Thus, this data collection is associated with burden for both participants and staff. Some programs might ask that the evaluation allow participants to complete the baseline data collection on their own. A self-administered baseline data collection is not expected to affect participant burden and might reduce staff burden.

2. **Follow-up survey (Attachment C).** The follow-up survey will be administered to 1,000 participants per program. If the study includes more than 1,000 participants per program, then the survey will be administered to a random sample of 1,000 study participants. We expect that 80 percent will complete the survey for a total of 800 respondents per program (approximately 2,400 across all three programs).

Implementation Study

This ICR seeks clearance for five data collection activities for the implementation study:

1. **Semi-structured staff interviews (Attachment D).** We expect to interview 66 program staff across all three programs (approximately 22 per program). Respondents will be selected purposively using organizational charts and information on each employee's role at the host organization and its partner organizations. Purposeful staff selection is appropriate because particular insights and information can only come from individuals with certain roles or knowledge. Program staff may include coaches, case managers, workshop instructors, job developers, supervisors, and managers. This attachment contains three different interview guides that pertain to three different types of staff (frontline workers, supervisors, and managers/program administrators).

2. **Staff survey (Attachment E).** The staff survey will be fielded to all management and staff involved in the coaching intervention at each program. Staff may include coaches, case managers, workshop instructors, job developers, supervisors, and managers. We expect to survey 48 management and staff.

3. **In-depth participant interviews (Attachment F).** In-depth interviews will be conducted with a subset of about eight participants per program (approximately 24 participant interviews across all three programs). These participants will be selected purposively from among the treatment group, ensuring we interview participants with different levels of engagement and lengths of time in the program, in order to capture a range of experiences.

4. **Staff reports of program service receipt (Attachment G).** We expect that 30 program staff will use the Random Assignment, Participant Tracking Enrollment, and Reporting system (RAPTER) to record information on case management and other program services that both the treatment group and the control group members receive, if the design allows control group members to receive these other program services over the duration of the program.

5. **Video recordings of coaching sessions.** A subset of coaching sessions at each program will be recorded. This subset will be chosen to include all sessions over a specific period of time with each coach and will capture multiple participants for each coach. We anticipate that 9 staff will record up to 90 sessions per program (approximately 27 staff and 270 sessions across all three programs).

B2. Procedures for Collection of Information

Impact Study

The data collection procedures for the impact study instruments are described below.

1. **Baseline data collection.** In all three programs, program staff will identify individuals eligible to participate in coaching services and enroll them in the study sample. Under most circumstances, when intake workers are ready to enroll an individual into the study, they will administer the consent form to the applicant (Attachment A) and enter baseline data in RAPTER. RAPTER is a secure, web-based system that program staff will use to

administer consent to participants, collect baseline data, and conduct random assignment. The use of check boxes, drop-down menus, and response categories will minimize data entry burden (Attachment G). Some programs might ask that the evaluation allow some participants to complete the consent and baseline data collection on their own as part of the evaluation intake process. In this case, some participants may complete a self-administered baseline survey, using the same web-based system. This is not expected to affect participant burden and might reduce staff burden. The baseline survey includes alternate language tailored to self-administration. For example, text transitioning between sections of the survey may read, "Now I would like to ask you some questions about the people who live with you" for the staff-administered version and, "The next questions are about people who live with you." for the self-administered version. After baseline data are collected, the program staff member will then use RAPTER to conduct the random assignment process. The entire intake process is expected to take 20 minutes to complete. Although we are not planning to administer the surveys in Spanish, if a program offers coaching in Spanish, then we will translate the instruments and materials.

2. **First follow-up survey.** The follow-up survey will be administered to treatment and control group members approximately six to 12 months after random assignment. If the study includes more than 1,000 participants per program, then the survey will be administered to a random sample of 1,000 participants. Study participants will be contacted approximately one week before the start of data collection by mail, to notify them of the upcoming survey request (Attachment I).

In addition to collecting data at baseline and at follow-up, administrative data will also be collected electronically on the full study sample. Administrative data will be collected from the National Directory of New Hires (NDNH), operated by the Office of Child Support Enforcement at HHS, and from TANF agencies for all study participants. The NDNH data will include quarterly earnings, unemployment insurance benefit amount, and start date of new job. The TANF data will include TANF benefits, the status of individual cases, information on whether study participants are exempt from work requirements, and details about whether they have participated in specific types of work activities. These data are already being collected and do not represent additional burden for respondents.

Table B.1 reports program-level minimum detectable impacts on outcomes obtained from survey data. We assume a study sample of 2,000 people per program (1,000 each in the treatment and control groups). We will survey a randomly selected subsample of 1,000 people (500 each in the treatment and control groups) per program. With an 80 percent response rate, the sample of survey respondents would include 800 people per program (400 in the treatment group and 400 in the control group).

Table B.1. Minimum detectable effects on survey-based outcomes, by size of survey sample

Sample size (treatment and control)	Minimum detectable effect
500	0.25
1,000	0.18
2,000	0.13

Assumptions: People are assigned with equal probability to the treatment and control groups. We assume that covariates in the regression model will explain 20 percent of the variation in the outcome measures. All power calculations are based on the following formula: $MDE = \left[T_{df}^{-1}(1-\alpha/2) + T_{df}^{-1}(1-\beta) \right] * \sqrt{(1-R^2)/np(1-p)}$, where T_{df}^{-1} is the inverse t distribution with df degrees of freedom, α is the significance level of the test, β is the level of Type II error, R^2 is the variance in outcomes explained by baseline characteristics, n is the number of participants after attrition, and p is the fraction of study participants in the treatment group. We assume $\alpha = 0.05$ and power is 80 percent $(1-\beta = 0.80)$. We assume 20 percent attrition in the survey data.

These samples are large enough to detect the expected impacts of the programs, even accounting for attrition in the survey sample. With a survey sample of 1,000 study participants (which implies an analysis sample of 800 people based on an 80 percent response rate), we will be able to detect an impact of 0.18 standard deviations. Standardized evidence reviews, such as the What Works Clearinghouse, consider effect sizes of 0.25 standard deviations or larger as substantively important (U.S. Department of Education 2014).

Implementation Study

This ICR, includes five information collection activities associated with the implementation study. The data collection procedures for these activities are as follows:

1. **Semi-structured staff interviews.** Interviews will be conducted with program staff during site visits approximately six months after study enrollment begins in that program. The interviewers will offer privacy assurance as part of the introduction to the interview. Interviews will take place individually or in small groups, depending on the staffing structure, roles, and number of staff in each role. All interviews will be conducted by a team of two study team members, one asking questions and the other typing close-to-verbatim notes capturing key quotes and responses on a laptop. With permission from respondents, site visit teams will use an audio recorder to record interviews to later confirm direct quotes and other details from the interviews. The discussions will be guided by a series of questions organized by topics and will range freely as study staff and informants engage in conversations exploring topic areas in depth. Three different interview guides were developed that pertain to three different types of staff (frontline workers, supervisors, and managers/program administrators).

2. **Staff survey.** A link to the staff survey will be emailed to respondents. The staff survey will be administered via the web approximately six months after study enrollment begins and is expected to take 45 minutes to complete. The introduction to the survey will inform management and staff that their participation is completely voluntary.

3. **In-depth participant interviews.** In-depth, in-person interviews will occur either at participants' homes or a place of their choice (but not at the program). Trained interviewers will obtain participant contact information from the program, schedule the interview with individual participants, and then record each one-on-one interview. Participants will be asked for consent to record the interview. For participants that do not provide this consent, notes will be taken in lieu of a recording.

4. **Staff reports of program service receipt.** Program staff will use RAPTER in each program to document service use by program participants. The evaluation team will train

all program staff on using RAPTER to enter data on program receipt for all participants over the duration of the evaluation.

5. **Video recordings of coaching sessions.** Program staff will video-record a subset of coaching sessions to collect data on the interaction between coaches and participants. The consent language participants agree to when enrolling in the study includes participating in these video recordings. We will provide each program a sufficient number of tablets and tripods so they can set up the video recording in the rooms used for coaching sessions. Program staff will be trained on how to record the interaction without being intrusive, upload the videos from the tablet to a secure file transfer site, and store the tablets in a secure location when not in use (Attachment M).

B3. Methods to Maximize Response Rates and Deal with Nonresponse

Expected Response Rates

Expected response rates for the information collection activities associated with the impact and implementation studies are discussed below.

Baseline data collection. Applicants eligible for study participation will only be enrolled in the study and randomly assigned if they complete the baseline data collection effort as part of the intake process. Therefore, the evaluation team anticipates that 100 percent of study participants will provide baseline data.

Follow-up survey. We anticipate an 80 percent response rate on the follow-up surveys based on our experiences conducting follow-up surveys with similar populations. In our evaluation of the Building Nebraska Families program (OMB control number 0970-0246), we achieved an 87 percent response rate on the 18-month follow-up survey and an 83 percent response rate on the 30-month follow-up survey. This program, which was conducted with a population similar to the current study, was designed to help TANF recipients and other low-income people enter, maintain, and advance in employment. For the Personal Responsibility Education Program (PREP) evaluation (OMB control number 0970-0398), we are on track to achieve response rates above 80 percent for the Healthy Families San Angelo program, a home-visitation program that targets a low-income population, similar to the current study. At this site, the cohorts for whom data collection is complete have a response rate of 85 percent on the one-year follow-up survey and 83 percent on the two-year follow-up survey. For the Parents and Children Together follow-up surveys, using the strategies outlined below, we achieved an 88 percent response rate for the low-income mothers and fathers in the healthy marriage program study (OMB control number 0970-0403). All of these examples demonstrate the usefulness of our responsive design strategies for achieving high response rates with low-income, at-risk populations. The combination of sound planning, using paradata and adaptive design, and our experience with at-risk populations produces balanced, high-quality data.

Semi-structured staff interviews. The evaluation team will target completing 66 semi-structured staff interviews (22 at each program). This is a reasonable target given that management and staff will already have agreed to participate in the evaluation.

Staff survey. Based on similar research projects, we expect a high response rate among management and staff (at least 70 percent). On the Job Search Assistance (JSA) Strategies

Evaluation (OMB control number 0970-0400), we achieved a 70 percent response rate to a web-based survey administered to TANF staff, which used the same mode and population as the current study.

Dealing with Nonresponse

All analysis of follow-up survey data will account for survey nonresponse using nonresponse weights. Weights will be calculated using standard techniques to estimate the probability of nonresponse as a function of baseline characteristics. The evaluation team does not anticipate significant item nonresponse based on prior experience asking similar questions with similar populations, as described in the studies above.

Some survey nonresponse is inevitable, although it will be minimized by providing incentives. The evaluation team will analyze nonresponse to assess whether the sample of follow-up survey respondents is representative of the full study sample. Using the data on participants' characteristics collected at baseline, Mathematica will conduct statistical tests (chi-square and t-tests) to gauge whether the treatment group members who participated in data collection are representative of all the treatment group members, whether the control group members who participated in data collection are representative of all the control group members, and whether there are systematic differences in the treatment and control group members who responded to the survey.

The evaluation team will use two approaches to correct for potential nonresponse bias in the estimation of program impacts. First, the regression models described in A16 will adjust for observed differences between the characteristics of treatment and control group respondents. Second, because this regression procedure will not correct for differences between respondents and nonrespondents in each research group, sample weights will be constructed so that the weighted baseline characteristics of respondents in the treatment and control group in each program are similar to those of the full sample (respondents and nonrespondents). These weights will be constructed using data from the baseline surveys.

Maximizing Response Rates

Impact Study

Methods for maximizing response rates for the impact study are discussed below.

Baseline data collection. The evaluation team will take the following steps to maximize response rates and data reliability:

- **Use a tested questionnaire.** The collection of baseline data has been tailored to the specific circumstances of this evaluation, yet is based closely on the Evaluation of the Supplemental Nutrition Assistance Program (SNAP) Employment and Training Pilots baseline survey (OMB control number 0584-0604), a U.S. Department of Agriculture-funded initiative that received OMB approval, was extensively tested, and was successfully fielded. The goal of the SNAP Employment and Training evaluation was to rigorously test innovative strategies for increasing employment and earnings among SNAP participants and reducing their dependence on SNAP and other public assistance programs. Thus the population and goal of the SNAP Employment and

Training evaluation was similar to the current study. A question-by-question justification for the items included in baseline data collection is presented in Attachment J.

- **Use a straightforward, undemanding questionnaire.** The baseline data collection effort is designed to be easy to complete. The questions use clear and straightforward language. The average time required for the respondent to complete baseline data collection and for staff to administer the questionnaire and enter it into RAPTER is estimated at 20 minutes.[1]

Follow-up survey. The follow-up survey is also a straightforward and undemanding questionnaire that was pretested with nine people. A question-by-question justification for the items included in the follow-up survey is presented in Attachment K. The evaluation team will also take the following steps on the follow-up surveys to maximize response rates:

- **Use incentives.** A two-tiered system will be used to mitigate the potential for bias by increasing response rates and minimizing differential response rates between treatment and control groups. Respondents will be offered a $35 gift card if they complete the survey, either online or by telephone, within the first four weeks after receiving the survey; respondents will receive a $25 gift card if they complete the survey after four weeks. This "early bird" model has proven effective on a 60-minute survey for the YouthBuild evaluation, which achieved an overall response rate of 81 percent at 12 months; 82 and 79 percent for treatment and control conditions, respectively. The YouthBuild study provided an incentive of $40 if respondents completed a 12-month follow-up survey within the first four weeks and $25 if respondents completed the survey after four weeks. This structure was approved by OMB (OMB control number 1205-0503). The efficacy of the two-tiered approach was also shown through an incentive experiment that was conducted as part of the Self-Employment Training (SET) Demonstration 20-minute follow-up survey (OMB control number 1205-0505). This experiment assessed the effectiveness of three incentive approaches: (1) offering a standard incentive of $25; (2) offering a two-tiered incentive, with an incentive of $50 if respondents completed an 18-month follow-up survey within the first four weeks and $25 if respondents completed the survey after four weeks; or (3) offering no incentive. This experiment found that the response rate was 37 percent for the sample members who were not offered an incentive, compared to 73 percent for sample members offered a two-tiered incentive. Based on evidence from SET and Project LAUNCH, which is discussed in greater detail in Supporting Statement A, we anticipate that without incentives, the survey response rate would be unacceptably low; it is likely to be less than 50 percent. Such response rates would put the study at severe risk of biased impact estimates. Over the course of the entire SET data collection (both with and without the incentive experiment), the two-tiered incentive approach achieved an overall response rate of 80 percent; 83 and 78 percent for treatment and control conditions, respectively.

- **Allow respondents to complete the survey in different ways**. The participants will be able to complete the survey either online (using a computer, tablet, or smartphone) or by telephone.

---

[1] As noted in A12, this time estimate includes the consent process.

- **Send reminder notifications.** The evaluation team will use a combination of letters, emails, texts, and telephone calls to encourage participants to participate. These notifications are included in Attachment I. For example, the advance letter (and insert) will be mailed to participants at the start of data collection. The email notification will be emailed to participants who have not yet completed the survey about three weeks after the start of data collection. The refusal avoidance letter will be mailed to participants who have not yet completed the survey and who we think will respond but are being avoidant or are delaying responding. A locating letter will be sent to participants who have not completed the survey after all available contact information has gone through a locating process (described below).

- **Obtain accurate, up-to-date contact information**. Detailed contact information will be collected at baseline that includes telephone numbers, addresses, and email addresses to aid in locating participants to complete the follow-up surveys (Attachment B). Detailed contact information will also be collected for three relatives, friends, neighbors, and past employers whom the participant selects and who may be able to help locate the participants if they move. The evaluation team will also request updates from project staff, if they have any. Before the start of the follow-up surveys, participant contact information will be updated through online database searches, and the evaluation team might request updates from participants via text message or email.

- **Use intensive locating methods, as needed.** Participants will initially be notified about the survey by mail and email and asked to complete it via the web, though they will also be able to complete it via telephone at that time (Attachment I). At that point, they will be offered a higher incentive to increase response rates and minimize differential response rates between treatment and control groups. After four weeks, the evaluation team will attempt to contact the participants via telephone at the numbers provided in the baseline data, in order to have them complete the survey via telephone. If participants cannot be reached by telephone, the evaluation team will contact the friends, family, neighbors, and past employers identified by the participant during the baseline data collection, for help in locating them. Customized, individual searches for contact information using specialized databases will be conducted next. Finally, if study participants still cannot be located, trained field locators will go in person to the study participant's home and neighborhood. If they locate the study participant, the field locators will lend him or her a smartphone to complete the survey.

- **Use paradata.** Data will be collected on each attempt to contact a respondent including the mode, time, date, interviewer, and contact results. Examining these paradata will help to identify the most effective calling times and interviewers. Paradata will also be used to determine which methods of contact (letters, emails, texts, or telephone calls) are proving to be the most successful in this study, so that the frequency and type of contacts can be adjusted to achieve high response rates.

- **Monitor response rates closely by group**. Response rates will be monitored closely throughout the fielding period, with an eye to any treatment–control differences that may emerge. If treatment–control differences are observed, then the locating efforts will be intensified for the group with the lower response rate to minimize differential nonresponse.

Implementation Study

Methods for maximizing response rates for the implementation study are discussed below.

Semi-structured staff interviews. Well before the site visits during which the semi-structured interviews will take place, the evaluation team will begin working with program staff to ensure the timing of the visit is convenient. The scheduling of specific interviews will be flexible to accommodate the particular needs of respondents and program operations.

Staff survey. The staff survey will also use a tested questionnaire. A question-by-question justification for the staff survey is presented in Attachment L. To maximize response rates and data reliability for the staff survey, we will use methods similar to those described above under the impact surveys. For example, we will:

- **Use a tested questionnaire common to all programs**. The staff survey has been tailored to the specific circumstances of the evaluation, yet it is based closely on the staff survey used in the National Implementation Evaluation of the Health Profession Opportunity Grants (HPOG) to Serve TANF Recipients and Other Low-Income Individuals (OMB control number 0970-0394). The HPOG survey was fielded to staff working with TANF recipients and other low-income individuals. Thus, the HPOG survey is similar to the current staff survey in terms of both content and population. The HPOG evaluation is ACF-funded, received OMB approval, was extensively tested, and was successfully fielded.

- **Use a straightforward, undemanding survey**. The staff survey is designed to be easy to complete and can be completed in more than one sitting. The questions use clear and straightforward language and are designed for close-ended responses. The average time required for the respondent to complete the survey is estimated at 45 minutes.

- **Use reminder emails and program liaisons.** To achieve a high response rate, the evaluation team will send periodic email reminders to respondents, beginning two weeks after the field period begins (Attachment I).

In-depth interviews with participants. As with the semi-structured staff interviews, the evaluation team will be flexible in scheduling specific interviews to accommodate study participants' schedules and needs. Respondents who participate in the in-depth interviews, which are estimated to take 2.5 hours on average, will receive a $50 gift card.

Staff reports of program service receipt. The evaluation team will monitor entry of program service receipt into RAPTER and will work with programs to encourage them to keep the records up to date to ensure high response rates. Data will also be pulled routinely to ensure completeness and quality.

Video recordings of coaching sessions. The evaluation team will monitor the video recordings and will provide assistance to program staff as needed to ensure high response rates for video recordings.

B4. Tests of Procedures or Methods to be Undertaken

Surveys. The baseline and first follow-up surveys were pretested on eight and nine people, respectively, similar to that survey's target population to estimate survey length, assess respondents' understanding of the survey questions, and identify improvements to the flow and structure of the instruments. We used cognitive interviewing and respondent and interviewer debriefings during these pretests.

RAPTER. RAPTER will be specifically designed for use by program staff. All functionality will be tested extensively by Mathematica before staff are trained on how to use RAPTER. The evaluation team will routinely examine the data staff enter into RAPTER to ensure quality.

Video recordings. During site visits, evaluation staff will test the video recordings to ensure they are being recorded and uploaded properly. The evaluation team will also monitor the data to ensure the recordings and upload procedures are functioning properly.

B5. Individual(s) Consulted on Statistical Aspects and Individuals Collecting and/or Analyzing Data

The individuals listed below consulted on the statistical aspects of the study to ensure the technical soundness of the research, or will be collecting and/or analyzing the data:

OPRE
Hilary Forster, Senior Social Science Research Analyst
Victoria Kabak, Social Science Research Analyst
Gabrielle Newell, Contract Social Science Research Analyst

Mathematica Policy Research
Dr. Sheena McConnell, Project Director
Dr. Quinn Moore, Deputy Project Director
Dr. Michelle Derr, Principal Investigator
Shawn Marsh, Survey Director

Abt Associates
Dr. Alan Werner, Principal Investigator
Dr. Bethany Boland, Senior Analyst

University of Chicago
Dr. James Heckman, Measurement Expert

References

U.S. Department of Education. *WWC Procedures and Standards Handbook*. Washington, DC: Institute for Education Sciences, March 2014. Available at http://ies.ed.gov/ncee/wwc/documentsum.aspx?sid=19. Accessed July 14, 2016.