Evaluation of Employment Coaching for TANF and Related Populations

OMB Information Collection Request 0970-0506

Supporting Statement Part B

April 2022

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 Officer: Lauren Deutsch

Part B

B1. Objectives

Study Objectives

This study is providing an opportunity to learn more about the potential of coaching to help clients achieve self-sufficiency and other desired employment-related outcomes. It includes the following employment programs: MyGoals for Employment Success in Baltimore (MyGoals Baltimore); MyGoals for Employment Success in Houston (MyGoals Houston); Family Development and Self-Sufficiency (FaDSS) program in Iowa; LIFT in New York City, Chicago, and Los Angeles; Work Success in Utah; and Goal4 It! in Jefferson County, Colorado. Together, these programs include Temporary Assistance for Needy Families (TANF) agencies and other public or private employment programs that serve low-income individuals. Each site has a robust coaching component and the capacity to conduct a rigorous impact evaluation.

This study is providing information on whether coaching helps participants develop self-regulation skills, obtain and retain jobs, advance in their careers, move toward self-sufficiency, and improve their overall well-being. The study objectives are to:

- Provide evidence of coaching interventions' impacts on participants' employment outcomes
- Provide evidence of coaching interventions' impacts on measures of self-regulation
- Assess whether coaching interventions are more effective for some subgroups of participants than others
- Assess how the impacts of coaching interventions change over time
- Investigate factors that could influence implementation of coaching and affect interpretation of impacts

To meet these objectives, this study includes an impact and implementation study, as approved by OMB. The approved impact study initially included two follow-up surveys at approximately 6 to 12 months and 21 to 24 months, respectively, after random assignment. The proposed third follow-up survey, to be conducted at least 48 months after study enrollment, will enable us to carefully trace how the impacts of employment coaching evolve over time for the outcomes the programs intend to affect, including self-regulation skills, labor market outcomes, public assistance receipt, and economic wellbeing. The proposed semi-structured interviews will enhance the implementation study, providing descriptive information about how coaches form trusting relationships with their participants and other key topics that have emerged as important in analysis of previously collected study data.

Generalizability of Results

The impact study is intended to produce internally valid estimates of the programs' causal impacts, not to promote statistical generalization to other sites or service populations. Findings from the impact study provide information about whether employment coaching can be effective in the context in which the programs participating in the study operate. The implementation study is intended to present internally valid descriptions of the service population and program implementation in chosen programs, not to promote statistical generalization to other sites or service populations.

Appropriateness of Study Design and Methods for Planned Uses

The impact study will continue to use an experimental research design. This design provides unbiased estimates of the effectiveness of each employment coaching intervention in improving employment-related outcomes, economic security, self-regulation, and other measures of well-being. People who are eligible for coaching were randomly assigned either to a "treatment group" at baseline and had access to the employment coaching intervention or to a "control group" and did not have access to the employment coaching intervention. The effectiveness of the employment coaching intervention. The effectiveness of the employment coaching intervention will continue to be assessed based on differences in outcomes between members of the treatment and control groups for the impact study. The implementation study will continue to use appropriate techniques for analysis of qualitative data to understand the employment coaching implementation challenges and solutions, such as how coaches form trusting relationships with their participants and implement coaching effectively. These techniques enable qualitative analysis of program implementation from a variety of relevant perspectives.

As noted in Supporting Statement A, this information is not intended to be used as the principal basis for public policy decisions and is not expected to meet the threshold of influential or highly influential scientific information.

B2. Methods and Design

To estimate the effects of coaching, we are using a rigorous experimental design. Participants eligible for the coaching services have been randomly assigned to one of two groups: (1) a treatment group offered coaching or (2) a control group not offered coaching. With this design, the research groups should be very similar in terms of their characteristics before receiving the intervention. Therefore, differences in observed outcomes can be attributed to the employment coaching intervention.

Target Population

Programs were selected for the evaluation after initial OMB approval in 2017 and have been participating in the previously approved information collections under this OMB number (OMB #0970-0506). Study enrollment is complete; the programs have recruited 5,026 people eligible for their programs to participate in the study with 2,514 randomly assigned to the treatment group and 2,512 to the control group. The target populations for these programs vary by program, as discussed in the next subsection.

For the impact study, we will attempt to administer remaining second follow-up surveys and third follow-up surveys to all study participants except those who have given adamant refusal to respond and survey participants who were confirmed deceased between rounds. For the implementation study, we will recruit interview participants based on convenience with information provided by program staff.

Sampling and Site Selection

Site selection. To be included in the evaluation, an employment coaching program needed to meet two broad criteria: (1) evaluating the intervention will inform the future development of employment coaching interventions; and (2) it is feasible to rigorously evaluate the intervention. To identify such programs, we solicited information from key stakeholders, consulted existing coaching research and literature reviews, and conducted web searches. We then conducted calls and site visits with programs of interest. Based on information gathered through this process, ACF selected programs to be recruited

to the evaluation. This process resulted in a set of programs that offer rich diversity in their coaching models, target populations, and geographic locations, as summarized in Table B.1.

Site	Distinguishing features of coaching model	Target population	Location of programs being evaluated	Type of implementing organization(s)
FaDSS	Offers self-sufficiency, domestic violence, and child development assessments, employment coaching, and referrals during home visits	TANF recipients with barriers to self- sufficiency	Seven local agencies located throughout lowa	State agency and local community- based organizations
Goal4 It!™	Offers employment coaching using a suite of tools in place of regular TANF case management	TANF recipients subject to work requirements	Jefferson County, Colorado	TANF agency
LIFT	Offers employment and financial coaching by volunteer students; financial support; workshops; and social activities	Parents and caregivers of young children	New York, New York; Chicago, Illinois; Los Angeles, California	National non-profit organization
MyGoals	Offers employment coaching by coaches trained on self- regulation skills; labor market information; and financial incentives	Unemployed adults receiving housing assistance	Baltimore, Maryland; Houston, Texas	City housing authorities
Work Success	Offers group employment coaching through a short-term, structured, time-intensive program	TANF participants and others job seekers at American Job Centers	Utah	State workforce agency

Table B.1. Key program features

Sampling—Follow-up surveys. We will attempt to administer the second follow-up survey to all study participants except those who have given adamant refusal to respond. We will attempt to administer the third follow-up survey to all study participants at the FaDSS, Goal4 It!, LIFT, MyGoals Baltimore, and MyGoals Houston sites, except those who have given adamant refusal to respond or are deceased. We do not plan to include Work Success in the third follow-up survey data collection because the program's service provision period is shorter than the other four programs and not expected to lead to different impacts at 48 months than at 21 to 24 months. Table B.2 reports program-level minimum detectable impacts on outcomes obtained from survey data, assuming a 75 percent response rate. These samples are large enough to detect the expected impacts of the programs, accounting for attrition in the survey sample. We will be able to detect an impact of 0.21 standard deviations or lower for each program. Standard is 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).

Site	Number enrolled	Minimum detectable effect for survey-based outcomes
FaDSS	863	0.197
Goal4 It!	802	0.204
LIFT	808	0.203
MyGoals, Baltimore	749	0.211

Table B.2. Minimum detectable effects on survey-based outcomes, by size of survey sample

Site	Number enrolled	Minimum detectable effect for survey-based outcomes
MyGoals, Houston	1.051	0.178
Work Success	753	0.211

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: $MDE = \begin{bmatrix} T_{df}^{-1}(1-\alpha/2) + T_{df}^{-1}(1-\beta) \end{bmatrix} * \sqrt{(1-R^2)/np(1-p)}$, 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 $\frac{(1-\beta=0.80)}{2}$. We assume 20 percent attrition in the survey data.

Recruitment—Semi-structured interviews. For the semi-structured discussions with program management, supervisors, and staff, we will select program staff and leaders purposively using organizational charts and information on each employee's role at the organization. Purposive staff selection is appropriate because particular insights and information available from individuals will depend on their perspectives based on their role at the organization. The results of the descriptive study are not intended to generalize beyond the program being studied.

For interviews of program participants, the project team will recruit approximately seven treatment group members from each program to complete the interviews among treatment group members who have participated in the program. These interviews are to provide narrative, in-depth context and experiences of program participants.

B3. Design of Data Collection Instruments

Development of Data Collection Instruments

Follow-up survey. The third follow-up survey was developed based on the approved first and second follow-up survey instruments. The third follow-up survey instrument includes only items necessary to achieve the study objectives. It removes items that are no longer relevant 48 months after study enrollment. For example, we have limited items related to service receipt relative to the first and second follow-up surveys because program group members would no longer be receiving services from the employment coaching program at the time of the third follow-up survey. We have pretested the instrument with three people similar to the 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 instrument. We have used cognitive interviewing and respondent and interviewer debriefings during these pretests. We plan to pretest the instrument with one more person. The average survey length to date is 42 minutes, consistent with our goal for survey length. Pretest respondents to date have not reported issues with understanding questions or with survey flow that would suggest revising the instrument.

Semi-structured interviews. The semi-structured interview instruments were developed by content experts at Mathematica and OPRE. The questions in the interviews are designed to collect information

to address gaps in our understanding of how coaching interventions are implemented based on previous rounds of interviewing.

Table B.3 presents a crosswalk between the data collection instruments and the study's objectives.

	Second Follow- Up Survey (Attachment N)	Third Follow- Up Survey (Attachment Q)	Semi-structured Management Interview (Attachment R)	Semi- structured Supervisor and Staff Interview (Attachment S)	Semi- structured Participant Interview (Attachment T)
Objective 1: Provide evidence of coaching interventions impacts on participants' employment outcomes	Х	Х			
Objective 2: Provide evidence on measures of self- regulation	х	Х			
Objective 3: Assess whether coaching interventions are more effective for some groups of participants	Х	Х			
Objective 4: Assess how the impacts of coaching interventions change over time	x	Х			
Objective 5: Investigate factors that could influence implementation of coaching and affect interpretation of impacts			х	Х	Х

Table B.3. Crosswalk Between Data Collection Instruments and Study Objectives

B4. Collection of Data and Quality Control

Follow-up surveys. The second and third follow-up surveys will be available to study participants to complete either through a self-administered, web-based survey or via phone by calling into Mathematica's Survey Operations Center and completing the survey with a trained interviewer over the phone. For respondents who continue to be difficult to reach we will send a trained field locator to facilitate the survey being completed over the phone with the trained interviewer from Mathematica's

Survey Operations Center. Mathematica will conduct trainings for telephone interviewers and field locators and hold refresher trainings as needed. Mathematica will provide staff with various tools throughout the study and periodically conduct refresher trainings as needed. Mathematica will listen to about 10 percent of all CATI interviews to detect inaccurate presentation of information on the study; errors in reading questions; biased probes; inappropriate use of feedback in responding to questions; and any other unacceptable interviewer behavior. Mathematica will provide staff with various tools throughout the study and periodically conduct refresher trainings as needed.

As discussed in section B7, the web survey and the telephone interview software include measures to improve data quality and consistency, such as real-time logic rules, enforced skip patterns, and data checks.

Semi-structured interviews. The project team members will interview program management, supervisors, staff, and participants. Interviews will be conducted over the phone or video and be recorded, with the consent of the staff and participants. The project team members will contact the program leaders to help identify staff to interview based on their involvement in the programs and to schedule the interviews. We will reach out to participants directly to invite them to an interview and then schedule a time that is best for them. To ensure quality and consistency in data collection, all interviewers will be trained. In addition, project leaders will periodically review completed interviews for quality and for missing information.

B5. Response Rates and Potential Nonresponse Bias

Response Rates

Follow-up surveys. The project team will calculate conditional response rates as the number of completed surveys or other data collection instruments as a percentage of the number of people asked to complete the survey or instrument.

The first follow-up survey had a response rate of 68 percent and a difference in research group response rates of 3 percentage points; the response rate is lower than initially expected because data collection activities were affected by the COVID-19 pandemic. The second follow-up survey has a response rate of 73 percent as of January 2022 and a difference in research group response rates of 1 percentage point; the response rate will increase as we continue related data collection activities.

The project team anticipates a 75 percent response rate on the third follow-up survey based on their experiences conducting the first and second follow-up surveys and based on their experience conducting longer-term follow-up data collection for other studies. To obtain these response rates we will continue the data collection procedures developed for the first and second follow-up surveys, adapting them as conditions warrant. Key procedures for the third follow-up survey include:

- Allow respondents to complete the survey in different ways. The participants can complete the survey either online (using a computer, tablet, or smartphone) or by telephone.
- 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 V. For example, the advance letter (and insert) is mailed to participants at the start of data collection. The email notification is emailed to participants who have not yet completed the survey about three weeks after the start of data collection. The refusal avoidance letter is 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 is sent to participants who have not completed the survey after all available contact information has gone through a locating process (described below). The text messages are sent to respondents who consent to receiving text messages and are sent at two weeks and four weeks after the advance letter is sent. We also send specific text messages to respondents on an as-needed basis during the data collection period.

- Obtain accurate, up-to-date contact information. Detailed contact information was collected at baseline that includes telephone numbers, addresses, and email addresses to aid in locating participants to complete the follow-up surveys. Detailed contact information was also collected for three relatives, friends, neighbors, and/or past employers whom the participant selects and who may be able to help locate the participants if they move. Before the start of the second follow-up survey, participant contact information was updated through online database searches. The study team also works with study sites to obtain participant contact information from the programs with a focus on updating contact information for nonresponding sample members.
- Use intensive locating methods, as needed. Participants are initially notified about the survey by mail and email and asked to complete it via the web, though they can also complete it via telephone at that time (Attachment V). When respondents are initially notified, they will receive a token of appreciation to generate goodwill (a \$5 pre-pay enclosed with advance letters, as proposed in Supporting Statement A, section A9). After completion of the survey, they will receive an additional token of appreciation for their time (a \$65 gift card, per Supporting Statement A, section A9). After four weeks, the evaluation team will attempt to contact the participants via telephone at the numbers provided in the most recent locating update, 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/or past employers identified by the participant, for help in locating them. Customized, individual searches for contact information using specialized databases will be conducted next.
- Use paradata. Data is collected on each attempt to contact a respondent including the mode, time, date, interviewer, and contact results. Examining these paradata helps to identify the most effective calling times and interviewers. Paradata is also used to determine which methods of contact (letters, emails, texts, or telephone calls) are 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 are 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.
- Other mitigations to address emerging issues in response rates. For the first and second followup surveys, as it became apparent that survey production in four sites would likely be insufficient to support unbiased estimates of program impacts, the contractor took additional steps to identify causes of non-response in these sites and to mitigate them. The contractor sent in locating experts to work sample in these specified geographic areas with a smartphone to have respondents complete the interview in person. The contractor analyzed the effectiveness of different notifications and mailing strategies and modified these strategies accordingly based on results of the various efforts. The contractor will apply the lessons learned from the first and second follow-up surveys to the third follow-up survey effort by utilizing locating experts and deploying similar notification and mailing strategies to survey non-respondents as needed.

Semi-structured interviews. The interviews are not designed to produce statistically generalizable findings and participation is wholly at the respondent's discretion. Response rates will not be calculated or reported.

Nonresponse

Follow-up surveys. We will conduct the analysis to account for the possibility that data missing due to survey or item nonresponse could introduce bias in the impact estimates and reduce statistical power to detect program impacts. Follow-up survey data could be missing because study participants do not respond to follow-up surveys or because survey respondents do not answer some survey questions.

To account for sample members who did not complete the follow-up survey, we will estimate all impact estimates using weights. The nonresponse weights will adjust the data to be representative of all sample members, not just those who completed the survey or could be matched to an administrative record. We will calculate the weights by estimating, for each program separately, the probability of nonresponse for study participants as a function of their baseline characteristics using regression analysis. We will adjust the standard errors of the impact estimates to account for the variability associated with these weights. In situations in which item nonresponse affects a subset of items used to create survey outcomes, we will use imputation. For example, if a sample member responded to at least two-thirds of the items on a scale, we will use the average scale score for that person based on the available items. We will not impute outcomes that are entirely missing.

In addition to these strategies, we will compare the baseline characteristics of those who are missing a given type of data and those who are not to assess selection into missing data status.

Semi-structured interviews. As participants will not be randomly sampled and findings are not intended to be representative, non-response bias will not be calculated. Respondent demographics will be documented and reported in written materials associated with the data collection.

B6. Production of Estimates and Projections

The estimates from this project will be released publicly following ACF review. The information collected is meant to contribute to the body of knowledge on ACF programs. It is not intended to be used as the principal basis for a decision by a federal decision-maker, and is not expected to meet the threshold of influential or highly influential scientific information.

Impact study. The impact study will estimate the effectiveness of each program in the study in improving outcomes of study participants. Any observed differences in outcomes between the treatment and control group members can be attributed to the effectiveness of the program. These differences are internally valid estimates of the mean impacts of the program, as delivered, on the corresponding outcomes for similar populations in the same environment. The analysis to produce these estimates will be guided by an analysis plan that summarizes data sources, identifies data elements to be analyzed, describes plans for merging data sets, describes statistical models for impact and outcomes analyses, and identifies potential challenges and solutions. The analysis plans for analysis of the first and second follow-up surveys have been pre-registered with Open Science Forum (https://osf.io/znkpu). We will update the plans and registration to include planned analysis of the third follow-up survey.

The project team will use the constructed sample weights described in Section B2 in the impact analysis 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). The project team will also address missing responses as described in Section B5.

The project team has tested for differences in means of key baseline characteristics and confirmed that random assignment successfully generated treatment and control groups with similar baseline characteristics, and that the treatment and control group respondents to the follow-up surveys are similar.

Impacts will be estimated for each program. The project team will use regression estimators to control for residual differences between the treatment and control groups and to construct more efficient estimators than the simple difference-in-means estimators.

To facilitate efficient archiving, we will conduct all data work with archiving requirements in mind. This will involve using systematic variable naming and labeling conventions and detailed documentation of data processing procedures.

Implementation study. The implementation study will use qualitative data methods to analyze the proposed semi-structured interview data, described in greater detail below.

B7. Data Handling and Analysis

Data Handling

Survey data. The web survey and the telephone interview software will use real-time logic rules, enforce skip patterns, and provide soft and hard checks. Soft and hard checks will be displayed for interviewers or respondents if the provided information conflicts with earlier responses or is out of range for expected values. Hard checks require resolution before continuing; soft checks can be suppressed. All CATI interviewers are subject to real-time or recorded monitoring to ensure they are correctly interpreting and entering respondent responses. Following data collection, the project team will conduct comprehensive data reviews and quality assurance reviews to ensure skip patterns are enforced and data are complete and within expected ranges.

During data processing and coding, the project team will conduct quality assurance reviews to ensure consistency and minimize any data processing errors. Specifically, coders will participate in a comprehensive training session, and the project team will monitor their work, perform quality control checks, and conduct quality assurance reviews of all weighting and imputation procedures. Any outliers, skip logic errors, or other recodes of survey data will be recorded in both internal programs and data editing spreadsheets.

The third follow-up survey will also be programmed with Mathematica's Confirmit software. Error messages will be programmed into Confirmit to alert respondents to inconsistencies between data elements, values beyond the expected range, and similar issues. Respondents will have an opportunity to correct such errors before the data are submitted. Surveys completed over the phone will be completed with trained Mathematica staff. The use of a web-based survey eliminates the need for an additional step for data entry, thus minimizing potential errors that may occur during that process.

Once a sufficient number of responses have been received, we will conduct an initial quality check to identify any potential issues with the data. Additional data quality checks will be conducted throughout the study. The study participants will continue to be randomly sampled.

Interview data. A small, trained team will code semi-structured interview transcriptions using qualitative analysis software. To obtain reliability across codes, all team members will code an initial set of documents, after which differences in their coding will be identified and resolved.

Data Analysis

Impact study. We will clean the data collected as part of the second and third follow-up surveys and merge them with appropriate baseline and earlier follow-up data from the Evaluation of Employment Coaching. We will construct the outcomes described in our analysis plans, estimate the long-term impacts of the programs, and conduct any additional analysis described in the plans.

Implementation study. Researchers will reduce the qualitative data to a manageable number of topics and themes for analysis using qualitative analysis software. Using the data collected from the multiple sources, the information will be summarized in tables that synthesize key findings.

Data Use

We will use the collected data to draft multiple briefs and reports. We plan to develop briefs on the results from the additional implementation study data collection. We will write a report on the findings from the impact analysis of data from each of the second and third follow-up surveys. We will also write some program-specific briefs on the results of the impact analysis.

We will develop a data archive that will allow other researchers to replicate and extend analyses conducted as part of the impact study; qualitative data collected as part of the implementation study will not be archived. We transform the analysis data into a format that can be used by a variety of statistical software packages and accompanying codebooks and supporting information to facilitate other researchers' use of the data. The final data files will contain enough information to allow duplication by other researchers of all the analyses we conduct.

B8. Contact Persons

ACF

- Victoria Kabak, Victoria.Kabak@ACF.hhs.gov
- Sarita Barton, Sarita.Barton@ACF.hhs.gov
- Lauren Deutsch, lauren.deutsch@acf.hhs.gov

Mathematica

- Quinn Moore, QMoore@mathematica-mpr.com
- Sheena McConnell, SMcConnell@mathematica-mpr.com
- Kristen Joyce, KJoyce@mathematica-mpr.com
- Shawn Marsh, SMarsh@mathematica-mpr.com
- Ayesha De Mond, ADeMond@mathematica-mpr.com

Attachments

Attachment N: Second Follow-Up Survey Attachment Q: Third Follow-Up Survey Attachment R: Semi-structured Interviews for Management Attachment S: Semi-structured Interviews for Staff and Supervisors Attachment T: Semi-structured Interviews for Participants

Attachment U: Third Follow-up Survey Question-by-question Justification Attachment V: Notifications