

## Evaluation Plan: Impact Evaluation for Hypothetical AmeriCorps Program

**Introduction:** This evaluation plan describes an impact evaluation for Financial Empowerment Corps (FEC), an AmeriCorps State and National (ACSN) program dedicated to improving the lives of vulnerable community members through financial education and empowerment. Through this evaluation, FEC hopes to learn more about the effect of its financial education and empowerment curricula on low-income clients.

**Program Background and Problem Definition:** Saving and managing money are important skills that many individuals lack. The inability to understand and manage debt, income, savings, and investments can make a difference between enjoying a comfortable life and living paycheck to paycheck. While some may be able to afford the services of a professional financial planner or manager, many Americans, especially those who are low-income, must manage their own personal finances. Knowing the basic principles of money management can help low-income or disadvantaged individuals climb out of poverty, insulate themselves from economic shocks, and plan for a comfortable future. Unfortunately, these skills are not taught in many schools and are not offered at an affordable price by lending institutions. FEC seeks to fill this gap with financial education and empowerment administered through local credit unions, lending organizations, and financial institutions.

FEC is a 10 site ACSN program in the greater Cincinnati region that places members in credit unions and local financial institutions to provide financial counseling to low-income individuals, retirees, and young people. Members also assist with financial seminars and informational fairs, and recruit experienced financial professionals to serve as volunteers in a credit counseling program. In 2014, FEC will have 5 full time AmeriCorps members at each site for a total of 50 members. **This evaluation will only focus on the financial counseling component of FEC members' work.**

Members are matched with clients on a first come first served basis by the site location's AmeriCorps program manager. Before their first meeting, clients take a short survey to determine their needs; this is administered by the local credit union or financial institution. Members then build a personalized education plan based on FEC's financial education and empowerment curriculum, which is administered over the course of twice weekly client meetings for 2-4 weeks. For example, many members work with clients on issues related to student loan and credit card debt. Using the core curriculum materials developed for student loan and credit card debt education, members work with clients to develop a savings and debt management plan, as well as develop realistic strategies to prevent further accrual of debt. According to a site survey conducted annually, most members work with clients to address issues of student loan and credit card debt, family budgeting, and basic saving for retirement.

If there is a severe or persistent need, members refer clients to FEC volunteers, who are recruited based on their careers in the finance industry, or host site staff for additional services. Each member has a maximum caseload of 10 clients at any given time.

**Commented [A1]:** Providing a thorough narrative description of your program is important in preparing any evaluation plan because it defines what will be studied. A clearly stated, well-defined program design will enable you to make decisions about what components of the program to focus on in the evaluation and to select evaluation objectives that are feasible. It sets a common understanding of what the program is (and is not), what resources and components are utilized, what products are produced or activities conducted, and what is expected to happen as a result of those activities.

In this example, we've given a high-level overview of the entire FEC program in the first two paragraphs, and then the third and fourth paragraphs provide more specific detail on the component of the program that will be the subject of the evaluation.

**Commented [A2]:** Defining the problem your program addresses rationalizes the intervention you implement. Your intervention should logically follow from your problem statement; alignment between your targeted problem and intervention is critical for producing the change you desire. In this example, the problem definition comes before the program background in order to make the connection between the problem and intervention more clear.

**Commented [A3]:** This sentence sets an important boundary around the evaluation, limiting the scope to just one component of the program. This is important for managing expectations amongst stakeholders and for focusing research questions. It also helps manage limited evaluation resources (both financial and human capital). For more resources on determining an evaluation's scope, see the CNCS Core Curriculum webinar on "Managing an External Evaluation".

## Past Research and Existing Evidence<sup>1</sup>

### Program Theory, Logic Model<sup>2</sup> and Outcomes of Interest

FEC's theory of change holds that individuals can most efficiently and effectively manage their money if they are educated and empowered with the relevant knowledge to do so. Low-income individuals in particular benefit, as they are not able to access these resources from professional financial planners and managers. Research shows that one-on-one, intensive counseling, using a curriculum constructed based on industry best practices, is the best way to educate and empower clients because it allows for personalized instruction that adapts to client circumstances and needs, challenges, and pre-existing knowledge. Multiple counseling sessions allow for the development of a learning plan that can comprehensively address clients' diverse and numerous needs, and allow for application of some of this knowledge in real time; clients can then troubleshoot emergent problems with their counseling member.

The short term outcomes targeted by FEC are that clients will increase their knowledge of basic components of personal finance; and that they will gain the knowledge and skills to address a current financial challenge they are experiencing. Medium and long term outcomes targeted by FEC are that clients will apply the education learned through the counseling program to make significant progress towards solving, or will solve, a current financial challenge.

### Research Questions to be Addressed in the Study

This evaluation will address the following questions:

1. Do low-income clients exit the program with increased knowledge of personal finance concepts relevant to their needs? Do they know how to use those concepts to address their financial challenges?
2. Are low-income clients that participate in member-led financial counseling through FEC able to better manage their personal finances than low-income individuals who did not participate in FEC counseling?

### Study Components

#### Evaluation Design and Rationale

<sup>1</sup> This is where you will record and describe past research conducted by your organization, academics and researchers, or similar programs that supports your intervention. Describing the evidence base of your program puts the present evaluation into context, situating your research questions and evaluation design amongst those already addressed. This information demonstrates how the current study will build on past work, avoiding duplication and providing useful information to your organization and to others with similar interventions.

<sup>2</sup> You should attach a logic model and theory of change to your evaluation plan. You may use the logic model submitted as part of your application for funding.

**Commented [A4]:** Including a summary of past research and existing evidence helps document and demonstrate how the proposed evaluation will build the program's evidence base. It also provides context for your research questions and design, especially if the proposed evaluation builds directly off of previous studies. Once you write a strong summary of evidence, you can copy and paste it into other evaluation plans, updating as needed.

**Commented [A5]:** A program's theory of change is the general underlying idea of how you believe your intervention will create change and why the desired change is expected to come about. Your theory of change articulates the assumptions underlying your choice of activities. A good program evaluation should test some part of your theory of change. It is important to note that in this evaluation, the program will only be evaluating a narrow segment of this theory of change.

**Commented [A6]:** Logic models can serve as a framework for your evaluation plan by helping you can make informed decisions about what to evaluate, when to evaluate, and how you will evaluate. By examining the different segments of the logic model, and their individual components, you can determine your research design, identify research questions, and set the scope of the evaluation. In this impact evaluation, the data collected will focus on understanding the relationship between the program's outputs (e.g., # of counseling sessions held) and outcomes (specifically, increased knowledge and changed behaviors). For more on using a theory of change and logic model to inform an evaluation plan, see the Logic Model Core Curriculum webinar on the AmeriCorps Resources page.

**Commented [A7]:** Research questions should be clear and measurable, and should reflect the desired scope of the evaluation. In this example, the program has chosen to focus only on impact questions, or questions that are concerned with the program's effect on outcomes. Impact evaluation questions seek to examine cause and effect, and focus on attributing the changes in outcomes directly to the intervention. Research question 1 is concerned with measuring a change in knowledge, while research question 2 seeks to measure a change in behaviors. These are separate questions because they are two different outcome areas that need to be assessed differently. For help on creating good research questions, see the CNCS Core Curriculum webinar on "Asking the Right Research Questions".

**Commented [A8]:** It is especially important when designing an impact evaluation that the selected evaluation design follow directly from your research question(s). Since impact evaluations seek to attribute results directly to the intervention, you need to include a comparison or control group, which helps to rule out other explanations for observed changes in the group that participates in your program. An evaluation that does not utilize a comparison or control group leaves open the possibility of alternative explanations for why change might have occurred. For more information, see the Core Curriculum webinar "Overview of Evaluation Designs".

The design chosen for this study will be a randomized control trial (RCT) using a pre and post-test assessment. Services to control group clients will be deferred for six months after being assigned to the control group; once that time has elapsed, these clients will receive regular counseling services.

An RCT is the strongest impact evaluation design available to answer our research questions because it effectively controls for client characteristics that may contribute to differential outcomes between the treatment and control groups. This means that any differences detected in the evaluation can be attributed to the FEC program. An RCT is preferable over a quasi-experimental design in this case because clients are not assigned via “triage” to a member based on severity of need<sup>3</sup>, and there are consistent waiting lists (of a minimum of four months) at each of the 10 sites. Because clients submit detailed contact information upon entrance to the program, they are easily tracked and will therefore be easier to follow over time. Additionally, the presence of a waiting list for services means that FEC will not need to identify, recruit, or incentivize comparison group individuals from across the region; we will simply use clients that would participate in the program anyway.

The study will span 12 months of data collection, and will use a pre/post-test to assess changes in knowledge and behaviors. Clients will fill out the usual intake survey and an application, as well as a supplemental set of items asking about their knowledge of basic personal finance concepts, including specific sets of questions regarding the three major content areas identified as critical for financial success through the site supervisor survey (student loan and credit card debt; family budgeting; and basic saving for retirement). The supplement will also ask them about current financial behaviors, such as saving, spending on credit cards, current levels of debt and loans, and investment in retirement accounts. Demographics and background information will also be solicited; we plan to ask clients about their preferred method of contact in order to maximize response to the post-test. The supplement will be administered at the end of their counseling period as well (approximately 1-2 months after pre-test is administered). After 6 months have elapsed after their counseling services end, treatment clients will receive a second post-test survey administered via phone. Each client will receive up to 2 phone call follow up calls from the evaluator to attempt the survey. Clients assigned to the control group will receive the same pre-test and intake survey at the time of assignment to the control group. They will also receive the post-test after 1 month and 6 months of being deferred.

The main strength of this design is that it makes efficient use of extant data collection tools and processes in place at FEC sites. It also makes efficient use of site supervisor and evaluator time, and minimizes the burden on clients. Finally, it contributes to building long term administrative data that FEC can use to analyze changes in outputs and outcomes over time.

<sup>3</sup> Triage refers to the practice of prioritizing which clients receive services first based on severity of need. FEC prioritizes clients on a first come, first served basis. Clients with severe or time sensitive needs are referred to a financial professional or legal counsel, and do not work with members.

**Commented [A9]:** A randomized control trial (RCT) is an experimental evaluation design that uses random assignment of participants to treatment (those who receive the intervention) and control (those who do not receive the intervention) groups. For this evaluation, pre and post-intervention measurements are taken of both groups at approximately the same times on relevant measures of outcomes, and the differences in outcomes between the groups, or lack thereof, are statistically determined. A well implemented RCT will allow you to attribute any effects on participant outcomes directly to your program. The main strength of an RCT lies in the use of random assignment, which allows any other potentially confounding factors between the two groups to be treated as occurring at random. In other words, random assignment removes any systematic differences between those individuals who are participating in the program and those who are not participating. Therefore, any detected changes in participant outcomes are attributable to the intervention and not to some other confounding factor.

**Commented [A10]:** For many programs, withholding or denying services to participants is either not possible or against the ethos of the organization. Nevertheless, you can still utilize randomization by temporarily delaying services. Rather than turn away clients, you can randomly assign clients to treatment or a waitlist; the waitlist clients simply get the intervention at a later date.

**Commented [A11]:** For this particular program, using randomization to form a control group is the most rigorous design feasible because it provides the necessary counterfactual (what would have happened to beneficiaries in the absence of the intervention) group of untreated individuals against which to measure the outcomes of the treated individuals. It allows inference of causality, and flows directly from the research questions. While the program could have chosen to implement a quasi-experimental design (QED) using a comparison group of similar individuals on the current waitlist or who might be referred by community affiliates, in this case, an RCT is preferred because is easier for the program to implement, will be cheaper than recruiting comparison individuals, and syncs well with the organization’s research and evidence goals. A well implemented QED would provide comparable results, but also requires more attention be paid to potential systematic differences between the treatment and control groups.

**Commented [A12]:** This is important not only because evaluations for CNCS must include at least one year of program data, but also because it allows for full implementation of the program from start to finish.

**Commented [A13]:** This was determined to be sufficient time for medium term outcomes (the behavioral changes related to implementing a budget, reducing debt, and investing in retirement accounts) to occur. It is possible that it could take longer though, and determining the appropriate measurement interval is critical to making sure you can detect an effect.

There are several limitations. First, the success of the RCT hinges on implementing the randomization process properly. We plan on reserving a few exceptions to the randomization at each site to build in flexibility and recognize emergent needs of clients. Our evaluator has determined the appropriate number of exceptions that will not affect the integrity of the final analysis. Additionally, members and site supervisors will be oriented to the evaluation process and trained to implement randomization, including emphasis on the importance of not deviating from the randomized assignments. Second, the diversity of client problems could be a challenge to accurately recording information on the pre and post-test. We have attempted to minimize issues by using historical data from our site supervisor survey to create 3 different sets of questions for the pre and post-test that capture the vast majority of need our clients report upon intake (student loan and credit card debt; family budgeting; and basic saving for retirement). A screening question on the pre and post-test surveys will ensure clients answer the most relevant supplemental questions related to their specific challenges. Another challenge will be the natural variation in member interaction with different clients. Though this is controlled to some extent by the limited number of one on one sessions and length of those sessions, some clients and members utilize their time together more effectively than others. Members will submit their activity logs, used currently for performance measures purposes, to track the intensity of interaction and time spent with each client. Finally, attrition from the study might be a problem, especially for those deferred from services as part of the control group. To minimize attrition from the control group, we are guaranteeing control group participants a spot in the program after 6 months has elapsed. We will also offer a supplemental financial planning kit that clients can utilize while they wait for program services. Given that the program typically has a waitlist of a minimum of four months, a fact well known among potential clients, we do not anticipate control group attrition to be a significant problem. Attrition from the program itself is extremely low, due to its popularity, and therefore we do not anticipate attrition from the treatment group to be a problem. We expect that informed consent procedures will address any fears that members of either group have that might otherwise cause them not to participate in the study.

#### *Sampling Methods, Measurement Tools, and Data Collection*

To achieve adequate sample size, all 10 program sites will participate in the study. Our external evaluator has conducted a power analysis and determined an adequate sample size of treatment and control group members<sup>4</sup>.

When the study begins, the evaluator will determine the mechanism for random assignment of clients. When the site supervisor conducts the intake screening, they will apply the random assignments to clients to sort them into the treatment or control group. Members of both groups will fill out the intake survey and pre-test. Treatment clients will begin counseling services immediately; control group clients will be placed on a waitlist and offered a financial planning kit to be mailed to their home. The kit will contain general information about financial services offered by credit unions and local financial institutions, as well as basic information on personal banking.

<sup>4</sup> It is important to conduct power analyses to determine how large of a sample you need for your study; you would want to include details on the sample size in your evaluation plan.

**Commented [A14]:** Typically, participants are randomly assigned to a treatment or control group using a random assignment generator or algorithm, such as is found in Excel or in online applications. Deviating from the assignments made by the algorithm, such as reassigning a control group member into the treatment group (or vice versa) introduces bias into the study because you are systematically altering the random assignment process. There may be important differentiating factors that then account for observed differences between the groups rather than the treatment. In this example, if the program deviates from the random assignment procedure by allowing several "neediest" clients to participate in the treatment, they may be biasing the treatment group with participants that are systematically different than the control group in terms of characteristics such as motivation or severity of financial distress. Those systematic differences might then influence the effect of the intervention on programmatic outcomes.

**Commented [A15]:** Despite the need to maintain the integrity of the random assignment process, in practice, exceptions may be needed to accommodate the realities of operating a social program. The number of these that can safely be accommodated without jeopardizing the integrity of the study can be determined by an experienced evaluator.

**Commented [A16]:** Supplemental qualitative information that comes from the activity logs will be helpful when it comes time to interpret the results of the quantitative analysis that is part of the RCT. Large variability in the intensity of interaction between members and clients could help to explain null (no effect) findings or variability in findings.

**Commented [A17]:** Incentives are often an effective way to engage comparison or control group participants who otherwise would not be inclined to participate in the study. In this example, the program offers a guarantee of service within a set timeframe, which is attractive given that the program has a long waitlist and does not accept all applicants. Note that since the minimum wait time for applicants on the wait list is 4 months, some control group clients will receive services after others on the waitlist. But, since 4 months is only the minimum (meaning some wait much longer), a guaranteed spot in 6 months is still an attractive incentive to participate in the evaluation.

**Commented [A18]:** A power analysis is a statistical procedure that helps in determining the appropriate sample size for your impact evaluation. After setting a few statistical parameters, a power analysis will produce a figure that represents the necessary sample size of respondents or study participants you will need to be in your study. Power is the ability to detect an effect in your statistical analyses. It is important to ensure that your impact evaluation has enough power so that you are able to detect any changes that are occurring because of your program. Conducting an underpowered evaluation presents the risk of failing to detect a statistical effect when it is actually present; that is, you could miss finding evidence of the program's effectiveness when it does in fact exist. Power analysis can be conducted using a number of software packages or tools, some of which are free and available online.

As described in earlier sections, we will be utilizing our intake survey as well as an existing pre and post-test survey developed a few years ago. Member activity logs will also be used.

## Procedures

### *Analysis Plan*

Survey data gathered via phone will be input into an excel spreadsheet by the evaluator and then transferred to a statistical analysis package.

Our evaluator will conduct the statistical analysis of the quantitative data using regression analysis for the overall treatment group as well as subgroups by type of financial challenge. The FEC program director will review member activity logs and flag clients that seem to have received very high or low levels of service from a member. That information will be sent to the evaluator to help explain any outlier cases.

### Institutional Review Board (IRB) Clearance

Our evaluator will submit an IRB clearance package to their firm's review board. We anticipate expedited clearance due to our rigorous informed consent materials and risk mitigation procedures that we plan to implement. Nevertheless, we have allotted 2.5 months for clearance in our timeline.

### Evaluator Qualifications<sup>5</sup>

### Reporting Results, Timeline and Budget Factors

#### *Timeline*

Our overall time allotted for this evaluation is 3 years. The first year will cover evaluation planning, hiring an evaluator, securing IRB approval, and pilot testing the survey supplements with a small sample of clients from each site. The second year will cover data collection, and the third year will cover analysis and reporting. Below is a more detailed timeline:

- Evaluation plan development- Aug. 2013-Dec. 2013
- Pilot test surveys and supplements
  - Dec. 2013-May 2014
- Hire an evaluator- Jan.-Mar. 2014
  - Vet plan with evaluator- Mar.-May 2014
  - Orientation to program, sites- May 2014
- IRB approval- May-July 2014
- Orientation to evaluation with members and staff- July 2014
- Train staff on randomization procedures - July 2014
- Begin data collection- late Aug. 2014

<sup>5</sup> You can either attach an evaluator resume or bio to the plan, or briefly describe their experience conducting evaluation work similar in size and scope to the current evaluation.

**Commented [A19]:** Regression analysis is a quantitative analytic technique frequently used in outcome and impact evaluations to estimate the relationships between different variables. Specifically, it is used to assess the effects of a defined set of variables (known as "independent variables") on the outcome variable of interest (known as the "dependent variable").

**Commented [A20]:** Running the analysis using only sub-groups of treatment and control participants, for example by gender, will show the effect on that smaller, more specific population. In this example, running the analysis on sub-groups identified by their specific financial challenge (debt; family budgeting; saving for retirement) will explore whether the treatment is more effective for any individual financial challenge. Sometimes, an intervention may be more effective for a more specific sub-set of a problem (e.g. debt as a sub-set of financial challenges), and will produce sub-group effects but not an overall effect. All of the proposed sub-group analyses must be specified in the evaluation plan at the outset of the study; it is not appropriate to "mine" the data by sub-group in order to try to find an effect.

- Continue data collection- late Aug. 2014-late Aug. 2015
- Close data collection- late Aug. 2015
- Clean data- Sept. 2015
- Analysis- Sept.-Nov. 2015
- Write evaluation report- Nov.-Dec. 2015
  - Submit to CNCS for recompetete- Dec. 2015
  - Generate derivative products for website, annual report, other funders- Jan.-Mar. 2016
- Lessons learned and reflections (with program staff and evaluation team)- Jan. 2016

*Budget*

Given the fact that we have a pre/post-test survey already developed, an existing intake form, and member activity logs, we expect costs for this RCT will be reduced. Our data collection and management system is robust and will not need to be modified or enhanced to handle the amount and type of data coming in from the surveys. We also have developed our evaluation plan with the help of volunteers from a local university and a professor working for a significantly reduced price. We anticipate paying mostly for the costs of data collection and analysis/reporting and have therefore budgeted \$150,000 for this evaluation. In addition to CNCS funding, we have secured a number of smaller private donations and unrestricted funds from our parent organization. Roughly 15% of our three-year evaluation costs are allocated to the first year of the study; year two is allocated 60% of costs; and year three is allocated 25%.

*Reporting Results*

We plan to develop an evaluation report to submit with our recompetete application in the winter of 2015 that documents the activities and results laid out in this evaluation plan.

FEC has a number of other constituents, including clients, sites, and other funders, that are interested in the results. A short brief will be developed by FEC staff to disseminate to private donors, members, and alumni. A short summary and infographic will be developed for potential clients and for FEC’s website. A summary of the evaluation report will be generated by FEC staff for attaching to grant applications. Additionally, a lessons learned report will be developed by the evaluator and FEC staff to reflect on the process and plan for the future. Finally, FEC staff will develop recommendations to follow up on the results generated by this RCT. We expect that to include a future study to examine member interaction with clients.

**Commented [A21]:** Note that this is a low-cost RCT due to the fact that the program is making use of a good deal of existing data collection infrastructure. If they did not have an existing instrument, data collection and management system, well-tested data collection procedures, and supplemental data collection efforts in place, costs could easily double. Note that this budget does not account for the cost of staff time required to manage the external evaluator and the evaluation itself. This information should be contained in a more detailed budget. For more information on developing a sufficient evaluation budget, see the CNCS Core Curriculum webinar “Budgeting for-Evaluation”.