Bundled Evaluation Methodology Webinar
Office of Research & Evaluation, Corporation for National and Community Service

BRAD: Welcome and thank you for standing by. At this time, all participant lines are in a "Listen Only" mode. After today's presentations, you will have the opportunity to ask questions and you may do so at that time by pressing star then one if you would like to ask a question over the phone. Today's conference call is being recorded. If you have any objections to this, please disconnect at this time. Now I will turn the call over to your host for today, Dr. Andrea Robles. Doctor, you may begin.

ANDREA: Thanks, Brad. Thanks so much to those on the phone and in the room for joining us in our first research and evidence webinar of 2018. As Brad said, my name is Andrea Robles and I work for the Office of Research and Evaluation at the Corporation for National and Community Service, what we refer to as CNCS. For those of you who are listening to one of our webinars for the first time, I want to mention that our office objectives are to support our agency's mission by building knowledge on civic engagement, volunteering and national service. We conduct in-house research but also fund research to a competitive branch to researchers, scholars, and dissertators at institutions of higher education and support research and evaluation of our programs and grantees. We strive to share and use our research findings in several ways, including research reports on our evidence exchange, which is a website on Nationalservice.gov, an annual research summit, and a quarterly newsletter.

Our webinar series is another way to share cutting-edge research, like what you would hear today. Before I move into introducing our speakers, I'd like to cover a few housekeeping details. All participants will be in "Listen Only" mode until the question and answer session following the presentation. As mentioned earlier, this webinar is being recorded and we will post it online after the presentation.

So for today's webinar, we will have introductory comments by Dr. Cheri Hoffman who wears many hats, including being the Director of the Children and Youth Policy Division at the Department of Health and Human Services, as well as the Chair of Interagency Working Group on Youth Programs, which is a collaboration between 22 federal agencies and offices with the goal of coordinating to improve these outcomes.

This will be followed by a presentation by Dr. Rebecca Frazier, a JDS Research Associate. She led the analysis and report writing for this Bundled Evaluation that you will hear today and provided technical assistance to six programs participating in the evaluation.

Finally, we will have concluding remarks from Linda Cook who is the Senior Program Officer for AmeriCorps State and National and works with opportunity youth grantees.

After the speaker’s presentations, we will conclude with a Q&A. One of our goals is to be able to make our research and findings acceptable and useful to academics as well as practitioners. So please feel free to ask any questions or provide comments. And you could do so through the chat box or, like I said, in the Q&A time.

So now, I will hand this over to Dr. Cheri Hoffman.

CHERI: Thank you so much, Andrea. Thank you to everybody for the invitation to join you here today. I'm very interesting in the concept of Bundling Evaluations. It was a new concept to me when CNCS, Mary Hyde, and others came to our Interagency Working Group on Youth Programs meeting and presented some information that was underway at the time. And it really struck a chord with me. I have some background as a program evaluator working in children's mental health programs before I came to the federal government. And the office where I sit in Health and Human Services is the office of the Assistant Secretary for Planning and Evaluations. And so we're kind of always looking for new ideas and new ways to improve things and have a lot of the same challenges that I think probably a lot of folks on the webinar would say that it's difficult to find large enough sample sizes and to really get to the power for some of the results that you want to be able to show when you're looking at things like effectiveness of programs that you're delivering.

And for the Interagency Working Group on Youth Programs, our audience is a lot of youth driven organizations that are out there. We operate the website Youth.gov, which is sort of a one-stop shop for a lot of information about our 22 partner agencies across the federal government and the things that they are doing with and for youth. And one of the things that a lot of those agencies tell us that their grantees struggle with is program evaluation.

And so we’re always looking for new information that we can put out and new ways we can do that. And also, we’re starting to think as an Interagency Working Group with all of these different partners -- we’ve got everybody from CNCS and HHS, lots of offices in HHS, to the Department of Labor, to the Consumer Financial Protection Bureau. All of us trying to coordinate and collaborate towards one goal in improving these outcomes. We’re starting to think about how do we measure that across all of these different agencies with these different missions and different activities that they undertake to do this.

How can we find common outcomes that we’re trying to achieve, common measurements, those sort of things?

So all of these ideas really started coming together and we got very interested in this. So I have been bugging our friends at CNCS for a while to hear about what’s going on. So we’re really looking forward to understanding more about how it was done. So we think that the methodology that we hope we can share across all of these federal agencies, as well as the youth serving organizations that we try to serve. So I’m looking forward to it and I will turn it over to Rebecca.

REBECCA: Thank you very much. I’m very excited to be here today and to finally have the opportunity to talk about the Bundled Evaluation Project. I want to start by thanking Adrienne Tomasso [ph.] who’s been working on the project for the last four years. This has really been a labor of love and we are excited to be able to share this methodology and in the forthcoming months to also share the results of the study, which will be posted on Evidence Exchange shortly.

But today we’re really going to focus on the methodology and I’ll talk a bit about how you can use the Bundled Evaluation Methodology to translate learning into action, both for grant makers and also for smaller programs and non-profits.

So I’m going to start by talking about what a Bundled Evaluation is. So we’ve thrown that term around a little bit and I want to clarify what I really mean by that. So a Bundled Evaluation approach bundles together small programs that have similar program models and outcomes into a single larger impact evaluation. And that evaluation approach is kind of a hybrid of traditional multi-site and clustered evaluation approaches. So traditionally multi-site evaluation approaches look at the same program model across these tool sites and just do one evaluation of that identical program model, where the clustered evaluation is typically a little bit more participatory and exploratory. So you’re looking at a clustered of similar program models in a more exploratory way.

So this is actually a combination of those two approaches and really designed to investigate impacts. And the way that it does that is by using propensity score matching to create a match group of treatment and comparison kids. In this case we recruited comparison groups both from community partners who are serving similar youth and also from program applicants who were eligible for the program but did not ultimately enroll.

We used that to create a match comparison group and then we used multi-level modeling to account for differences across the site and the nested nature of the data.

This approach is really informed heavily by participatory evaluations. The CNCS from the very beginning wanted to make this an evaluation capacity building exercise. So now just a rigorous evaluation by an outside third party, but also an evaluation learning opportunity. One in which programs could build evaluation capacity through experiential learning. So we had sites participate in comparison group recruitment and also in the data collection process throughout the course of the evaluation.

So for this particular evaluation, CNCS chose to focus on efforts with opportunities. So we’ve defined opportunities as young people between the ages of 16 and 24 who are low income and either homeless, in foster care, involved in the juvenile justice system, unemployed, or not enrolled in or at risk of dropping out of an education institution. So formerly, these might have been called disconnected youth. But we really view these as youth who present an opportunity for change.

And according to recent estimates, there are more than 6.7 million opportunity youth in the U.S. currently. And CNCS really has begun recruiting those opportunity youth as AmeriCorps members and engaging them in national service projects. And so we were interested in looking at the impact of that participation in national service on use outcomes. And so for the study we wanted to investigate whether or not opportunity youth who participate in

AmeriCorps show improvements in three domains: in education, employment, and in civic engagement. And whether those improvements were observed relative to a matched comparison group?

So to participate in the study, we actually recruited 19 AmeriCorps sites from across the country. And I heard, Linda, that you deserve some credit for that recruitment process. Thank you so much for enlisting the support of our programs. We actually have a mix of urban and rural sites. And program sites ranged in size from five opportunities members up to 85 opportunities members. So kind of a pretty diverse pool of sites for just stating in the evaluation. And we had around 1100 youth who participated in the evaluation. And we administered the survey at three time points so youth participated in the survey at the start of their term of service, at the conclusion of their term of service, and then six months after their term of service.

So this process really took place over a four-year period. So we started in 2014 with a feasibility study, really just looking at whether this was a viable approach. CNCS started the process really interested in building evaluation capacity and figuring out a solution to this problem of smaller sites who had limited evaluation capacity and resources. So we did this feasibility study. It seemed like sites were interested. The opportunity youth bundled seemed like a good candidate. And so we moved forward in the next year with planning and so all of the sites were involved in the planning process. They provided input on measures for the study. They also helped to develop plans for how they would implement the survey with their youth and also comparison group recruitment strategies. And they began the pre-test data collection in 2015. Again, with the pre-testing and match rate, our youth were beginning their term of services with AmeriCorps.

In 2016, we had mastered the post-test so that was used for as youth are finishing their term of service. And also administered six-month follow-up data collection. So reaching out to youth, seeing how are things going in your life now that you've exited AmeriCorps?

And then in 2017 we completed the follow-up data collection, cleaned the data, analyzed it and authored the final report, which, as I mentioned, is under review at CNCS and will be available shortly.

Okay, so now that I've told you a little bit about our general approach, I want to talk some more broadly ono how you can take this approach and assess "Is this a good fit for me and for my organization? And how could I implement this approach to really promote learning among my grantees or among other organizations in my community?"

So I'll talk about how you can assess if it's a good fit first for the evaluator because this is definitely a project you're going to want to do with an external evaluator. And then how to assess if it's a good fit for a potential program who could be a part of the bundle? And then we'll talk about some strategies for using it to promote learning both by involving programs in the evaluation design and through the use of learning communities. And I'll share a few specific tools that you could use to promote evaluation capacity building and learning.

So how do I assess if the bundled evaluation is a good fit? Well, first off, like I said, you're going to need a strong evaluation partner. And so if you are an evaluator, these are some questions you might think of for yourself. If you happen to be a funding agency, these are some questions that I would encourage you to think about as you're hiring an evaluator, which is really one of the first steps of this process.

So, first and foremost, you want to have somebody who has really good statistical and subject matter knowledge because you're dealing with data from lots of sites, with very different interests. You need to be able to statistically control for those differences so you want someone who has that statistical knowledge. You also want someone who has a good working understanding of the interests of the programs. So if you’re dealing with opportunity youth, it helps to have somebody who’s really familiar with that population and the challenges or the sites you’re really going to be encountering on the ground. Someone who’s very relatable in that way.

Additionally, when you choose an evaluator, you’re dealing with a lot of sensitive information, essentially personally identifiable information. So you want to have someone who has access to really good data management systems. We’re talking about really good longitudinal data tracking of personally identifiable information. So good software in place for that. And then also good access to statistical analysis software to run the kinds of complex multi-level models that you’re going to need for this type of evaluation.

Also, flexibility is super important. So we have 19 sites. Everybody has a different interest, different challenges that they’re experiencing on the ground and so rally being able to adapt to that in the moment is really important. And then I think having a team is also really important. Because this is really -- like I said, there's a lot of complexities to the data collection with a multi-site evaluation like this. And so you want to have folks who can do that follow-up, who can quality administer the surveys, and then also having like your more advanced subject matter experts and your statistical folks as well. So having like a real team effort.

And then last but not least, a lot of time. So like I said, this was a four-year process. In selecting an evaluator, you want to look for somebody who you can work with for an extended period of time.

Now once you select an evaluator, you need to be able to shift gears a little bit and think about, "Okay, how do I figure out which program to put in my bundle?" Right? And the first and probably most important thing is to really figure out programs that are similar. It doesn't make sense to have these programs in one bundle, right? And we said that this approach looks at programs that are similar but have slightly different program models. So they need to be able to at least agree upon a set of shared outcomes. And so all the sites who participated in our evaluation worked together to build a collective logic model and to agree on a set of core outcomes that they were willing to track collectively across the bundle. And so it’s really important that there be similarity across the program.

Also, this is something that we hadn’t really anticipated at the beginning, but it really helps to have programs that have consistent start and stop dates across your bundles because logistically it can be very challenging to manage data collection if folks are entering the program and exiting the program on a rolling basis. And so, we think this methodology is great, but it seems to be less suited to programs that have a little bit more consistency so that they can manage the data collections. Because places with rolling enrollment really struggle with data collection because it was a large burden for them to be constantly administering the survey.

Also organizational support and capacity -- I cannot emphasize this enough, like really having buy-in from all levels of the organization is essential to success for this type of evaluation. Because you have so many different stakeholders, you'll really need to have a commitment not only from the executive leadership, but also like from the folks who are going to be on the ground working with the youth. This is something that's valuable and important and that we're all in this together. So really making sure that you have that organizational support. And also capacity. It does take a lot of time to implement this. And if you want programs to learn through the experience, they have to have the time and money to devote to the experience. To be able to participate in the data collection themselves and that takes time.

Also, I think for programs, it's really important to have a clear member management and tracking system. So this is a longitudinal study and so we need to be able to find them six months later. And so you can't work with programs who don't really have systems in place to do that. And also, we are interested in being able to understand how aspects of program implementation influence impact. And so it’s important that programs are also documenting what types of activities and programs individual members are receiving. So what did Bob actually do? Did they do job skills training or not? And they need to have a system in place to do that so that when we look at impacts, we can say kids who participated in X, Y and Z activities are doing better than kids who did not. And not all of the programs that we worked with were really prepared for that level of member data management.

And then last but certainly not least, it really helps to have a clearly agreed upon comparison group. So in our study we used two different types of comparison groups. We used program applicants and youth from community partners. So for sites that are working with program applicants, we really encourage you to be sure that they have a documented history of over subscription. Not just saying like, “Oh, sure. I totally have over subscription.” Because when push comes to shove, a lot of sites thought that they would have more members applying but actually didn't really have a sufficient pool for comparison group improvement. If they don't have that pool really feeding to programs honestly about the importance of over recruitment and having sites commit to being willing to over recruit, even if that means that they have to turn some youth away.

And then for community partners, really making sure that 1) the youth who are at the community partner are really program eligible youth. So they are youths who would be eligible to participate in your program. So are they similar enough to really be a valid comparison group? And then does that community partner have the staff capacity and commitment to implement data collection themselves? And that was really challenging because those sites weren't receiving any incentives to proceed in the study. And so trying to figure out how to get the community partners to commit to data collection over a longitudinal period was really challenging. So definitely thinking about that from the get-go is helpful.

Okay, so once you figure out how to get your evaluator and you have all the programs in your bundle, how do you make it participatory? How do you make it more than just an external evaluation that somebody is telling you to do? And so I think the way to do that is to involve programs in the evaluation design and the execution. So in terms of design, we invited programs to provide feedback on the survey measures and we also in terms of execution really set clear expectations and check-ins where we were constantly communicating with the programs so that they had input throughout the evaluation process.

We had each of the sites identify their own comparison groups. So they said, you know, “Here’s a similar group of kids in our community that we think is a good fit for us.” And they also developed their own individualized plans for survey implementation. So “Here’s when we’re going to collect the data. Here’s how we’re going to collect it.” And they participated in the data collection process. So they were on the ground building their own evaluation capacity by participating in the data collection.

Additionally, we were really interested in creating a learning community for evaluation participants. So we had group TA calls where we joined all the participants together to ask questions about how the evaluation was going, to share our challenges, and also ideas for strategies to improve recruitment or attention throughout the study. And we also provided folks with individualized feedback about how they were doing a survey progress. So we sent them email updates where we said, “Here’s how much data you’ve collected. Here's how much you have to go. Here’s how you compare to everybody else in the evaluation.” So they could sort of see how they were doing and feel like they were a part of this larger evaluation process.

Additionally, we created individualized results reports for each of the sites. So after each study time point, so after the pre-test, the post-test, and the follow-up, each individual site got a results report, which I'll share in a second, which showed information about their youth to participate in the study. So that they weren't just throwing their data into the collector holder. Also getting some feedback about how they were doing specifically.

And then last but not least, we facilitated connections between the programs. So we said, "Hey, Bob and Susie, you guys are both really working with a lot of rural youth. Let's figure out are there ways that you could combine your efforts to serve that population more efficiently? So facilitating connections throughout the process so that they could build communities that would last longer than the evaluation.

And then I wanted to share a few of the specific methods that we used for providing individualized feedback and evaluation capacity building services. So as I mentioned, we developed these individualized results reports. So this is just like the first page of one of those results reports. But it had a very [unint.] demographics for all of these who participated in the study so they could say, “Okay, like how does this compare to what I thought my members looked like? Does this seem representative for me?” And we also had individualized calls with each of the programs to review the results reports with the evaluators so that they could get some feedback about, “Okay, does this look like what you’re expecting it to look like?” And actually, a lot of the programs use this information to try to actually inform their future recruitment efforts. So when they got the feedback, some of them said, “Oh, actually I thought I was serving more African-Americans than I am. I’d like to really [unint.] that out.”

So we developed those individualized results reports and they also included information about outcomes. So looking at changes from pre-test to follow-up, there were no inferential statistics with this because they’re pretty small sample sizes. But we did want to at least share some individualized information back with each of the programs about how their kids were changing over time.

And then we also gave programs this checklist of evaluations and sustainability strategies where we said, “Tell us what you’re really interested in. What do you want to do? What are your evaluation priorities moving forward?” And we had them select which topics they were most interested in and then we provided individualized TA on each of those topics specific to the program.

Now just a little chart here where you can see like of the topics like what things people are most interested in. And most of the TA that we provided was really around data collection and maintenance. So folks were really interested in building out these data collection systems. Like I said at the beginning, like really establishing a data management system for their program. And then the second most common topics were around recruitment and retention, specifically around opportunity. So a lot of programs were new to the opportunity population and wanted to understand how they could intensify their recruitment and retention efforts for that group.

Okay, so just to put it all together and recap what I’ve talked about today. So how do you assess if it’s a good fit? So first, when you think about an evaluator, you want to select an evaluator that has the relevant statistical and subject matter knowledge, someone who has organizational and interpersonal skills and the right tools and capacity to implement this kind of complex evaluation. And then you want to select programs that have similarity, organization support and capacity, an agreed-upon comparison group, consistent start and stop dates, and then well-established member management and tracking systems.

So once you decide that it’s a good fit, how can you use it to promote learning? So the things we talked about today would be involving sites in the planning and execution of the evaluation, creating a learning community, providing individualized feedback throughout the evaluation process, helping programs to identify their own ECB goals, and then providing relevant resources to help sites meet the evaluation capacity building tools.

So thank you all so much for joining us today on the call. I'm going to hand things over to Linda to provide some concluding remarks. And then we will open the floor up for discussion and then Q&A.

LINDA: Thank you. I appreciate the opportunity to provide a few brief remarks. I like the concept of building learning communities. In any funded environments, whether it's public or private money, it is critical for programs to be able to demonstrate impact. It's wonderful to create opportunities where we can learn best practices and share resources as we are today.

National Service creates a wonderful opportunity for communities to solve pressing community needs. Everyone has something to offer. Programs that recruit and involve opportunity youth play a dual role. They provide mentoring and support to a young person, who in turn uses that growth to impact their communities. If done well, the impact can be enormous. Through service opportunity youth who receive mentoring, coaching, or skills training are positioned for success after their service ends.

The lessons we learn this evaluation will give us valuable insight into how we can continue to support this important service opportunity. Thank you.

AMDREA: Thank you. Thank you, everyone. So what we’ll do is we’ll open it up for Q&A and any comments from people in the room. And so, Brad, would you let everyone know what they need to do?

BRAD: Certainly. If you would like to ask a question over the phone at this time, please press star, then one, please unmute your phone and record your name at the prompt. If any time while you’re in the queue and your question has been answered, you can remove your request by pressing star two. Once again, that is star one for questions over the phone at this time.

ANDREA: Thank you. While we want for any comments, is there someone in the room who would like to follow-up? There are probably a few questions like on that screen that [unint.] that we can share. If you want to start with those or we can –

CHERI: I have one question. This is Cheri. You mentioned that it was really important to have the common outcomes that folks were measuring, but how important was it that the program activities were really aligned towards the outcomes that people were trying to achieve?

FEMALE: Yeah, I think that’s a really good point. So I think definitely having alignment and program activities is essential because otherwise if they’re looking at common outcomes, but they’re not actually doing the same thing, then you can’t make causal claims about what are the impacts of what they’re doing. So definitely having alignment. But within the bundle, for example, there were some sites that maybe were a little more focused on employment and others that are maybe a little more focused on educational outcomes, for example. And so I think it’s important to have like a minimum criteria of like okay, yes, this is something that we’re working on and we do have some shared activities. But I think it’s also okay for there to be some variability in terms of how those activities are implemented across sites.

ANDREA: Thank you. Okay so now I see the questions. So let's see, here's a question from Stephanie Shulie [ph.]. Are there existing member management and tracking systems that you would recommend for this type of effort?

REBECCA: We actually -- so as a part of this process, we developed a handout with resources for conducting a bundled evaluation. We can make that available on [unint.]. And that handout has some really great information about member tracking systems. I don't want to recommend a particular brand because we're not at a position to endorse that. But I do have some articles that are on that list that would highlight pros and cons of different member management and tracking systems. There's a great article from the Bill and Melinda Gates Foundation that talks about pros and cons of different member management tracking systems for non-profits. And so we can share that. That would be great to share.

ANDREA: Thanks. So our second question here is what is an estimated cost for such an evaluation, total and for agency? I think you said there were 19 different sites or 19 programs.

REBECCA: Right. There were 19 different sites.

ANDREA: Do you want to talk about the cost of the evaluation?

REBECCA: I'm sorry, say the question again.

ANDREA: So what is an estimated cost for such an evaluation total and for agency? I think you said there were 19.

REBECCA: 19 sites, total and for agency. I'm not sure what the distinction is between that.

FEMALE: I think they're talking about the program level cost for the 19.

REBECCA: Right. So the programs didn't get offered an evaluation because this evaluation was funded exclusively through CNCS, although I will say that programs did contribute in terms of time and effort.

FEMALE: Do you have an estimate of how much time it took staff?

REBECCA: We actually did ask them that on our last group call with the programs and they said I think the estimate was around 10 hours per week during data collection periods. So sometimes a little bit more than that when they were in the midst of the data collection. But 10 hours on that route.

FEMALE: And just to the cost question, I would say that, as Becca said, the costs came out of the budget of the Office of Research and Evaluation and costs varied by year. So certainly, the years in which the intensive data collection was going on and the years in which intensive technical assistance was happening were more, obviously, costly than years when it was more about planning or writing up final reports.

Overall, I’d say over the four-year period, it was about a million dollars that was the cost of the project and, again, financially the participating sites did not have to pay for that, but certainly people had to invest their time. It was, as Becca said throughout her presentation, designed to be experiential and participatory. And as such, it really did take time. So there was certainly cost on that end. But financially, that was a cost that was borne by the agency.

ANDREA: Brad, do you have any questions on the phone?

BRAD: At this time, we have had no questions on the phone, but, once again, if you would like to ask a question, please press star then one.

FEMALE: I have a question -- actually, I have more than one question, but I can hold as other questions come up. So, one, I wanted to know of the a) fascinating talk, thank you. I've always been hearing this bundling and I wanted to know more about it. Really neat. So of the 19 sites that participated, did they stay with you all through the four years? Did you have any that dropped off? Was 19 the final number or the number you started with and ended with?

FEMALE: Yeah. So we had 19 sites which started and ended. Right?

FEMALE: There were ones that came back [unint.].

FEMALE: Yeah, but I will say that as the process went on, there were some sites that became kind of less willing to engage and communicate. I mean, that is something, I think, you know, making sure when you select the sites at the beginning like really communicating like this is not just like a super fast process and that there is an expectation that you're going to be involved over an extended period of time. I think it's really important. But all 19 sites did stay through the end of the evaluation.

FEMALE: That's very incredible.

FEMALE: Yeah.

FEMALE: And then if we don’t have any other questions. I was just curious is you could maybe talk about I’m sure there have been umpteen challenges that you came across in this whole process that maybe like one or two of the biggest challenges you guys encountered while doing this?

REBECCA: I think someone actually asked a similar question on the online chat. So having a question saying what were the major challenges in the bundled evaluation? How did you overcome them? And I think honestly -- well, okay, so there are a lot of things that are challenging about this type of multi-site evaluation. One of the things that I think was a big challenge was working with opportunities, they have a tendency to move, to drop out, to not be in contact. And so retention was a challenge in the current evaluation and trying to maintain that continuity over time.

I mentioned the number of management systems were also sometimes not as developed as we had hoped for sites. And so having really good tracking in place to stay in contact with youth and encourage them to participate, especially for a comparison group sites, that was really challenging. So it's one thing if you're an AmeriCorps grantee and we're working with you really directly. You have an incentive to participate in the evaluation. If you're a community partner and you're not receiving any funding, the incentive to go track down Joe Smith is not necessarily there. So that made it challenging to maintain sample sizes over time, and particularly among comparisons. So that was really tough.

Another thing I think is the rolling enrollment piece is really challenging. I alluded to that. That's one of the things that we were like "We would probably not do this with rolling enrollment." Which is unfortunate because they're doing great work. It's not that we think that we don't want to evaluate them. But just logistically, if you are a non-profit and you have two staff members and I'm asking one of those two staff members to administer a survey to everybody who applies to your program over a six or seven-month period, that's a big ask. So I think it would be much easier for a program to have a really clear start and stop date. So it's like, "Okay, we're doing this." So only having to do it once. You could do it maybe twice.

So those are a couple of the challenges. Adrienne, do you have other ones that you wanted to --

ADRIENNE: The only one that I would add and it's related to what you just said is organizational capacity. So when we started this, we obviously wanted to select grantees who had a basic level of capacity to be able to participate, but not so much capacity that they wouldn't benefit from the experiential piece. And so in trying to get that sweet spot, I think we may have included some programs who didn't meet that baseline criteria of having the capacity to really participate well. So, as Becca mentioned, programs with two staff or programs with one staff, programs with really high turnover, it's a problem in non-profits in general. But I think it tends to be a bit of a bigger problem for some of our AmeriCorps national programs. The infrastructure within these organizations a lot of times are tracking data, selecting members, keeping good recruitment practices, systematic recruitment practices. All those kinds of things, I think, needed to be very tight for this evaluation. And, unfortunately, we found out in the middle of data collection that some of these practices weren't quite as rigorous as we had hoped and needed them to be.

ANDREA: I have another one online. What participation rate did you get from the control group participants? How did you incentivize participation from those participants?

REBECCA: So everyone who participated in the study was offered a financial incentive for their participation. So for the first two study periods, so for the pre-test and the post-test there was a $10 incentive, and for the follow-up it's a $20 incentive because it's much harder to find them when they're not actively enrolled in the program. Within the comparison group -- I don't have it in front of me, but I think -- I'm not sure exactly what their rates were, but they were lower among comparison youth participants and that definitely was a challenge in terms of incentivizing that. So one of the ways that we addressed that was by working directly with community partners. A lot of our programs mentioned having to spend some social capital for the evaluation, so really working with partners that they had had existing relationships with and saying to them, “This is really important to us. Please help us out. Help us find these kids. Communicate to them why this is important and get them to be engaged with the evaluation.” So with comparison youth, I think that was a lot easier for sites, for comparison youth that were in a site. Whereas like for the comparison youth that were program applicants, who never had any contact with any program, I think it was especially challenging to track them down and encourage them to participate.

But we did a lot of hands-on follow-ups. So everybody who participated in the study, we sent them emails and then we also sent them text reminders throughout the study period. We sent them postcards and said, “Hey, by the way, we’re going to have this study coming up in a couple of months. We’d really encourage you to participate. If your address has changed, let us know.” So kind of facilitating that communication throughout the process and really just trying to make sure that we have up-to-date contact information on the kid. And then also following up over the phone. So we did like three rounds of email outreach and then phone outreach for everybody as well. And we also solicited information from their primary and secondary contacts. So we had like their parents or friends and family. So we would contact those individuals and follow-up with them if we still hadn't heard from the youth.

ANDREA: So this is similar to what you're talking about, but just to maybe dive into this a little bit more. Did you do anything specific to incentivize programs to remain involved in the process throughout the four years? So not the participants or the programs. Did participating programs formally sign on in a way to guarantee their involvement from start to finish?

FEMALE: Did they formally sign on?

FEMALE: No.

FEMALE: I wasn't involved in that.

ADRIENNE: So we used a rubric. I think I mentioned when I was talking before that we had a pretty systematic recruitment process. We used a rubric to assess who we thought would be good participants. They consented to participate, but they didn’t sign a formal MOU. I think it was pretty clear to everyone involved this was a really valuable opportunity not only to participate in a free evaluation, but to participate in a really rigorous free evaluation.

And also to give credit to our grantees, when we offer these kinds of capacity building resources, they tend to be very receptive and uptake tends to be very high and positive. And so I think the fact that we’ve retained pretty much every grantee that participated speaks to their willingness to participate and the value they place on the capacity building work that we do.

ANDREA: I have a few more online. I just want to look around the room and see if there’s any specific questions. I’ll continue reading. And on the phone are there any?

BRAD: We do have two responses on the phone.

ANDREA: Oh, can you go ahead and ask them to place their calls?

BRAD: Certainly. Our first question is from Shadong Jang [ph.]. You may go ahead.

SHADONG: It was already addressed. It's the one challenges.

ANDREA: Okay, thank you.

FEMALE: Thank you.

BRAD: And then we do have another one from Madeleine Chisolm [ph.]. Your line is open.

MADELEINE: Hi. Yes, my question has to do with looking at an umbrella agency and whether or not you could use the same method to look at internal programs? So applying it, rather than looking at multiple agencies, looking at multiple programs within one agency. Thank you.

ADRIENNE: I can take this one. So I would absolutely think you could. The cool thing about this methodology is that it doesn't necessarily have to be an impact evaluation. You could use other methods too. You could do an [unint.] evaluation. You could do case studies. The important part is though that you just need to make sure that the outcomes that the units would be targeting would be the same, that the approach, the intervention is the same if not identical if possible. I think there's actually some precedent for other federal agencies using similar methodologies to this bundling approach. They used it with -- it wasn't an umbrella organization, but it was a grant making program that made a series of like intermediary grants that used an approach like this. But you could definitely do it with an umbrella organization.

ANDREA: Cheri, do you know of any others? So any more on the phone?

BRAD: We have no more questions on the phone at this time.

ANDREA: Okay. I have a few more online. For an organization interested in convening a group of programs for a bundled evaluation and serving in a role that CNCS did in this case, do you have any recommendations for how the convening organization can solicit funding to support the work? There's a lot of people thinking. And if anyone on the phone has any suggestions, please let us know.

ADRIENNE: So this might be good to get some clarification. I think one reason why this worked out is because we have the funding. I mean, the Office of

Research and Evaluation. We were able to spend time developing the idea because we knew that we would be able to budget for that. But I don’t know how it would have gone had we just had the idea but no funding. I suppose you could potentially look for evaluation grant money. I don’t know if there’s too many foundations or organizations that give out grant money to do evaluation work to where you could then fund this. So I guess our experience may be a little unique because we had the funding ready to go if I’m not misunderstanding the question.

FEMALE: I had a similar question about when my first [unint.]. You all were doing whether there would be foundations out there that are interested in capacity building for non-profits. And I think that those exist but we haven’t done any research to sort of identify who those folks are, but it might be something that we would look into given interest across a number of grantees from our different federal partners.

ADRIENNE: And I do know that there are other non-profit organizations and foundations that do intensive capacity building work in an experiential way, but they control their own funding in that case. I'm thinking like Propel Max [sic].

FEMALE: [unint.] evidence.

ADRIENNE: Mm-hmm. They are in control of that. They're not an intermediary seeking funding to implement a bundle.

ANDREA: Okay. I have another. Are there any resources or directives for creating a special education toolkit for non-profits to fall more quickly under the gifted and talented side of the spectrum versus the disability side to empower the opportunity of growth within your mission?

FEMALE: I don't know if I understand the question.

ANDREA: I'm thinking that not only right -- so, Jessica, you posed that question. Can you clarify a little more what you mean by --

FEMALE: Is the question on the target population?

ANDREA: I think it is about the target population.

FEMALE: I mean, I think the methodology would apply across regardless of the target population if that's the question.

ANDREA: Right. So I have a question in terms of you had a lot of participation from the different sites in terms of getting input for the measures you mentioned. I believe in the data collection. Can you describe a little bit how that input strengthened or challenged the way you did the data collection or how the sites were able to use their knowledge to increase their response rate? And how they created the data collection techniques?

FEMALE: I can’t think of a few specific examples, but I could give you -- I mean, I think, in terms of the first part, so in terms of providing feedback on the measures, I think that was really useful in the sense of like really just making sure that we are measuring the right things for their programs. So having conversations about is this something that you would really expect to see change on. And it’s interesting because a lot of the more useful conversations that we had out of that came actually a little bit later. So when we collected the first round of data, we had a lot of really interesting conversations around the individualized feedback reports where sites were like, “Oh, I really thought that kids would be improving in self-efficacy, but actually their self-efficacy is off the charts of baseline.” And so it turns out like maybe this really isn’t the best measure for our program. And so that did have an impact on sort of how we thought about the results moving forward and which measures we really chose to emphasize as impact measures in the outcomes’ study.

And so I think that was really useful is having that feedback throughout the process to hear from programs to say, “Okay, like actually the data looks a little different than what I thought, but this is changing the way that I think about how our program operates.” And it’s not just about improving kids’ self-efficacy, but actually we were really expecting to see change on like some of these more behavioral outcomes. Like we’re really thinking that because they’re going to go out and be getting a job. And so having those kinds of conversations was really helpful.

There’s also a lot of feedback about language and so like trying to make the language more accessible for teams. And a lot of it was as we were having those initial planning conversations, I think when sites were able to be involved in the process, they often developed some of their own mechanisms for increasing participation. So it was interesting, as the process was going on, some of the sites started to talk about how they would introduce the study, which wasn't something that we had required them to do. We told them like, "Make sure that you have a really clear testing environment." But we didn't require them to say like why the study was important. But a lot of the sites started doing that on their own. They would sit the kids down and say, "This is really important to us as an organization and we care about you. We want to have your honest feedback about your experience." And when they did that, participation rates really skyrocketed. And once some of the sites started telling us that they were doing that, we were able to recommend that as a best practice to others sites. And I think that really allowed for a lot of adaptive learning in terms of best practices to improve recruitment and retention for the study.

ADRIENNE: Consent was another big one that came up. I remember a lot of the youth were very nervous about sharing data, especially because there were some sensitive things you were asking about. And so being able to have the program have the dialogue with them like, "This is what we're going to do with your data. This is how you're going to see this come back to you. And this is when the report will be ready. Like Becca said, coming up with language that made sense or like a little intro that made sense about what is the point of this and you can consent or not consent was a very interesting style. I think it came up a couple of times from some programs.

REBECCA: Yeah. And I think that like around that also a lot of folks had questions about confidentiality versus anonymity was a real like learning point. So we spent a lot of time talking to programs about, "No, your data's not anonymous because we have to track you over time so we have to know that it's you. But it is confidential. We're not going to share it back with the program." So that was a real point of learning and discussion for a lot of sites.

FEMALE: Sorry, I keep asking questions. But I had a question on so you used surveys, right? You collected survey data. And I know the sites participated in developing the survey. Yes. So I wanted to know if they have the option of adding additional questions if they wanted to track additional stuff that they were tracking and then reporting back the core set of questions to you? Or was it the exact same survey that had to be used by everyone?

FEMALE: Everybody in the study took the same survey. But one of the things that came up is that many of the sites kind of already had existing data collection procedures. So sometimes they would view like our survey and they would have like their own survey that they would also do or like a survey for another funding agency, for example. And so they weren't like adding questions to our survey, but they had kind of separate surveys. And that actually I would not recommend that as a practice because what ended up happening is in a lot of cases the youth kind of felt the survey fatigue and burnout from first [unint.] and all these data questions. And especially because, I mean, AmeriCorps already has surveys that folks complete as a part of their exit process. And so they would do the AmeriCorps survey and sometimes the programs would set it up where like "Today is survey day. Here's your AmeriCorps survey.

Here’s the bundled survey. Here’s your survey from this other funder.” It gets to be really frustrating because they’re doing the same information over and over again. And so I think like one thing that we’ve talked a little bit about is in the future really trying to consolidate those. And I mean, I think that there is a possibility to maybe add in a few questions that are specific to programs. But more broadly, also just thinking about ways that we could streamline the data collection or share the data. So if we’re already collecting it for AmeriCorps, is there a way that we could get the consent process so we could like facilitate that data sharing so they’re not having to provide their name and their age and their gender like 12 times?

FEMALE: Can you talk a little bit about how you analyze the data?

FEMALE: Yeah. So let’s see, so we’ll share the full report. I’m not allowed to talk about the -- not yet. We’re not talking about the actual results yet. But for the actual data, what we did is we used a multi-level model. And so we basically accounted for -- the first thing we did is we matched the kids. I’m using a process called Propensity score matching. And so for those of you who might not be super familiar with that, basically it’s a statistical process where you account for a bunch of different characteristics of the kids that might impact the probability that they are in the program or not. And then you basically give each kid a score. Like this kid seems really likely to be in the treatment versus the comparison groups. And then you match individuals who have similar scores. And so what that did is that took our full sample into a much smaller sample of just the final match sample. And not all the sites ultimately ended up in the final match sample because some of the sites didn’t have enough similarity in the treatment and comparison kids or didn’t have enough of a sample at host sites. So like I said, retention was definitely a challenge. And so for some of the sites there just weren’t enough people to be able to do that final match.

But for the subset of sites, I think we had 15 sites that were included in the final match sample. So once we take that data, we ran basically either logistic or linear multi-level mixed effects models, which I won't get super technical about. But we are predicting each of the individual outcomes based on characteristics at baseline. And then also accounting for the nesting of individuals within sites. And so that's treated as a random effect in the model. So it allows us to say, "Okay, yes, we've created groups that are equivalent at baseline, so that's the matching. And now that we have these groups that are equivalent at baseline, are they different at the final follow-up? And after we control for all their characteristics at baseline, do we still see this difference as a result of treatment participation?"

And in the models, we also did some analyses looking at interactions of baseline characteristics and treatment for key outcomes. So for example, like looking at whether kids who had some criminal history at baseline, were those kids more or less likely to have criminal involvement at post-test based on their participation in the program?

Sorry, that’s kind of a long answer, but I hope -- so linear and logistic [unint.] effects models.

ANDREA: So we’re almost out of time, but I want to ask Brad, is there anybody else on the phone?

BRAD: We’ve had no further questions on the phone.

ANDREA: Okay. So I want to give Linda and Cheri an opportunity if you have any thoughts on what you heard in terms of the program or the federal agency perspective?

LINDA: Sure, I mean, I think that it’s not very surprising that the programs all wanted to participate because this is such an opportunity for them to learn more about improving their systems as well as to insure that they are reaching the types of kids that they wanted to reach with the program. So that was wonderful to hear. And what other wonderful way would we have to be able to have hands-on experience, hands-on learning experience? So it’s always great to hear that programs are really taking advantage of that. So I’m excited to see the results.

ANDREA: Thank you.

CHERI: Yeah, the same. I think my interest stays as piqued as it has been. [unint.] hear the results. Yes. I would continue to thank people for that. But I also think that it’s something -- I see the reason why any one of our partner agencies wouldn’t be able to do the same things with their grantees. And some of the questions that you’ve been talking through in terms of sharing the data. What are we already collecting when we could do sort of a bundled evaluation without even really having to add a whole other layer evaluation and programming? And I think that’s sort of where our conversation needs to go. And whether we could do that across these agencies is a bigger question. But I think, certainly, within each of the agencies, this is a capacity that we could help add into their toolbox. So I think it’s really useful.

ANDREA: Okay, that’s great. Mary, any final words? All right, well thank you so much for everyone on the phone and in the room and we look forward to you all calling in maybe in our next webinar, which should be in March. So anyway, thank you. Have a great afternoon.

BRAD: Thank you for your participation on today’s conference call. At this time, all parties may disconnect.

END OF FILE