



T1. Evaluation: Learning More, in Order to Do More Good

Tuesday, June 4, 2019

3:30 p.m. – 4:45 p.m.

Moderator:

- Dr. Samantha Illangasekare, Senior Social Science Research Analyst, Office of Planning, Research and Evaluation, Washington, D.C.

Presenter:

- Sharon Rowser, Vice President, MDRC, Oakland, California

Dr. Illangasekare: Welcome everyone to the afternoon breakout sessions. My name is Samantha Illangasekare and I am with the Administration for Children and Families at the Office of Planning Research and Evaluation. We work closely with several offices that you heard from this morning. The Children's Bureau and the Office of Child Support Enforcement and the Office of Family Assistance are responsible for research and evaluation and implement our learning agenda. I am the team lead for the Healthy Marriage and Responsible Fatherhood research and evaluation that we do within OPRE at ACF. I'm very excited to moderate this session because this is a topic that is obviously very near and dear to my heart. I am happy to introduce our speaker today, Sharon Rowser, Vice President at MDRC. Sharon has more than 30 years' experience developing demonstrations that serve disadvantaged populations, including single parents and low-income couples. Sharon leads MDRC's partnership with Public Strategies and provides training and technical assistance for the Healthy Marriage and Responsible Fatherhood program, and the integrated innovative employment and economic stability strategies.

Sharon Rowser: Thank you Sam. Welcome to the session. As Sam said, my name is Sharon Rowser and I'm with MDRC. We are a mission driven nonprofit that conducts policy research and demonstration programs, usually at the national level. We recruit, train, and provide technical assistance and evaluation to the staff of the programs implementing the intervention that we're studying. First of all, I'm assuming that most of you are not evaluators. How many evaluators do we have in here? Three, okay. This session is intended for program leaders who are not doing evaluations themselves, but work with evaluators. It is geared more toward questions like "what is this thing?" "Why are we doing it?" "What do these different things mean?" Because I don't typically do the research myself, my job is really to help organizations onboard evaluations. We're going to talk about program evaluations starting with why programs choose to do evaluations. We're going to focus on process evaluations and outcome evaluations. We will be talking about research questions and what they are. How do you get one [a research question]? Logic models, and how you can use those. Lastly what makes a strong evaluation proposal.





First, why do an evaluation? The most important reason is because it's going to answer questions that are important to your program. It will also provide your program with information for continuous program improvement. It shouldn't be something where you don't hear anything about the evaluation until the very end. You want an evaluation to be answering questions you care about, and also providing you with information about how your program's performing so that you can continue to improve.

Another thing to consider is that participants in your program are busy people. You're asking them to give you some of their time. An evaluation can help answer questions about whether participants are better off as a result of participating in the program, or if your services are delivered as intended so that they're getting high quality for their commitment of time. Some programs commit to these multi-site evaluations because they're contributing to a larger public policy that connects them to a larger community. In addition to learning about the effectiveness of your program, you can also look at whether or not this program is helpful across communities, across different kinds of organizations, or across populations. This audience tends to be national or regional. Therefore, you're working with a different type of audience.

This leads to the next bullet that an evaluation can be a sustainability tool. I don't know how many of you have other funding aside from federal funding, but often funders are interested in working with organizations that are learning organizations, and by opening yourself up for an evaluation, you're saying that knowledge development is important. Even if all the results of the evaluation aren't positive, it's still a positive to participate, particularly if you can point to what you've learned from this evaluation, and what you've done to change your program to make it better. What's not very sellable is getting to the end and saying, "Well, the evaluation's wrong, it didn't really answer the right questions, or they didn't do it right." So that's not a sustainability tool.

The last bullet says, sometimes evaluations are required with the receipt of funding. This is not a selling point. It's true, it could be required. But whether or not it is required, evaluation is a use of resources. No one wins when those resources aren't expended in a wise way. If it's simply one way to get funding, it's not the best use of funds. You want to have a win-win situation: "I have an evaluation, it's going to tell me things I would like to know."

I'd like to go back to that first bullet, which is that evaluation will answer important questions and help programs. These are the reasons that I think of and the things that I tell organizations when I'm asking them to participate in evaluations. Are there other things that would be a reason to evaluate your programs that maybe I've left out?

Participant: To gauge that you are accomplishing your process outcomes, that you are in line with what you said upfront.





Participant: To be persuasive for more support. Maybe from decision makers that are higher up. Not necessarily just for funding, but also to convince them to invest.

Participant: To get an evaluator to help write evaluation sections of a grant proposal.

Participant: To change behaviors, talking specifically about staff, so you can demonstrate that you need to do something in a particular way.

Sharon Rowser: Right. It can be a sustainability tool, but it's important to use it in a way that's going to make your program stronger. So, the bottom line is it helps you when you're evaluating your program and it helps you better understand if you're doing well, not just [that you are] meaning well. That's something to think about when you think about the people that you're serving. Let's look at different types of evaluations. There are a lot of different methodologies that people talk about in evaluations. I'm not going to describe all of these but ethnographic, qualitative, quantitative, implementation, impact, experimental, quasi-experimental, cost benefit, etc. Most of them can be categorized into two broad types of evaluations: process and outcome. At times, some of the methodologies seem to straddle these two.

Process evaluations mostly answer questions that begin with the word how. How many? How well? Who? What? They're mainly descriptive and try and tell a story about what happens in the program. So, what is this program? What is it? What does it contain? How is it delivered? Who are the people served? And what did they receive? How did they feel about the services they received? These are all kinds of questions that get answered in a process evaluation.

Specific examples. Probably the most common are things such as characteristics of participants. I know if you're currently doing an evaluation in your program for the federal government, you're collecting characteristics of participants. People want to know who it is that are in this program. Some studies also collect characteristics of staff. Why would you want to know about these characteristics?

Participant: To find out whether or not the program was successful or not, it could be a factor.

Participant: Evaluate factors of who's delivering the program to find out what factors may contribute to success.

Participant: Can women run a fathering program? It may be a good effective program if men are the facilitators. It's possibly a gender difference.

Sharon Rowser: Yes





Participant: What types of impact did the pilot or whatever you're doing have on staff such as burnout.

Sharon Rowser: At the most basic level, you might want to have some information that helps you in the future figure out the type of staff you want to recruit and what kind of training they need. Some larger studies really look to see if you have enough staff. You could look at some of the characteristics and see if there are particular curricula that work better with some type of staff than others. One reason to do participant characteristics is so you understand that you're serving the groups that were your targets, which seems like a pretty basic issue. I actually didn't participate in this, but MDRC did a study many years ago where we were working with a program that was providing immediate employment for individuals recently released from prison. It was providing income, trying to reduce recidivism, etc. We did a study on it and after we had enrolled for several months, we realized that the people who were coming into the program had mostly been out of jail for a year or more, which was not their target audience.

They didn't really know that. They didn't stop immediately, but they had to retool their program and think, what are we doing wrong that we're not getting the population that we said is our key population? They had to retrain their referral partners so that they were getting referrals of the right group of people. Describing the frequency, length, and duration of services helps us understand how participants experience the program. Do you deliver your 12-hour workshop over six weeks or all day on Saturday and half day on Sunday? Is your 40-hour program delivered over two weeks or 10 weeks? Do the other services extend the program over a month, three months or a year? Those are all things that a process evaluation will document. You want to make sure that what's happening is what you thought was happening. But it's a way of documenting these things. In addition, programs should develop their service and delivery strategies with some theory behind why they made these choices. There should be some reason as to why you chose to do it over two weeks or a year. You know your intent. Your program offers so many hours of program services over a particular number of months delivered by certain staff and partners. But, how much of these services do your participants actually receive? What services are they receiving? How much are they? How long do they remain active in the program? Recruiting and enrollment are often the most challenging parts of delivering programs, but keeping people engaged is equally challenging. You can't always tell how well your program is doing in engagement by just looking at who is in the room. I've been told often when talking to programs that it's not a problem for them because at the end of the day, they have a hard time getting people to leave the program when program services are over.

The problem is they're looking at people who stayed. How many of the people who started are people who stayed? That is what an evaluation can help you understand. Are you losing people? If you are, is there a particular place where you're losing them, which is important to know because that may be another opportunity for retooling and strengthening your program. The qualitative methods can help you understand if your program is operating as it should and how





staff or participants feel about the program. Those are different types of measures. But do your staff feel like they're adequately trained and supported by leadership? Do participants feel like they have something of value from the program? Those are more qualitative measures of how people feel about what they have or what they're doing, or what their involvement is in your program. These types of evaluations and methods can tell us a lot about what is happening, and they're usually less expensive to conduct than outcomes evaluations, but they don't tell you what difference the program makes.

Participant: Can you clarify dosage?

Sharon Rowser: Yes. When I think of dosage, I think how much? You could think of it in two ways. How many hours of service do they get? Or you could also think of it as your 40-hour program, what percentage did they get? There are different ways of looking at the same thing. When we think about describing how people participate in the program, how much they participate in becomes really important.

Participant: If you have a federal grant, being part of a federal evaluation requires an evidence-based program. Many of those require a longer time, 12 weeks for example. That is too long, so we try and shorten up the programming, but as you talk about dosage, how is that affecting our dosage?

Sharon Rowser: Dosage is not how much time is spent in the program, but the amount of services received. You could have someone who stays in a program for a year, but they show up when they want to. Dosage is how much did they get, not how long did you maintain contact. When I think of dosage, it's that we have this program that's supposed to deliver this amount of service, for these number of hours. How much of those hours did the person actually get? Whether it's all lumped together in a couple of weeks, or if it's over a long period of time, it really is just how many hours of the actual service did they get. So, I don't think it's a matter of whether it's a federal evaluation or not.

Participant: The developers should be part of the implementation of that. You can be in a program, select the curriculum, then that curriculum developer should be a part of that.

Sharon Rowser: I think that's an interesting conversation. So often conversations between program deliverers and evaluators and funders are confusing. Because you're talking about using the same terms and not necessarily meaning the same things. As I said, dosage is how much of something you received. Most evaluators don't think about whether you finished the program or not. They want to know how many hours you received, did you get a large enough sample, people who received more hours, did they have better results than people who had less? Program operators and funders think everyone who starts should finish. What finished means can be very different. It could be that you have a 30-hour program and maybe decide that as long as people





get 20 or 24 hours of it, that's not really finished, you'd like them to get all of it, but they've gotten a significant amount. We think of that as dosage. You might think of it as a way of saying this person qualifies to me as having completed. But that's something you have to work out with your funder because we're all talking about the same people doing the same thing. But we're using different terms for it.

Participant: We have 13 weeks, but if we get men who attend 10 of the 13 weeks, they could get a completion. Because we know there are going to be men who may not be able to come one day, possibly due to a medical situation. We work with our funders, as long as they attend. And there are some men who attend all 13 weeks.

Sharon Rowser: That's great, as long as your funders agree to it.

I'm going to move on to **outcome evaluations**. These are generally numbers driven and often use comparative measures. Often questions start with how many or what percentage of participants accomplished x? Generally, something that's a programmable goal simply states what happened. For instance, you have an employment program and 60 percent of your participants find employment while attached to your program. That's a program outcome. It says what happened. But it doesn't tell you what difference the program made. More rigorous models that are outcomes based will tell you that.

Earlier, I spoke about when people come to your program and participate, they're giving you their time, they're using their time in this particular way. If they weren't in your program, they'd be doing something else with that time. Or maybe doing nothing. They may have gone and enrolled in another program, or they may just be looking for work on their own. You don't know if they hadn't participated in your program, how many of those people came because you said it was an employment program, would have found employment in some other way. What you're really trying to measure in a rigorous way is what percentage? Or how many of those people that found work, found it because of your program? How big a difference did your program make over what else they could have been doing? You might want to know if participation in the program led to finding a job or how big a difference did your program make over what else they could have been doing with their time. That's really the basis of random assignment evaluations. They are harder to do, which is why you get more points for doing them, but they are considered the gold standard of evaluations, and considered to give the most reliable evidence. Because this one is difficult to do, it's probably the most expensive. What makes it expensive is in order to look at what difference your program made, it usually requires a comparison of groups of similar people in a similar place at a similar time who are not being served by your program. You track both groups and see what happens to that group that wasn't served by you and the group that was and the difference between that. If, depending on how rigorously you do that, the difference is what the impact of the study is.





Participant: Will an incarcerated father fall into one of those categories?

Dr. Illangasekare: Is the question can you do random assignment studies with incarcerated fathers?

Participant: Yes. Because with the incarcerated fathers that we serve, we can only provide them workshops, we can't provide them services that our grant evaluates us on. So, we have to look at incarcerated fathers separately. However, they make up the largest percentage of our population.

Sharon Rowser: If you're serving two different populations, it may be the case that you're evaluating one group and not the other because they're in completely different situations. I'm not saying you can't do a random assignment study with an incarcerated population, you can, but you need a comparison group of similarly incarcerated fathers. I don't think the question is whether or not they're incarcerated. It's what your program is doing with the populations it's serving and what your evaluation is saying. It sounds like you need a one-on-one with that.

Participant: Can you speak to any tools that are out there? For instance, I use the adult adolescent parent inventory for self-reporting data for people that use my program. Programs often use self-reported data, but funders may see that as a weaker measure.

Sharon Rowser: There is an evaluation resource guide for responsible fatherhood programs on line.

Participant: If you are an OFA grantee and I wanted to get a grant, would it be a good idea to familiarize yourself with that document?

Sharon Rowser: Yes. That's who it's written for. But you need to have a relationship with your evaluator. You need to talk to each other. From the program point of view, you can say this is our program, this is what we're doing, this is what we're trying to accomplish.

Participant: I have reviewed grants for OFA in the past, and the evaluation sections give a certain point value. So, you have to write something.

Sharon Rowser: Yes, but you also need to have some assistance from an evaluator. Because one of the key things is it has to be relevant. I would say if you're writing, there are opportunities you could also set aside for the evaluation and write into the proposal. Because, to some organizations, information matters enough that they have a greater chance of funding the grant based on that. There may be people who have that gift without being a professional, but you have to look at it based on what your team looks like. Larger organizations at times may have somebody in their organization that's very good at that. But it really depends on the team and the level, versus maybe a local foundation that may be a little more lenient, versus a first-





time grantee. But the larger grantee sometimes will have an evaluation tool because they've been doing it for nine years.

Sharon Rowser: Some of you should have some kind of an advisory group. I know implementation teams are one of the things required. There should be somebody on there that can help you with this. You shouldn't be completely in the dark about what's going on. You need people giving you advice because that's not what your expertise is. But the two should be connected to each other. So, when you're putting together an advisory group think about whether you have someone that understands evaluations.

Participant: Do you have recommendations for evaluating groups that are familiar with the field of fatherhood? I know there are evaluators, but they don't understand the work that we're doing. Their standards aren't realistic for fatherhood programs.

Participant: I agree because I've had evaluators that come with education backgrounds and now we're doing maternal child health. I find that I'm having to teach them to learn how to evaluate my program and advising them not to give me outcomes when I need process development in a lot of my programs.

Sharon Rowser: There are organizations that do this. But you need to think about how to select an evaluator. What is their field of study? What is their experience? What evaluations have they done before that would lead you to believe that they can evaluate your program?

Participant: And just where to find evaluation wording. Because that's where the money is.

Sharon Rowser: You build that into the proposal that you submit, so the funds are coming from the board. You get points for it, but you also need to have a budget. You won't get points for having a great design and no money in the budget. It comes with the OFA dollars. Otherwise, if you aren't going for a federal grant, foundations may be another source for evaluation funding.

Participant: A good evaluation will get you money.

Sharon Rowser: Yes, a sustainability tool. National evaluations are another option but a small number of grantees. There are different ways of doing this. If what you're talking about is OFA funding, you don't need to find the money on your own. It's part of the whole process. You have to build it in.

Participant: Once IRB boards find out they are going to evaluate a program do they contact your office?





Sharon Rowser: No. I actually serve on an IRB. Usually, if you're doing a local evaluation, and especially if the evaluator's attached to universities, universities will have an IRB. The purpose of IRBs is to protect human subjects. Anytime a study is being done that involves people, and it's treating them and collecting information about them and doing things with them, IRBs review to see that they are being treated ethically. Do they know what's happening to them? There are a few other things they're doing that generally happens at the local level. If you're in a national evaluation, then the national evaluators may have their own IRB. But, if you're doing a local evaluation, I don't think they would contact you at all.

Dr. Illangasekare: Yes, usually with the local evaluations that are happening for the OFA grants, each of the local evaluators working with the programs are working with a local IRB. So that might be like whether it's a university or there's another local IRB that they can work with.

Sharon Rowser: If you're getting technical assistance, there's technical assistance with local evaluations that come as part of the whole process. Those are questions that you could ask the technical assistance provider. You need to know what's going on in your particular community.

We will now move on to outcomes evaluations. Examples of measures that can be addressed in outcomes evaluations include improved parent-child relations, decreased reports of homelessness in fathers who attended, increased knowledge and use of positive parenting practices, and increased months of employment. When deciding on reasonable research questions, ask: Is it of interest to my program, to my funders, to the policy community? Does my logic model suggest it could be answered by my program? Will the data be available to answer the question? Will I have the time, money, and sample to answer the question? Can the outcome be assessed in the time you have? Do you have the grant period when your evaluation is due? Because, if it's part of this grant cycle, it may be due at the end of five years. Do you have the sample? And this is a tricky one. More complicated questions need larger samples. Meaning more people, the sample is the number of people in the study. So, there's a couple of things you have to think about with that one. Are you depending on what that number is? Are you confident you can get that number of people into your study? Are you confident you can get them in in time because if the evaluation is due in the last year, the last year of enrollment is not likely to be in there, because there isn't enough time for that. So, you're really talking about the first four years.

Next, we will talk about logic model. Logic model is a visual representation of your program, but it's different than a program flow diagram. Is everyone in here an OFA grantee? The reason I ask that is because I know in this round, all of the grantees were required to do a program flow diagram. Program flow diagrams show what happens from beginning to end and in the flow, meaning beginning, middle, and end.

A logic model looks at the expected outcomes. What are the milestones that you expect to happen to people as a result of being in it, so it's not just about you? Because of that, it includes





four categories: inputs, activities/interventions, outputs and outcomes. The inputs are resources. What resources do you have? What is it that you have that you're offering? Program components, your staff, money? Those are the resources. What are the activities or interventions that are happening as part of your program? What are the outputs? These are immediate results. Let's say you have a parenting program. An output would be the people who participate have a better understanding of parenting. An outcome happens after. The important thing about outcomes is when you identify them, you want to know if they're short term, intermediate, or long term. With long term outcomes you have to ask if you have the time, money, and is it logical to expect this outcome?

Participant: Is there a place where we could find some good logic models that might be fatherhood related?

Sharon Rowser: There's a whole section on logic models in OFA's guide.

Participant: Is that a required component when you apply for an OFA grant?

Sharon Rowser: I looked at the RFP and OFA highly recommends it, but it is not required. It provides a roadmap, builds consensus, monitors your program, and helps find inconsistencies in your program. I really think it requires input from you and from the evaluator. But it needs to be consensual.

This is a logic model that shows how program elements drive outputs/mediators and outcomes. The language might be different across logic models, but they all show the flow of a program. I didn't pick one for fatherhood. It's a different type of program. This is a financial aid program for community college students. The theory was, Pell Grants are given in lump sums and people get them and, because they don't know financial management, they spend the money, they run out of money before school's over. Then they either drop out or start working, or cut classes and work. This was giving them the money and giving them messages about financial management. You can say the financial aid and communication are the inputs. So, information about financial aid. They call these mediators, because they're not really activities. How is it changing their college experience? Now they have an ongoing, continual flow of cash, they have some information about how to manage money, etc. So, these are the outcomes, and you see they have it within a semester, that's short term, and then long term. They don't all use the same terminology, but they all sort of have the same flow.

Let's do a thought exercise. Your organization is focused on parenting but you're doing a fatherhood program. In thinking about the way your programs are organized, what are the inputs, the resources that you have?

Participant: Community partners





Participant: Curriculum

Sharon Rowser: You have staff. You have curriculum, you have case managers, and you have money.

Participant: Fathers/participants.

Sharon Rowser: Yes. those are resources and inputs. So, what are the activities? There's the parenting component, the relationship component, and there's employment. All of those are resources and there's various activities in those. Some of those activities, regardless of those components, might have outputs that are related to being a better parent. What kind of things might you think of as outputs of those activities?

Participant: Employment like job readiness classes

Participant: Time spent with family.

Sharon Rowser: Yes, exactly. More time spent with family. Maybe you have father-child events or activities where they're spending one-on-one time with their children. That's an output. We went into this with a goal, what do we think this will likely lead to? Some immediate things might be spending more time with your children, it might be improvements in parent child relationships which also affects your ability to parent.

Participant: I'm working on a program that many times we see knowledge gained.

Participant: Knowledge is power.

Sharon Rowser: And that's an output.

Participant: You use knowledge as a measurement?

Sharon Rowser: The logic model isn't going to tell you what you should have as your research question, but it should point you in a lot of directions. Do we want to measure this? Is this an outcome? Is this something we want to look at? And I would say pre-post is not the same as impact. So even though they seem the same, it's like we'll see what they know in the beginning and what they know at the end.

Participant: On the second path, I'm trying to see how you see it?





Sharon Rowser: Pre-post says, we'll give people a test the first day of class, and then we'll give them a test the last day of class and we'll see how much they've learned. What you don't have there is the people who dropped out. And you don't know what would have happened if they hadn't been in your program. They may have done some other thing that was related. The other choice is, this is what impact evaluations are. You pick people and half of them you're going to serve and half of them are in a comparison group. You survey both of them at the end and look at the difference in knowledge. How much knowledge has this one group gained? That's very, very short term. But again, you're somewhat limited based on time. You're going to have to decide what kind of things you want to evaluate. But those are two different things. Pre-post is saying you started here, you ended here. Here's where both of us are. And what's the difference between this group that didn't get our services and this group that did get our services?

Participant: How do you find the comparison group?

Sharon Rowser: Programs should recruit twice as many people as they can serve. Typically, there's more people in the community that need your services than you can serve. With a random assignment study, it's like a lottery. Everyone who is interested and eligible coming in has a chance, but only half of them will actually get the services.

Participant: Can you compare two curricula or does the comparison group have to get nothing?

Sharon Rowser: The comparison group doesn't have to get nothing. Doing a head-to-head comparison wouldn't be valuable. That's the fallacy. People think you're the group that you don't serve gets no services at all. We often are surprised in random assignment evaluations at what people do who are in the control group. People are doing things. It just may not be your thing. I'm not sure you'd want to do a head-to-head test within a fatherhood program. Because, if you think you're that much better, why would they want to be in the evaluation?

Participant: Maybe I'm looking at board incentive. We're looking at the common goal of the work that we're doing. That if yours is doing A mine is doing B and what is the best service for the overall [inaudible] that they're doing.

Sharon Rowser: That's a different type of study. You can do a differential study, and you can do that within your program. I don't know if anyone here is in the B3 [Building Bridges and Bonds] study, which is one of the national fatherhood evaluations.

Dr. Illangasekare: B3 is an example of this. Some programs are testing a specific intervention that they've added to a program. They are comparing people that are getting this kind of additional intervention. For example, the parenting intervention, Just Beginning, involves having dads bring their young children in and doing monitored sessions, like play sessions. And then they're comparing them to the dads in the program. So, no dads are getting nothing, and all the





dads served by programs are getting basic services. They're comparing basic services to dads that are getting basic services plus this enhanced Just Beginning. That's another design.

Sharon Rowser: Right. That's what you don't want to compare, because there's too many differences. The whole thing about randomized trials is your control. What it's saying is that you've controlled a certain sample size for all the differences between people, staff, programs, location, and community. But if you then pick two different programs, you haven't controlled all those things. So, what you're talking about is what B3 is. You might try offering your services in two different ways, maybe intensively versus over a long period of time. Those are called differential studies.

It's not that you don't have a no services control group, you just have a different service when you're comparing these things. This is not going to tell you how to write an evaluation design, but here are the key things that are considered when looking at the evaluation plans in the fatherhood proposals. The rigor of the design. How convincing is the information? The feasibility of the design. If you don't have money for your evaluation, this is an infeasible design. So that's money, time, and sample. Not only does that matter in getting it approved, but it matters in keeping it going. Not everyone is able to do the things they said they were going to do in their evaluations, and they've had to stop.

The research question. The impeccable source I got this from said there should be one. Apparently, some people submit evaluation proposals that don't have a research question. I don't know if that immediately invalidates you, but it's not going to help. You need to have what it is you're going to learn. What is your question that you're trying to answer in your evaluation? What is it that you're trying to learn? Research evaluations start with a hypothesis. The hypothesis is that if you do x, y, and z, q, r, s will happen. And the question is, are we going to find that in this study?

Participant: If I get the fathering curriculum then they're not going back to jail.

Sharon Rowser: Was that a short-term outcome, intermediate outcome, or a long term?

Participant: A long term outcome.

Sharon Rowser: I know, but can you measure that within the time period?

Participant: That they didn't go back to jail?

Sharon Rowser: There are problems that we're trying to solve that we often think are long term outcomes. But if incidence is not high, you will need a large sample measure. By that I mean even though we know our fathers are involved with criminal justice and over a long period of





time there is some likelihood that they could become involved again. But, in any one year, how likely is that? Or in any one six-month period? And if it's not very likely, it's a probability thing. This is statistics and you really need to talk to an evaluator, but you don't want to measure. You don't want to have an outcome that isn't happening very often in a short period of time in the general population because otherwise, you won't be able to tell the difference.

Participant: If all these fatherhood programs were successful, would you then assume increased income? Then to his hypothesis, if they didn't go back to jail, then they would make more income?

Sharon Rowser: Again, is that a short term? How much time do you have to measure it? Is making more income a short term, intermediate, or long term?

Participant: Well that could be short term or weekly.

Sharon Rowser: If you have a very strong connection to employment, let's say you're putting people into internships. Where you have a social enterprise, then perhaps. But if it's sending people out, you'd be surprised. I worked on the first, it's still considered one of the seminal evaluations for fathers, Parents Fair Share, and the control group income was pretty comparable to the programs for a couple of reasons. One is, we had them in programs, while the ones that weren't in the program, were free to look for jobs. And the other is, men can get work if they're looking. It's hard, I'm not saying they can keep it or that it pays well. Those types of programs with short, quick things having to do with income are more effective for women, because women have always had a harder time earning and finding jobs. Going back 25 years, I think we're closing the gap somewhat. But the thing about income is it's a hard measure with men in the short run, because whether or not you provide the services, they can get work if they're looking for it; it's are you helping them stay connected to jobs longer or better jobs? Or where they earn more, but those aren't short term outcomes.

Participant: I worked for the Jobs Plus program for a very long time.

Sharon Rowser: Oh yes, we did that.

Participant: You all get involved with the Jobs Plus program which gives the evaluation and we measure it over short term.

Sharon Rowser: So, you need long term outcomes.

Participant: And, if you had more partners who helped the women increase their income, and men, if you put the two together then you start to see a significant increase in income.





Sharon Rowser: Is it still only in public housing?

Participant: Yes.

Sharon Rowser: Okay. So, if you're working with public housing, let me tell you one thing about the list of terms I created. These are not research terms. These are the terms that we frequently miscommunicate between program operators and evaluators. As you start thinking about measures and collecting data you should be asking your evaluator what they mean by these terms. The one that I want to point out is the easiest to point out and that is “participant.” So often programs think a participant is anyone who shows up. I read the brief on the PACT evaluation that Mathematica did. Their definition of “participant” is people who received 50 percent of the program services. Evaluators’ definition might be you have three components, a participant is someone who received some of all three components. So, you could think you have an 85 percent participation rate. And your evaluation could say you have a 30 percent participation rate, if you are calling those things different. It's fine, as long as you know that, and you understand what they're doing. But you don't want to be surprised at the end.

