TRANSCRIPT

HOW TO MONITOR PERFORMANCE BASED ON A THEORY OF CHANGE

FEBRUARY 28, 2019

Presenter: Jerome Gallagher (PPL/LER)
Welcome, everyone, to this webinar on how to monitor performance based on a theory of change. I’m Jerome Gallagher, and I’m an M&E Specialist on the Institutional Support Services contract, the Office of Learning, Evaluation and Research and the PPL Bureau here at USAID. So this webinar is the second part of a two-part webinar series on theories of change. The first in the series was the webinar just two days ago on how to develop a theory of change led by my colleague Carissa Page in PPL/SPP. If you didn’t attend that webinar, don’t worry. You can still attend this one; you won’t be lost. In that previous webinar, Carissa used an example of a youth crime-reduction project to talk about theories of change. And in this session, we’re going to pick up where she left off and talk about monitoring performance based on a theory of change.

And I’m going to use the same example that she used of a youth crime-reduction project, although I’m going to take it down to the activity level. Now in that earlier webinar, the hero of the story was the project designer who goes through a series of steps to develop a theory of change and its associated logic model to address the development problem. In this webinar, the hero of our story is the M&E specialist, and in particular it’s the M&E specialist of the implementing partner who has been given the theory of change and its associated logic model and has been asked to come up with a monitoring plan for this theory of change. So for this webinar I want to put ourselves in the shoes of this M&E specialist and try to work out how we might monitor performance. Now you might not be an M&E specialist and you’re certainly not on this webinar an M&E specialist for implementing partner.

But that’s all right, because many of us collaborate with our M&E partners on performance monitoring plans, or I certainly hope we do. But even if you’re new to performance monitoring, I hope this webinar will be of use to you. So in the past, our M&E specialists may have been presented with a project or activity design in the form of a log frame. And here I’ve put up a log frame, or a partial log frame, for the youth crime reduction project that was discussed in the previous webinar. Hopefully this will look familiar
to you. For each result listed on the left, we have a column for indicators that we would come up with for each of these results and data sources. And the M&E specialists would add those indicators and data sources in those columns. Nowadays, our M&E specialists might see something more like this.

This is the logic model that was developed for the reducing youth crime theory of change that was discussed in the previous webinar. As Carissa noted, a logic model is just a visual depiction of our theory of change. But how should our M&E specialist use this to help devising a plan for monitoring performance for this activity? That’s what we’re going to discuss today. So here’s our agenda for the webinar. First, I’m going to start with a brief introduction, two logic models and performance monitoring concepts. And then we’re just going to go right into trying to develop a performance-monitoring plan for the reducing youth crime activity. And then we’ll end hopefully, we’ll have some time for questions and responses. I can’t promise I’ll have answers but at least I’ll try to give you some responses.

All right, let’s start with an overview of logic models and performance-monitoring concepts. Again, logic models are visual depictions of our theory of change. So for a generic theory of change, here is a generic logic model, a very basic type of logic model. In the center, we have our activity. It starts with inputs: the money, the people, the time that we put into an activity. These inputs are transformed by some intervention or process into outputs. When we deal with our implementers, we often call these outputs deliverables. And hopefully these outputs lead to outcomes. First, some intermediate outcomes, which are just the changes in some condition in the world. And then ultimately the end outcomes, which are those outcomes that we ultimately care about, the reason we are doing the interventions.

So pretty simple. As you probably noticed, this logic model depiction also includes this round blue oval that I have labeled context. So the context just represents the context that your program operates in, or you might call it the system in which you operate. It’s all the stakeholders, their roles, the rules, norms,
relationships, resources, conditions, etc., all the factors that may affect your programming. As noted in the previous webinar on theories of change, the context is very important, and understanding the context is very important for developing the theory of change, right. So we need to understand the context in order to determine how we’re going to intervene in it. But the context is also important during implementation. And indeed, one of the things we should be doing during implementation is what we like to call context monitoring.

I’m not going to discuss context monitoring in this webinar, but if you find this webinar useful, maybe we’ll do another webinar on context monitoring using this example. But what I do want to note about the context is as you move through this logic model from outputs to intermediate outcomes to end outcomes, those external factors in the context generally have a greater effect on your outcomes of interest. So while we expect the outputs to be well within the control of our implementors, those end outcomes usually depend on a wide variety of factors outside of USAID’s control or implementor’s control. So with that, let’s move on to performance monitoring. So here I’ve added the definition of performance monitoring.

And the first thing to note is that performance monitoring is ongoing, right. It’s something that we do throughout the life of the activity. Next, performance monitoring includes quantitative and qualitative information. That’s right, performance monitoring is not just indicators. Yes, performance indicators are required; they’re particularly for our PPR. But if you’re monitoring performance just with indicators, you're probably not doing it right. We really want to add some qualitative information in our monitoring. The next part of the definition says what performance monitoring is doing. And first, performance monitoring is for determining if implementation is on track. So that’s this first half of our logic model, the inputs to interventions to outputs. And even more importantly, performance –

monitoring is about whether expected results are being achieved. That’s really the heart of performance monitoring, are we getting
those results that our theory of change predicts that we will get? And that’s this half of the logic model, our outputs to intermediate outcomes to end outcomes. So if you take nothing else from here, just remember performance monitoring is really focusing on ensuring whether we are tracking whether we are achieving these expected results. Two final points I want to make about performance monitoring before we move on. The first is that performance monitoring is not just monitoring the performance of our implementors. I think this sometimes gets people confused. If we wanted to just monitor the performance of our implementors, we would really just look at outputs.

[0:09:00]

But we want to monitor not just the performance of our implementors but the performance of our theory of change. So that requires us to look not just at the outputs that are in the manageable interest of our implementors but also those intermediate outcomes and end outcomes that our theory predicts will happen as a result of our theory of change.

And then the second point I want to make is that while performance monitoring is more than just monitoring the performance of our implementors, it’s not everything. So I’ve already noted that there’s a distinction between performance monitoring and context monitoring. There’s also other types of monitoring that we might do throughout the life of activity. We monitor those finances of implementors, right, make sure they’re spending their money in the proper ways. We also might do some monitoring of compliance with USAID regulations.

[0:10:00]

For instance, regulations on branding and marking. These are important parts of monitoring your activity, but they are not performance monitoring.

All right, so let’s move to our example of reducing youth crime project and activity. So here’s the logic model that Carissa shared on Tuesday’s webinar. Let’s just recap. In that webinar, Carissa described a development problem where we had many youths who had committed crimes being sentenced to incarceration, being sent to prison as a result of their crimes. And these prisons are not helping these youths to rehabilitate. They just made life
worse for them, right, and this led to these youths committing more crimes.

[0:11:00]

So we wanted to break this cycle and we created a theory of change in which we would intervene in this development problem to help solve it. The way we did that was we had a couple of interventions. First, some policy interventions with the national legislature and with the local governments to change the incentive structure for sentencing youth. We also had some interventions with ARCs, so and that abbreviation stands for alternative rehabilitation centers, so places where we could sentence juvenile offenders to instead of sentencing them to incarceration. And hopefully in these alternative rehabilitative centers, these youth would get services that would help them reenter their community and not go back into crime.

[0:12:00]

And then finally, we have an intervention with our judges, to encourage these judges to sentence juvenile offenders to these alternative rehabilitation centers instead of sentencing them to incarceration. And then this theory poses that through these actions we will lead to lower levels of youth crime recidivism and eventually a lower youth crime rate. For this webinar, I want to focus on just part of this theory of change and its associated logic model. So we're just going to look at two sets of interventions, our interventions in helping building the capacity of these alternative rehabilitative centers and also our interventions with judges to help change their behavior to sentencing juvenile offenders to alternative rehabilitative centers instead of sentencing them to incarceration, sending them to prison.

[0:13:00]

All right. So that's our theory of change and its associated logic model. I'm going to go through five steps in how I think one might go about trying to monitor performance of this theory of change and its associated logic model. I don't want to make it sound like this is a required way to do it. This is how I would go about doing this and I hope you'll find it useful as well. So step one, step one is really two steps in one but I think we can handle it. So first step is just to review the theory of change, right. Everything starts from our theory of change. We want to monitor the theory of change,
and we have to understand that theory of change. So we want to
go to the program designer or activity designer.

[0:14:00]

Pick their brains. Understand why they’ve built this theory as they
did, why they think these interventions will lead to these outputs,
why they think those outputs will lead to the various outcomes,
why this whole theory eventually leads to a lower youth crime
rate. Once we’ve understood that theory of change, we want to
go to the logic model that visually depicts this theory of change.
And it’s this second part of the definition of this step one that I
think is just very critical and something I don’t think we see
enough of. So logic models come in all shapes and flavors, right.
Sometimes they’re very detailed. Sometimes they’re very
summarized. A logic model that is used by a project designer or a
logic model that is shared by stakeholders might be summarized
and not very fully articulated.

[0:15:00]

When we want to do performance monitoring, we really want to
unpack that logic model and make it more on the end of the
spectrum of being really fully articulated. So let’s look at the logic
model that we are currently working with. We’ve got a couple of
interventions listed at the bottom, and then we’ve got three
boxes that show some outcomes we want to achieve in order to
get our end outcome of the youth crime rate reduced. So I think
we can do a little more than this; I think we can unpack this and
make it a little more detailed. So a couple of things I like to have
in my logic model when I’m looking to do some performance
monitoring, first I want to label the different elements of the logic
model, really understand which parts describe interventions,
which our outputs and outcomes.

[0:16:00]

I really want to make sure that my logic model in particular
includes outputs and outcomes, ’cause those are the things that
we really want to monitor as part of our performance monitoring
plan. And then finally this last point, really want to specify that
sequence of outputs to intermediate outcomes in detail, right.
The more articulated our logic model is, the more helpful it will
be to us in performance monitoring so that we can make the best
choices about what to monitor in this theory of change. All right.
So going back to our logic model, let’s unpack it. Let’s expand it. Let’s open it up and put more details in. And that brings us to this, a revised version of our logic model. As you can see, it’s big. It’s going off the slide even. On the left, we have our interventions.

Our capacity building and grants to our alternative rehabilitative centers. And then below that we have our interventions for judges. We’re going to do workshops. We’re going to do tours of alternative rehabilitative centers for judges. That’s going to lead to our outputs, which is just the results of those interventions. And then those outputs lead to a whole series of intermediate outcomes that stretch all the way through a whole series of intermediate outcome steps leading eventually towards the end, towards a decrease in juvenile offenders who commit crimes again, right, a decrease in juvenile offender recidivism, to our end outcome of the youth crime rate being reduced. We’ll go in through these steps in a little more detail as we proceed, so don’t worry if it doesn’t all make sense to you yet. All right, so that’s step one, understanding our theory of change and unpacking our logic model.

Step two, and for me this is the real fun part. And if you’re an M&E specialist, I hope this is the fun part for you too, right. It’s brainstorming, right, how to monitor each of the elements in this fully-articulated logic model. So let’s go back to our logic model and just do it. Let’s go brainstorm how are we going to monitor each of these elements of our logic model theory of change. I like to start from the bottom up, or in this case from the left. So I’m going to start with our interventions and outputs for our alternative rehabilitative centers. So we’re going to provide some TA to these ARCs and some grants. How do we monitor that? Well with outputs we want to think about quantity and quality. So we’re going to maybe do some narrative reporting.

Just describe what was the technical assistance that we provided to these ARCs. We might ask about the quality of that technical assistance by just getting some feedback from these alternative rehabilitative centers on the quality of the technical assistance provided. For our grants, we’re going to count the number of
grants we provided and the total dollar amount of those grants. Now those outputs are going to lead to these intermediate outcomes, right. So how are we going to monitor these intermediate outcomes? Well let’s brainstorm. So are the ARCs implementing improved practices? Well we might count what percent of the technical assistance recommendations that we made were actually implemented by these ARCs. We might do some direct observations or ask these ARCs for feedback on the implementation of these improved practices. The grants, they’re supposed to use the grants to improve training or improve infrastructure for their staff. So we’re going to count the number of ARCs who use the grant money to train staff and the number of ARCs who use grant money to improve infrastructure.

These intermediate outcomes should lead to the ARCs providing quality alternatives to detention services, right. How are we going to monitor that? Well we might calculate the percent of the ARCs who are providing quality alternatives to detention services. Right away when I start to brainstorm, I think okay, I’m going to have to define what it means to provide a quality alternative to detention services. So I might have to come up with a rubric. I might look at some international standards of what a quality alternative to a detention service is. I might do some qualitative observations by visiting those ARCs and seeing what they’re providing, or I might even do a trained observation in which I’m looking specifically at practices against a protocol of good practices. Let’s go down to our interventions for judges. We’re going to do some trainings for judges and we’re also going to do some tours.

So how do we measure these outputs? Well we’re going to count the number of judges that we train. We might also get some feedback. What do the judges think about the quality of our training? We’re going to count the number of judges we provide tours to and get some feedback on that. We might even ask the ARCs for feedback on the quality of these tours. This is going to lead to hopefully improved knowledge about alternatives to detention and more positive use about alternative to detentions by these judges. How are we going to monitor that? M&E specialists, I hope you immediately thought oh we'll do a pre and post-test, right. Test judges' knowledge before they took our training and then we'll test afterwards to see if our training helped
them learn more about alternatives to detention, help them have a better view about alternatives to detention. What’s that going to lead to? Well hopefully it should lead to judges increasing a percent of juvenile offenders who are assigned to alternatives to detention rather than to detention.

How are we going to measure that? Well, we’ll measure the percent of juvenile offenders who are sentenced to alternatives to detention as opposed to the percent assigned to incarceration. All right, here’s the second half of our logic model. What are some of these next intermediate outcomes? The first set of outcomes you can see here on the left are ones that we just talked about. But what do they lead to over in the middle? So here are these middle outcomes. We are expecting that as a result of going into these alternatives to rehabilitation that they’ll respond positively. They’ll learn some rehabilitation lessons, some life skills. And their overall wellbeing will improve while they’re in the alternative rehabilitative centers.

And that after when they leave these alternative rehabilitative centers, then they will sustain this wellbeing; they’ll sustain the lessons that they learned in their communities. Well how are we going to monitor that? Well, while they’re in the alternative rehabilitative centers we might do some surveys of these youths. We might do check-ins at the alternative rehabilitative centers to see how these practices are going. And then after they leave the alternative rehabilitative centers, we might do some surveys or check-ins with these youth in their communities. Finally, we get close to our end outcome. We expect all of this to lead to a decrease in juvenile offenders committing crimes again, right.

And then ultimately to a reduced youth crime rate. How are we going to measure that? Well here we might look towards government data, right? We’ll calculate the juvenile offender recidivism rate, and we’ll also look towards government data for the youth crime rate. All right, so we made it through. We went through all the elements of our logic model. We brainstormed different ways that we might monitor the theory of change. We tried to think of some quantitative measures. We also tried to
think of some qualitative measures for each one of these elements of our logic model. And hopefully with this fully-articulated logic model it really provided us with this roadmap, to guide our brainstorming about these different ways of monitoring.

All right, that brings us to step three. Now I think we want to consider a little bit about analysis and use of this data. How might we analyze some of these ideas we have for monitoring and how might we end up using some of these measures? So here are a couple of ideas. One is we might want to start thinking about disaggregate. How are we going to disaggregate some of these measures into subpopulations: male/female, right? Judges in one part of the country versus another part. And then we also think about other kinds of comparators, so are we going to compare some of these measures over time? Are we going to compare some of these measures against international standards or best practices? So start thinking about ways we might analyze this data. And then next we want to do some kind of scenario planning, how are we going to use this data. Is this going to be useful for us in making decisions, right?

If we have a measure and if the results of that measure don’t have any impact on how we might make decisions about this program or a program that follows, then maybe that’s not something that we need to monitor. I’m not going to go through all our logic model needs, but let’s look at a couple examples, particularly on the issue of disaggregates. So one of the intermediate outcomes that we wanted to measure was the percent of judges who are sentencing youth offenders, juvenile offenders, to alternatives to detention. How might we want to analyze that data? Well we might disaggregate it by each judge? We’re going to train judges. We want to track each one of those judges and see which judges are actually sentencing more youth to alternatives to detention and which judges are not. Maybe there’s some regional variation. Maybe there’s a more conservative part of the country where it’s harder to get the judges to sentence the kids to alternatives to detention.
We might look at the type of juvenile crime. Are they sentencing juvenile offenders who commit lesser offenses to the alternatives to detention? That's something we probably want to disaggregate our information by. So we start to think about these ideas of how we might analyze this data. Similarly, when we look at these outcomes towards the end of our logic model, we also want to think about some disaggregates. For the recidivism rate, we definitely want to look at the recidivism comparing those youths who are sentenced to alternatives to detention compared to those who were sentenced to incarceration or detention. Our theories suggest that youth who are sentenced to alternatives to detention are going to be more likely to not commit crimes again. So we'll want to disaggregate by that.

So just a couple of thoughts about how we might analyze this data. So that brings us then to step four. And for me this is the not very fun part of planning for monitoring performance. I kind of wish we could monitor everything, but unfortunately, we don’t have unlimited time. We don't have unlimited budgets, so we want to start thinking about winnowing our plan to what is most important to monitor and what is feasible to monitor. So these are kind of vague terms, so let's start with determining what's important to monitor. How should we determine out of all these ideas we had about monitoring what are some of the most important? Well, we can ask our stakeholders, right? We can ask the program designer or our project designer.

We can ask government stakeholders. Which of these aspects of our logic model are the most important to monitor? We can think about our theory of change and which of these expected results in our theory of change are the most necessary, right. Are they sort of key links in this change for our theory of change to work? We might want to monitor those especially? Or we might want to look at what are those aspects of our theory of change where we think there's the most uncertainty; we’re really not sure of the assumptions behind our expectation that one intermediate outcome is going to lead to the next intermediate outcome. We also want to think about what is required, right? We have certain indicators that are required for certain programs. And then finally, we might have evaluation questions or learning questions that we have planned.
And we might think about how our performance-monitoring data will help support those evaluations and learning activities, and that might encourage us to keep those indicators. So just an example, starting with the first half of our logic model, one intermediate outcome judges increasing, right? Their percent of juvenile offenders assigned to alternatives to detention, for me this is a really key outcome that we'll want to monitor, right? If the judges aren't changing their behavior in how they sentence youth, then the rest of our theory of change really falls apart. So this is to me one of those imitators that is really important; I really want to make sure I monitor even if it's going to be an expensive or difficult thing to monitor. If we go to the second half of our logic model, we might come to this result.

We're expecting juvenile offenders to respond positively to the rehabilitative services in our alternative rehabilitative centers. And this might be an area where we're not sure about the quality of these alternative rehabilitative centers. This might be an area where we think it's a really weak link in our chain of results, and so that might encourage us to focus on this as something we really want to monitor because of the uncertainty that we have about whether we can achieve this result. Okay. So in addition to importance, we also want to think about feasibility. How feasible is it to monitor some of these things? And when we think about feasibility, we're really thinking about how expensive will it be to monitor something. How ethical, right? Are there ethical constraints on what we want to monitor? In our case, we're dealing with youth, right. So we need to be especially careful when we're monitoring youth or doing surveys with youth.

There's certain ethical constraints and challenges we need to address. There might be technical challenges. That could just be if we're planning to monitor based on a survey. Surveys have a lot of technical requirements we want to make sure we meet. Or it could be just there are some places that are very far geographically to reach, and that makes it difficult to monitor. And then we also want to think about the quality. It might be possible to get data but it might be infeasible to get real high-quality data, so it's something to keep in mind. And there's of course all kinds
of other reasons that might make it difficult or not feasible to monitor some of these ideas that we had during our brainstorming session. Again, just a couple of examples.

[0:33:00]

One of the intermediate outcomes we thought we might monitor is judge’s knowledge, right. After we train them, we thought we’ll do a pre and post-test. We’ll test their knowledge before our workshop and then we’ll test after the workshop to see if judge’s knowledge and attitudes changed. This is a pretty standard thing we’d want to monitor. However, in this case, we’re talking about judges, and I know from personal experience that sometimes asking judges to take a test is not something you want to do. These are high-status individuals. So they might be insulted to the idea that they have to take a test. So the chief of party of this activity might say, “No way, we’re not going to give judges a test. They won’t take it. This is just not something we’re going to be able to monitor.” So in that case, we’re not going to be able to do it. Similarly, when we go to the second half of our logic model, we talked about maybe doing some surveys or check-ins with youth in their communities after they leave the alternative rehabilitative centers.

[0:34:00]

This is something that’s going to be tough to do, right. Tracking these children after they leave the alternative rehabilitative centers to different communities will be difficult, time consuming. In addition to doing surveys of youth means we have to get consent, possibly institutional review board to review it. So it could be very expensive, very difficult. It might be something that we think is not feasible for this particular activity. All right, and that brings us to step four. After we’ve winnowed. We’ve brainstormed. We’ve thought about how we might use the data, and then we’ve winnowed our plan. We’ve removed some of the things that we thought about monitoring. Now we want to just kind of step back, think about the big picture, right. And so here I think we want to look at our overall coverage in monitoring these elements of the theory of change.

[0:35:00]

Do we have enough things that we’re monitoring so that we really have a good idea throughout the life of this activity of whether
our theory of change is working as we expected it to, right? So we'll look at overall coverage. We'll also look at the mix of quantitative and qualitative measures, and then we'll talk about timing. I'll talk about timing in a little bit, but let's think about those first two points, overall coverage and mix of quantitative and qualitative. Here's our logic model again, and I'm just going to add golden scatterplot icons for all of those elements of our theory of change where we're thinking of doing some quantitative monitoring. There I've added some golden text boxes to indicate where we're doing some quantitative monitoring. And then again, this is just to give us an idea do we have good coverage. Do we have a mix of quantitative and qualitative monitoring?

And I think we do. We might not be testing the judges and monitoring whether the judges learned about alternatives to detention from our workshop, but we are going to be monitoring whether their behavior changes as a result. So we still have some pretty good coverage. And then looking at the second half of our logic model, again, we see we have a mix of some quantitative monitoring, some qualitative monitoring, and we have pretty good coverage overall. And then finally, I want to talk about this last point about thinking about timing. So I love logic models. I think logic models are great. They're very helpful roadmaps in thinking about plans for performance monitoring. And while they show the sequence of events, they don't really tell you anything about when events are timed over the life of an activity, right, so when particular results or changes in outcomes you can expect at different times.

So I think it's worthwhile when you're thinking about your overall performance-monitoring plan to step back and think about the timing of the results over the life of your activity. Here I just put a timeline in and noted when we're starting different interventions, right. We might start with our interventions with the alternative rehabilitative centers and then we move on to training the judges, right, and providing the tours for them. And when we start to look at some of the monitoring of juvenile offenders, particularly these intermediate outcomes, right, we might not see the behavior change in judges, starting to sentence the juvenile offenders to the ARCs towards the end of year two, right, which means that we have from the beginning early year three to the –
end of year five of our activity, to determine for a juvenile offender who was sentenced to an alternative rehabilitative centers we can look over at that whole timeframe of two to three years of whether that juvenile commits a crime again. So that’s fine for juvenile offenders who are sentenced soon after we trained judges. But we also might have a judge who assigned a juvenile offender to an alternative rehabilitative center towards the end of year four or in year five. And so in those cases, we won’t have a long amount of time to really track that youth to see if they’ve committed a crime after being in these alternative rehabilitative centers.

So unless we do monitoring after the end of the activity, we don’t really have a good sense of whether that juvenile committed a crime again. So I think thinking about that timing can really help us consider how well we’re going to be able to monitor some of these end outcomes, right, or these later intermediate outcomes, right, the decrease in juvenile offenders who commit crimes again. How does our timing affect our ability to really monitor that adequately? Will we see changes in indicators as a result of our programming during the life of our activity in those later intermediate outcomes or end outcomes? So timing of it is just something you want to consider that really you need to take a step outside of your logic model to consider.

And that’s the end of these five steps. And just to review, we want to really make sure that we understand our theory of change and unpack it to get a fully-articulated logic model. We want to brainstorm how we might monitor all those elements of our theory of change. We want to consider analysis and use of data. We want to winnow the plan to what’s important and feasible, and we want to look at the big picture. Now that of course is not the end of the journey for our hero, the M&E specialist. There’s lots more to do. Refining the specific indicators that we’re going to monitor or developing those qualitative protocols for some of that qualitative monitoring.
But I think these five steps are something that I think in particular USAID can really collaborate with our partners and be part of this process in thinking about how to monitor performance based on the theory of change and really be involved in these five steps. Again, this is not a required set of steps; this is just how I would approach it. I hope it is useful for you in thinking about how you might monitor performance based on a theory of change.

[End of Audio]