Core Components Analysis

September 8, 2022

2:00pm-3:00pm ET

Webinar Transcript

Russell Cole

Thanks all for coming. We're excited to again talk with all of the TPP Tier 2 grantees and evaluators. It's been about a month since our last webinar on analysis planning, so we're happy to be in front of you all again.

A quick moment on housekeeping...we tried to mute everyone on entry or are in the process of muting everyone on entry. We're going to try to keep lines muted to minimize any issues with feedback and echoes.

If you have a question, please submit it into the Chat at the bottom of the WebEx screen. We have some team members who are watching that, and we'll also have a Q&A session at the end. So we'll try to address any tech issues that you have immediately via the Chat, and we'll save some of the substantive questions until the end. Please do stay muted throughout the presentation to eliminate any tech interruptions, and I think we're going to begin now.

Next slide, please.

Here's a high-level presentation of what we're going to try to do today. In a moment, I'll spend a couple of minutes doing introductions; and then we'll get to the main content. We'll start with a quick reminder of some of the background material from the introductory webinar we did on this topic last year. Then we'll spend some time defining what components are, how we should think about them specifically in the TPP context since there's a lot of different content areas and they all have different definitions and thoughts about how to define components of programs.

We're going to briefly discuss the tool that we've been working on for OPA to enable program developers, such as yourself, to disaggregate TPP programs into their individual components. This is definitely going to be useful for reporting and for presentation purposes, and we're going to make that component checklist available to you all in the future if you'd like.

The latter two-thirds of the presentation is going to focus more on analytic approaches to link components to outcomes. If you have any questions, submit them; and we'll touch on that at the Q&A session at the end. We'll

follow up with links to the slides and a recording of this presentation in the future.

Next slide.

So here's who is speaking today. Emily is a researcher at Mathematica and a member of our TPP Eval TA Team. She has experience using component analyses to identify the drivers of substance use prevention programs, and we're going to have a cite for one of her articles at the end of the presentation. She's leading our team that's developing a series of resources for describing components of TPP programs, including the checklist and instructions that we'll talk about today.

I'm Russ Cole. I'm the PI on this project. I've been doing eval TA for over a decade with OPA grantees. I've been working with Emily on this program components work for OPA, and I've developed some of the earlier versions of the TA briefs that we'll talk about as some resources.

All right, so let's begin. So why are we here today?

There's an existing body of evidence about effective TPP programs, and you're in the midst of a rigorous impact evaluation right now. So that's good. There's more evidence about effective programs, and that's always welcome. That being said, there's much less known about the drivers of effectiveness within these programs, like what's the secret sauce that makes some programs work and others not so much. Is it the content? Is it the mode of facilitation? Is it a program needs to be a certain length? Is a combination of multiple things?

The National Academy of Science has completed a recent report where they investigated the research about components of effective programs for improving optimal health. One of the things that they highlighted is that there isn't a whole lot of research here. More research into the components of TPP programs is really warranted.

So OPA asked Mathematica to help develop its research agenda on components of TPP programs to try to address this gap. One of the first things that OPA is hoping to see is more systematic documentation of the components of TPP programs so that there's more information about these ingredients. To support this work, we developed a components checklist to standardize the process of documenting components; and we'll talk about that. It's something that you might want to use as part of your reporting that you're on today. In addition, linking components to outcomes will help create some of the necessary evidence to fill the gap identified by that National Academy of Science's report.

So in sum, there's some work to be done here. That's the kind of motivation, and I'm going to pause here and turn things over to Emily for a little while.

Emily LoBraico   Thanks, Russ.

So we're going to spend a lot of time later talking like nitty-gritty, specifically about different types of component analysis. But for now, we're going to talk about it a bit more kind of high-level and general.

Component analysis offers a different way to learn about programs when compared to traditional impact evaluations. Traditional impact evaluations assess the effects of the whole program, so the program in its entirety and whether it's related to better outcomes for those who receive it versus some kind of comparison group. Whereas a component analysis is a way to learn about the smaller pieces of the program or its components and figure out which pieces or components actually make up the program and how they're related to different outcomes.

For example, the program shown in this logic model on the left-hand side, an impact evaluation might test the impact of the program on the proximal, so those shorter-term outcomes, and the distal outcomes based on some kind of underlying theory or some other evidence. Of course traditional impact evaluations are valuable; they're very important; but they don't tell us everything about a program and the way that it affects these different outcomes of interest. What we don't learn in a traditional impact evaluation is anything about what's actually inside that program and how those components are related to those short-term and long-term outcomes.

For example in this made-up TPP program, there are three different distinct components. By doing a component analysis, we can look at what those components are and describe them and define them and evaluate how they're associated with short-term and long-term outcomes.

Before we conduct a component analysis, we have to know what our components are. In a lot of ways, this might be simple. They're just the ingredients of the program. They're all of the things that go into the program. They're usually defined in a manual or some other kind of program documentation. But it's actually pretty complex because there's so many different ways to think about components and different ways to break down a program into its component parts. Although none of them are wrong, it's important that we're using kind of the same language across programs so that we can look at the components not just within a program but across programs.

In order to be systematic in this process of describing and analyzing components of programs, we have to be using the same language. With a

lot of support from the National Academy of Science report which dove into this topic, we have articulated seven types of program components that together can describe entire programs and the way that they're intended to be delivered. We went over these in the last webinar, but we're going to dig into them a little bit more right now and talk about the checklist and how it describes and uses these component types.

As I said, there are seven different component types. The first one is the one that gets a lot of attention. It's the content component. This is the intended subject matter of a program. An example of a content component is something like information about condoms. The delivery mechanisms are the intended principles and practices by which the content is provided, so a lecture is a common delivery mechanism. The formats are the intended structure and organization by which that content is delivered, so in-person and virtual formats are commonly discussed.

The staffing component is the training and characteristics of the people who deliver the content, and this can cover things both like the developer required training in order to facilitate a program. It can also cover preexisting training or experience, so for example if teachers are the intended staff for a program. The dosages are the intended duration, frequency, and intensity of the program; and this could be at a super zoomed-in level, such as at the activity level, maybe a 25-minute activity. But it can also be across the whole program, such as six two-hour lessons.

The environments are the intended settings or locations where the program occurs. For example, a health class would be an environment. Then, the intended population characteristics are the characteristics of the intended population who receive the program. Again, this can be a little bit more broad, like high school students; or it can be very specific, like LGBTQ youth in the seventh grade.

Together these seven different types of program components are what form the youth experience or the intended experience of the youth who receive a TPP program. Even those I just described these each as separate things, it's really the combination of program components that describe how a program is meant to be implemented.

For example, if we think about the single activity from a multi-session TPP program, there are several different components happening at the same time. So this is a 20-minute small group activity with high school students during health class featuring a discussion about communication and healthy relationships. We can actually pull out six different components.

There's a dosage component, there's a format, there's a population characteristic, an environment, a delivery mechanism, and a content. So

even though we would implement the whole activity this way with all six components happening at the same time, it's possible and actually likely that the developer or some other key stakeholder, like an implementer who implements this activity often, has thoughts about some of these components being more essential to the program's impact on program participants. So these are what we call core components. In other words, if you think about all the components in a program, there's some subset of those components that we, or someone, thinks of as core; and these are hypothesized to be driving the program effects.

A lot of times we don't have much evidence beyond theory or other studies to support these ideas. We could do rigorous effectiveness evaluations to produce evidence about which core components affect outcomes if the study is specifically designed to test the component impacts; but a lot of times we're already in the middle of implement a program or collecting data, or we have already implemented it and we have data. So there's other ways to generate useful preliminary evidence using this data about components without building tests of components into the study design, and that's a lot of what we're going to discuss later on in this webinar.

This is a screenshot of the checklist tool that Russ was referring to earlier. Using this tool, we can disaggregate programs into their components. So there's a place – or each of the component types listed. This is obviously a screenshot. There are all seven component types listed; and there's a place to indicate whether the component is present in the program; where in the program it's present, like a page number; whether or not that program component is core; and if there are allowable or unallowable adaptations specific to that component.

There are a few other helpful materials that accompany this checklist, like an instructions document and a framing brief, that provide a little bit more context. The idea is that by using this checklist to document the components in your program, you can get to a place where you're outlining your program very, very simply – like this slide presents – to explain it very succinctly to someone else.

There's lots of benefits of being able to describe program components in this more efficient manner. So for developers, being able to communicate their programs' components and which are core. For implementors, they'll be able to document implementation and OPA-approved adaptations very clearly, kind of alongside with the developer described as the intended program implementation. Then for people who are potential users of the programs – so they're selecting between two or more programs – they'll be able to compare across programs and make program selection decisions much easier than having to kind of like flip through several pages of manual documentation.

Then more broadly, the ability to compare TPP programs and their components to each other will support research efforts that can lead to stronger evidence and informed program improvement. So this type of work can answer questions like which components are most common across TPP programs, and which components are associated with specific behavioral outcomes.

I'm going to turn it over to Russ, and he's going to dig into specific different types of component analysis.

Russell Cole   Thanks so much, Emily.

Let's go to the next slide.

So that was really helpful. Now we have a way to think about program components, thinking about core components. We've got a sense of the checklist that we can use to operationalize this work and how it can help with clear dissemination about the ingredients of the programs.

I'm going to switch gears and try to talk about some empirical work. In this section, we're going to try to explore ways to understand which components of a program are the ones that appear to be influencing outcomes.

Just to say this...some folks do these types of analyses in a really rigorous way. They randomize folks to get different component experiences. I'm going to present a way for us to kind of get into this in a simpler manner, building off of some of the work that you're probably already doing as part of your typical program implementation in CQI. So really, this is work that's feasible for you to do with the data that's already being collected as part of your ongoing rigorous impact evaluations.

A couple of notes about a word on this slide, and it will show up throughout the remainder of this presentation. I'm going to use the term "influence" here as a shorthand for talking about the relationships between components and outcomes. Everything that we're going to talk about here is going to be relatively exploratory. We're not necessarily identifying a causal connection between components and outcomes through a rigorous design. There's going to be more on that in a bit, but I just wanted to highlight that nuance about language here.

I also just wanted to say I'm going to largely be talking about core components here to try to narrow our focus to the subset of components that we hypothesize to have the strong relationship to outcomes. This work is definitely more tractable, and we try to narrow our focus and refine

ourselves to really just thinking about the small number of potentially core components of a program.

Next slide, please.

Here's the four steps that you'll need to take on if you're interested in going down this path. I'm going to talk about each of them over the next few slides, but here's the sequence in a nutshell:

First, you need to hypothesize what the core components of your intervention are.

Then, you're going to need to do a pathway diagram that shows how the individual core components are going to be related to individual outcomes of interest.

Third, you'll need to do some measurement both on how folks are receiving the components that you're interested in and also on the outcomes that are in your pathway diagram. Again, I think you're probably already doing a lot of this already.

Finally, you'll have to do some analyses to link the variation in implementation experiences to variation in outcomes. That will help you to tell that story about promising components.

So let's dig in. Next slide...perfect.

The first step is really all about defining the core components of your intervention. There's a couple of ways to do this.

First off, there's the checklist that Emily was talking about. That's a pretty straightforward way to do this. It helps you to lay out all of the components of your program and the subset of them that are core. So that's one possible way of doing this. Some folks may want to just skip this step because you have some prior preconceptions about what the key or the core ingredients of your program are.

So maybe your hypothetical program has three structural elements that are intended to address a known need for your population, and they're shown here. There's a personal safety and consent discussion. There's a condom demonstration and STI lecture. These are the key ingredients of your program that you think are going to be driving outcomes. If this is where you're at, then you're done. You know what your core components are already; you can move on. I'm going to continue with these three core components as a working example for the rest of the presentation.

Next slide, please.

The next step asks us to create a pathway diagram to link components to outcomes. This is really kind of like intersecting our understanding of the general program logic model with a more fine-grained thinking about core components. If we previously thought that we had a sense of which outcomes our program as a whole is going to affect – and again, that would be like a programmatic logic model – now we're going to go one layer deeper and think about which of the individual core components of the program might affect each of those individual outcomes of interest.

As Emily mentioned, we're going to be thinking both about proximal outcomes -- those are outcomes that are really well-aligned with the content and the activities of a program -- as well as the more distal outcomes, things that are more downstream from the proximal stuff. That's oftentimes like the more policy-relevant outcomes. Just to say this...in many interventions, proximal outcomes tend to be things like attitudes about sex or intentions to have sex or feelings of self-efficacy. They can also be measures of process, things like satisfaction or engagement. The more distal outcomes might be the behavioral outcomes, things like initiation or risky sex or pregnancy. Some folks even think about proximal/distal in terms of short- versus long-term.

The idea here though is that we're going to create a diagram that shows this visually, where core components are going to influence immediate or proximal outcomes; and those immediate and proximal outcomes will ultimately influence longer-term distal outcomes.

Next slide, please.

This is a slide of what this looks like. It's the same example from Emily earlier. We've got the three core components shown on the left-hand side: personal safety and consent discussion, condom demonstration, and STI lecture. We've got two proximal or intermediary outcomes that these core components are supposed to affect, attitudes about sex and knowledge about pregnancy and STIs.

Now, here's an interesting thing. It may not be the case that we believe that each core component is going to affect all of these outcomes. For example, maybe we think of the personal safety and consent discussion is going to affect attitudes about sex; but it won't have any content to affect knowledge about pregnancy and STIs. That's why there's only a single arrow from this component to only one of the proximal outcomes. We can make these types of hypotheses about how each of our components affect outcomes and test them later on throughout this process.

Just to say this...finally, you might assume that through the improvements in some of the proximal outcomes, we might see improvements or delays in some of the distal outcomes. The idea here is really to kind of come up with an articulation of the outcomes that the intervention might be affecting at a really fine-grained level. Specifically, we're trying to identify which outcomes are going to be affected by each of the individual components of interest.

Next slide.

The first part of Step 3, we really need to think about the means or the mechanisms by which program participants are going to receive the intended core components; or, thinking about it from an alternate perspective, what are the potential barriers that might impede a program participant from receiving the intended core component. So here's some examples – illustrative but definitely not exhaustive features of implementation.

There's attendance and dosage. So if youth aren't attending program sessions, it's going to be impossible for them to experience the core components.

Engagement – like if youth aren't paying attention or are only superficially engaged, the components won't affect them; and therefore outcomes shouldn't be expected to improve.

Quality of delivery – maybe youth are attending and are engaged, but the material is being delivered terribly. In that situation, then the content might be adequately presented in a way that youth outcomes will be affected.

In sum, the idea here is that we're trying to identify the avenues and the barriers that might help or hinder participants from receiving the core components. So really, this is like listing the features of implementation that matter. What we want to actually collect data on is for us to be able to understand youth experiences of each component, so we want to really identify data elements that are feasible to collect and the things that are going to matter most in terms of measurement of youth experiences with each of these components.

So the idea here is really that we want to have those data to see the extent to which youth are having different component experiences. That sets us up for the next thing on Step 3, to collect data on outcomes in our pathway diagram probably through like surveys or other data collection sources. We really want a measure or a series of items for all of the outcome constructs that are in our pathway diagram. Why? Because like now we're going to have the necessary ingredients to explore the extent to which variation in

component experiences – that would be based on the variation and things like attendance or quality or engagement – is associated with variation in these outcomes that were hypothesized to be affected by the individual components.

Just to say this...while the outcome data are really required for this type of analysis, having baseline measures of these outcomes really helps the storytelling and the credibility of the analyses. I'll come back to that in a moment.

What we're going to really need to do at this point is link the implementation data to the outcome data. We want to know how Johnny's implementation experience in terms of his attendance, in terms of his quality, in terms of his engagement, for each of these components is linked to their survey data to enable these types of outcomes and these types of analyses to occur.

Next slide, please.

So here's a slide that shows what this might look like. You'll have rows for each observation or each participant in your dataset, and you'll have columns of information that will capture variation in exposure to individual components. So this is like on the left-hand side you'll see attendance as a way to capture variation in component experiences. We've also got outcomes in the logic model on the right-hand side of this illustrative dataset. All of these data are linked at the individual level to enable these analyses to occur.

Next slide, please.

So there's lots of ways to link variation in implementation experiences to variation in outcomes. In the remainder of today's presentation, I'm going to showcase three different approaches that kind of answer three different classes of research questions:

One is going to focus on a quasi-experimental approach. It will answer the research question about the effect of being exposed to a component on an outcome.

One is going to be a correlational approach. This will answer the research question about which component or components are best predictors of an outcome.

The last one focuses on a structural equation modeling approach, which answers the research question about how everything in your logic model is related potentially sequentially.

Just to say this...I mentioned this earlier, all of these analyses can be done without using data from a comparison group. We're going to be able to leverage variation and exposure to components as a way to understand how they influence outcomes without actually manipulating components. That's certainly feasible, and it's potentially useful to do in future evaluation projects; but what we're trying to talk about today is ways to use the existing data that you have in hand to try to answer some of these questions.

So let's go to the first approach, the quasi-experimental one. Next slide, please...great.

So the quasi-experimental approach allows researchers to estimate the difference in an outcome between youth who received a component of the intervention compared to those who did not. Receipt can be defined as being exposed at all to a component, whether they received the component as it was intended. There's a lot of different ways that we can slice this. The main idea here is that it is particularly useful for understanding the effect of a specific core component.

For the next working example, I'm going to focus on whether youth attended a condom demonstration activity on outcomes.
Next slide.

The first step in this process requires us creating an indicator, a yes/no variable, for whether or not an individual experienced a given program component. How you define that indicator is going to really define or determine what research question that you're answering. Perhaps you want to know the effect of taking up any aspect of a particular component, or perhaps you want to know the effect of getting the entire component, or perhaps you want to know the effect of getting a dose of a component relative to some established threshold. For example, maybe the developer said that there's a minimum attendance requirement for a particular aspect of a program.

In sum, you're going to create an indicator to differentiate folks who got a program component according to your definition relative to those who didn't. So this is kind of like effectively making a treatment and a control group for this component exposure outcome or this component exposure variable.

All right, next slide.

So we've got an indicator that we could use to compare folks with different exposures to a component, like the condom demonstration activity; and in theory, we can certainly look at whether those two groups have different outcomes. However, a skeptic would probably argue that our two groups

might differ in something above and beyond their exposure to this component and that that difference might be what's truly causing the difference in outcomes.

So to address this, we strongly recommend assessing the degree to which the two groups are equivalent on key variables that you think might influence the outcome of interest. For example, you can compare the demographic profiles of your two groups. Are they comparable? You can use baseline survey data to compare the pre-intervention measures of the outcome of interest. Hopefully, you'll see that you have two groups that look relatively comparable on these variables.

In a perfect world, you'll also have some measure on your survey of something that could proxy for motivation. Maybe you have grades or persistence in other programs. Again, skeptics might argue that more motivated youth are the ones who attend components and, thus, any difference we observe should really be attributed to differences in motivation rather than differences in attendance of this component or not. So if you have a direct assessment of motivation or a proxy for this variable, you can show equivalence of the sample to convince a skeptic.

What happens when you assess equivalence, and you see that the groups look extremely dissimilar on an important variable of interest?

Well, you can exclude outliers. You can do a matching analysis to ensure that you're making a credible comparison. This would be similar to what you would try to do in a standard analysis to try to understand the effect of your complete program when you've got substantive based on inequivalence problems.

Next slide, please.

So once you've done enough work to demonstrate that your two groups are relatively comparable, that they differ only in terms of their exposure to the core component, it's appropriate to then compare them in terms of their score on the outcome of interest. Just to say this...we probably also want to adjust for baseline differences on the key variables that we're worried about just to improve credibility of this approach.

At the end of the day, this analysis is going to produce a finding that describes the difference in outcomes associated with component exposure, along with like a p value for that statistical significance of that difference. So you might be able to say something like the illustrative example findings shown here, like individuals who attended the condom demonstration answered 20% more items correctly on the sexual health knowledge assessment; and there's a p value shown there. If that's the kind of finding

that you'd like to have in terms of describing the role of a core component, I'd recommend planning for the quasi-experimental approach.

I'm going to say this here, but I'm also going to say this again later. These types of exploratory component analyses where we're focusing on like just the treatment group and their experiences, variation in experiences of a particular component, they're going to have less power than what our main analyses have. So we may not get statistically significant findings. We've actually got John on the call here; but it's useful for us to be thinking about Bayesian interpretation approaches for thinking about components to try to highlight things that look promising, even if our p values are greater than 0.05.

So John Deke and Mariel Finucane are going to talk about Bayesian approaches to interpretation of our overall study in a couple of weeks in a webinar. So I just wanted to quickly plant that seed. That's it for the quasi-experimental approach; so I'm going to shift gears and talk about the correlational approach, the second of these methods.

So this approach allows researchers to estimate the relationship between all measured core components and a single outcome of interest. This approach can identify which components are most or least influential in terms of moving participant outcomes. The core components are the horses in this analogy. You want to see which one wins as being most influential in terms of pushing outcomes.

This approach is, again, very exploratory. We're going to pit a bunch of variables against each other, and we're going to write about what we find. What kinds of research questions does this approach lend itself towards? Research questions where we're interested in all core components and a single outcome that they influence.

So for example, which core component – personal safety and consent discussion or condom demonstration or the STI lecture – which of those plays the biggest role in influencing attitudes about sex?

Next slide, please.

The analytic approach to be used here is going to be like a flavor of multiple regression. The dependent variable will be the outcome of interest, for example, participant refusal skills, self-efficacy. The predictor variables in this regression analysis are going to be the variables that represent implementation of each of the core components.

In our example, I'll have dosage or attendance measures for the consent discussion, for the condom demonstration, and for the STI lecture. As with

the quasi-experimental approach, we're going to do what we can to try to guard against omitted variable bias. So we should try to adjust for all of the x's that we've got measured at baseline that might influence our exposure to components; for example, measures of motivation or things that might influence outcomes, so things like a baseline measure of the outcome and demographics as well. So doing that additional adjustment will help us to feel more confident that the regression coefficients that we observe for our key predictors – that those are the measures of implementation for each of our components, that we can interpret them as the components' influence on outcomes.

So let's look at this visually, perfect.

This diagram is a way to visualize the analytic approach. Puzzle pieces represent our components of interest, and folks vary in terms of their dosage for each of them. The green baseline variable box represents our attempt to address omitted variable bias by adjusting for lots of stuff that might influence outcomes. The dotted line around everything acknowledges that everything on the left-hand side is correlated. So we're allowing the baseline variables, like motivation and the baseline measures of the outcomes, to covary out of the estimated relationship between components and outcomes.

So the arrows from those individual components to outcomes, like attitudes about sex, are the beta coefficients from our regression model. They would tell us the strength of the relationship that we're interested in. In this illustrative example, we see that personal safety and consent was the component that had the strongest relationship, the thickest arrow. That means it was our strongest predictor, the one that had the highest partial r2 statistic. That's the total amount of variation in the outcome that's uniquely explained by this predictor.

We can also pull out the beta from this regression and its p value to help with interpretation, and I've shown an illustrative finding on the right-hand side here. Personal safety and consent attendance was the core component that played the largest role in influenced attitudes about sex, and I showed a partial r2 statistic. Attending that lesson was associated with a 0.12 standard deviation with a p value improvement in attitudes. The idea here is that we used regression to pit the components against each other and found the one that had the strongest relationship with outcomes, and we pulled some numbers from this analysis to help make that case.

Again, this is a really high-level presentation. There are more details in some of the briefs that we showcase that we'll showcase at the end. But that's the correlational approach in a nutshell.

All right, so let's do the last approach – the structural equation modeling. The structural equation modeling approach is really an extension of the previous correlational approach. It allows the full pathway model to be tested, including multiple sequential outcomes, instead of just looking at one outcome at a time. So it's really an enhancement over the previous one, but I will say that this is more complex to operationalize.

In addition, instead of being a purely exploratory approach, it actually allows us to test the fit of our model to the data that we have in hand to see whether our pathway diagram is an appropriate way to link core components to outcomes; and we actually get a fit statistic out of it.

If you remember the example research question from the correlational approach, we looked at how three core components influenced a single outcome. It was self-efficacy. In this approach, we can extend that and ask which core components – so personal safety and consent discussion, the condom demonstration, or the STI lecture – plays the biggest role in influencing both proximal outcomes, like attitudes and knowledge, as well as distal outcomes like risky sexual behavior.

Next slide, please.

So to estimate a structural equation model like this, you can either use a specific structural equation modeling software package – I listed a few here -- or you can find the right procedures or syntax for your general statistical package to do this type of analysis. I'm not going to spend too much time getting into the specifics of how to do this. It's definitely beyond the scope of the remaining time that we have in this presentation. Again, the brief that we have that I'll point to in a bit in the references section has some references that will be useful.

In a nutshell though, it's an iterative process where you're going to estimate a model. You're going to tweak that model by adding or deleting pathways; and then really re-estimating it until you have a refined, a final pathway diagram that actually aligns well with the data that you have in hand. You'll note the model aligns with the data by looking at fit statistics at each step of this iterative process and seeing whether adding or deleting pathways helps to improve your model fit.

As with the correlational approach, you're probably going to want to try to adjust for all of the background x's that you have to try to account for omitted variable bias. Once you've got a model that fits the data pretty well, you'll have the quantitative results that you can interpret in a comparable way to what was presented earlier in the correlational analysis.

So you'll be able to identify which core components have the greatest influence on each of the proximal outcomes – so again, be comparable to our previous approach. But results for each outcome are going to be found in a single model. In addition, you'll be able to see how the core components influence distal outcomes as well and the degree to which the proximal and distal outcomes are related. So it's really an assessment of whether your path diagram is justified by the data. So if you're really looking to try to quantify your underlying logic model, the structural equation modeling approach may be for you.

Next slide, please.

This figure helps to showcase what your package will show – hypothesize all the potential relationships as arrows in the model showing how your components influence sequential outcomes, again, aligning with your logic model developed above. It's not shown visually here but implied by the dotted box on the left-hand side, you'll also allow all of the components in the background or baseline x's to covary just like in the correlational model.

Next slide, please.

This slide is intended to show a subset of the hypothetical structural equation modeling results for this working example. To simplify presentation for this slide, we only showed a subset of the path estimates here, those that were statistically significant in this hypothetical example.

So for program components, the hypothesized path model indicates that the personal safety and consent discussion have had a positive effect on knowledge about STIs. Like a 1.0 standard deviation increase in the dosage of this discussion is associated with a 0.38 increase in knowledge of STIs. As the figure shows, this analysis produced p values that can be used to indicate whether the relationship is significant or not.

In addition to reporting and interpreting the direct effects of program components on proximal outcomes, these SEM results can provide more nuanced information about how outcomes are related to each other. So one notable feature is estimating the extent to which proximal outcomes mediate distal outcomes, the indirect effects of these program components.

So here we see some suggested evidence that one of the proximal outcomes of the intervention, knowledge about STIs, has a strong negative impact on subsequent frequency of sex. So this finding would corroborate our assumption that increased knowledge about STIs is going to decrease the likelihood that youth would engage in frequent sexual behavior, since that might put them at risk for STIs. So it provides a contribution about the role of these proximal variables as key mediators. Again, so this is really just

like a hand-waving illustration. There's a lot more detail and references and guidance in the brief I'll show in a moment.

Let's take a moment to recap.

If you've got data on implementation of each component showing that participants vary in terms of their experiences of components and you have data on outcomes, you can answer these research questions that link components to outcomes. We really briefly outlined three approaches, each of which is really tailored to answering a broad type of research question shown here. Hopefully, one of these approaches sounded like it might be a useful thing for you to explore in the future. If this is the case and you think doing this type of analysis will help you unpack the black box of your intervention, it's important to know how to present this information.

We know that the RCT evaluation that you're currently doing is going to show the effect of the full intervention; that is, all of the components of your program. That will probably be the main story, the main piece of your final report or journal article. On the other hand, these types of component analyses are useful supplements to try to unpack. They try to give you a better understanding of which pieces of the intervention matter the most. But typically, they should be presented as supplements to the main analysis, especially if you're doing these types of exploratory analyses that I presented here.

Again, we'd want to talk about these as supplementary to the main analyses based on the key limitations that we're going to see in these types of work. So here are some of the things. Right now, we're in the space where we're talking about leveraging available implementation data as predictors of outcomes. We didn't manipulate them, so we really do need to be very careful in labeling these types of analyses and findings as exploratory.

A second one that I talked about earlier is the limitation of sample size. We're typically doing these types of analyses looking solely at the treatment group, so we've got probably half the sample size of our broader impact study. So we probably want to look beyond statistical significance to try to highlight findings that are useful. We shouldn't just use a p less than 0.05 as the sole barometer of whether or not a finding is useful or not.

So in sum, there are some limitations; but this is really a useful way to use available data to understand how core components play a role in the outcomes. If you find something, you can certainly use this information to set the stage for a more rigorous evaluation of individual components. For example, if you found that a key component was hugely important, you really have a strong argument, a strong justification, to try to do a future RCT where you assess the impact of that component alone.

Next slide, please.

So hopefully something that you heard today is going to be useful for you all to explore. If you're digging into your analysis plans, the exercise of thinking about core components is all about really getting into the weeds of your program and documenting its ingredients. That's the first step. Then, using the data that you probably have to better understand how core components of your program are affecting outcomes is a great next step.

This information really can provide insight into how core components are implemented and received by youth, and that offers insight into how a program could potentially be adapted in the future. Maybe there are certain components that should get more attention in future adaptations, or there are certain components that could potentially be dropped because they have no appreciable relationship with any of the outcomes of interest. This is definitely something that can help tell more of the story about your program beyond your overall impact analysis that shows the effectiveness of your program as a whole.

Next slide.

So if you're interested in learning more about this, there are two briefs on this topic that OPA has available on its Eval TA webpage. One is about structural elements of an intervention. One is more about analytic approaches for linking components to outcomes.

There's the example journal article that Emily wrote here, looking at how individual components affect outcomes using a really credible analytic approach for estimating that relationship.

As Emily mentioned earlier, please keep your eyes peeled for three additional resources. There will be a checklist for documenting the components of programs, instructions about doing this work, and a framing document that provides some definitions and helps situate this work within the broader literature and OPA's research goals.

I think that that's pretty much it. The next slide – these are our contact information for Emily and I. Of course, we're also reachable through your TA liaisons.

We're under time, so we've got plenty of time for Q&A if folks have any questions or comments about anything that they heard here.

	If anyone does have any questions, feel free to submit them via the Chat. Or we're a small group; so if you'd like to just go off mute and talk, we're happy to talk that way as well.
Emily LoBraico	I'm kind of curious to see if anyone had some ideas that popped in your mind about certain components that you might want to explore. I know I have some ideas from some of the grantees I work with, but I'm curious to hear if anyone has any thoughts.
Katie Henley	This is Katie Henley. I can pop in. One of the things I was thinking of for our study was Boston Children's Hospital, which we're looking at – it's called the MARSSI intervention. It has in my mind three core components. There's a main counseling session that's delivered virtually. It's an app – 28 days of app use on their phone, and then a booster counseling session. So those are kind of the three component parts.

One of the things that I know I'm curious – and I know Brittany is on the line too, so if she has anything to add as well – that I think we would like to explore is we've had really good attendance at the main counseling session but not such great attendance at booster counseling sessions with the app use.

So it's not that attendance should necessarily drive the decision about what's important, but I am curious in exploring when we get to the point of doing our impact analysis whether it was sufficient to attend the main counseling session to impact outcomes of interest -- hopefully we see an effect, I'm assuming we see an effect – or what type of influence on outcomes we see with a combination of the main counseling session and just limited app use or the main counseling session and just the booster session without app use. So that's just an example from one of our interventions of how I might approach this.

Russell Cole	Yeah, that sounds great. It sounds like you are already aware of variation in implementation experiences of your participants. It sounds like you've already got some data to suggest that there is non-uniform take-up of all of these key components. Is that right, Katie?
Katie Henley	Yes, that's right. We do monthly data monitoring. So we do keep our finger on the pulse of kind of which components people are completing, so we do have a sense of that now.
Russell Cole	Yeah, so I feel like you're 100% in the right space. You've got your rigorous impact evaluation that's going to be able to provide an estimate of the overall effect of offering this program to everyone; and you've got these implementation data that are telling you, hey, folks who actually were offered MARSSI substantively differed in terms of their actual implementation experiences. Let's explore that variation in MARSSI

experiences across these three components and see to what extent outcomes differ for folks with different implementation experiences.

It sounds like you've got the data, and you're interested in understanding the extent to which those variations in implementation experiences are associated with outcomes. So I really think that this is a great opportunity for you to add that as part of your planned analytic approach and to work with your TA liaison. Emily and I can also weigh in and chip in if you have any questions about how to do that type of work.

Are there any things that you'd like to talk about specifically here since we're all together? I know that some questions are coming in through the Chat, so we can jump to those as well. But if there are specific questions that you have, Katie, we'd be happy to try to answer them right now.

Katie Henley   That's a great – I'm trying to think if I have any specific thoughts yet. Beyond us knowing kind of what dosage looks like, I don't know if I've given it much more thought than that. Nothing particular popped up. So maybe go to the other questions, and I'll see if I can think of anything using the example I just gave before the time is up.

Russell Cole   Great.

So the first question that I'm seeing is from Jonathan: Do you recommend methods that compare, for example, outcomes for the highly engaged participants to the control or comparison group?

So this would be an alternate analytic approach, I think, for identifying the subset of folks in the treatment group who had a good component experience. It sounds like they were highly engaged -- maybe they attended a lot or something like that – to a subset of the comparison group. I think that the key thing in that analytic approach is going to be to try to convince yourself that you've got the right folks to compare your highly engaged treatment group against in the comparison group.

So first off, we're going to definitely want to make sure that we're characterizing this as exploratory because we know that we didn't manipulate these folks to these conditions. The highly engaged folks are kind of self-selecting into being engaged, or there's something else that's going on there. But I really think that the key thing would be to try to think about what are the variables that are going to drive someone to be engaged in the program and the variables that are going to drive outcomes.

So again going back to some of the stuff that we talked about earlier, I think it would be great if you had some kind of measure of motivation, since motivation might be associated with engagement and the types of variables

that we think are strongly predictive of the outcomes of interest. So certainly a baseline measure of the outcomes might be the single best predictor that we have, and we know that demographic differences tend to have a strong relationship with outcomes. So looking for equivalence on our eventual samples on demographic characteristics would be important.

I don't know if that's addressing your question, Jonathan. So please, feel free to follow up with more in the Chat or just unmute yourself. But hopefully, that was at least a little bit helpful.

Others, please jump in if there are other thoughts that you have.

I'm going to turn to the third question that came in: "Not a question but a thought...over the years, we have thought and rethought about the effectiveness of condom demonstration activities. Given some youth's past resistance about participating/logistical challenges, particularly during virtual implementation, et cetera, that may be something we'd want to explore."

So it sounds like, Archana, you're thinking about to what extent is there effectiveness around condom demonstration as a particular component. I'm pretty sure -- just to say this -- that we, Mathematica, recently did a review of the literature around condom demonstrations. Jean isn't here right now, but I think that she actually was part of the team that did some of that work. So we can certainly follow up and share some of the things that we found about the literature that specifically looked at the effectiveness of condom demonstrations as a particular component of interest.

We can certainly follow up about that. But I don't know, Archana, if that's answering your question, or if you're thinking about it particularly as a component of a study that you're thinking about right now.

Archana I figured I'd just unmute. It wasn't really a question as much as this is something that has come up for a lot of our studies, just because of our experience actually conducting those and having resistance from youth. Then, we end up having to kind of adjust a lot of times to different kinds of ways of implementing it to get by. But I don't know that we particularly have ever really done that particular analysis, so I'd be interested in the research for sure. Thank you.

Russell Cole Yeah, I mean just to say this, Archana, if you have those data that like kind of acknowledges the fact that youth have had substantively different experiences of that key lesson and some youth, I'm guessing, have fully participated in it; some youth have refused it; some youth have had a partial implementation of it. I think the argument here is try to understand the

extent to which variation in that critical lesson is associated with variation in outcomes.

Now, it's going to be important for you to do that necessary homework of trying to understand why are youth having different experiences. Is there some sort of background characteristic that we can try to ensure looks really – are there enough background characteristics that we can look at to ensure that we're feeling like these folks are reasonably comparable and that we can attribute any differences in outcomes to the differences in their condom demonstration experiences. If you can get to that point, that's a real contribution that you can make to showcase how the condom demonstration is a critical component in terms of affecting youth outcomes.

Emily LoBraico   Then before you move on to Jenn, just one more thing for Archana – and for everyone – is you can think about components in various different ways; and one of them is delivery. So like you mentioned, you referred to the fact that there are some cases where you've been delivering the program virtually versus in-person. That's definitely something that you could explore using these components analyses.

Russell Cole   Great, Jonathan, glad that that was helpful. The big thing was we're just trying to make an argument for trying to think about how to construct the right frame for thinking about who the "no component" group is.

All right, we've got another question from Jenn and two more minutes: "We have core components related to goal setting, SRH, and healthy relationships in that order; but attendance dropped off over the course of the program. So those who received the last sessions may be the most motivated or had the least distractions or difficulties with attendance."

So I'm just going to try to – so it sounds like the content of the program was structured where goal setting was offered first, SRH content was offered second, and healthy relationships content was offered third – if I'm understanding you correctly, Jenn. So it sounds like if you were interested in understanding the roles of each of these individual content pieces, it's going to be difficult to identify that because you have this kind of downward trend in attendance over time.

 Am I interpreting your question/comment correctly?

Jenn   Yeah, I think it's more of a comment that knowing those who attend that last session might be very different from those who didn't in a different way than those who did and did not attend the first seven sessions.

Russell Cole   Yeah, that's a good question – that there's like this ordering effect that's going to contribute there. I don't have a perfect answer on the tip of my

	tongue that I can efficiently say right now especially; and it's three o'clock, so we're out of time. How about this, Jenn? Can I follow up with you by e-mail, and we can try to think more about this afterwards?
Jenn	That would be great, thanks.
Russell Cole	All right.
Emily LoBraico	I think it's just lifting up this issue of, like, this is all correlational; and there's a lot of other things that might explain why you might see these differences. So even though these are important analyses, thinking about those kinds of things are really important when you're conducting the analysis and writing up the findings.
Russell Cole	Well said, yeah, I think the general takeaway here is this is just an opportunity for us all to do more with our data and try to tell more of a story with our data. But with that key, caution that this is exploratory -- that there's a lot of limitations around the things that we're trying to do here. This is really a supplement to our main analysis to try to better understand some stuff.
	But hopefully, this is helpful. Again, please let us know if you have any follow-up questions; and we'll loop back with these slides and with the recording and other materials in the future.
	Thanks, all. Hope everyone is doing well.