

Tip Sheet: Evaluation Strategies for Virtual Implementation in Response to COVID

This tip sheet provides guidance on conducting virtual implementation and data collection in response to COVID-19 closures.



TIP SHEET: EVALUATION STRATEGIES FOR VIRTUAL IMPLEMENTATION IN RESPONSE TO COVID

As schools and community organizations across the country have closed in response to COVID-19, Teen Pregnancy Prevention Program (TPP) grantees have had to quickly assess the feasibility of virtual implementation and start planning for it. Transitioning quickly to online delivery can be challenging, but it's also an opportunity to test new approaches to providing content through a new platform. This tip sheet has some points to consider in transitioning implementation and evaluation online, including assessing need and demand, demonstrating feasibility, conducting continuous quality improvement, documenting lessons learned, and undertaking data collection. Also included are spotlights from TPP19 grantees who are in the process of or have already transitioned their programs online.

Assessing need and demand

The first step in going virtual is to determine whether it's feasible. This means reaching out to (1) your curriculum developer, as applicable, to understand whether it's possible to deliver the program online; (2) your partners, to assess their needs and desire to continue programming during the disruption; and (3) your participants, to determine their ability to join virtual sessions (both in terms of technology access and general stability of circumstances). Once you've made that initial assessment of feasibility, consider the following tips to help shape the design of your online program:

- Survey your participants about their constraints. Consider using this [tool](#) to document information they give you about constraints and any possible accommodations that could be made.¹ If participants are lacking Internet access, try to reach out via phone or mail. Though you may not reach everyone, try to collect responses from a good sample of your group. Be sure any data collection activities comply with your Institutional Review Board's guidelines. Information to collect includes (but may not be limited to):
 - What devices will they use to access the program?
 - What is their comfort level with the platform you're considering?
 - Do they have reliable Internet access or cell phone service?
 - What's their learning environment like? This is important given that young people are especially likely to need privacy to talk about these subjects. Will other family members be present or nearby? Are they able to find a private space to engage in sensitive conversations?
 - What is their ability to commit to the schedule? Will other family members be accessing the computer or phone during the day?
 - Other considerations might include access to supplies they need for activities.

¹ This design tool, along with additional resources, can be found here: <https://uxdesign.cc/how-to-leverage-design-to-reimagine-your-class-online-f9b53de28c6a>

- Take steps to ensure your partner is comfortable with your delivery approach. Depending on the circumstances, this could involve co-designing an approach or simply taking the time to share your plans with them. For an example of how San Diego Youth Services involved partners in its planning, see Spotlight 1.
 - If you've been partnering with schools and they can't accommodate your program via their distance learning, explore other ways your partner may be able to help. For instance, they may be able support your program by promoting it to their students and encouraging youth to attend outside of the school's distance learning hours.

Grantee Spotlight 1: A collaborative approach to going virtual

San Diego Youth Services (SDYS), a TPP19 grantee in southern California, partners with schools to deliver teen pregnancy programming. The team is moving to an online delivery format and will coordinate with their school partners to implement it on the same virtual platform used by each school. One of SDYS's challenges is ensuring its partners are comfortable with the implementation plans, so the team is planning **five-minute Zoom webinars** with individual partners to ensure school staff are comfortable with the logistics and to coordinate on planning. Check out the SDYS video [here](#).

Some teams will assess feasibility and determine that continuing with the program is not feasible. In that case, many teams have invested in professional development opportunities to keep their staff engaged.

Others are planning for Year 2 of the grant, laying the groundwork for recruitment or reviewing their data as part of a continuous quality improvement plan. Still others are investing time in building their social media presence online to expand messaging in the community about teen pregnancy and program offerings. Also important is staying connected with youth participants to help them feel supported. For an example of how Mission West Virginia is building out its supplemental components to continue serving the community while its core program is on pause, see Spotlight 2.

Grantee Spotlight 2: Being creative with supplemental content

As some grantees work to move their content online, others have assessed demand and logistics and decided that virtual implementation is not possible at this moment. **Mission West Virginia**, a TPP19 grantee in rural West Virginia, arrived at this decision because teenagers do not have equal access to laptops and Internet in its delivery area. In the meantime, however, the team is focusing its energy on beefing up its supplemental component, including its educational component for parents and community members. MWV will develop an **e-learning site with archived webinars on topics of interest for parents**—for example, recordings made by its health educators and information on recognizing signs of drug use in children. Although MWV recognizes that not all parents will be able to access the site, the team sees its flexible supplemental programming as an opportunity to do what it can to continue serving the community. MWV is also developing an e-learning site for its staff, with resources for skill building and professional development.

Planning a delivery strategy

Planning a delivery strategy includes making decisions about the mode of delivery and technology platforms, developing plans for training facilitators on the new delivery approach, and determining how you will engage participants via the online version of your program.

Delivery mode and technology platform

- Programs may choose a synchronous or asynchronous approach to online delivery. Synchronous approaches involve live communication, such as videoconferencing, chat forums, or text-based communication. With asynchronous learning, participants won't necessarily be learning at the same time, but they can access materials when they are able to. This type of learning features produced videos, worksheets, discussion boards, or emails. While many teams are using online platforms to share materials, where Internet access is an issue, some teams have delivered hard copies of packets to young people. Grantees may also use a combination of approaches.
- The approach you choose is likely to depend on your partners' preferences for their ongoing learning approaches. For an example of how PATH decided on an approach, see Spotlight 3.
- Similarly, if your partners are already conducting online education or programming, it is ideal to use the same technology, so participants don't have to toggle back and forth between multiple platforms. If you need to select a delivery platform or tools for enhancement, see this [crowdsourced list](#) compiled by OPA or [this list](#) from UNESCO.

Grantee Spotlight 3: Selecting a delivery approach that meets partner needs

Though many grantee teams are considering delivering content in real time via synchronous learning like videoconferencing, another option—if it's possible to maintain fidelity—is to deliver prepackaged programming content, such as videos and worksheets (asynchronous learning). **PATH**, a TPP19 grantee in northwest Indiana, has chosen **an asynchronous approach** that works well for its school partners, because teachers can include online videos and worksheets related to the program as part of students' weekly assignments. To measure engagement, the team will examine key metrics, including the number of students who view the videos, and will ask students to self-report on whether they completed lessons.

Preparing staff for online delivery

- Make sure your facilitators are comfortable with the technology before going live. Ideally, you'll have time to pilot virtual delivery with a small group of young people. For an example of how LEADS piloted its virtual approach, see Spotlight 4.
- Facilitators could take time to adapt to the new online format, and technology glitches are inevitable. Be prepared to address concerns and issues as they happen. For example, Keiki O Ka Aina, a TPP19 grantee, checks in with its staff three times a day for five-minute sessions that focus on topics to troubleshoot problems or tools to enhance delivery.
- Another option to prepare facilitators are teach-backs, a strategy in which the person delivering information asks participants in the training or workshop to explain back the concept or skill that was just taught. This is a strategy that could be used with young people in your sessions but can also be used in trainings with facilitators on the new technology to ensure they understand processes and expectations and to correct anything that's not clear. [This blog](#) presents a succinct overview of teach-backs.

Grantee Spotlight 4: A gradual approach to going virtual

LEADS, a TPP19 grantee in St. Petersburg, Florida, successfully launched virtual programming, but not without first implementing a pilot to work out the kinks. LEADS recruited two young people who had previously participated in an in-person session and **piloted virtual implementation**. The aim was not just to get a young person's perspective, but to hear feedback from youth who had experienced the program in person and could talk about how similar the online experience was. When LEADS was ready to launch, the team started with a group of six youth, then increased it to seven. Going forward, LEADS has decided to maintain a group size of 10 or fewer to allow opportunities for engaging discussions in the new online format.

- If you need help learning about some of the common tools being used, see the linked resources below:
 - Google developed [a site](#) with training modules and resources to use its tools for distance learning. This includes a one-pager on Google Meet, a platform commonly used by schools.
 - Many programs and schools are using Zoom. [This site](#) provides tips and tricks for Zoom. Given the recent privacy concerns about Zoom, the [company’s blog](#) discusses how to protect your meeting. This [Google doc](#) also has steps you can take to keep your meeting safe.
 - If you’re using [Microsoft Teams](#), check out this training site that includes a 5-day guide for shifting to distance learning.

Helping participants to adjust online

Young people could adjust instantaneously to online delivery, but that’s not a given. To make the transition smooth, consider the following tips:

- Be upfront and clear about how much time and participation you expect.
- Don’t assume that because young people tend to be technologically savvy, they won’t need support on the new platform. One best practice from the field of e-mentoring is to conduct a training or walk-through of the platform with participants before launching it.
- Use virtual tools to create an engaging virtual experience. Adjusting to group discussion online can take time. You can help participants feel comfortable engaging by incorporating easy-to-learn tools, such as Kahoot (quizzes), Animoto (animated videos), Padlet (post board), and Mentimeter (polling).
- Think about the things that made your program a success when you were delivering it in person, and do your best to bring them into your online version. With so little time to plan for online delivery, programs might decide to focus on transitioning just the core curriculum. But if key enhancements or supplemental components were critical to the success of your program, you’ll enhance your chances of success by bringing them online. To read about Keiki O Ka Aina’s success in maintaining its parent component, see Spotlight 5.

Grantee Spotlight 5: Maintaining an essential component online
Keiki O Ka Aina (KOKA), a TPP19 grantee in Hawaii, designed its in-person program to kick off with a two-hour parent orientation to describe the content of the program, connect with parents on a personal level about how to have conversations with their children, and respond to any questions or concerns. In moving to a virtual platform, the team strongly believed that **the parent component was essential to its success**, and adapted it to a 30-minute Zoom version, in which the aim is still to prepare parents for their role as trusted adults in supporting participants and also to cover some of the logistics of online delivery. As a result, instead of having concerns about how parents will react to the TPP programming once it’s brought home, KOKA has found that parents have been supportive partners in encouraging their children to attend sessions each day. KOKA’s experience speaks to the importance of maintaining critical enhancements or programmatic supplements in online formats, even if it takes extra planning.

- For thoughts from educators about maintaining engagement and building community online, see the following resources:
 - [Maintaining relationships](#)
 - [Keeping students engaged](#)
 - [Building community](#)
 - [Using culturally responsive practices](#)

- [General tips for online success](#)
- [Resources and tips](#)

Continuous quality improvement (CQI) and lessons learned

Programs have not had much time to plan for online programming, so they will probably have to go through an iterative process to develop a strong approach. To adopt a continuous improvement approach to online delivery, consider the following:

- Document what you're doing, and assess your online strategy. As early as your first lesson, it's a good idea to reflect on what went well and what could be improved on. You may want someone to observe the first few sessions to provide an independent perspective on how things are going. Some specific questions you could ask are:
 - ***For live delivery***
 - Are participants comfortable with the technology? Are they comfortable participating? Are online conversations moving along well, or are people talking over each other?
 - Was the facilitator set up with technology in advance so the session could start on time?
 - Were there any technological glitches? Does the facilitator make good use of platform features, like muting participants, allowing them to raise their hands, or incorporating chat features or polling?
 - ***Documents and videos***
 - How easy is it to access the documents and videos? Have staff who didn't set up the site explore and give feedback. If they have trouble finding course materials, participants might have trouble also.
 - ***General***
 - Does the facilitator give adequate instructions for how to use the learning platform?
 - For a template for conducting observations focused on improvement (not fidelity), see [this resource](#).
- Put a mechanism in place to get rapid feedback from your participants. It's valuable to get feedback in real time as participants are experiencing your lessons. If you're using live delivery, you could opt to set aside five minutes at the end of each lesson to ask how things went or use built-in polling features with software like Zoom. If you're not delivering live or would prefer to get feedback anonymously, you could send a short feedback survey at the end of each lesson. Try to ask about various aspects of the online experience. Here are a few suggestions from [Educause](#):
 - Technological issues: discomfort with the technology; unclear expectations about which applications to use; access to reliable Internet; access to reliable digital devices.
 - Learning issues: finding time to participate; concerns about whether lessons are translating well to remote environment; difficulty focusing or paying attention.
- While gathering feedback after every session is valuable, if it's not feasible, you could also incorporate questions about engagement and satisfaction online onto an existing participant or partner satisfaction survey, in order to gather feedback at the conclusion of each program cycle.

Data collection

Figuring out how to implement your program virtually is often the first step to transitioning to online delivery, but it's also important to consider how you will move and adapt your evaluation activities to an online platform. Below we share tips and methods for data collection to assess your virtual program. However, it's important to keep in mind that, as this is a time of transition, you don't need to collect everything at once. For instance, understanding whether youth are connecting to the virtual program – both in terms of technology and engagement - is critical; collecting pre-post information to understand changes in attitudes and knowledge is useful but could be phased in as you determine a plan for collecting higher priority data first.

Online survey

If you weren't already administering surveys online, you will need to choose an online survey program from the many that exist. Beyond taking this first step, you should also think through how online administration can be different from other forms of survey administration. A few options:

- If you're delivering lessons via live video, you could administer the survey verbally during the session with participants using an online survey form. This could be a good approach if you think you need to explain the questions to participants.
- You could send a link to the survey form in an email, and have participants complete it on their own.
 - Many of the online survey programs, such as SurveyMonkey® or Qualtrics, allow you to track who completed your survey if you take this approach. SurveyMonkey, for example, allows you to track the number of people who opened your email, read through the survey, and completed it. Even if you choose to track survey completion rates, you can still ensure the survey data are anonymous because tracking is tied to the emails and not the data responses.
 - With this approach, you will need to plan a strong process for following up with participants to ensure you get good response rates. Strategies include regular tracking to see who completed the survey, following up with nonresponders and emphasizing the importance and anonymity of the survey, and, if feasible, offering incentives for completion.
 - If your data collection was intended to be in the classroom, you may not have participant contact information. In that case, you may need to request surveys be sent out via the school or classroom teacher.
- You could also send a survey with a text messaging with SMS software. This is a potential approach for shorter surveys (up to 12 questions).

If you're administering pre- and post-surveys, you may have developed a solid system for linking surveys pre-to-post. But now that you're moving online, what are some options you have for linking?

- Many survey platforms allow you to link surveys when you send the links via email.
- Another option is to create a unique code for each participant and keep a separate list of the codes in an Excel sheet. You will need to send the unique code to participants when you share the link to the survey, and should set the survey up so they use the unique code to get into it.

Observations and attendance

Data collection may look much different for virtual programming when compared to in-person. For instance, assessing things like participant engagement, quality of discussion, and facilitator poise during observations may involve additional or different data points than what was used for in-person observations. Measuring fidelity when programming is delivered asynchronously will require rethinking what implementing with fidelity looks like in the new format.

- Programs should revisit their observation and data collection tools to reflect the virtual environment.
- Consider using website analytics to supplement assessments of live delivery or as proxy engagement measures for asynchronous programming. Although analytics available for tracking can vary by platform, a few to consider include:
 - **Live delivery**
 - Attendance: You should track attendance just as you would for in-person delivery. The benefit of online delivery is that the attendance rosters keep a record for you!
 - Retention: How many participants stayed for the entire session? It's easier for participants to disengage from and leave an online session, so tracking retention during a session is important.
 - Interaction: You can observe the quality of discussion, but you can also supplement this assessment by noting the number of participants who responded in the chat feature or polls during the session.
 - Downloads: If you offered documents or worksheets to support the lesson, programs will often allow you to track the number of downloads.
 - **Asynchronous programming**
 - Prerecorded videos:
 - Video views
 - Number of unique viewers
 - Number of impressions (or the number of times the page with the video was loaded)
 - Percentage of video watched
 - Worksheets/packets:
 - Number of downloads
 - Create your own mechanism (for example, PATH (spotlighted in Box 3) had students respond to facilitators about whether they completed a lesson, thereby creating a tracking mechanism.
 - Discussion board
 - Number of participant posts
 - Quality of participant posts (assessed by evaluator)
 - Number of peer-to-peer responses
 - For guides on analytics and reporting, see report guides for [Google Meet](#) or [Zoom](#).

Fidelity

You should work with your curriculum developer, if applicable, to ensure your plans for virtual delivery allow you to maintain fidelity to the evidence-based program. Many developers have put out resources and guidelines for virtual implementation. Some additional considerations:

- You may need to make changes to your fidelity logs to include new fields. For example, was the session delivered in person or virtually? Was content delivered synchronously or asynchronously? It's also important to note adaptations that you made to program delivery given the new online platform (e.g., rather than small group work you did the activity as a large group.)
- Be sure to document any other factors that may have been different; for example, if a cohort of youth started with one facilitator in person but finished online with a different facilitator.
- To the degree that you have flexibility to adapt online, a strong approach is to consider the core objective and purpose of the activity being translated and think about how to fit it to a virtual environment. For example, an activity may call for youth to use beads, but now you can't provide beads to participants. Think about what the bead activity is attempting to convey and consider how to share that message without using beads.
- For a few examples of supports offered by curriculum developers, see the [Dibble Institute's toolkit](#) on online teaching or [ETR's](#) series of resources on transitioning to virtual program delivery.

Focus groups

Many grantee teams planned focus groups to get feedback from participants, parents, facilitators, partners, and community members at the end of Year 1. Focus groups provide valuable feedback for Year 2 programming. It is possible to conduct successful focus groups virtually using video conferencing. To plan such a group, consider the following tips:

- Make them smaller (six to eight participants) than an in-person group would be.
- You are likely to discuss sensitive topics, so be sure participants have a private place to participate.
- Ask participants to enter their first names only when they enter virtual sessions to maintain anonymity.
- If you need to collect consent forms, think about the most efficient method and check with your IRB about your plans (for example, email and scan, verbal consent via the phone, or digital signature with a program like DocuSign).
- For in-person groups, common roles are facilitator and note-taker; another useful role for online focus groups is technical lead, or someone who can help troubleshoot technical issues and run polls or other applications, if you intend to use them. You may also want to have a second facilitator to monitor and engage with participants via the platform's chat box.
- Tips for keeping participants engaged during virtual focus groups include using slides to communicate expectations and pose questions; asking round-robin questions, deploying online polls, and using the chat feature.

- Use emailed gift cards as incentives.
- Many programs will produce chat transcripts, which can supplement or serve as an alternative to audio recording.

For an example of how Project Vida's evaluation team conducted a focus group online, see breakout box 6. For more on virtual focus groups, see [this guide](#) from MDRC.

Grantee Spotlight 6: Moving focus groups to a virtual setting

Evaluators for Project Vida, a TPP15 and TPP19 grantee in El Paso, Texas, typically conduct focus groups with key project stakeholders at the conclusion of each year to inform programming. This year, Helix Solutions, the evaluation team working with Project Vida, moved its focus group with six TPP15 program facilitators onto Zoom. They typically use dot voting for in-person groups to gather information anonymously and found that the activity could be easily recreated online using Zoom's polling feature. Key takeaways included: 1.) **Encourage participants to use video rather than calling in**, as it allows observation of body language and fosters greater connection, 2.) **Do a trial run** with a few volunteers to test how any features (e.g., polling) work on different types of devices, 3.) Ensure privacy by **requiring a password to enter**. The evaluators are planning a similar group with TPP19 facilitators and is considering a student group, though noted a session with students would have added complications due to the sensitive nature of the conversation and questions of whether students have a private place to participate at home.

Other considerations

- Institutional Review Boards: Notify your IRB if you need to make revisions to research protocols and instruments. Revisions could include adding questions to your survey or adding a new data collection activity.
- Privacy: To keep privacy concerns to a minimum, especially with live videoconferencing, consider the following suggestions:
 - If you're thinking about recording sessions:
 - Consider the purpose for recording and whether there are ways to achieve this through other methods.
 - If you plan to record to share a session with your evaluation team, could the evaluator join the live sessions instead?
 - Are you recording sessions in order to share them with participants who were not able to attend? You can avoid privacy concerns by recording only the parts where the facilitator is speaking. Or, a tip from [this article](#) is to refrain from recording live sessions, but have the facilitator go into Zoom outside of the session and record themselves delivering the lesson.
 - If you decide to record a live session where youth participate, make sure you are following the privacy rules required by any school partners and your IRB.
 - For more guidance, [this site](#) contains tips related to virtual youth programming, including information about recordings and screen captures.
 - Be transparent with participants during live sessions:
 - Share upfront whether the discussion for the day will include any sensitive discussion topics.
 - Make sure participants are aware of who else is in on the conference with them.
 - When considering which software to use for program delivery or data collection, make sure the company has a privacy policy and review it. [This post](#) has a list of questions to ask when evaluating an app or website. [This site](#) reviews and rates the privacy practices of popular applications.

Suggested citation

This report is in the public domain. Permission to reproduce is not necessary. Suggested citation:
Buonaspina, A. (2020). Tip Sheet: Evaluation Strategies for Virtual Implementation in Response to COVID, Washington, DC: Office of Population Affairs, Office of the Assistant Secretary for Health, U.S. Department of Health and Human Services.

Tier 1B Grant Implementation Study Planning

This brief guides grantees through the initial steps of designing an implementation study, including selecting and prioritizing research questions, mapping data sources, and developing study timelines.

EVALUATION TECHNICAL ASSISTANCE BRIEF

for OAH Teenage Pregnancy Prevention Grantees

September 2017

Tier 1B Grant Implementation Study Planning

The Teen Pregnancy Prevention (TPP) Tier 1B grant program of the Office of Adolescent Health (OAH) funds grant projects focused on scaling up evidence-based teen pregnancy prevention programs in multiple settings using inclusive and trauma-informed practices. OAH Tier 1B grantees are encouraged to use a holistic approach for their implementation and scale up of TPP programming, to address goals across a variety of key activities. This holistic approach includes the following grant activities that are implemented concurrently: (1) community mobilization; (2) engagement of youth and families; (3) evidence-based program (EBP) implementation in multiple settings to scale; (4) project delivery in safe and supportive environments; (5) project implementation using trauma-informed approaches, inclusive of lesbian, gay, bisexual, and transgender (LGBT) youth, and uses a positive youth development approach; and (6) establishment and maintenance of linkages and referrals to youth-friendly health care services. Tier 1B grantees are required to conduct an implementation study to evaluate the planning, implementation, and lessons learned for the grant activities listed above, but may also choose to evaluate additional components of the grant project, such as the selected dissemination strategy or sustainability efforts.

The implementation studies will identify strategies that support quality implementation of the Tier 1B grant project that can be used to improve current and future programmatic efforts. For example, these studies can improve future program delivery by identifying key strategies for applying positive youth development practices, or providing information on how best to engage stakeholders and community partners.

This brief will guide grantees through the initial steps for designing an implementation study that meets OAH requirements and will provide useful formative information to implement grant programs. It first discusses how to identify, prioritize, and select meaningful research questions for an implementation study. Next, it describes how to align research questions with data sources. Finally, it discusses how to develop a timeline to guide implementation study activities.

Identifying and prioritizing research questions

Research questions for implementation studies have two essential roles at study startup: (1) they guide study planning and (2) they structure and organize study activities. In the context of the Tier 1B grant project, research questions should set the stage for understanding the implementation experiences for the full array of activities conducted as part of the Tier 1B grant project.

To begin identifying research questions, the first step is to define the implementation study's goals. Clearly stated goals will define the focus and scope of the implementation study. Possible goals include: (1) documenting the grant project design and how grantees planned to address the expected activities of the holistic approach; (2) describing the extent to which each element of the grant project was successfully implemented; (3) assessing fidelity and quality of grant project implementation; and (4) describing challenges, successes, and lessons learned through grant project implementation. It is also important that study goals—and the resulting research questions—are useful and appropriate for the target audience, such as program staff or community stakeholders. Grantees can use several strategies to align the study's goals with agency and stakeholder interests.

An **implementation study** is a scientifically valid analysis of how providers implement programs and how youth receive them. It can assess successes, challenges, and lessons learned from implementation, which can be used to identify facilitators of and barriers to program delivery. For example, an implementation study can measure the degree to which staff established linkages with youth friendly services, as well as the extent that a grantee built partner capacity to scale-up EBPs in the community. Implementation studies typically focus on the process of implementation and lessons learned rather than outcomes such as behavior change (Werner 2004).

Two approaches, which can be used independently or in tandem, are backward mapping and logic model review.

- **Backward mapping.** Backward mapping is a technique used to identify the most valuable research questions for an evaluation. It begins with key stakeholders discussing what they hope to learn from the implementation study and how that

information will be used. For example, program staff might want to know whether using community and youth advisory groups was beneficial in the planning process. On the other hand, community members might want to know whether youth were receptive to programming. Using stakeholder input as a guide, evaluators can design study goals and research questions that are relevant for each intended audience and ensure the implementation study findings will address the primary areas of interest. Backward mapping also aligns well with the grant project’s expectation for community mobilization by engaging adult and youth advisory teams to learn what they would like to learn from the implementation study and how they plan to use study findings (Rossi et al. 2004).

- **Logic model review.** Reviewing logic models can help grantees identify areas in the implementation process that would be beneficial to evaluate (see Figure 1). Implementation studies that build off a logic model typically evaluate whether a program’s intended activities resulted in the intended outputs. For example, the study could assess whether facilitators were able to conduct the expected number of workshops (activity) or whether the target number of youth attended EBP programming (output). Notably, given the complexity of scaling up and replicating multiple EBPs, it may be necessary to review multiple logic models, including a model specifying all of the efforts taking place under the Tier 1B grant or models for specific EBPs to identify goals for the evaluation.

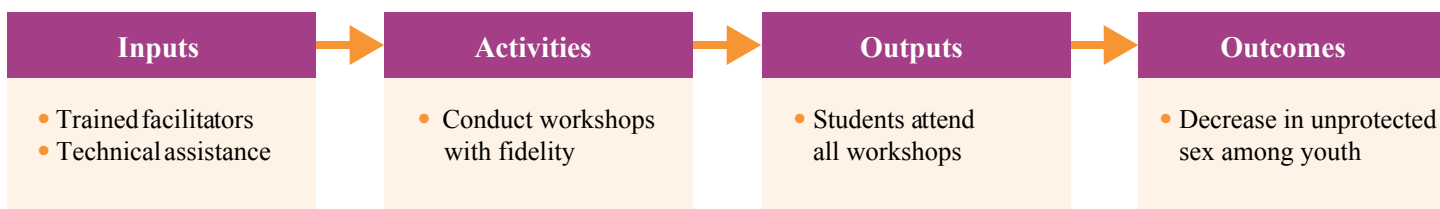
Once defined, study goals serve as a guide for developing research questions. For example, if a goal of the study is to describe the design of the grant project, grantees could ask questions such as, “How was the grant project designed to engage and mobilize local stakeholders?” or “How did the grant project intend to provide a safe and supportive environment for youth in each setting?”

Research questions should be specific and concrete, reasonable for the implementation study to answer using data that can be collected and analyzed, and likely to produce findings to improve program quality or identify lessons learned. For example, a research question that assesses how a grant project incorporated community mobilization could include, “How did the grant project recruit and engage adult and youth leadership teams to plan and coordinate community-wide TPP efforts?” This question focuses on a particular aspect of community mobilization, in this case recruiting and engaging community members. This is also a practical question for the study team to examine with data that are likely readily available, and the findings will be useful to assess the success of community mobilization strategies. On the other hand, a research question such as “Was community mobilization successful?” is not specific nor likely to identify lessons learned and thus should be avoided.

Figure 1. Logic model overview

A **logic model** is an illustration of the relationship between a program’s inputs and activities (the planned work), and its intended results. Logic models can be depicted in multiple ways, but the common components include:

- **Inputs.** Human, financial, organizational, and community resources that support programming.
- **Activities.** Processes, tools, events, technology, and actions that are an intentional part of the program.
- **Outputs.** Result of program activities; may specify types, levels, and targets for program delivery.
- **Outcomes.** Short- and long-term changes in participants’ behavior, knowledge, or attitudes as a result of program participation.



For more information, please see W.K. Kellogg Foundation (2004).

The following are additional examples of appropriate research questions for each grant activity that is part of the holistic approach for a hypothetical “Program Y”.

- 1. Assessing community mobilization efforts.** What was the process for establishing the Youth Leadership Council (YLC) and Community Advisory Group (CAG)?
- 2. Understanding the level of engagement of youth and families.** What strategies did Program Y use to engage youth and families in the planning, implementation, and evaluation of the grant project?
- 3. Implementing evidence-based TPP programming to scale with fidelity and quality in at least three settings.** What were the key successes and challenges that staff and program facilitators experienced with scaling-up EBPs in their community?
- 4. Delivering the grant project in a safe and supportive environment for youth and their families.** What approaches did Program Y use to create a safe and supportive environment for grant project implementation?
- 5. Implementing a grant project that is trauma-informed, LGBT inclusive, and uses positive youth development approaches.** What training was provided to staff on trauma informed approaches? What adaptations to grant project content, delivery, or organizational policies were necessary for Program Y to make the grant project LGBT inclusive? What were the key positive youth development approaches used during project implementation?
- 6. Establishing and maintaining linkages and referrals to youth-friendly health care services.** How was the youth-friendliness of services assessed?

After developing an initial list of research questions, it may be necessary to prioritize questions, if there are too many to address within the scope of the study. Ideally, the questions that are examined in the implementation study will yield useful information for grantee agencies, OAH, and key stakeholders. For example, some research questions could be interesting but might only be of interest to a specific stakeholder. In such situations, grantees will need to assess whether they have enough resources to include this question in the implementation study, especially if other questions provide more information to a broader audience. Grantees will also need to prioritize questions based on the remaining time in their grant period. With the shortened grant period due to the funding truncation there may not be sufficient time left to answer all questions of interest. Other prioritization considerations include deciding which questions will provide the most useful information for program improvement and which are most feasible to answer. The data sources needed to answer a research question also will affect

whether the research question is feasible to answer, which we discuss in detail in the next section.

Aligning research questions and data sources

The next step in the planning process is to identify data sources that are available for the implementation study and to cross-walk the research question(s) with each data source to ensure appropriate and relevant data are being collected. The process of selecting research questions and mapping data sources is iterative and grantees should approach it with flexibility. For example, grantees may need to revise a research question to align it with available data or drop a research question if the data sources are unavailable or too expensive to pursue.

Data sources fall under two broad categories: quantitative and qualitative. The two types of data complement each other and can provide a fuller picture of implementation when used together in a mixed-method approach. Each data source has advantages and disadvantages, and these must be weighed against the evaluation’s goals. (Note: Tier 1B grantees are already collecting some of these data elements for other purposes, such as data collected for performance measures).

Quantitative data sources provide information that can be measured, such as summaries of program dosage, number of implementation settings, or number of partner organizations. Quantitative data can be important to understand the extent that programming is implemented with fidelity and the reach of programming; however, it cannot provide context for the findings or lessons learned from implementation.

Quantitative data sources include, but are not limited to:

- **OAH performance measures.** Tier 1B grantees are required to submit measures of reach, dosage, fidelity, quality, etc. to OAH on an ongoing basis. These data are valuable to understand the extent and usage of grant project services. A benefit of these data is that they are already being collected as part of this grant, and therefore this does not require any additional grant resources; however, grantees will likely need to supplement performance measure data to answer research questions on all the components of the holistic approach.
- **Surveys.** Surveys of program staff, youth participants, partners, advisory group members, and other key stakeholders can provide a systematic assessment of perceptions of grant project implementation, quality, training, and administration. Surveys conducted for implementation research are different in purpose from outcome evaluation surveys, as they focus on respondents’ perceptions of implementation rather than knowledge, attitudinal, or behavioral outcomes. The level of effort for survey administration is largely determined by the number of respondents and survey length. Therefore, if larger

samples are a concern due to resources or logistics, it may be feasible to systematically select a subset of youth or to conduct brief surveys with staff rather than youth.

- **Fidelity logs.** Fidelity logs are an EBP-specific account of whether facilitators implemented core program components as intended by the developer. They provide useful data on EBP content coverage, quality of implementation, adaptations, dosage, and youth engagement. Facilitators should be completing fidelity logs as part of the Tier 1B fidelity monitoring plans. Some limitations of fidelity log data are the potential for self-report bias when facilitators complete the logs and possible inconsistencies over time. To minimize this issue, provide early and ongoing training to facilitators and occasionally check the data quality (for example, have an observer code the same data and compare).
- **Other administrative data.** Other administrative data sources such as school attendance data or data from partners on youth service utilization might be beneficial to collect. These data can be used to provide insight into program attendance and dosage of programming received. However, accessing these data may require a data use agreement or school research board review.

Qualitative data sources provide information on the quality of programming and context of implementation. For example, focus groups with youth provide insights into their responsiveness to the holistic approach, such as assessing healthcare linkages, community mobilization, or efforts to ensure safe and supportive environments during project implementation. Discussions with staff can deepen understanding of staff perceptions of the success of programming. To collect qualitative data, grantees should work with an independent evaluator or staff not involved in direct program activities to ensure objective data collection and analysis. Staff conducting qualitative research should have experience developing protocols, conducting interviews or focus groups, and analyzing qualitative data. If a grantee agency do not have staff with these skills, they might need to contract with experienced staff to do the work.

Qualitative data sources can include:

- **In-depth interviews.** Interviews provide valuable insight on grant project activities such as planning, youth and family engagement, grant project implementation and EBP scale-up, and lessons learned. The information obtained in interviews is more nuanced and often more detailed than what is available through staff surveys. The study team may conduct interviews by telephone or in person in either individual or group settings, and can include program staff, advisory group members, or youth, for example. Group interviews can be beneficial when staff members have similar or complementary roles and can each provide input on interview questions.

- **Focus groups.** A focus group is a facilitator-guided discussion with multiple respondents. Common topics for focus groups are discussions on grant project planning, staff and participant perceptions of particular aspects of the grant project, and lessons learned. Focus groups tend to promote discussion by allowing respondents to react to one another's responses and share personal opinions. However, the group setting can make some respondents less candid due to confidentiality concerns. Most often, focus groups take place in person, but they can be conducted by telephone, videoconferencing, or online approaches.
- **Observations.** Observations can provide an assessment of service delivery including EBP implementation fidelity, the use of programming that is LGBT inclusive and trauma-informed, quality, and youth engagement. OAH requires observation of at least five percent of program sessions during the course of the grant, so these data are readily available provided the protocol covers the elements of interest. It is important to consider the limitations and benefits of observation data, including subjectivity of the observer and the risk that having an observer in the classroom changes the facilitator's and youth's behavior. It is possible to minimize these limitations by observing video-recorded sessions of program delivery.
- **Meeting minutes.** Meeting minutes provide an account of content covered during staff meetings and YLC/CAG meetings such as understanding how the community mobilization and youth leadership processes were planned and implemented. Reviewing minutes can help grantees understand the process of program planning and the reaction of staff and advisory group members to the grant project once it is underway. The consistency of minute-taking and the level of detail contained in the notes will impact how useful they are for the implementation study.
- **Grant project documents.** Grant project documents may include organizational policy documents, curriculum materials, facilitator manuals, program protocols, fliers, and so on. Organizational policy documents can highlight the changes that occurred within organizations involved in the grant project to integrate positive youth development and trauma informed approaches, promote LGBT inclusiveness, or create safe and supportive environments. Curriculum materials and facilitator manuals outline how the developer intended the program to be delivered, which is essential to assess fidelity. Outreach materials such as fliers can provide context about how youth were recruited into programming. These materials will not paint a complete picture of actual implementation or outreach but will shed light on what optimal implementation with fidelity might look like.

One strategy to select appropriate data sources for each research question is to create a table. First, grantees can list each research question, ideally organized by grant activity, as a separate row in the table (see Table 1 as an example). Then, use a list of available data sources as columns in the table. Finally, grantees can look across each research question and data source combination and identify whether the data source(s) could help answer each question. If no data source is available to address a research question, it might be necessary to revise the question or drop it from the study.

Most research questions will rely on several data sources. For example, to answer the research question, “What adaptations to grant project content, delivery, or organizational policies did Program Y need to make to make the grant project inclusive?” grantees could use data from the OAH inclusivity review, staff focus groups, classroom observations, grant project documents, and fidelity log data. Using multiple data sources is typical of implementation studies and enables a broad assessment of the research questions being evaluated. Grantees need to identify the best use their resources when deciding which data sources to use in their implementation study.

Table 1. Tier 1B grant activities, example research questions, and data source mapping

Tier 1B grant activities and potential implementation study research questions	Potential data sources								
	Quantitative				Qualitative				
	OAH performance measures	Surveys	Fidelity logs	Other administrative data	In-depth interviews	Focus groups	Observations	Meeting minutes	Grant project documents
Assessing community mobilization efforts									
What strategies were most effective in gaining community participation?					✓	✓		✓	✓
What was the process for establishing the Youth Leadership Council (YLC) and Community Advisory Group (CAG)?					✓	✓		✓	
How did the grant project use the YLC and CAG in planning for scale-up in target communities?					✓	✓		✓	
How did the grant program use community members to plan and coordinate community-wide TPP efforts?					✓	✓		✓	✓
What was the extent that the grant project used Advocate for Youth’s “Strategies Guided by Best Practice for Community Mobilization” to guide community mobilization activities?		✓			✓	✓		✓	
Understanding the level of engagement of youth and families									
What strategies did the grant project use to engage youth and families in the planning, implementation, and evaluation of the grant project?					✓	✓			✓
To what extent were youth and families actively engaged in implementing the grant project?		✓	✓		✓	✓	✓	✓	
What were the successes and challenges of engaging youth and families in grant project implementation?		✓			✓	✓			✓

Tier 1B grant activities and potential implementation study research questions	Potential data sources								
	Quantitative				Qualitative				
	OAH performance measures	Surveys	Fidelity logs	Other administrative data	In-depth interviews	Focus groups	Observations	Meeting minutes	Grant project documents
Implementing evidence-based TPP programming to scale with fidelity and quality in at least three settings									
Implementation of individual EBPs									
To what extent were selected EBPs a good fit for the target community?	✓	✓	✓		✓	✓		✓	
What were the key successes and challenges staff and program facilitators experienced with EBP delivery?		✓	✓		✓	✓	✓	✓	✓
To what extent were youth engaged in EBP delivery?	✓	✓	✓	✓	✓	✓	✓		
Scale-up of EBPs									
What were the key successes and challenges the grant project experienced with scale-up of EBPs in the target community?		✓			✓	✓		✓	✓
What were the key strategies the grant project used to assess EBP fit across the community?					✓	✓		✓	✓
To what extent did the grant project build partner capacity to deliver EBPs with quality and fidelity?					✓	✓		✓	✓
Delivering the grant project in a safe and supportive environment for youth and their families									
What approaches did the grant project use to create a safe and supportive environment for grant project implementation?		✓			✓	✓	✓	✓	✓
What positive youth development approaches were integrated into the grant project?		✓	✓		✓	✓	✓	✓	
To what extent did grant project staff and partners implement the grant project using positive youth development approaches?		✓	✓		✓	✓	✓		
Implementing the grant project that is trauma-informed and LGBT inclusive									
What adaptations to grant project's content, delivery, or organizational policies were necessary to make the grant project inclusive?		✓	✓		✓	✓	✓	✓	✓
How were trauma-informed approaches integrated into grant project activities?		✓	✓		✓	✓	✓	✓	✓
To what extent did the community perceive grant project activities to be trauma-informed and LGBT inclusive?		✓			✓	✓	✓	✓	✓

Tier 1B grant activities and potential implementation study research questions	Potential data sources								
	Quantitative				Qualitative				
	OAH performance measures	Surveys	Fidelity logs	Other administrative data	In-depth interviews	Focus groups	Observations	Meeting minutes	Grant project documents
Establishing and maintaining linkages and referrals to youth-friendly health care services									
How were partnerships with youth-friendly health care services established and maintained?		✓			✓	✓		✓	
How were providers assessed to determine whether they were youth friendly?					✓	✓		✓	✓
What partnerships were essential to ensuring linkages to a wide range of youth friendly-services, not just reproductive health services?	✓	✓			✓	✓			

Developing a timeline to guide implementation study activity planning

After identifying research questions and data sources, grantees should develop a timeline for implementation study activities that outlines the timing of data collection, analysis, and reporting. Planning these activities from the outset will help the study stay on track and within budget.

The timing of data collection depends on the types of data included in the study and the ways they will be used. For some data sources, it makes sense to collect data on an ongoing basis, whereas others are best collected at specific points. Ongoing data collection might be useful for quality improvement; grant project monitoring (for example, fidelity, adherence, and participant engagement); and technical assistance. For example, grantees are required to continuously collect performance measure data on the reach and quality of grant project delivery, which might be use to improve implementation in real time. Point-in-time data collection occurs at specific intervals, typically aligned with milestones of grant project implementation. For example, it might make sense to plan interviews with staff for both before and after grant project delivery to collect data on planning and training efforts, community mobilization, key partnerships, early implementation challenges, and the strategies

used to address those challenges. Similarly, youth focus groups might take place after youth have had a chance to experience programming but not too long after implementation has ended. Both ongoing and point-in-time data collection give program staff and evaluators a well-rounded assessment of grant project implementation activities and outcomes.

Grantees might be able to conduct interim analyses throughout the grant period; however, in the early phases of planning they are encouraged to consider the time required for data analysis and reporting. This will help them allot adequate time for these activities, as data collection will need to end before analysis begins. In general, grantees should allow nine months for implementation data analysis and reporting at the end of the grant period.

Using these guidelines, grantees can begin to plan the timeline for data collection while considering the level of rigor desired and resource limitations. There are several ways to conduct an implementation study, and grantees will need to balance evaluating the required components of programming with available resources, grantee agency needs and those of the community, and the benefits and constraints of various data collection approaches. Timelines should also factor into the grant’s analysis and reporting requirements.

Conclusion

Implementation studies conducted by Tier 1B grantees will provide valuable information on the successes and challenges of implementing all components of the Tier 1B grant project's holistic approach; however, there is no one-size-fits-all approach to implementation study planning. Beginning the planning process by establishing implementation study goals, selecting key research questions, and mapping data sources will guide evaluation activities from the outset and assist in building the framework for later data collection, analysis, and reporting activities. Using this approach, grantees can plan implementation studies that are comprehensive and useful for their agency, stakeholders, and OAH.

References

- Rossi, P., M. Lipsey, and E. Howard. *Evaluation: A Systematic Approach, 7th Edition*. Thousand Oaks, CA: Sage, 2004.
- Werner, A. *A Guide to Implementation Research*. Washington, DC: The Urban Institute Press, Rowman & Littlefield, 2004.
- W.K. Kellogg Foundation. *Logic Model Development Guide*. Battle Creek, MI: W.K. Kellogg Foundation, 2004.

Example Active Parent Letter and Consent Form

This document is an example of an active parent consent letter and consent form, illustrating the types of information to include in these documents.

Dear Parent or Guardian:

Hello! The [SPONSOR NAME] is conducting an important study of the effectiveness of ways to reduce teen pregnancy and STD infection. The [STUDY NAME] evaluation will provide communities like yours with sound scientific findings on program effectiveness. [SCHOOL] is taking part in this study and students, including your child, are invited to participate. We are requesting your permission for your child to participate in the study.

The lead evaluator for this research, [EVALUATOR] will survey youth on attitudes, beliefs, and activities three times over the next few years, beginning with a first survey in [MONTH]. These surveys will ask about families, friends, communities, and schools, and about attitudes, knowledge and activities in school and with their peers, including sexual activity, drug use and alcohol use. Your child will receive a [INCENTIVE] as a thank you for their time once they have completed each survey. The study team will also gather information from participants' school records, including grades, attendance, and test scores. The researchers may also invite your child to participate in a focus group to discuss his/her experiences in the program.

All information collected through the surveys and from school records will be kept private to the extent allowed by law. However, if your child reports that they may hurt themselves or someone else, or that someone is hurting them, the study team may be required to report this to someone who can ensure your child is safe. If you choose to let your child participate, the information from your son/daughter will be combined with information from other youth to determine the effectiveness of pregnancy prevention programs. Names are not kept with the answers to the surveys. Results will be reported in the aggregate. For example, 20% of youth reported they had smoked a cigarette. No personally identifiable information will be shared with anyone outside of the study team. All data collected through the study will be stored in a secure location, accessible only by approved study team members.

The only risk to your child is that he/she may be uncomfortable answering some questions in the surveys. If that happens, your child can refuse to answer those questions.

Participation in the study is voluntary. If you agree that your child can participate, you or your child can still choose later to stop participating, and your child can decide to complete only parts of the surveys. But we hope you agree with us that it is important to learn about effective ways to prevent teen pregnancy through studies like this. It is important that you let us know whether you will allow your child to be in the study. Please complete and sign the attached form, noting whether you consent or do not consent to your child's participation, and return it to your child's teacher **within a week**.

If you have questions about the [STUDY NAME] or about your child's participation, please call [CONTACT NAME], toll-free, at [EVALUATION ORGANIZATION] at [NUMBER] between the hours of 9 a.m. and 5 p.m. eastern time, Monday through Friday.

Sincerely,

[NAME]

Study Director

[STUDY NAME]
 Parent Consent Form
[SCHOOL]/[SCHOOL DISTRICT]
Sponsored by [SPONSOR NAME]

I have read the attached information sheet describing the study. By signing this form, I am:

giving my permission not giving permission

for my son or daughter, _____, to participate in the study.

Print Child's Name

I understand that, as part of the study, information for all study youth will be collected from school records and through surveys. By signing this form, I am giving my permission for my child to participate in the surveys and for information to be gathered from my son/daughter's school records for use in a study that will be conducted by [EVALUATOR]. I further understand that all information on my child will be kept private to the extent permitted by law and used only for the purposes of the study. If I have questions about my child's rights as a research volunteer, I can call [CONTACT NAME] at [IRB], toll-free at 1-800- [NUMBER].

Parent/Guardian Signature: _____ Date: _____

Child's Name: _____ Child's Date of Birth: ____ / ____ / ____

Month Day Year

Please fill in the following information. We will use your contact information only if we need your help in completing a survey with your child. Thank you.

I give permission for the school to provide access to the following administrative data for my son or daughter (please check yes or no for each):

Administrative Data for 2009–10, 2010–11, and 2011–12	Yes	No
School attendance record	<input type="checkbox"/>	<input type="checkbox"/>
Course grades	<input type="checkbox"/>	<input type="checkbox"/>
Standardized test scores	<input type="checkbox"/>	<input type="checkbox"/>
Student directory information (school address and phone)	<input type="checkbox"/>	<input type="checkbox"/>

Parent/Guardian Name: _____

Street Address: _____ Apartment: _____

City: _____ Zip Code: _____

Telephone: (_____) _____ - _____ Home Email: _____

(_____) _____ - _____ Work

(_____) _____ - _____ Cell

Other Contact Name: _____

Street Address: _____ Apartment: _____

City: _____ Zip Code: _____

Telephone: (_____) _____ - _____ Home Email: _____

(_____) _____ - _____ Work

(_____) _____ - _____ Cell

WE NEED YOUR ANSWER, WHETHER IT IS YES OR NO. PLEASE RETURN THIS FORM TO YOUR CHILD'S TEACHER WITHIN A WEEK. THANK YOU!

NOTE: These documents were created by Mathematica and are intended to serve as examples for TPP grantees. Grantees should work directly with their IRBs to ensure they meet the necessary requirements regarding the content of letters, forms, and other materials developed for their study.

Example Combined Parent and Child Form

This document is an example of a joint active parent and child consent form, where information for both parent and child consent are contained in the same form.

STUDY LABEL



[STUDY NAME] PERMISSION FORM
[Program Name] and [Evaluator]

Check here if you would like us to mail you a copy of your signed form

PARENT'S OR GUARDIAN'S PERMISSION TO BE IN THE STUDY

I have read the enclosed Frequently Asked Questions and accompanying study materials. My daughter and I have been given the opportunity to ask any questions we may have. I understand that:

- My daughter's name will be in a lottery to see if she gets picked for the [PROGRAM NAME] program or not. If she is not picked for the [PROGRAM NAME] program she will still be an important part of the study but she will not be able to take part in any [PROGRAM NAME] programs and activities for the next five years. She may quit the study with no penalty.
My daughter will be asked to fill out four questionnaires during the next five years. Each questionnaire will take about 45 minutes to complete. The questionnaires will ask about her friends, family, school, her opinions about sexual activity, and whether she has had sex or experimented with drugs or alcohol. There are no right or wrong answers and she does not have to answer a question if she doesn't want to. Her questionnaire answers will be combined with other girls' answers. Only the study team will be able to see her answers—not her teachers, parents or anyone else. My daughter may also be asked to volunteer for small group discussions lasting about an hour.
I understand that [EVALUATOR] will get my daughter's school records, including school attendance, test scores, participation in the National School Lunch Program, and grades during the study period (even if she changes schools). I understand that [EVALUATOR] will also get my daughter's [PROGRAM NAME] attendance records if she is in [PROGRAM NAME] programs. These records will be used to see if girls make different decisions depending on their background, how they do in school and whether they are in the program.

I understand as a parent or guardian that both my daughter and I must circle "yes" and sign this form for my daughter to participate in the study.

- A. Do you give permission for your daughter to be in the study? YES NO
B. Parent must sign here Parent's (or Guardian's) Signature
C. Parent (or guardian) print name here Parent's (or Guardian's) Printed Name

- 1. What is today's date? MO-DAY-YR
What is your daughter's address?
2. Street
3. City State Zip Code
4. Phone 1
5. Phone 2
6. Daughter's e-mail address

DAUGHTER'S CONSENT TO BE IN THE STUDY

I have read the enclosed Frequently Asked Questions and accompanying study materials. My parent or guardian and I have been given the opportunity to ask any questions we may have. I understand that:

- My name will be in a lottery to see if I get picked for the [PROGRAM NAME] program or not. If I am not picked for the [PROGRAM NAME] program I will still be an important part of the study but I will not be able to take part in any [PROGRAM NAME] programs and activities for the next five years. I may quit the study with no penalty.
I will be asked to fill out four questionnaires during the next five years. Each questionnaire will take about 45 minutes to complete. The questionnaire will ask about my friends, family, school, my opinions about sexual activity, and whether I have had sex or experimented with drugs or alcohol. There are no right or wrong answers and I do not have to answer a question if I don't want to. My questionnaire answers will be combined with other girls' answers. Only the study team will be able to see my answers—not my teachers, parents or anyone else. I may also be asked to volunteer for small group discussions lasting about an hour.
I understand that [EVALUATOR] will get my school records, including school attendance, test scores, participation in the National School Lunch Program, and grades during the study period (even if I change schools). I understand that [EVALUATOR] will also get my [PROGRAM NAME] attendance records if I am in [PROGRAM NAME] programs. These records will be used to see if girls make different decisions depending on their background, how they do in school and whether they are in the program.

I understand that both my parent (or guardian) and I must circle "yes" and sign this form for me to participate in the study.

- A. Do you want to be in the study? YES NO
B. Daughter must sign here Daughter's Signature
C. Daughter print name here Daughter's Printed Name

- 7. Daughter's First/Middle Names
8. Daughter's Last Name
9. Daughter's Nickname (if any)
10. Daughter's Date of Birth MO-DAY-YEAR
11. Do you have more than one daughter who is completing an application for this study? YES NO
If yes: What are the other girls' (sisters') names and dates of birth?
Name Date of Birth MO-DAY-YEAR
Name Date of Birth MO-DAY-YEAR

NOTE: These documents were created by Mathematica Policy Research and are intended to serve as examples for TPP grantees. Grantees should work directly with their IRBs to ensure they meet the necessary requirements regarding the content of letters, forms, and other study materials.

Please Write Girl's Initials Here |__|__|__|__|

What grade is your daughter in? |__|

For questions, contact [CONTACT PERSON]:
[CONTACT PHONE AND EMAIL]



This page should only take a few minutes to complete. Please fill it out with as much information as you can. This information will allow us to send you information for the [STUDY NAME]. Some of the information you give us may help us locate you in the future if you move and we no longer have your current mailing address. We also need information on up to two other relatives or people who would know how to reach you in the future; for example, an aunt, uncle, or grandparent, or an adult friend who would know where your family would be living if you moved. We will also use this information to send you a brief report of the results at the end of the study.

<p>Parent (or Caregiver) Information</p> <p>First Name _____</p> <p>Middle Name _____</p> <p>Last Name _____</p> <p>Relationship to Daughter _____</p> <p>1st Parent e-mail Address _____</p> <p>1st Parent Date of Birth __ __ MO- __ __ DAY- __ __ __ __ YR</p> <p>If 1st Parent is working please provide:</p> <p>Work Phone _____</p> <p>Company Name _____</p> <p><input type="checkbox"/> CHECK HERE IF ADDRESS AND PHONE IS SAME AS DAUGHTER'S (YOU DON'T NEED TO RE-WRITE THIS INFO.)</p> <p>1st Parent (or Caregiver) Address</p> <p>Street _____</p> <p>City _____</p> <p>State _____ Zip Code _____</p> <p>Phone 1 _____</p> <p>Phone 2 _____</p> <p>OPTIONAL INFORMATION:</p> <p>1st Parent Social Security # __ __ __ _ - __ __ _ - __ __ __ _ </p>	<p>Second Parent (or Caregiver) Information</p> <p>First Name _____</p> <p>Middle Name _____</p> <p>Last Name _____</p> <p>Relationship to Daughter _____</p> <p>2nd Parent e-mail Address _____</p> <p>2nd Parent Date of Birth __ __ MO- __ __ DAY- __ __ __ __ YR</p> <p>If 2nd Parent is working please provide:</p> <p>Work Phone _____</p> <p>Company Name _____</p> <p><input type="checkbox"/> CHECK HERE IF ADDRESS AND PHONE IS SAME AS DAUGHTER'S (YOU DON'T NEED TO RE-WRITE THIS INFO.)</p> <p>2nd Parent (or Caregiver) Address</p> <p>Street _____</p> <p>City _____</p> <p>State _____ Zip Code _____</p> <p>Phone 1 _____</p> <p>Phone 2 _____</p> <p>OPTIONAL INFORMATION:</p> <p>2nd Parent Social Security # __ __ __ _ - __ __ _ - __ __ __ _ </p>	<p>Other Contact 1</p> <p>First Name _____</p> <p>Middle Name _____</p> <p>Last Name _____</p> <p>Relationship to Daughter _____</p> <p>Street Address _____</p> <p>City _____</p> <p>State _____ Zip Code _____</p> <p>Phone _____</p> <p>E-mail Address _____</p> <p>Date of Birth __ __ MO- __ __ DAY- __ __ __ __ YR</p> <p>Other Contact 2</p> <p>First Name _____</p> <p>Middle Name _____</p> <p>Last Name _____</p> <p>Relationship to Daughter _____</p> <p>Street Address _____</p> <p>City _____</p> <p>State _____ Zip Code _____</p> <p>Phone _____</p> <p>E-mail Address _____</p> <p>Date of Birth __ __ MO- __ __ DAY- __ __ __ __ YR</p>
--	---	---

THANK YOU FOR YOUR TIME AND INTEREST! Please have your daughter return this form to [CONTACT NAME/PLACE] by [DATE].

NOTE: These documents were created by Mathematica Policy Research and are intended to serve as examples for TPP grantees. Grantees should work directly with their IRBs to ensure they meet the necessary requirements regarding the content of letters, forms, and other study materials.

Example Student Assent Form

This document is an example of a student assent form, illustrating the types of information to include when requesting assent for study participation.

Guidance for Developing Parent Letters and Parental Consent Forms

This document serves as a guideline for developing consent forms and letters. It suggests key items to include in letters to parents and in consent forms.

DEVELOPING PARENT LETTERS AND PARENTAL CONSENT FORMS

This document is intended as a guideline for developing consent forms and letters. Grantees should work directly with their IRBs to ensure they meet the necessary requirements regarding the content of letters, forms, and other materials developed for their study.

Although letters and forms need to include a lot of information, best practice is to develop straightforward, clearly worded materials and to limit the length to one page. Try to avoid using technical terms, such as “random assignment”. It is recommended that forms be written at no higher than an 8th grade reading level.

Suggested items to include in parent letters and consent forms:

- Name of study sponsor
- Purpose of the study, why it is important to parents, and any value it will have for their children, the school they attend, or their community.
- Specific information on what is required of their child – include all activities for the length of the study (number of surveys, when surveys will be administered, focus groups, interviews, collection of administrative records, etc.) so parents can provide fully informed consent.
 - Note: though focus groups should be mentioned in the study consent form as part of potential study activities, you may want to request consent separately for focus groups (at a later time) since this usually involves only a small subset of youth.
 - For the surveys, you may also want to include a brief description of the topics addressed.
- Privacy language – describe how participants’ data will be protected and any limitations to that, if applicable.
- Handling of data – who will have access to participants’ data, and how the data will be stored.
- Mention of incentives for returning consent forms (if applicable) and any other incentives that will be offered for the surveys
- Any additional benefits to study participants
- Any potential risks to study participants
- Note that participation in the study is voluntary, and whether they allow their child to participate will not affect services received, grades, etc.
- Date to return the consent form and to whom/where the form should be returned.

- Request for contact information – for the parent, the child, and others who may know where to find them if they move during the course of the study. Request as much information as possible (cell phones, emails, addresses, etc.) and include a statement that the information will be used only to notify them of study activities

Developing Parent Letters and Parental Consent Forms

- How to obtain more information about the study – contact name, number (toll-free, if possible) and email address; list the study website, if applicable.
- Provide postage paid envelopes if you want parents to mail signed forms back.

Additional materials: You may want to consider including a sheet (one page) of Frequently Asked Questions or a study brochure, if available. You may also want to consider including an additional letter of support from your partner school/clinic/ community organization – whoever the parents are familiar with that would lend some legitimacy to your study (or, ask them to co-sign your letter).

Guidance on Developing MOUs for the Evaluation

This document illustrates key considerations for developing MOUs with sites. MOUs with study sites serve several functions. They establish a record of what the evaluator expects of the sites and of what the evaluator will do. They also describe how the program will be implemented.

Developing a Memorandum of Understanding (MOU) for the Evaluation

MOUs with study sites serve several functions. They establish a record of what the evaluator expects of the sites, and what the evaluator will do. They describe how the program will be implemented. An MOU can identify the possible consequences of either party's failure or inability to fulfill its commitments or objectives. While an MOU is not a binding contract, it does serve to deter study attrition, especially at times of leadership transition (such as a change in school principals). It is important for the study team to remember that this is not a legal document, because treating it as one can raise the sensitivities and concerns of the study sites. In fact, the evaluation team might instead refer to these written understandings as "work plans," instead of MOUs.

To be durable enough to survive turnover in site leadership, MOUs should commit the organization, not just current staff. Termination provisions typically include an option to withdraw from a study, but the formality of the withdrawal process can deter such action when leadership changes occur. The formality of the agreement can also protect against internal expressions of dissatisfaction that might emerge long after the agreement is made.

MOUs are particularly important for evaluations assigning clusters (such as schools, clinics, community-based organizations, or housing projects) to the program or control/comparison condition. For experimental design studies, these MOUs should be signed prior to random assignment. The loss of study sites can be detrimental to the evaluation because, in clustered randomized controlled trials, site loss is assessed as sample attrition, as is student sample loss among sites that remain. For school sites, separate agreements with the individual schools and the school districts are recommended.

An MOU should contain the following information:

- Describe the evaluation, including its scope and duration.
- Define the program and the counterfactual condition, and indicate whether and when the control/comparison group will receive the program.
- Identify what the evaluator and the site will do throughout the consent, random assignment (if applicable), and data collection processes.
- Specify clear benchmarks for continued study participation (if applicable) and consequences if they are not met.
- Explain the terms of any financial compensation.
- Include the confidentiality expectations and the terms for acquiring additional data. These data include attendance, facilitator logs, interviews, and observations.

The TPP Evaluation TA website contains a sample MOU for your reference.

Guidance on Developing MOU for the Evaluation - Sample MOU

This document is an illustrative, example MOU, showing roles and responsibilities of evaluation staff and participating sites.

SAMPLE MEMORANDUM OF UNDERSTANDING (MOU)

Evaluation of [PROGRAM NAME]

Study Agreement for [SITE NAME]

NOTE FROM TPP Eval TA: *This MOU contains very specific details for an agreement between a TPP grantee and evaluation schools. Names and other identifying information have been removed. Study teams should refer to this as a sample template and develop their MOUs to reflect their specific program implementation and evaluation plans.*

Study Description. As part of the [EVALUATION NAME] this study will assess the effectiveness of the [PROGRAM NAME] program, a promising intervention offered by [ORGANIZATION NAME] for reducing teen sexual activity and its consequences. The program will be implemented and evaluated in two “waves” to test the effects of a two-year program and a three-year program. There will thus be two cohort studies, focusing on [PROGRAM NAME] effects on 7th-graders. Cohort 1 youth who receive the program will participate for two years (7th and 8th grades), and Cohort 2 for three years (7th-9th grades).

The combined study will include about 18 school districts. The districts will be randomly assigned to deliver either the [PROGRAM NAME] program or an alternative “control” diet and exercise program. The study teams will conduct a baseline survey of each cohort of 7th-graders in the participating schools, and then two follow-up surveys at the end of 8th and 9th grades, and determine the impact of the program by comparing outcomes from survey data for youth in the program and control schools. They will also carefully document the services received and how they are delivered. The evaluation of [PROGRAM NAME] is part of the research and evaluation efforts at the U.S. Department of Health and Human Services (DHHS) focused on preventing risky sexual behavior and pregnancy among unmarried adolescents.

Main study activities:

The key activities of the study, along with school responsibilities related to each, are described below, and Table 1 presents the expected timing of key activities.

Parental Consent: No data can be collected for the study without explicit written parental consent. In early fall 2011 (and again in fall 2012 for Cohort 2) the study team will coordinate with schools in the [DISTRICT/COUNTY NAME] district to distribute consent forms to the parents of 7th-grade students in the selected schools. These materials will ideally be sent to parents in school packets at the start of the school year. The consent forms ask for permission for the students to complete the surveys. Parents will be asked to return the consent form to the school as soon as possible, indicating whether or not their child may complete the surveys. The study team will work with school personnel to identify the best process for distributing and collecting the signed parent consent forms, including the possible use of incentives to encourage high rates of returned forms.

Random Assignment of Schools: A total of approximately 18 “clusters” of middle schools and the high schools attended by their graduates will be included in the study. It is thus estimated that about 22 middle schools and 18 high schools will participate. (Implementation of the study is contingent on recruitment of the target number of school clusters.) To produce scientific evidence on the effects of [PROGRAM NAME], [EVALUATOR] will randomly assign these clusters (or in some cases the individual middle schools), using a

Sample Memorandum of Understanding (MOU)

process like a lottery or coin toss, to either the [PROGRAM NAME] (treatment) group or the diet and exercise curriculum (control) group. Both groups of schools are vital to the study.

Intervention Cohorts: Each of the participating schools, whether they deliver [PROGRAM NAME] or the alternative diet/exercise program, will involve two study cohorts. They will deliver the programs to Cohort 1 beginning in fall 2011 for two years, during 7th and then 8th grades. Follow-up surveys will be conducted by the study team at the end of 8th and 9th grades, in the spring of 2013 and 2014. These same schools will deliver the programs to Cohort 2 (7th graders in fall 2012) over three years, during 7th, 8th, and 9th grades, and the study team will conduct follow-up surveys at the end of 8th and 9th grades, in the spring of 2014 and 2015. Participation of high schools is essential so that 9th grade surveys can be conducted, and because the third program year for Cohort 2 will occur in 9th grade.

- **[PROGRAM NAME] schools** will deliver a curriculum-based program model for middle school students (and for ninth-grade high school students in Cohort 2). [PROGRAM NAME] is an interactive, activity-driven program that uses “hands on” strategies to deliver its core teen pregnancy prevention message. The program covers eight units which include a range of topics from consequences of teenage sex to conflict resolution. Boys and girls are separated into different classrooms during program delivery.

Teachers in these schools deliver the program. Teacher training and planning will begin in fall 2011. Teachers will deliver [PROGRAM NAME] to all 7th-grade students starting in fall 2011 for Cohort 1 (typically as part of a required health or physical education class) and in fall 2012 for Cohort 2. The program involves eight hours of class time and is typically delivered 8 days in a row, but can be delivered over a period of up to eight weeks if necessary. All study schools must agree that if assigned to the [PROGRAM NAME] group they will implement the curriculum in its entirety and strive to do so in a thorough manner, with fidelity to the program model as designed.

- **“Diet and exercise curriculum schools”** will implement a diet and exercise education curriculum provided by [ORGANIZATION NAME], over a period comparable to the period in which [PROGRAM NAME] is delivered in the program schools. Training will be provided for teachers who will deliver this program. These schools act as a control group and are crucial in helping the study team gain an understanding of the impact of the [PROGRAM NAME] program. The diet/exercise program also has the potential to bring positive benefits to participating students.

District and school leaders will be notified of their randomly assigned status ([PROGRAM NAME] or the diet and exercise group) after collection of parental consent has been completed at the start of the school year in fall 2011. Schools, with supporting materials provided by [EVALUATOR], will notify parents of which program their children will be receiving. Schools will be required to abide by the results of random assignment and help to facilitate the study team’s efforts to obtain parental consent and conduct data collection.

Teacher Training: There will be two types of training for teachers:

- *Supporting Consent-Gathering.* In late August or early September of 2011, [EVALUATOR] will provide webinar training for all study schools on how to assist in collecting parental consent for the evaluation. Several webinar sessions will be held to accommodate different schedules.
- *Implementing [PROGRAM NAME].* In mid- to late October 2011, teachers in schools assigned to deliver [PROGRAM NAME] will participate in a two-day training session offered by [ORGANIZATION NAME]. The training will be offered at several central locations and at multiple times to accommodate varying schedules. In fall 2012, additional training will be provided, as a refresher for teachers who will continue teaching [PROGRAM NAME] the next year, and for new teachers. Districts will be reimbursed (see below) for the cost of substitute teachers filling in for regular teachers who participate in the training.

Sample Memorandum of Understanding (MOU)

Teachers delivering the diet and exercise curriculum in control schools will receive similar training opportunities. Schools will determine which teachers will be involved in delivering the programs.

Program Materials: [ORGANIZATION NAME] will provide all of the materials used in the [PROGRAM NAME] curriculum and the diet/exercise program. Multiple copies of the materials will be provided at no cost (and in bound, published form) to participating schools, and the schools may keep the materials at the conclusion of the study.

Data for the Study: The study includes the following data collection activities during the 2011–15 school years:

- 1. Student Surveys:** Youth from both the [PROGRAM NAME] and diet/exercise groups will be asked to complete a questionnaire (roughly 30 minutes in length) three times, at the school, during the school day. These questionnaires will pose questions about attitudes about abstinence and sex, relationships with peers and parents, types of sexual experience, and other risk behaviors. For Cohort 1, the first survey will be in fall 2011, after parental consent is provided, but before delivery of [PROGRAM NAME], or the diet/exercise classes. Students entering 7th grade will be assembled, and trained [EVALUATOR] staff will administer the **baseline survey**. [EVALUATOR] will also administer **two follow-up surveys**, one at the end of the 8th- grade year (May 2013 for Cohort 1) and a second at the end of the 9th-grade year (May 2014). The study team will administer the surveys to Cohort 2 at the same times in the school year, but on a schedule one year later (fall 2012-May 2015). The study teams will coordinate with a designated contact person in each school to schedule survey administration, including a main survey administration session and make-up sessions as needed for youth who are absent for the first session.¹
- 2. School Site Visits.** To learn how [PROGRAM NAME] and diet/exercise classes are delivered, members of the study team will visit some of the schools participating in the study, and this may include one or two visits to [DISTRICT/COUNTY NAME] district schools. The purpose of these visits will be to observe program classes and conduct short interviews with key staff. Staff in some schools delivering the diet and exercise classes will also be interviewed, so the study team can learn about the types of health education provided there.
- 3. Participant Focus Groups:** During one of the site visits, the study team may invite a small group of students in some schools to participate in a focus group discussion about their experiences with [PROGRAM NAME]. The study team will coordinate this activity with schools in advance, and ensure that parent permission has been obtained.
- 4. School Records/Administrative Data:** The study team will request copies of the roster of Cohort 1 and Cohort 2 7th-grade students in each participating school, to identify the student sample and updated rosters prior to each follow up survey. The study team will request administrative data on these students for each of their three years in the study, including attendance, grades, standardized test scores, and student directory information (including contact information for the parents and guardians of all 7th grade students).

¹ The surveys will also be available for youth to complete via web or by telephone if they miss the group survey administration sessions.

Sample Memorandum of Understanding (MOU)

Table 1. Study Timeline:

Activity ^a	Planned Timing (Cohort 1)	Planned Timing (Cohort 2)
Parental consent collection	Late 2011 August-late September	Late 2011 August-late September
Schools notified of random assignment result	Mid- to late October 2011	Mid- to late October 2011
Teacher training	Mid- to late October 2011	Mid- to late October 2012
Refresher training	Mid- to late October 2012	Mid- to late October 2013
Baseline data collection	Mid-Oct. 2011 to mid-Nov. 2011	Mid-Oct. 2012-mid-Nov. 2013
Program delivery	Early Nov. 2011 to early Jan. 2012 Early Nov. 2012 to early Jan. 2013	Early Nov. 2012 to early Jan. 2013 Early Nov. 2013 to early Jan. 2014 Early Nov. 2014 to early Jan. 2015
Site visits/staff interviews	August 2011-January 2012	n/a
First follow-up data collection	May 2013	May 2014
Second follow-up data collection	May 2014	May 2015

^a Each activity is applicable to both [PROGRAM NAME] and diet and exercise schools.

Confidentiality of Data. The study team will protect the confidentiality of all information collected for the study and will use it for research purposes only. No data will be released by the study team in any manner that would allow schools, staff, or students participating in the study to be identified or associated with any data, except as may be required by law.

Recognition of Burden and District Reimbursement. Recognizing the burden of participating in the study on schools, [ORGANIZATION NAME] will compensate participating school districts for each student completing all 8 program hours, including the control group. Additionally, schools will receive [PROGRAM NAME] materials such as t-shirts, pens, and backpacks. Reimbursements will be based on the individual needs of schools, and be contingent on school and district cooperation in scheduling and providing space for survey administrators. Specific details of school reimbursement are:

- [FILL IN AS NEEDED]
- [FILL IN AS NEEDED]

Commitment to School Participation. District-level commitment is essential to study success, but sustained commitment of schools is critical as well. The [DISTRICT/COUNTY NAME] district will secure commitments in writing from the principals of the selected middle schools and high schools to fulfill their role in the [PROGRAM NAME] study, and in the event of turnover at the school level, the district will engage the continued participation of the school under its new leadership.

Term of Agreement. This agreement will run from May 1, 2011, through May 30, 2015. If the target number of schools is recruited for the study, then the parties will pursue implementation of the evaluation as described in this agreement. However, if the target number of schools is not recruited successfully, then the agreement will be terminated.

Communication: [CONTACT NAME] at [ORGANIZATION NAME], will be the primary contact person from [ORGANIZATION NAME]. Please contact him/her with questions related to this agreement or the responsibilities described in it.

Sample Memorandum of Understanding (MOU)

[CONTACT NAME]
[ORGANIZATION NAME]
[CONTACT INFORMATION]

The primary contact at [EVALUATOR] will be [CONTACT NAME]. Please contact him/her for more information about the evaluation.

[CONTACT NAME]
[EVALUATOR]
[CONTACT INFORMATION]

Your district has named the following people as contact for discussions related to the study:

[TO BE ADDED BY DISTRICT BEFORE SIGNING AGREEMENT]
District Level: Contact Name Address
Telephone
E-mail Address

[TO BE ADDED BY DISTRICT BEFORE SIGNING AGREEMENT]
School #1 Level: Contact Name Address
Telephone
E-mail Address

[REPEAT FOR EACH PARTICIPATING SCHOOL]

Agreement to Proceed. This agreement accurately describes the plan for including [DISTRICT NAME] in the [PROGRAM NAME] study, which is part of the [EVALUATION NAME]. Please sign below and complete the short checklist to indicate your agreement to proceed. Then, please keep a copy for your records and return the original in the enclosed Federal Express envelope. We appreciate your assistance with this important study, and we look forward to working with you.

_____ Date

_____ Date

_____ Date

Checklist of Key Items:

I have examined the baseline survey instrument _____ Initial here

I have discussed this agreement with the relevant parties in my district and participating schools. By signing this agreement, it is confirmed that any approvals required for participation in this study have been received (unless otherwise noted in the space below). _____ Initial here

Additional notes (if applicable): _____

Suggestions for Boosting Response Rates for Consents and Surveys

This document provides evaluators with suggestions on how they can improve consent and survey response rates.

SUGGESTIONS FOR BOOSTING RESPONSE RATES FOR CONSENTS AND SURVEYS



Consents

- Host a parent night or attend one that is already scheduled at the school/center/clinic to explain the study and answer questions. Have consent forms on-hand for parents to sign on the spot.
- Send consent forms home with other important information (application to program, report cards, permission forms, etc.) the school/center is already planning to send home. Try to make them stand out by using colored envelopes and address them “To the parents of...”.
- Work with your point of contact at the school or center to obtain parent contact information to send forms directly home (using school envelopes) or to conduct phone calls home. If they are hesitant to provide the contact information to you, work with their staff to see if they are willing to conduct the mailing for you (you would assemble materials, but they would address the envelopes). Possibly offer incentives to staff that assist with the mailing.
- Offer incentives to parents and/or sample members for returning consent forms and offer incentives to staff assisting with consent collection. Consider offering larger incentives for earlier returns or offering graduated incentive payments (to staff or classrooms) for higher response rates.
- Post reminder fliers at the school/clinic/center.
- For school-based programs, ask if the school will include reminders in their daily announcements or on message boards.
- Consider alternate modes of consent (ex: verbal) with IRB approval.



Surveys

- Organize group administrations at the school/community center/clinic for baseline or make the survey part of a program application package. Consider group administrations for follow-up surveys as well, where possible.
- When conducting group administrations, consider administering the survey on Saturdays or Sundays (for certain target populations).
- Post fliers at the school/clinic/community center prior to the administration to remind students about the survey (for make-up dates as well).
- For school-based programs, ask if the school will include reminders in their daily announcements.
- Send reminder emails, texts, conduct reminder phone calls (include link to web survey in email, if applicable). Consider sending reminders through social media, such as Instagram or Snap Chat.
- Offer alternate modes to complete the survey (if resources allow). Post reminders on the study website, if applicable.
- Maintain contact with your sample in between survey administrations. Send contact update postcards, birthday cards, holiday cards; host events for the study (think about whether you need to host separate events for treatment and control sample members to avoid contamination).

Suggestions for Incentives

This document lists incentives that can be used to potentially boost survey response rates.

INCENTIVES

The following is a list of suggested incentives. Grantees should confirm proposed incentives are allowable with their Project Officer and make sure they are approved by the study IRB.

- Gift cards – any amount can be an incentive. Keep in mind what works with your specific target population and be sure any card you offer is a place your sample members are likely to access. Work with your point of contact to see if they have suggestions about places that would be appropriate. Remember some cards have an expiration date and incur fees over time and try to avoid these cards.
- If you offer gift card incentives in smaller amounts (\$5 or \$10) consider places where this would go far (Starbucks, Dunkin Donuts, Subway, McDonalds, iTunes, etc.) while being considered acceptable to the local community. Schools and programs should always be consulted on the selection of the type of retail outlet/product you are considering.
- Cards to places like Target or Wal-Mart can be used widely and for a variety of items. Consider offering cards to grocery stores for certain populations (ex: pregnant/parenting teens).
- Enter student names into a drawing once they have returned a consent form or completed the survey. This may allow you to offer larger items as incentives (but fewer of them).
- If you offer online gift cards –where students complete the survey online and are automatically shown a link to gift cards they can use online, consider including a back-up option to mail a gift card for those that may have access issues or do not want to incur shipping costs (ex: They are not an Amazon Prime member and you choose Amazon).
- If you have multiple follow-up surveys, consider increasing the incentive amount each round. You can also consider offering greater incentives for earlier returns.
- Host a pizza party or other class/group event when they reach a target response rate; if possible, offer to pay for a field trip (organize this activity so that it does not provide an opportunity to contaminate the control group through interaction with the treatment group).
- For surveys conducted outside of schools, consider offering meals during the survey and reimbursing respondents for transportation, if necessary.
- If it is applicable for your sample, consider reimbursing them for childcare expenses. Consider offering t-shirts, pens, pencils, accessories for cell phones, and other items with the study name (not the program name). In some instances, you may want to avoid using the study name (ex: if it contains the word “pregnancy”).

Teen Pregnancy Prevention Evaluation Technical Assistance - Analysis Plan Frequently Asked Questions

This document provides guidance on approaches to producing credible estimates of the impact of a TPP program.

**Teen Pregnancy Prevention
Evaluation Technical Assistance
– Analysis Plan Frequently Asked
Questions**

May 17, 2013

Russell Cole
John Deke
Susan Zief



MATHEMATICA
Policy Research

FREQUENTLY ASKED QUESTIONS ABOUT PREPARING AN ANALYSIS PLAN

Preparing a complete and detailed analysis plan before beginning impact analyses is particularly useful in the context of program evaluation for three main reasons:

1. **Enhances objectivity and credibility.** An analysis plan identifies all primary research questions, outcomes, and analytic approaches before examining the data. Thus, it provides an evaluation *structure* that specifies key hypotheses on intervention effects and the methods for testing them up front. This increases the objectivity and scientific nature of your approach and, therefore, the credibility of the results. Everyone will be better assured that the reported findings are not influenced by knowledge of the outcomes.
2. **Secures stakeholder buy-in.** The analysis plan is an agreement that you have with key stakeholders, such as the grantee and the program staff. It is a shared road map that guides everyone from data to findings. If key stakeholders agree on the approach, there is little room for debate when the findings emerge.
3. **Makes efficient use of resources.** Because an analysis plan provides researchers with a clear guide for analysis and reporting, it enables you to move more efficiently from raw data to findings. In addition, much of the description and information presented in the analysis plan can be recycled into evaluation reports and journal articles.

This document is designed to support the requests from the Office of Adolescent Health (OAH) and Family and Youth Services Bureau (FYSB) for an analysis plan from its Teen Pregnancy Prevention (TPP) and Personal Responsibility Education Innovative Strategies (PREIS) grantees. In particular, this document addresses the more frequently asked and technical questions about preparing your analysis plan. These questions and answers complement the detailed outline that OAH and FYSB released to their grantees and is posted on the Evaluation Technical Assistance SharePoint site [link]. In the questions below, we have included a link to the sections of the analysis plan template that they are able to address.

- *In a blocked/stratified/paired random assignment design (such as assigning schools to condition within districts), what should I do when there is only one unit within a given block/strata/pair? (Section 3.c.iii of analysis plan template)*
 - Multiple strategies exist for addressing this situation, and different strategies could be preferable in different contexts. One strategy is to combine all singletons in the study into a new block. However, that strategy works only if there are both treatment and control singletons. Another strategy is to add singletons to blocks containing similar (in terms of key pre-intervention variables) units. Evaluation teams should consult with the TPP Evaluation technical assistance (TA) team to develop a strategy that is best suited to their context. Regardless of the strategy selected, sensitivity analyses should be conducted to determine whether alternative strategies would substantially change the study's findings.

- *Some of my outcomes are not direct measures from the survey (that is, they are constructed responses). How should I describe the constructed measures? (3.f)*
 - In many cases, the outcomes of greatest interest will be based on the direct responses to survey items collected at the focal follow-up data collection point (for example, whether the respondent indicated that he or she engaged in a risky sexual activity in the past 30 days). However, for outcomes such as a constructed measure of sexual risk, it will be necessary to provide detail about the creation of this construct (for example, which survey items were used in the construction of the risk scale, how the survey items ultimately were pooled to create the risk index scale, whether the pooling of the survey items was determined based on a previously developed scales, how the creation of the scale treated missing data for items, and so on). In addition, please provide information on how a reliable construct will be determined (for example, by using factor analysis in the construction and computing an alpha coefficient for the final measure).
- *How should inconsistent and seemingly inaccurate responses be scored during data preparation? (4.a)*
 - Examples of inconsistent responses might include an individual who indicates at baseline that he or she had sex in the past three months; then, at survey follow-up, the respondent indicates that he or she never had sex. Alternately, within a single survey, a respondent might indicate that he or she has never had sex on one question, but then provides a non-zero count for the number of times that he or she had sex in a separate question.

There is no one “best” approach for handling inconsistent survey responses, since any approach requires the researcher to make assumptions and assertions about why data are inconsistent. At a minimum, the issue of inconsistent responses should be acknowledged and a systematic approach and clearly defined decision rule should be applied to all inconsistent data in the study. The appropriateness of the rule should be justified and offer an approach that has face validity to a critical reviewer. For example, the researcher should provide data about the prevalence of inconsistent responses by treatment condition and by survey mode in justifying an approach.

One plausible (conservative) decision rule for handling inconsistent responses is to treat all responses for an individual with inconsistent data as missing and handling the resulting missing data under the general missing data strategy used in the study (see questions in Section 4.c.iv). The justification for this rule would be that one cannot know which of the responses is correct and, therefore, it is advisable to remove any potentially inconsistent and uninformative responses from the data. Another possible decision rule is outlined in the response to the question below (i.e., treating the “first response” as the “correct response”), but the use of this rule with outcome data requires more assumptions, and in particular as the intervention could influence the first response.

Importantly, results should always be examined and presented according to a variety of approaches for handling inconsistent data. In particular, when there are inconsistent responses in the follow-up survey (or between the baseline and follow-up surveys), it is necessary to conduct sensitivity analyses to assess whether the intervention affected a respondent’s propensity to respond inconsistently (see a

suggested sensitivity analysis in Section 4.c.vii), and whether results are robust to different methods for handling the inconsistent responses.

- *Are there any alternate approaches for handling inconsistent data that occurs within the baseline survey only? (4.a)*

One potentially plausible rule for inconsistent data, and in particular on the baseline survey, (for example, a respondent indicates that he or she has never had sex and then provides a non-zero count for the number of times that he or she had sex in a separate question on the baseline survey) is to use the first of the inconsistent responses as the “correct” response, and to impute subsequent responses based on this “correct” response. In a computerized administration of an assessment that utilizes skip patterns, the first response would presumably skip any possibility for subsequent inconsistent responses. Therefore, the justification for this rule would be to use an approach that mimics the data that would have been obtained using a computerized assessment. This justification is especially compelling if the study uses multiple modes for data collection, including a computerized assessment as one mode, in that the inconsistent data in the non-computerized assessment are revised to comply with the mode used in the computerized assessment. Under such a decision rule, the justification might be that in a randomized controlled trial (RCT), we would expect a similar number of inconsistent responses across treatment and control groups before the start of the intervention.

As with the recommendation above, sensitivity analyses should be conducted that use multiple methods for handling inconsistent baseline covariate data, to determine the robustness of the findings to these decisions.

- *How should the frequency of a given sexual behavior be scored when there are respondents who indicate that they did not engage in sexual activity (for example, at follow-up, a participant indicates that he did not engage in sex in the past three months and, therefore, did not respond to the frequency of sexual activity item in the survey)? (4.a)*
 - When individuals do not provide frequency data for sexual behavior questions, *but do indicate that they did not engage in this behavior*, we recommend imputing a zero count for the frequency data and reporting this decision rule in any write-up of the results. When using such an approach, analyses of the frequency data can be conducted on the full sample, instead of focusing on those who engaged in a certain behavior (which might be an endogenous subgroup).¹

This is considered a logical imputation and is a reasonable way to fill in certain item nonresponse gaps in the survey when the underlying assumption is reasonable. In this example, the underlying assumption is that the response to the initial question about ever engaging in sexual activity is correct and, therefore, the zero count of episodes of sexual activity is a logical result.

Although logical imputation is seemingly appropriate in this example of a missing response to a survey question, *we suggest caution when attempting to assert a logical*

¹ See the brief by Colman (2012) that addresses the issue of estimating impacts in endogenous subgroups.

imputation process for handling inconsistent data (such as the preceding example in which there is inconsistent information about whether the respondent ever engaged in a sexual act across two time periods, or across questions in a single survey administration). In a situation in which there is inconsistent information, we cannot know which responses are correct; therefore, it is inappropriate to choose one version of the inconsistent data for imputation in favor of an alternate response. As a result, our suggested approach in such a situation is to treat the inconsistent responses as missing data. Note: information on approaches for handling missing data are presented in Section 4.c.iv.

- *What are some of the recommended approaches for demonstrating the equivalence of the analytic sample at baseline? (4.b)*
 - Regardless of the level of randomization (individual or cluster), we recommend conducting multivariate regression analyses to demonstrate baseline equivalence, regressing the baseline variable of interest on a treatment indicator and any stratification variables (but no other variables), and reporting the adjusted means and p -values of the differences in the baseline variable of interest for the treatment and control groups. If the baseline variable of interest is dichotomous, we recommend using a linear probability model rather than a logit or probit. However, if the logit or probit is used, regression-adjusted mean proportions should be reported for the treatment and control groups—not odds ratios.
- *How should I conduct baseline equivalence calculations in a situation in which clusters, rather than individuals, were randomly assigned to condition? (4.b)*
 - For designs in which clusters are assigned to treatment or control conditions, standard errors should reflect the clustered nature of the design. This can be achieved either by estimating a mixed-effects model (for example, a hierarchical linear model [HLM]) or by adjusting standard errors for clustering using the sandwich estimator of variance (Williams 2000). Be sure to report the p -value for the test of the difference based on this analysis that adjusts for clustering (which is likely to be larger than the p -value that would be estimated if clustering were ignored).
- *Is it necessary to conduct a multiple comparison adjustment when calculating the statistical significance of baseline equivalence across multiple variables or outcome measures? (4.b)*
 - We do not suggest using a multiple comparisons adjustment when calculating statistical significance in baseline equivalence calculations. Multiple comparison adjustments keep the probability of a Type I error low but increase the probability of a Type II error. That is, these adjustments increase the probability that we mistakenly conclude there are no differences between the treatment and control groups. In the context of baseline equivalence testing, this is not a desirable tradeoff.
- *What is a benchmark analysis and how do I define an analysis as a benchmark approach? (4.c)*
 - The benchmark analysis is the combination of analytic approaches deemed most appropriate in the context of a given study. To conduct impact analyses, researchers are required to make decisions about statistical models, estimation methods, the treatment of missing/inconsistent data, and so on. Some of these decisions involve choosing between multiple valid approaches to addressing an issue. While each

approach may be valid, each approach may also result in slightly different findings. At the analysis plan stage, it is advisable to propose approaches that seem the most appropriate and defensible for the context of the given study.

Findings from the benchmark analysis should serve as the lead story when you report how you conducted the analyses. As suggested elsewhere in the FAQs, the robustness of findings to the choice of approaches used in the benchmark analysis should be assessed with sensitivity tests (Section 4.c.vii).

- *How should I define the analysis sample for this evaluation? What information do I need to provide about the description of the analysis sample? (4.c.i)*
 - The analysis sample is defined as the sample members who contributed outcome data at the focal follow-up data collection period described in the research question of interest. In some evaluation designs, this sample might involve data pooled across multiple cohorts of implementation, or across multiple study sites. In this section, provide information on how the analysis sample will be defined. In particular, indicate whether the analysis sample used to answer the primary research question (at the focal time point) will include all individuals participating in the study or, for example if the analysis is limited to a subset of cohorts. In general, the analysis plan should carefully delineate how the analysis sample was obtained from the initial sample.
- *How should I handle dichotomous outcomes in my impact study? (4.c.ii)*
 - For the purposes of presenting a meaningful and interpretable effect, the TPP Evaluation TA team recommends that all grantees use a linear probability model to estimate program impacts in RCTs for dichotomous outcomes; the team also recommends conducting sensitivity analyses using logit/probit models to check the robustness of the results. As communicated in the 2012 presentation on estimating interpretable and accurate impacts, conducting the analysis using a linear probability model will produce treatment effects that can be interpreted as percentage point differences in the prevalence rates of the outcome of interest, which is a more interpretable and understandable effect size than an odds ratio typically reported from analyses of these outcome variables.²
- *What would a linear probability model look like for an individual-level RCT? (4.c.ii)*
 - For individual-level RCTs focused on a dichotomous outcome, the model specification for dichotomous outcomes using the linear probability model will likely look like the following:

$$(1) \quad P(Y = 1) = \beta_0 + \beta_1 * T + \sum \beta_k * X_k$$

² When conducting an analysis under a linear probability model framework, we advise calculating heteroskedastic-consistent standard errors when estimating impacts. Alternately, if impacts are calculated using a logit or probit instead of a linear probability model, we recommend reporting the regression-adjusted mean predicted probabilities for the treatment and control groups and conducting statistical testing for the difference between those two values using the correctly calculated standard error for that difference (which is not the standard error associated with the odds ratio).

where $(Y = 1)$ represents that the outcome of interest was observed at the focal follow-up data point (for example, “had sex at the 12-month follow-up period), β_1 is the difference in the observed prevalence rates of the outcome (in percentage points) across the treatment and control conditions, T is a treatment indicator variable, and X_k is a vector of covariates.³ Note that in this specification, the impact estimate β_1 represents the impact of the intervention at the focal follow-up time period. When a baseline measurement of the outcome of interest is included as a covariate, then this impact represents an impact at the focal time point, after adjusting for baseline differences in the outcome, which is often a more precise impact estimate in an RCT. Other baseline covariates X_k can also be included to improve precision in the impact analysis (under the assumption that the baseline covariates are uncorrelated with treatment assignment, they will explain only residual error variance in the outcome, thereby shrinking the standard error of the treatment effect).

- *What would a linear probability model look like for a cluster-level RCT under an HLM framework? (4.c.ii)*
 - For clustered RCTs focused on a dichotomous outcome, the base model might look somewhat similar; however, the model will be expressed in an HLM framework:

$$(2) \text{ Level 1: } P(Y_{ij} = 1) = \beta_{0j} + \beta_{1j} * T_{ij} + \sum \beta_{kj} * X_{(k)ij}$$

$$\text{Level 2: } \beta_{0j} = \gamma_{00} + u_{0j}$$

...

$$\beta_{kj} = \gamma_{k0}$$

The main difference in this specification is that individuals i are nested within clusters j , and the impact analysis at a minimum includes a random effect u_{0j} to adjust for the nonindependence of individuals within clusters. The analysis can be considerably more complicated, with cluster-level covariates included to improve study precision or additional random effects, but, at a minimum, the model specification for a hierarchical linear probability model has to demonstrate how the impact analysis will adjust for clustering through the inclusion of a random effect for each site’s intercept. As with the linear probability model presented earlier, this specification will look largely the same if a continuous outcome were used as the dependent variable, instead of the probability of the outcome.

³ The model specification for a continuous outcome would be largely the same, but instead of $P(Y = 1)$ as the outcome of interest, the continuous outcome would be presented as the dependent variable.

- *Would other model specifications be appropriate for a cluster-level RCT without using HLM? (4.c.ii)*
 - Another approach for accounting for the nonindependence of observations within a cluster is to use a sandwich estimator for the treatment effect, which appropriately adjusts the standard errors for the treatment effect. (Williams 2000). For the purposes of the analysis plan, it is sufficient to provide a model similar to the individual-level RCT and just indicate that the model will be estimated using cluster robust standard errors.⁴
- *Aside from providing information on model specification, what other information should be included in this section? (4.c.ii)*
 - It is necessary to describe the software package (Stata, SAS, R, HLM, and so on) that will be used to estimate program impacts, as well as the procedures within these packages that will be used for estimation

Finally, this section should specify the criterion against which the results will be deemed statistically significant. For the purpose of the U.S. Department of Health and Human Services (HHS) Pregnancy Prevention Research Evidence Review (PPRER), findings are considered statistically significant if the p -value is less than 0.05, using a two-sided test. We recommend using this criterion for asserting results as statistically significant.

- *Do I have to include blocking or stratification variables as dummy (indicator) variables in my impact analyses? (4.c.iii)*
 - Blocking is often performed to improve the precision of impact estimates and to ensure treatment and control group balance on key baseline variables. To be consistent with the sample design and to obtain the precision gains, indicator variables for each block should be included as covariates in all impact models. Alternately, impacts should be estimated separately within each block and then a weighted average of the block-specific impacts should be used to generate an overall impact estimate. However, when there is a single unit within a block (as described above), including the indicator variable for that block would result in the singleton unit not contributing to the impact estimate (which might not be desirable). Similarly, if the goal is to estimate block-specific impact estimates with a single unit within a block, it will not be possible to calculate a block-specific impact; thus, this block would not contribute to the overall impact estimate. The loss of these blocks from contributing to the overall impact estimate could potentially produce a biased overall impact estimate; therefore, it is important to consider various strategies for handling this issue.

⁴ See Schochet (2009) for a discussion of estimation procedures for cluster RCTs and suggested procedures for assessing the sensitivity of findings using different estimation methods.

- *Should the list of covariates vary across different outcomes used to answer primary (and secondary) research questions? (4.c.iii)*
 - We advise that you include a common set of core baseline covariates across impact models. In general, the core set of baseline covariates should include variables that you expect to correlate with the outcomes of interest described in the primary and secondary research questions, have strong face validity, and that adjust for significant differences at baseline. Importantly, in order for a study to receive the highest evidence rating in the HHS PPRER, core baseline covariates must include demographic variables and the baseline assessment of the outcome of interest that are significantly different at baseline.

Whenever possible, we recommend using the baseline assessment of the outcome(s) of interest as a covariate. For example, if the outcome of interest is “had sex in the past three months,” we recommend including the baseline assessment of that variable as a covariate. If a secondary research question focuses on “had sex in the past three months without a condom” as the key outcome of interest, we suggest using the baseline version of that variable as a covariate, instead of (or in addition to) “had sex in the past three months.” In summary, we suggest using a common set of covariates across all outcomes, and either varying only the baseline assessment of the outcome used as a covariate in each impact model or including the pre-intervention assessments of *all* outcomes as covariates in each impact model. As noted in the sensitivity analysis section, we also recommend conducting analyses with and without covariates to verify whether the inclusion of covariates substantively affects results.

- *What types of approaches for handling missing data are recommended by the TPP Evaluation TA team? (4.c.iv)*
 - The TPP Evaluation TA team has written a brief on suggested approaches for dealing with missing data on both the outcome variables and covariates to use in the impact models. This brief will be released in the immediate future and we will send out an announcement when it is available for download on the TPP Eval TA SharePoint site.
- *In my study, I don’t have baseline data for all members of my analytic sample (the members of the sample that have observed post-test data). What can I do to show baseline equivalence of my analytic sample? (4.c.iv)*
 - We suggest reporting the proportion of observations in the treatment and control groups missing each baseline variable. This in itself is a way to assess baseline equivalence. We also suggest reporting baseline differences for the nonmissing sample and conducting a sensitivity analysis that calculates impacts for the sample of youth with both baseline and follow-up data. If researchers desire a complete data set without any missing baseline data, we describe an approach for imputing missing baseline (and outcomes) data in the following section.
- *Can I just use multiple imputation to generate baseline and follow-up data for anyone who was observed at either data period (i.e. observed at baseline but not at follow-up, or vice-versa), and use this resulting (imputed) sample for baseline equivalence and impact analyses? (4.c.iv)*
 - Yes, this approach is valid as long as you conduct imputation separately for the treatment and control groups. If imputation is conducted without regard to

treatment status, then missing values in the treatment and control groups will be imputed to a common mean, which will result in attenuation bias when calculating differences between the treatment and control groups. Again, we suggest conducting a sensitivity analysis that presents baseline equivalence and impacts for the sample with complete baseline and outcomes data.

- *When might sample weights be needed in an impact evaluation, and how would they be calculated? (4.c.v)*
 - In some impact evaluations, researchers create sample weights for sample members, to enable the members of the sample to more generally represent a population of interest, to adjust for nonresponse bias in impact estimation, or to account for variation in the probability of treatment assignment across sample members. Note that this refers to situations in which the intended, not the observed, probability of treatment varies across sample members. For example, consider a hypothetical multisite RCT in which there is a need to assign half of the participants in one site to treatment status and, in a separate site, to assign two-thirds of the participants to treatment status. In such a situation, the research team might calculate inverse probability of treatment weights to reflect the difference in treatment probability (or they could just use site fixed effects in their analysis to also account for the nonequal probability of treatment assignment across sites).
- *What is a simple method for adjusting for multiple comparisons? (4.c.vi)*
 - The easiest adjustment method to implement is the Bonferroni (1935) method, but it is also the most conservative (meaning that it will be harder to detect effects using this method). To implement the Bonferroni method, simply divide the α criterion used to determine statistical significance by the number of hypothesis tests being conducted and use the smaller version of α as the barometer used to establish statistical significance. For example, if four impacts are used to answer a study's primary research question, the revised version of α used to determine statistical significance will be $\alpha/4$. In most situations, researchers will want to use $\alpha = 0.05$ for asserting statistical significance; in this example study, the criterion used to characterize a p -value as statistically significant after adjusting for multiple comparisons would be $\alpha = 0.0125$.
- *What is a less conservative approach for adjusting for multiple comparisons? (4.c.vi)*
 - The Benjamini-Hochberg (1995) method is more difficult to implement (see the following algorithm), but when a large number of tests are being conducted it can sometimes noticeably increase the number of statistically significant impacts.

This description of the Benjamini-Hochberg adjustment are adapted from Appendix D of the What Works Clearinghouse evidence standards, version 2.1.

Rank order statistically significant findings within the domain in ascending order of the p -values, such that $p_1 \leq p_2 \leq p_3 \leq \dots \leq p_m$, with m being the number of significant findings in the primary research questions.

For each p -value (p_i), compute: $\alpha'_i = \frac{i^* \alpha}{M}$, where $i = 1, 2, \dots, m$, $M =$ the total number of primary research questions being examined, and α is the prespecified criterion used to assert statistical significance (likely $\alpha = 0.05$).

Identify the largest i —denoted by k —that satisfies the condition: $p_i \leq \alpha'_i$. This establishes the cutoff point and enables us to conclude that all findings with p -values smaller than or equal to p_k are statistically significant, and findings with p -values greater than p_k are not significant at the prespecified level of significance ($\alpha = 0.05$ by default) after correction for multiple comparisons.

Here is a hypothetical example that applies the Benjamini-Hochberg adjustment with four primary research questions, three of which show a statistically significant difference at the 0.05 level before adjustment. Assume that the ordered p -values are 0.01, 0.02, 0.04, and 0.09 (note: only the first three are originally significant at the 0.05 level). That is, in this example, there are three initially significant findings ($m = 3$), and four total contrasts ($M = 4$).

Study-Reported p -Value (p_i)	p -Value Rank (i)	$\alpha'_i = \frac{i^* \alpha}{M}$ $= i^* 0.05 / 4$	$p_i \leq \alpha'_i ?$	Statistical Significance After BH Correction ($\alpha = 0.05$)
0.01	1	0.0125	Yes	Significant
0.02	2	0.025	Yes	Significant
0.04	3	0.0375	No	Not significant
0.09	-	n.a.	n.a.	n.a.

n.a. = not applicable.

In this example, although three contrasts had p -values less than the 0.05 criterion, only the first two would be considered statistically significant at the 0.05 level after adjusting for multiple comparisons using the Benjamini-Hochberg adjustment. That is, when reporting results for the study, it would be appropriate to report all of the study-computed p -values, but then to report that only the contrasts with 0.01 and 0.02 p -values were found to be statistically significant at the 0.05 level, after adjusting for multiple comparisons.

Note that using the Bonferroni adjustment, only the first impact in this example would remain statistically significant after adjustment. Also note, however, that although the Benjamini-Hochberg procedure can increase the number of statistically significant impacts relative to Bonferroni, it does not affect the probability that *at least one* effect will be significant (this is because the Benjamini-Hochberg procedure applies the Bonferroni adjustment to the impact with the smallest p -value).

- *Should I conduct a multiple comparison adjustment if I am testing the impact of two different interventions against a comparison group (for example, T_A versus C and T_B versus C) as primary research questions? (4.c.vi)*
 - Yes, you should perform a multiple comparison adjustment for all comparisons tested as primary research questions. In this case, in which there are multiple comparisons

performed across different group contrasts, there is still a threat of increased Type I errors; therefore, it is advisable to use a multiple comparison adjustment procedure to guard against false positive results. In this context, alternate multiple comparison adjustment procedures might be advisable, such as the Tukey (1953) or Dunnett (1955) tests.

- *What types of sensitivity analyses should be conducted? How should they be presented? (4.c.vii)*
 - Sensitivity analyses are used to assess whether the study’s findings would change substantially if alternative, but defensible, methodological choices were made when calculating impacts. Some examples of sensitivity analyses include the following:
 - **Estimate impacts with and without baseline covariates.** There is some debate as to whether regression adjustment for baseline covariates is appropriate in the context of an RCT (Freedman 2008; Schochet 2010), so we recommend estimating impact models both with and without covariates, although in most cases the regression-adjusted findings will be the benchmark (main) findings.
 - **Estimate impacts using different approaches to handling missing data.** One approach to conducting impact analyses in the face of missing outcomes data is to use listwise deletion and conduct the impact analyses using the available sample. Other approaches include multiple imputation, maximum likelihood, or nonresponse weights (Puma et al. 2009). We suggest assessing the sensitivity of the benchmark impacts to at least one alternative approach for handling missing data.
 - **Estimate impacts using different approaches for handling inconsistent survey responses.** As noted earlier, when individuals provide inconsistent data within a given survey administration period or across multiple survey administrations, the research team should apply a well-defined rule for handling the inconsistent data and conduct the impact analysis following the application of this rule. However, given that there are multiple ways of handling inconsistent data, we advise re-estimating the results using other, alternate decision rules for handling inconsistent data. These sensitivity analyses can lend credibility to the results being robust to different ways of handling inconsistent data.

In addition, if the intervention affects the consistency of survey responses, then estimates of program impacts on outcomes of interest could be biased. For example, some interventions might lead youth in the treatment group to provide more socially desirable (but inaccurate) answers to sensitive questions than youth in the control group, creating the illusion of an impact on an outcome of interest. To examine this, one could calculate impacts on the proportion of logically inconsistent answers at post-test to assess whether the intervention might have affected the accuracy of survey responses.

As previously noted, the main text in the analysis plan and subsequent reports should present full impact results under the benchmark (preferred) assumptions and methods. The sensitivity analyses should be presented in appendix tables (or be made available upon request), to convince readers that the benchmark impact results are robust to different, equally valid methods.

REFERENCES

- Benjamini, Y., and Y. Hochberg. “Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing.” *Journal of the Royal Statistical Society, Series B (Methodological)*, vol. 57, no. 1, 1995, pp. 289–300.
- Bonferroni, C.E. “Il calcolo delle assicurazioni su gruppi di teste.” In *Studi in onore del Professore Salvatore Ortu Carboni* (pp. 13–16). Rome.
- Colman, C., “Estimating Program Impacts for a Subgroup Defined by Post-Intervention Behavior: Why is it a Problem? What is the Solution?” 2012.
- Dunnnett C.W. “A Multiple Comparison Procedure for Comparing Several Treatments with a Control.” *Journal of the American Statistical Association*, vol. 50, 1955, pp. 1096–1121.
- Freedman, D. “On Regression Adjustments to Experimental Data.” *Advances in Applied Mathematics*, col. 40, 2008, pp. 180-193.
- Puma, Michael J., Robert B. Olsen, Stephen H. Bell, and Cristofer Price. “What to Do When Data Are Missing in Group Randomized Controlled Trials.” NCEE 2009-0049. Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education, 2009.
- Schochet, Peter Z. “The Estimation of Average Treatment Effects for Clustered RCTs of Education Interventions.” NCEE 2009-0061 rev. Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education, 2009.
- Schochet, Peter Z. “Is Regression Adjustment Supported by the Neyman Model for Causal Inference?” *Journal of Statistical Planning and Inference*, vol. 140, January 2010, pp. 246–259.
- Tukey, J.W. “The Problem of Multiple Comparisons.” Unpublished manuscript. Princeton, NJ: Princeton University, 1953.
- Williams, R.L. “A Note on Robust Variance Estimation for Cluster-Correlated Data.” *Biometrics*, vol. 56, 2000, pp. 645–646.

TTP Eval TA: Analysis Tip Sheet

This document provides a number of tips and suggestions for what to consider at the data preparation and analysis stages of an evaluation in order to ensure that the observed evaluation findings are credible, robust, and appropriately interpreted.

TPP Eval TA: Analysis Tip Sheet

This document is intended to provide a number of tips/suggestion for considerations at the data preparation and analysis stages of an evaluation. The TPP Eval TA Team recommends reviewing this list of tips to make sure that the evaluation findings you observe are credible, robust, and appropriately interpreted. For technical guidance, please review the Analysis Plan FAQ and other documents linked below.

Data preparation:

- Check the descriptive statistics for all key variables of interest. Examine Ns, means, standard deviations, minimum and maximum scores numerically and graphically. Make sure that the data for each variable has face validity (e.g. that there are no out-of-range scores, or that the scores seem reasonable).
- For most grantees, the dataset that will be used for analysis will include a single observation for each individual. This dataset will include variables for the outcome measures (the dependent variable for most analyses), variables for the treatment indicator (the key predictor variable for most analyses), variables for demographics and baseline measures of risk (baseline covariates necessary for statistical adjustments and for improving the precision of impact estimates), and any necessary indicators for stratification or clustering used in the design of the study.
- Create multiple versions of the analytic dataset that can be used to assess the sensitivity of findings to different ways of handling inconsistent survey responses (e.g. a student reports having had sex at baseline and then at follow-up reports never having engaged in sex). For example, consider these options for creation of datasets:
 - Inconsistent survey responses recoded as missing
 - Inconsistent survey responses recoded to use first response in the given instrument as the “true” response.
 - Inconsistent survey responses left untouched (i.e. all inconsistent data represent random, and not systematic errors).
- Create multiple versions of the analysis dataset that can be used to assess the sensitivity of findings to different ways of handling missing data (see the Missing Data brief for suggestions on how to handle datasets with missing data). For example, consider these options:

TPP Eval TA: Analysis Tip Sheet

- Complete data: A dataset where there are no missing records for any key variables (outcomes, baseline measures of interest, demographic characteristics, etc.). That is, this dataset eliminates any observations where any key variables are missing.
- Complete outcome data: A dataset where there are no missing records for any outcome data, but there may be some missing data for some baseline measures. In this dataset, a simple approach for imputing missing data is the “Missing Value Dummy and Impute to Constant Value” procedure described in the brief above.
- All observed data: The raw dataset that contains all observed data, which may contain missing baseline and outcome scores. For this type of dataset to be used, a more complicated analytic procedure (e.g. multiple imputation or maximum likelihood estimation) may be required for analysis.

Reviewing initial descriptive statistics at baseline and follow-up:

- Examine baseline measures and demographic characteristics of a given dataset described above across intervention and comparison arms. Specifically, calculate means and proportions for intervention and comparison groups, to get a sense of the equivalence of these groups. Be sure to note any significant or substantive differences in the samples and plan on adjusting for them in the impact estimation.
- Compare follow-up survey responses across intervention and comparison arms, using the same datasets. Again, get a rough sense of the crude impact of the intervention, without any fancy statistical adjustments.
- Examine correlation coefficients among baseline and outcome scores. Make sure that the correlation coefficients have face validity (i.e. aren't too high or too low – which might suggest a data preparation error).

Conducting inferential analyses:

- We recommend using the linear probability model for ease of estimation and interpretation when the outcome is dichotomous (e.g. “ever had sex”, or “engaged in risky sexual activity”). Of course, estimating impacts using a logit or probit model is fine too – however, it may be more difficult to produce

TPP Eval TA: Analysis Tip Sheet

- interpretable program impacts or the appropriate model-based prevalence rates from such analyses.
- When the linear probability model is used as the method for dichotomous outcomes, the beta parameter for the treatment coefficient represents the difference in prevalence rates across the intervention and control conditions (e.g. a $-.07$ parameter estimate for an analysis on a “risky sexual activity” outcome would imply that the intervention group engaged in this activity 7 percentage points less than the comparison group).
 - When calculating impacts and standard errors, the analytic model should align with the design of the study.
 - If the study conducted random assignment using stratification/blocking variables, we recommend including strata/blocks in the analytic model as fixed effects. For example, if students were randomly assigned to condition separately in each of 4 schools (i.e. schools served as their own strata), then the analytic model could include school dummy variables as fixed effects.
 - If the study conducted random assignment where clusters, not individual, were randomly assigned to condition, then the analytic approach must adjust for clustering in order to obtain the correct standard errors. See the Clustering FAQ for suggestions (common approaches are using Huber-White clustered standard errors, estimating the model using a random effects model, etc.)
 - We recommend using an analytic approach for baseline equivalence analysis that is similar to the approach used for impact analysis.
 - For the analysis of baseline equivalence, we recommend regressing the variable of interest (e.g. age, gender, baseline assessment of the sexual activity variable) on a treatment indicator, any strata/blocks used in the design, and adjusting for clustering if necessary. Most statistical packages can provide adjusted means (or proportions) for the intervention and comparison groups, which will allow for the most intuitive way to show the equivalence of the groups at baseline.
 - For the analysis of program impacts, we recommend regressing the outcome of interest on a treatment indicator, any strata/blocks used in the design, a set of covariates (e.g. gender, age, race, baseline measures

TPP Eval TA: Analysis Tip Sheet

of sexual activity) and adjusting for clustering if necessary. Most statistical packages can provide adjusted means (or proportions) for the intervention and comparison groups, which will allow for the most intuitive way to show the impact of the program at follow-up.

- In order to receive the highest rating under the HHS Evidence Review, impact analyses must adjust for statistically significant differences in sexual behavior measures, as well as statistically significant demographic difference in the analytic sample. As such, we recommend that variables that are significantly different at baseline be adjusted for in the impact analytic model.
- When estimating impacts across a number of outcomes, we recommend using a common set of covariates across models. Doing so helps to convey to the reader that a systematic approach was used to estimate impacts across outcomes, rather than signaling that a different analytic procedure was used in each model.

Attrition scenarios:

- We discussed a variety of attrition scenarios for cluster RCTs (e.g. high cluster attrition with low subcluster attrition, or low cluster attrition with high subcluster attrition), and how they should be handled in terms of the final report. The recording of the first call describes suggested approaches, and this information is also presented in the forthcoming research brief on sample attrition, to be released in the very near future.

Multiple comparisons:

- The Analysis Plan FAQ provides suggestions on how to implement both the Bonferroni (1935) and the Benjamini-Hochberg (1995) multiple comparison adjustment procedures. This document also provides references to two other multiple comparison procedures that may be useful: Tukey (1953) or Dunnett (1955).

Analysis of site-specific impacts, as well as subgroup analyses:

- Some evaluations may attempt to estimate site-specific impacts, or impacts for particular subgroups. The TPP Eval TA team has two cautions about these types of analyses:
 - In many cases, examining impacts for sites or subgroups should be presented as exploratory, not confirmatory impact analyses. These

TPP Eval TA: Analysis Tip Sheet

analyses may be picking up on chance impacts, depending on the number of sites or subgroups examined.

- In addition, particularly when the overall impact is zero, we recommend proceeding with caution in examining subgroup impacts. For example, consider a study that shows no impact for the full sample, but presents information that suggests that the program is especially effective for female participants. A savvy reader will worry that in order for the program to have a zero overall impact for the full sample and a positive impact for female participants, this would imply that the program is actually harmful for male participants. As such, we recommend that if subgroup impacts are to be examined, then results should be presented for all related subgroups (e.g. males and females).