Summative evaluation approaches for rigorously testing a promising, well-defined intervention

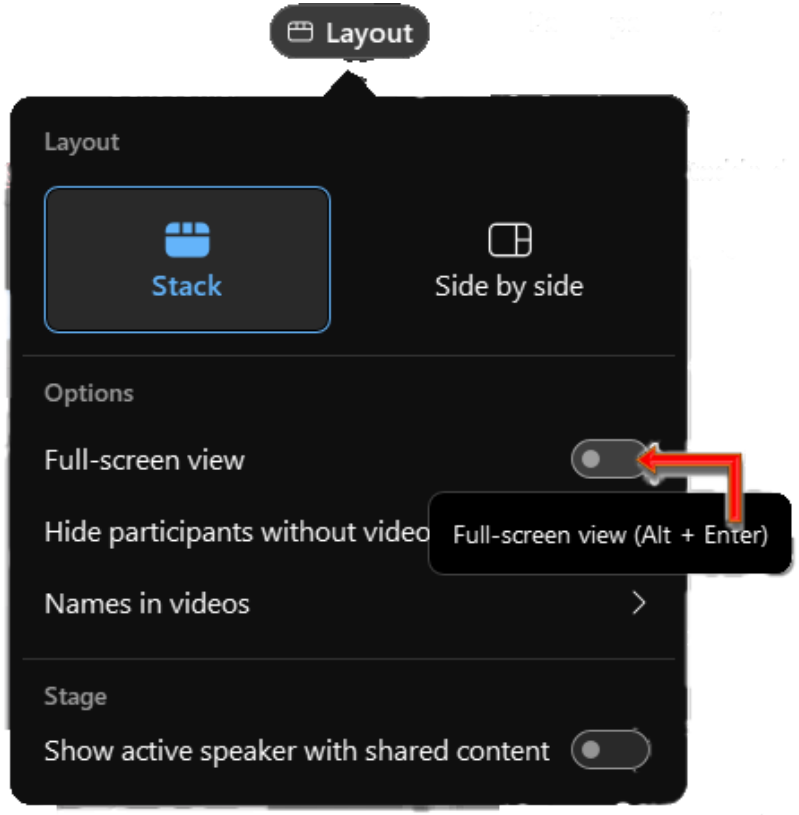
March 27, 2023

- Tara Rice, Office of Population Affairs
- Lauren Scher, Mathematica

OASH | Office of Population Affairs

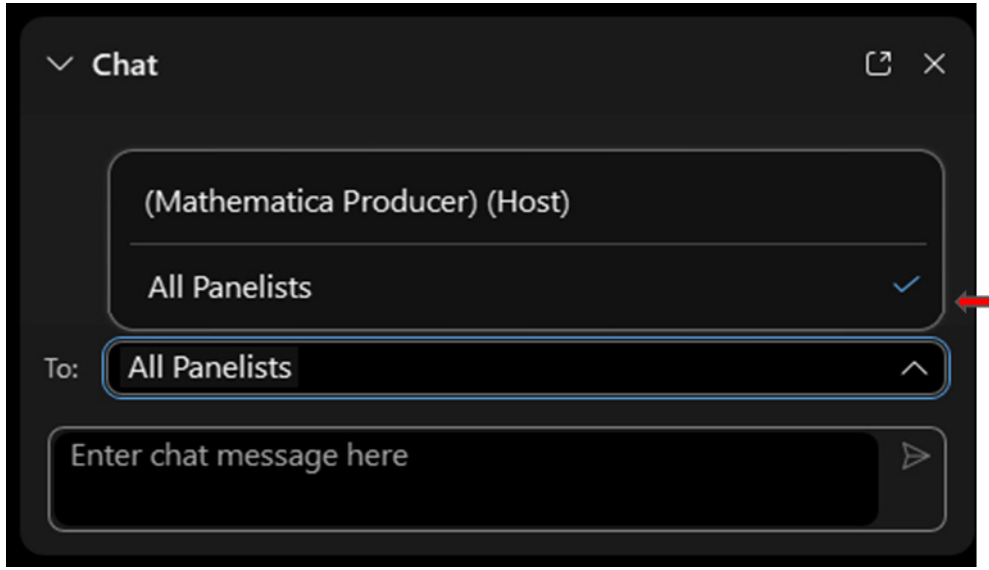
Event Window Adjustments

- All attendees are muted for the duration of today's webinar.
- To enter or exit full-screen mode, toggle the "Full-screen view" switch via the Layout button in the top right corner of your window.
- To change the size of the text and shared content on the screen, use the + and – buttons at the top of your screen.

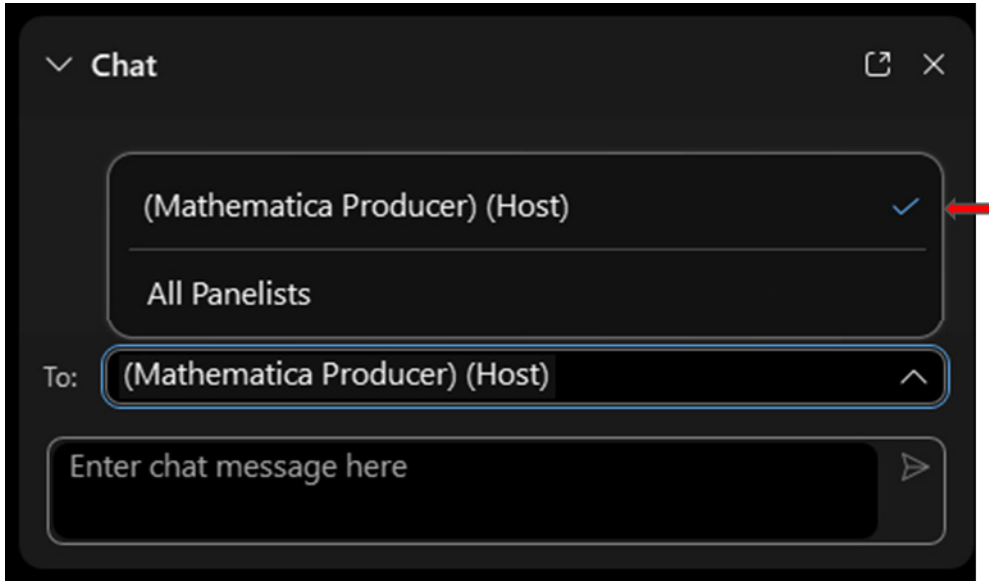


Q&A and Technical Assistance

- To submit questions for presenters, please use the Chat. Select “All Panelists” from the “To:” dropdown list, type your message, and press enter.



- To request technical assistance, please select “Host” from the “To:” dropdown list before sending your message.

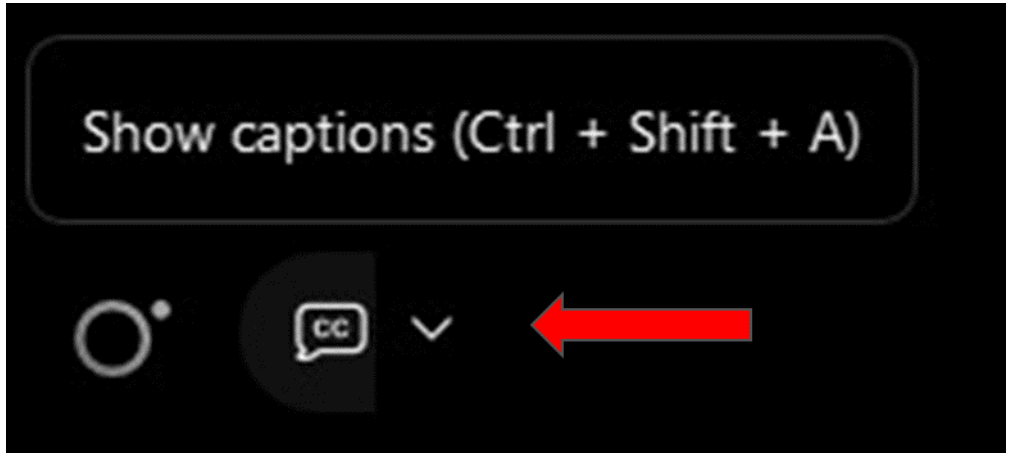


Event Materials

- Slides are available in the Webinar materials section of the meeting information page.

Closed Captioning

- Closed captioning may be turned on or off by clicking the CC icon in the lower left corner of your screen.



Welcome
From Tara Rice, OPA

Your presenter

Lauren Scher, senior researcher at Mathematica

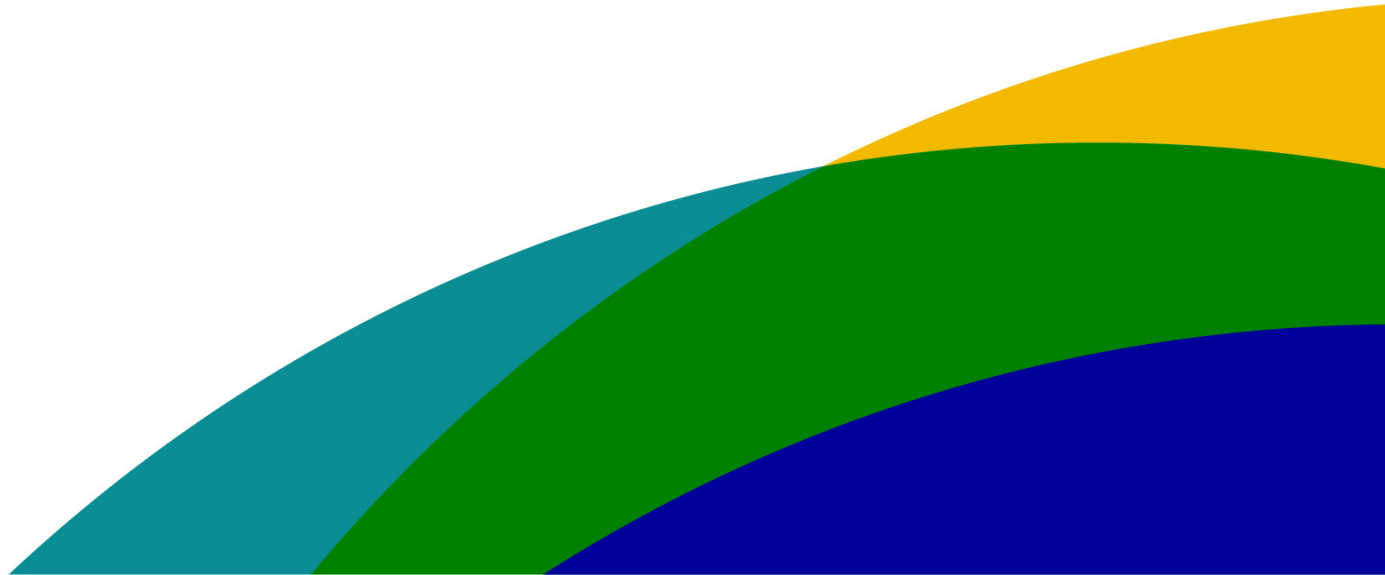
- Evaluation technical assistance (TA) provider for the Teen Pregnancy Prevention (TPP) program since 2010
- Current role: Direct eval TA for TPP 20 grantees
- Provide direct, individualized support to grantees
- Supervise TA liaisons and manage the development of products, toolkits, and group TA activities for TPP grantees and evaluators
- Direct an effort funded by the U.S. Department of Education to develop and disseminate practice guides to support evidence building and use
- Designed and implemented impact evaluations; conducted systematic evidence reviews

Agenda

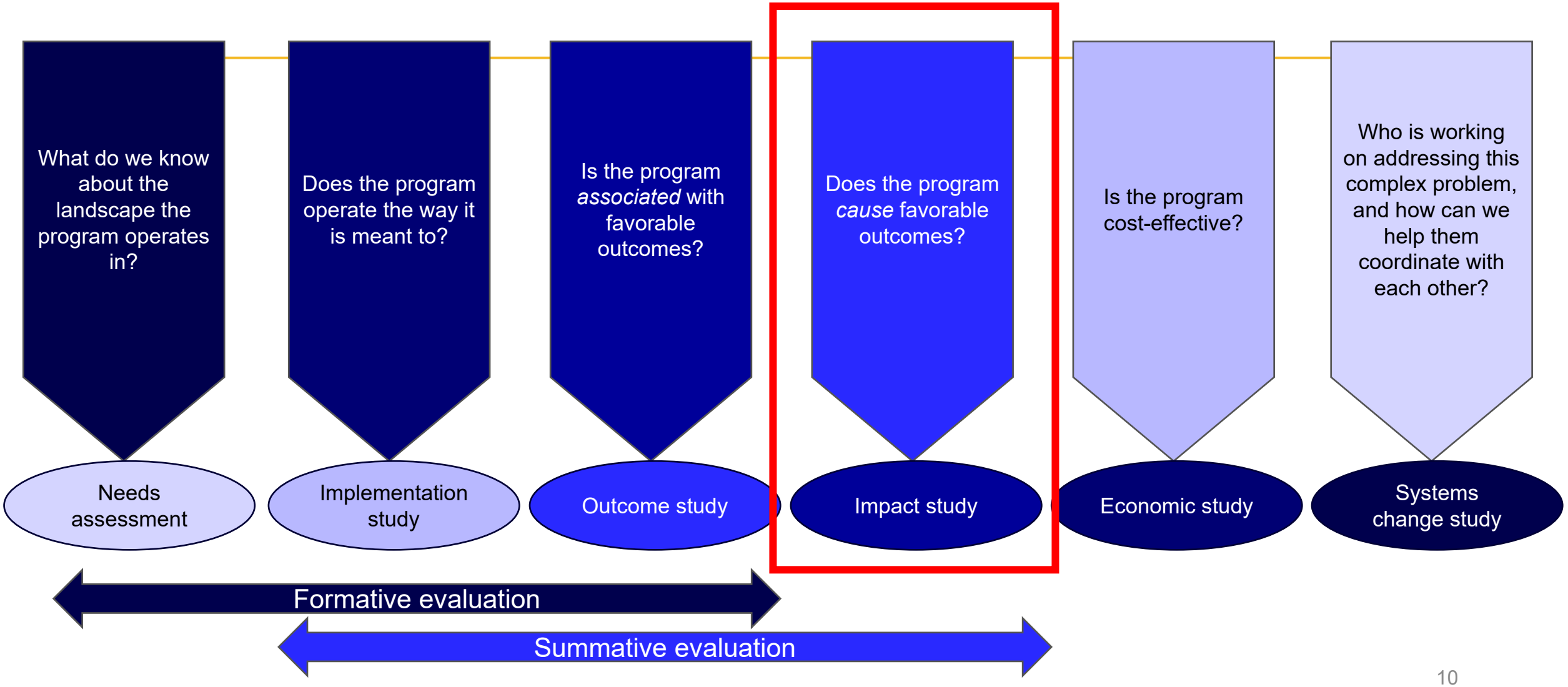
- Reminder: Continuum of evaluation activities
- Readiness for an effectiveness evaluation
- Design considerations for impact studies
- Analysis and reporting considerations (brief touch)

Today's Focus:

Testing the effectiveness of an intervention



Where are we on the continuum of evidence?



Why conduct an impact study?

- Produce evidence that an intervention is responsible for improving outcomes
  - **Testing a promising intervention that has not yet been tested rigorously**
  - **Assessing an intervention that has been tested rigorously but has been substantially adapted**
  - **Learning whether an intervention works for a different population or in a different setting**

When is an intervention ready for an impact evaluation?

- ✓ Environmental scan suggests need and shows community lacks similar programming
- ✓ Participants (or caregivers) say they are satisfied with programming
- ✓ Interested parties find it compelling and a good fit
- ✓ The innovation is clearly defined and implemented with fidelity and quality
- ✓ Near-term outcomes improved (and logic model suggests longer term behavioral outcomes will, too)
- ✓ Potential program partners are on board for an evaluation

Design considerations for rigorous evaluations

Different kinds of impact study designs

- Randomized controlled trials (RCTs) and quasi-experimental designs (QEDs) are common group designs in the TPP field
- To select a design, consider
  - **Level at which intervention is implemented (e.g., policy vs. individually provided)**
  - **Willingness of partners to participate in the activities the design calls for**

Planning an impact study with a group design

- ✓ Clearly specify research questions
- ✓ Choose a study design to answer research questions
- ✓ Prepare to demonstrate baseline equivalence
- ✓ Clarify the contrast you plan to test
- ✓ Design study to avoid confounding factors
- ✓ Determine an appropriate sample size
- ✓ Plan how to recruit an eligible sample
- ✓ Consider strategies for obtaining consent
- ✓ Specify data collection and analysis procedures
- ✓ Plan how to disseminate to a range of stakeholders

Clearly specify research questions

- Research questions should focus on estimating impacts (a difference in outcomes across conditions)
- Outcomes should align with logic model
- Specify primary and secondary outcomes, time frame, and population of interest
- Be realistic about what you can test in a specified time period (such as near-term outcomes versus distal outcomes)

- EXAMPLE research question: “What is the impact of ‘INNOV’ on sexual initiation rates six months after program end for females ages 12-15 attending middle school?”

Choose a study design to answer research questions

- A well-implemented RCT will provide the most credible estimates of effectiveness
- Questions to consider in determining whether an RCT is feasible
  - **Will there be fewer available slots than eligible participants?**
  - **Can individuals be randomly assigned? Classrooms? Clinics? CBOs? Schools?**
  - **Are partners on board with randomization? Communication is key!!!!**
- If random assignment is not possible, consider how to identify a QED comparison sample that is as similar as possible to the intervention group

Examples of QED methods

- Selecting a well-matched comparison group (e.g., schools/ clinics) based on measured characteristics (like school type or size, race/ethnicity, financial hardship rates)
- Using statistical methods such as propensity score matching or weighting to ensure equivalence on observed outcomes
- Problematic QED designs:
  - **Benchmarking to national-, state- or county-level data**
  - **Focusing solely on comparability of comparison group (without considering similarity of available services)**

Design considerations in an impact study

- Prepare to demonstrate baseline equivalence
  - **Pre-specified characteristics (such as age, race, ethnicity, gender)**
  - **Baseline measures of the outcome(s) of interest**
- Clarify the contrast you plan to test
  - **Make sure you are offering the intervention in a setting that is different enough from the business-as-usual condition**
  - **Document the effective contrast in services being tested**
- Design study to avoid confounding factors (e.g., one comparison cluster, different data collection procedures)
- Determine an appropriate sample size

Determine an appropriate sample size

- Be realistic and conservative
  - **Use information from pilot studies to develop assumptions about expected recruitment, sample attrition rates, and expected impacts**
  - **Account for planning, follow-up data collection, analysis, and reporting when determining the enrollment window for your study**
- Conduct power analysis that takes into consideration:
  - **Clustered nature of the design (if applicable)**
  - **Strength of the contrast between intervention and comparison groups**
  - **The expected size of the minimum detectable effect given the anticipated sample size**
  - **Existing literature examining similar interventions**
  - **There are tools and resources you can use to conduct power analyses (linked at the end of this presentation)**

Planning for implementing a rigorous evaluation

- Consider strategies for obtaining consent
  - **Consider Institutional Review Board requirements and partners' comfort with consent and assent**
  - **Develop clear procedures and protocols**
- Specify data collection procedures
  - **Determine what data you will need to collect versus what might be available**
  - **Develop plans to create, pilot, and refine data collection instruments**
  - **Consider challenges of collecting data from hard-to-reach populations**
  - **Design protocols and procedures for collecting implementation data**

Considerations for collecting implementation data

- Important implementation data to collect and examine in combination with conducting an impact study
  - **Fidelity**
  - **Dosage**
  - **Feedback from key interested parties (e.g., interviews, focus groups, surveys)**
  - **Intervention contrast**

Planning for analysis and reporting

- Consider analytic methods that will present credible estimates of program impacts
  - **Consult with Teen Pregnancy Prevention Evidence Review standards**
  - **Consider applying Bayesian methods for interpreting impacts**

- **Disseminating learnings throughout the lifespan of a study**
  - **To support understanding among partners and the community that is part of the study**
  - **To support program improvement**
  - **To support broader learning**

Resources

Topic	Link
Formative evaluation	Formative Evaluation Toolkit
IRB approval	Tip sheet
Implementation study	Reporting implementation findings How to conduct qualitative analysis
Impact study	Impact Evaluation Toolkit Study power calculator brief and companion Excel tool Evaluation start-up considerations Planning evaluations to meet scientific standards Using a Bayesian framework to interpret findings from impact evaluations Recruiting and retaining districts and schools Recruitment through social media
Economic study	Overview of economic methods
Systems change study	Webinar, Fact sheet, and TPP example study

Questions?