A Strong Foundation: Definitions of Terms for Teen Pregnancy Prevention Grant Recipients

Foundational terms in program innovation, implementation, and evaluation

Establishing a common set of program- and evaluation-related definitions for the Teen Pregnancy Prevention (TPP) field is central to effective communication and shared understanding within and between the Office of Population Affairs (OPA), its grant recipients and partners across funding tiers, and technical assistance providers. This set of definitions is meant to facilitate conversations within and across teams. Additional terms related to teen pregnancy prevention grants are available in the Notice of Funding Opportunities, TPP Evidence Review Standards, and on Connect. Teams may choose to define additional terms specific to their projects. The definitions are organized by program and evaluation type as defined in Box 1, although some terms apply in a variety of contexts, as shown in the index that immediately follows Box 1.

Box 1. Definitions in this document are organized by program and evaluation type:

- **TPP program** terms are relevant to recipients of all TPP funding streams.
- **Innovation** terms describe activities related to developing, testing, and refining new programs or practices. These terms are likely most relevant to OPA Tier 2 Innovation Hub recipients awarded in 2023 and their subrecipients.
- **Program implementation** terms include processes and features related to delivering a program. These terms are relevant to all OPA grant recipients awarded in 2023.
- **Formative evaluation** terms describe activities related to testing for feasibility, appropriateness, acceptability, desirability, and short-term effectiveness while programs are still under development. These terms are likely most relevant to Tier 2 Innovation Hub and Tier 1 grants.
- **Summative evaluation** terms describe activities related to implementation studies, outcome studies, and impact studies. These terms are likely most relevant to Tier 2 Rigorous Evaluation grants.

Index of defined terms and their relevant program and evaluation types

Terms	TPP program	Innovation	Program implementation	Formative evaluation	Summative evaluation
Accelerator		X			
Attrition					X
Baseline equivalence					X
Comparison or control group					X
Confounding factor				X	X
CONSORT diagram					X
Continuous quality improvement		X	X	X	
Core components	X	X	X	X	X
Counterfactual					X

Terms	TPP program	Innovation	Program implementation	Formative evaluation	Summative evaluation
Dosage	X	X	X	X	X
Effect size					X
Environmental scan		X	X	X	
Equitable evaluation		X		X	X
Evidence of effectiveness	X				X
Evidence-based programs	X		X		X
Fidelity		X	X	X	X
Fidelity monitoring system			X	X	X
Formative evaluation		X		X	
Hybrid		X			
Implementation context			X	X	X
Implementation quality			X	X	X
Incubator		X			
Innovation		X			
Innovation Development Teams		X			
Innovation Hub		X			
Institutional review board approval (or exemption)				X	X
Intent-to-treat principle					X
Intervention	X				
Logic model		X		X	X
Minimum detectable effect					X
Monitoring and improvement		X	X	X	
Needs assessment		X		X	X
Outcome evaluation (or pre-post design)				X	X
Performance measures	X	X	X	X	
Pivot		X			
Preliminary evidence		X		X	X
Project	X				
Prototype		X			
Quasi-experimental design					X
Randomized controlled trial					X
Rapid cycle learning			X	X	
Root cause analysis		X		X	
Statistical power					X
Summative evaluation					X
Systems thinking		X		X	
Systems-level evaluation		X		X	X
Theory of change		X	X	X	X
Treatment contrast					X
Treatment group					X

Definitions

TPP program terms *(relevant to all TPP tiers)*

Core components. The parts of a program or its implementation that are critical to the program's ability to produce outcomes. Developers might consider a component to be core because they hypothesize that components are related to outcomes based on existing theories or frameworks or because there is evidence from research that indicates the components influence program outcomes (Forrester and Cole 2022).

Evidence of effectiveness. A program or intervention is determined to have credible evidence of effectiveness based on standard elements like the quality and execution of the research design, the equivalence of treatment and comparison groups, the adequacy of sample sizes, the validity and reliability of outcome measures, and the appropriateness of statistical analyses and reporting (United States Government Accountability Office 2009). The Teen Pregnancy Prevention Evidence Review (TPPER) updates and applies evidence standards to reviews of new publications that regularly report effects from randomized controlled trials (RCTs) and quasi-experimental designs (QEDs). Programs with evidence of effectiveness are those that have been shown in a credible impact evaluation to favorably (and statistically significantly) improve sexual behavior outcomes, without having any adverse impacts. Evidence of effectiveness is different from preliminary evidence.

Intervention.* Innovative, promising, or evidence-based programs, models, components, curricula, products, approaches, and strategies implemented or evaluated by the TPP grant recipient; what the treatment group of an impact study receives.

Performance measures.* A set of measures of program implementation that OPA requires each grant type to use as a tool for program monitoring. Examples of performance measures include the reach of an evidence-based program, average program attendance rates, observer-reported overall quality, and program fidelity as reported by facilitators.

Project. All of the interventions or innovations and strategies (for instance, referrals) a grant recipient is implementing as part of a TPP grant.

Innovation terms *(likely most relevant to OPA Tier 2 Innovation Hub grants and their subrecipients)*

Accelerator.* Accelerators are a type of Innovation Hub that supports cohorts of Innovation Development Teams (IDTs) as they prepare their innovations for evaluation, dissemination, and new funding opportunities. IDTs that are supported by accelerators must have existing prototypes with preliminary evidence indicating their promise. Accelerators primarily support IDTs in conducting early summative evaluations to assess and disseminate information about an innovation's preliminary evidence of

---

* Indicates definition is based on the 2023 TPP Notice of Funding Opportunities (AH-TP2-23-001; AH-TP1-23-001; AH-TP2-23-002)

success. This includes documenting its functioning, fostering its refinement, and potentially making a case that the innovation would benefit from a pivot, is ready for dissemination, or is ready for rigorous impact evaluation.

Hybrid.* Hybrids are a type of Innovation Hub that support cohorts of IDTs through both the incubator and accelerator phases.

Incubator.* Incubators are a type of Innovation Hub that supports cohorts of IDTs in exploring gaps in the fields of adolescent sexual and reproductive health, positive youth development, and health equity and user needs, and using participatory methods to develop innovations that address those gaps. A key goal for Incubators is to support IDTs as they develop prototype innovations. In addition, Incubators support their IDTs as they focus on formative evaluation activity—encompassing a needs assessment and development and subsequent refinement of prototype innovations. Incubator IDTs also assess whether an innovation is feasible, appropriate, desirable, and acceptable, and can assess innovation functioning and some short-term outcomes while the innovation is being developed and improved.

Innovation.* For the TPP program, innovation encompasses the broad spectrum of new or adapted programs, strategies, approaches, interventions, policies, practices, and products designed to help teenagers avoid unintended pregnancy, to prevent sexually transmitted infections (STIs), and to promote positive youth development.

Innovation Development Teams (IDTs).* Identified and supported by Innovation Hubs, IDTs are made up of diverse individuals including those with lived experience of the problem to be addressed by the innovation they are developing. IDTs are identified according to standards and processes set by the Hub. They are usually small teams of three to five people.

Innovation Hub.* Adolescent sexual health Innovation Hubs are intermediaries—funded as grant recipients through TPP innovation funding—that identify and support cohorts of Innovation Development Teams (IDTs) that generate and test their own unique innovations. Hubs can take three forms: accelerator, hybrid, or incubator.

Pivot.* Deliberate changes in an innovation or changes in the approach to developing an innovation. A pivot is the result of identifying a problem, selecting a course of action to address the problem, and executing and monitoring the change. Pivots may be small or large in nature, and may include dropping or overhauling an innovation. Hubs and IDTs are expected to implement and share what they learn from pivots.

Preliminary evidence.* Findings from studies such as formative or implementation evaluations that show the promise of an innovation. The findings might show that a program addresses a known need, or that it was acceptable to the intended population, or that it was implemented as expected, or that outcomes of interest changed during program implementation. Preliminary evidence differs from evidence of effectiveness.

Prototype.* A model built to test a concept with users. A prototype helps users understand, explore, and communicate what it feels like to engage with the innovation in real working conditions and not theoretical conditions. Prototypes can be used to test and refine innovations. Hubs may have different expectations for what a prototype entails depending on which phase of innovation an IDT is working on. Incubators may look for prototypes that are under development. Accelerators may look for prototypes that have already gone through one or more rounds of testing.

Program implementation terms *(likely most relevant to OPA Tier 1 and Tier 2 RI grants)*

Dosage. The amount of an intervention the participant should optimally receive, including the duration, frequency, and intensity of a program (Forrester and Cole 2022). For example, the optimal intended dosage of a program might be 10 sessions lasting 45 minutes each, twice a week for a total of five weeks. Attendance records are a common source of information about the dosage participants actually received.

Evidence-based programs.* Programs that have been proven through rigorous evaluation to be effective in reducing rates of teenage pregnancy, behavioral risk factors that can lead to unintended teenage pregnancy, or other associated risk factors. The Teen Pregnancy Prevention Evidence Review (TPPER; refer to Box 2) lists evidence-based TPP programs on its website.

Fidelity.* The degree to which an organization implements a program or curriculum as the developer intended (including implementing core components, and not using any unallowable adaptations). Implementation with fidelity can be a key factor in a program's success and whether it favorably moves youth outcomes (Keating 2020).

> **Box 2. Teen Pregnancy Prevention Evidence Review**
>
> The TPPER is a systematic review of published manuscripts that present impact study findings about TPP programs and components. This systematic review is a tool to help policymakers, practitioners, and other decision makers identify evidence-based TPP programs. At the federal level, the U.S. Department of Health and Human Services has used the findings in part to determine eligibility for federal grant funding for TPP programs. The review findings are also intended as a broader resource for states and local communities.

Fidelity monitoring system. A system of measuring and analyzing the degree to which a program is implemented as the developer intended (Keating 2020). Often, fidelity monitoring tracks and documents features of implementation quality, use of adaptations (planned or unplanned), and aspects of dosage. It may include independent observations of a subset of sessions to get the perspectives of individuals who are not the study participants.

Implementation context. The circumstances that form the setting for the delivery of a program, service, or intervention (adapted from Nilsen et al. [2019]). This may include features such as the attitudes toward teen sexual and reproductive health held by administrators in the school where are a new program is being implemented.

Implementation quality. The extent to which a program, service, or intervention is delivered with fidelity, including delivering the expected material and actively engaging participants in the activities and

discussions. Setting benchmarks to represent high quality implementation is recommended, particularly for evaluation studies.

Formative evaluation terms *(likely most relevant to Tier 2 Hub and Tier 1 grants)*

Continuous quality improvement. The systematic process of identifying, describing, and analyzing problems with aspects of program implementation (or process), and then testing, learning from, and revising potential solutions to those problems. Continuous quality improvement (CQI) is the label for this ongoing cycle of collecting data and using them to make decisions to gradually improve program processes (OPA, 2016).

Environmental scan. The process of systematically gathering, reviewing, and interpreting data across many sources to identify issues and opportunities for a particular type of activity or approach related to designing or implementing a TPP program (adapted from CMS n.d.[b]). Environmental scans may involve conducting literature searches, interviews, and focus groups to understand the context an organization operates in and the needs of a population.

Formative evaluation.* An evaluation approach that assesses whether a program (or practice, innovation, etc.) is feasible, appropriate, acceptable, and desirable. The goal of a formative evaluation is to improve and refine a program, and it is an appropriate process while the program is still being developed and finalized.

Logic model.* A visual that describes how a program or practice is expected to affect outcomes of interest. A logic model details the program's inputs (for example, the characteristics of the population or the context where it is to be offered), activities (for example, motivational interviewing or using an app), outputs (for example, participation rates or program completion), and short- and long-term outcomes (for example, condom use or STI rates). A logic model articulates what is needed to deliver an intervention and the changes that should be seen as a result. A logic model is related to a theory of change in that a theory of change explains why and how the inputs and outputs in the logic model relate to each other.

Monitoring and improvement.* The process by which grant recipients use performance measures and other relevant data, including feedback from youth and stakeholders, to monitor progress in meeting approved project goals and objectives. Monitoring and improvement is part of the CQI process (Ruiz and Adamek 2023).

Needs assessment* (including ongoing needs assessment). A process of collecting and analyzing information to determine whether a defined population or service area requires certain services, and how much of those services they need (National Family Planning & Reproductive Health Association 2017). A needs assessment may follow an environmental scan and be focused on the services identified in the environmental scan.

Rapid cycle learning. An evaluation approach that quickly tests a new innovation or intervention, potential changes to an intervention or a component of an intervention, quickly extracts what was learned

through the test to refine the innovation or intervention, and tests the refined version (Baumgartner and Eddins 2020).

Root cause analysis. A structured team process used to uncover underlying factors that set a cause-and-effect reaction in motion, leading to a problem(s) (ASQ n.d.; CMS n.d.[a]). For example, a root cause analysis might involve investigating a key problem, such as when clinicians avoid talking about contraception with teens, that acts as a roadblock to teens accessing sexual and reproductive health care. A root cause analysis would define this problem, gather data, identify causal factors, determine the root causes, and recommend potential solutions. A root cause analysis may incorporate human-centered design approaches and could be used to inform CQI or answer systems evaluation research questions.

Systems thinking.* An approach to problem-solving that considers the overall system instead of focusing on specific parts of the system in isolation. For example, when attempting to address disparities in sexual health care access between different populations of teens, systems thinking would involve investigating and addressing the upstream source of the disparities (such as the local policies, practices, and perceptions related to sexual health) rather than solely training the teens on ways to access health care.

Theory of change.* An ongoing process of reflection to explore changes caused by a TPP program, and how those changes unfold, and what that means for a particular program in a particular context, sector, and/or group of people. A theory of change is related to a logic model in that a theory of change explains why and how the inputs and outputs in the logic model relate to each other.

Summative evaluation terms *(likely most relevant to Tier 2 Rigorous Evaluation grants)*

Attrition. Loss of study sample from an impact analysis. Attrition results when an outcome variable is not available for all subjects initially assigned to the intervention and comparison groups of a randomized controlled trial (What Works Clearinghouse n.d.).

Baseline equivalence. A statistical demonstration of the similarity between individuals in a treatment group and comparison group before the introduction of a program. Baseline equivalence is often examined on demographic characteristics and baseline measures of the outcome of interest to show that these key measurable characteristics are balanced before the intervention starts. Under TPPER standards, both randomized controlled trials with high attrition and quasi-experimental designs must establish that the treatment and control groups used in the analysis were equivalent on observable characteristics at baseline to be eligible for the moderate evidence rating. (See the TPPER protocol for more guidance on this standard.)

Comparison group. The counterfactual condition in an RCT or QED, representing an absence of the intervention under evaluation. The comparison group may receive business-as-usual services or a different intervention that may or may not include similar content and dosage to the intervention of interest. Sometimes people use the phrases "comparison group" and "control group" as interchangeable

synonyms; however typically the term control group is reserved for use in the context of a randomized controlled trial.

Confounding factor. A feature of a study that is completely aligned with one of the study conditions. For example, a study may have one facilitator implementing a program to all health classes assigned to the treatment group condition and a different facilitator implementing business-as-usual services to all health classes assigned to the control condition. In this case, it is impossible to separate how much of the observed difference in effects between the two conditions was due to the particular facilitator delivering the program and how much was due to the program itself. (What Works Clearinghouse n.d.).

Consolidated Standards of Reporting Trials (CONSORT) diagram. A flow diagram of the sample loss through the phases (that is, enrollment, random assignment, baseline, follow-up) of a randomized controlled trial of two (or more) groups done in parallel (Schulz et al. 2010). The information provided in a CONSORT diagram can be used to calculate sample attrition.

Counterfactual.* The counterfactual in the context of an experiment represents what would have happened if the program being tested were not available. The counterfactual is typically operationalized through the use of a well-matched comparison group; by comparing the outcomes of a treatment group relative to this comparison group as a counterfactual, it is possible to estimate the effect of a program.

Effect size. The difference in average outcomes across a treatment and comparison group. Effect size differences can be unstandardized, and represent a difference in the unit of the outcome measure (for example a percentage point difference in prevalence rates for sexual initiation across conditions). Alternately, standardized effect sizes are rescaled so they represent differences in standard deviations of the outcome (Moreno and Cole 2014)

Equitable evaluation.* An approach researchers use to evaluate programs in a way intended to benefit those who receive programmatic services. Equitable evaluations address power imbalances between researchers and participants, engage community partners, reflect assumptions and biases, consider different worldviews, ensure that evaluation methods are multiculturally valid and oriented toward participants, strive for intentionality in data collection and analysis, and aim for accessible and actionable findings (Parekh et al. 2023). Equitable evaluations are intended to yield incremental changes that contribute to dismantling systems or processes that hold inequities in place.

Impact evaluation. An evaluation approach designed to provide evidence about the effect of a program. An impact evaluation compares the outcomes of a treatment group with the outcomes of a comparison group, and attributes the differences in the outcomes to the program being evaluated. RCTs and QEDs are study designs commonly used in impact evaluations.

Institutional review board (IRB) approval (or exemption). IRBs review study plans to ensure the study protects the welfare of human research subjects. Researchers must obtain approval from a relevant IRB before beginning a study if their planned study meets the definition of research and is not eligible for an exemption. Examples of the reasons for exemption may include analyzing data that are already collected

in schools through normal education practices. An IRB needs to determine if a study qualifies for exemption, not the individual researcher. If a study does not meet the definition of research, it does not require an exemption (and thus does not need to be submitted to the IRB [Office for Human Research Protections 2018]).

Intent-to-treat principle.* A framework in which the impact analysis is conducted based on the conditions individuals were originally assigned to. In other words, the treatment group in the analysis consists of all participants who were assigned to receive the program, regardless of whether they actually did. Likewise, the comparison group in the analysis contains all study participants who were assigned to the comparison group, even if some managed to take up the treatment.

Minimum detectable effect (MDE). A calculation of the smallest true treatment impact that the study is likely to detect as statistically significant, measured in terms of effect size (Moreno and Cole 2014). The MDE is largely a function of the study's sample size, desired statistical power (often assumed to be 80 percent) and the levels selected for describing findings as statistically significant (often $p < 0.05$, two--tailed).

Outcome evaluation (or pre-post design). A study design that quantifies how participants' outcomes change over the course of a study, a pre-post design can document the change in outcomes between a period before programming (pre) and a follow-up period after programming (post). This design does not require a comparison group (Lee and Cole 2020). Because you don't know what would have happened in the absence of the program (i.e. the counterfactual), changes in outcomes cannot necessarily be attributed to the program. Note that all Tier 2 RI grant recipients are expected to conduct impact evaluations, not outcome evaluations.

Quasi-experimental design (QED).* An impact evaluation design that forms a comparison group by means other than random assignment. Unlike in an RCT, where equivalence on measurable and unmeasurable characteristics is achieved by design, in a QED, equivalence on unmeasurable characteristics is not ensured. To produce a credible estimate of a program's effect, a QED should demonstrate baseline equivalence on observed characteristics likely to influence the outcomes of interest (Adapted from What Works Clearinghouse n.d.).

Randomized controlled trial (RCT).* An experimental design that assigns program participants to one of at least two distinct groups at random: the treatment group, which receives program services, and the comparison group, which does not. The comparison group serves as the "counterfactual," or the condition in which the program or intervention is absent. Random assignment ensures the treatment and comparison groups are initially similar and do not differ on any measured or unmeasured characteristics. Random assignment thus creates an evaluation design in which any observed differences between the two groups after the program or intervention can be attributed to the program or intervention with a high degree of confidence, provided the study has low levels of sample attrition and no confounds.

Statistical power. The probability that a study will detect a given magnitude impact as statistically significant, given the observed sample size. A statistical power analysis can estimate the sample size

necessary for an impact estimate of a given magnitude reflecting the difference between treatment and comparison groups to be detected as statistically significant. As noted in the definition of minimum detectable effect, statistical power is related to MDE calculations.

Summative evaluation. Evaluation approach that assesses whether a promising, well-defined intervention works. Summative evaluation approaches are appropriate when programs (or practices, innovations, etc.) are well defined and have some early indications of promise, often based on findings from a formative evaluation. Implementation studies, outcome studies, and impact studies can all be types of summative evaluation (Rice and Scher 2023).

Systems-level evaluation. Evaluation approach that uses systems thinking to assess the conditions and dynamics that hold complex systems in place. Systems change evaluation is the assessment of the success of an attempt to intervene in those conditions and dynamics to alter complex systems, with the results including structural, relational, and transformational changes (Koleros and Forrester 2022). A systems-level evaluation might assess, for example, whether a required training for health care center administrators on destigmatizing sexual and reproductive health among adolescents is related to a change in frontline staff members' comfort with addressing use or non-use of contraceptives with adolescent patients.

Treatment contrast. In impact evaluations, the difference in the experiences (content, dosage, activities, etc.) of those in the treatment group and those in the comparison group (Hamilton and Scrivener 2018).

Treatment group.* The group assigned to receive the intervention in an evaluation. Sometimes referred to as the intervention group.

References

American Society for Quality (ASQ). "What Is Root Cause Analysis (RCA)?" n.d. https://.org/quality-resources/root-cause-analysis. Accessed October 10, 2023.

Baumgartner, S., and K. Eddins. "Using Rapid-Cycle Evaluation to Inform Policy Decision Making." *Mathematica Insights,* March 17, 2020. https://mathematica.org/blogs/using-rapid-cycle-evaluation-to-inform-policy-decision-making.

Centers for Medicare and Medicaid Services (CMS). "Five Whys Tool for Root Cause Analysis." n.d.(a).https://www.cms.gov/medicare/provider-enrollment-and certification/qapi/downloads/fivewhys.pdf. Accessed December 11, 2023.

CMS. "Measure Conceptualization." n.d.(b). https://mmshub.cms.gov/measure-lifecycle/measure-conceptualization/information-gathering/conduct-environment-scan. Accessed December 11, 2023.

Domitrovich, C. E., C. P. Bradshaw, J.M. Poduska, K. Hoagwood, J.A. Buckley, S. Olin, L. H., Romanelli, P. J. Leaf, M.T. Greenberg, and N.S. Ialongo. "Maximizing the Implementation Quality of Evidence-Based Preventive Interventions in Schools: A Conceptual Framework." *Advances in School Mental Health Promotion*, vol. 1, no. 3, 2008, pp. 6–28. https://doi.org/10.1080/1754730x.2008.9715730.

Forrester, E., and R. Cole. "Core Components of Teen Pregnancy Prevention Programs." OPA Brief. Office of Population Affairs, Office of the Assistant Secretary for Health, U.S. Department of Health and Human Services, 2022.

Hamilton, G., and S. Scrivener. "Measuring Treatment Contrast in Randomized Controlled Trials." MDRC, June 2018. https://www.mdrc.org/work/publications/measuring-treatment-contrast-randomized-controlled-trials.

Keating, B. "Fidelity Monitoring Tip Sheet." Office of Population Affairs, Office of the Assistant Secretary for Health, U.S. Department of Health and Human Services, 2020. https://opa.hhs.gov/sites/default/files/2021-05/fidelity-monitoring-tip-sheet-july-2020.pdf.

Koleros, D., and E. Forrester. "Introduction to System Change Evaluation in Teen Pregnancy Prevention." Office of Population Affairs, Office of the Assistant Secretary for Health, U.S. Department of Health and Human Services, 2022.

Lee, J., and R. Cole. "Pre-Post Outcome Study How-To Guide." Office of Population Affairs, Office of the Assistant Secretary for Health, U.S. Department of Health and Human Services, 2020.

Moreno, L., and R. Cole. "Calculating Minimum Detectable Impacts in Teen Pregnancy Prevention Impact Evaluations." Mathematica Policy Research, December 2017.

National Family Planning & Reproductive Health Association. "Crafting Your Best Needs Assessment." 2017. https://www.nationalfamilyplanning.org/file/Needs-Assessment_ResourceGuide.pdf.

Nilsen, P., and S. Bernhardsson. "Context Matters in Implementation Science: A Scoping Review of Determinant Frameworks That Describe Contextual Determinants for Implementation Outcomes." *BMC Health Services Research,* vol. 19, no. 189, 2019. https://doi.org/10.1186/s12913-019-4015-3.

Office for Human Research Protections. "Human Subjects Regulations Decision Charts: 2018 Requirements." U.S. Department of Health and Human Services, 2018. https://www.hhs.gov/ohrp/sites/default/files/human-subject-regulations-decision-charts-2018-requirements.pdf.

Office of Population Affairs (known as the Office of Adolescent Health at the time of publishing). "Continuous Quality Improvement Part 1: Basics for Pregnancy Assistance Fund Programs." U.S. Department of Health and Human Services, 2016. https://rhntc.org/sites/default/files/resources/oah_cqi_basics_resource_2016-09-20.pdf.

Office of Population Affairs. "Notice of Funding Opportunity: Teen Pregnancy Prevention Tier 2 Adolescent Sexual Health Innovation Hubs." U.S. Department of Health and Human Services, 2023.

Office of Population Affairs. "Notice of Funding Opportunity: Teen Pregnancy Prevention Tier 2 Rigorous Evaluation Cooperative Agreements." Office of the Assistant Secretary for Health, U.S. Department of Health and Human Services, 2023.

Office of Population Affairs. "TPP Performance Measures Snapshot." Office of the Assistant Secretary for Health, U.S. Department of Health and Human Services, 2023. https://opa.hhs.gov/sites/default/files/2023-03/TPP-performance-measures-2021-2022.pdf.

Parekh, J., S. Ciaravino, K. Welti, C. Ragonese, L. Lapointe, J. Manlove, J. Gilbertsen, and R. Katz. "Equity-Focused Strategies in a Federally Funded Evaluation of a Sexual Health Program." Child Trends, 2023. https://doi.org/10.56417/4593q1135m.

Rice, T., and J. Knab. "Evaluation 101." Webinar, February 22, 2023. https://rhntc.org/resources/evaluation-101-webinar.

Rice, T., and L. Scher. "Summative Evaluation Approaches for Rigorously Testing a Promising, Well-Defined Intervention." Webinar, March 27, 2023. https://rhntc.org/resources/summative-evaluation-approaches-rigorously-testing-promising-well-defined-intervention.

Ruiz, J., and K. Adamek. "Monitoring and Improvement Approaches Appropriate When Implementing a Teen Pregnancy Prevention Program." Webinar, February 23, 2023. https://rhntc.org/resources/monitoring-and-improvement-approaches-appropriate-when-implementing-teen-pregnancy.

Schulz, K.F., D.G. Altman, D. Moher, and the CONSORT Group. "CONSORT 2010 Statement: Updated Guidelines for Reporting Parallel Group Randomised Trials." *BMC Medicine*, vol. 8, no. 18, 2010. Published March 24, 2010. doi:10.1186/1741-7015-8-18.

United States Government Accountability Office. "Program Evaluation: A Variety of Rigorous Methods Can Help Identify Effective Interventions." 2009. https://www.gao.gov/assets/gao-10-30.pdf.

Vardatira, S. "The Path to Program Sustainability Webinar Series: Part 2: Five Steps to Assess the Environment." Webinar, June 27, 2019. https://rhntc.org/resources/five-steps-assess-environment-tpp-programs-webinar-part-2.

What Works Clearinghouse. "Glossary of Terms." Institute of Education Sciences, U.S. Department of Education, n.d. https://ies.ed.gov/ncee/wwc/glossary.