

Using EVALUATION to Strengthen Practice

A Practical Toolkit for IPV/SV Programs

Prepared by Tara K. Prendergast, MCRP, and Rachel Wildfeuer, PhD
Center for Policy Research (CPR)

For the Tennessee Coalition to End Domestic and Sexual Violence (TCEDSV)



Table of Contents

HOW TO USE THIS TOOLKIT	2
PART I: UNDERSTANDING EVALUATION.....	4
WHY EVALUATION MATTERS.....	4
WHAT EVALUATION IS (AND IS NOT).....	4
TYPES OF EVALUATION	5
<i>Performance Monitoring</i>	5
<i>Process Evaluation</i>	5
<i>Outcome Evaluation</i>	5
A NOTE ON EVALUATION METHODOLOGY	5
BUILDING A CULTURE OF EVALUATIVE THINKING.....	6
PART II: DOING EVALUATION	7
STEP 1: CLARIFY YOUR PURPOSE AND ENGAGE STAKEHOLDERS.....	7
<i>Engaging Stakeholders</i>	8
STEP 2: DESCRIBE YOUR PROGRAM — THEORY OF CHANGE AND LOGIC MODEL	8
<i>Theory of Change</i>	8
<i>Logic Model</i>	8
STEP 3: DEVELOP YOUR EVALUATION QUESTIONS AND PLAN	9
<i>Developing Your Evaluation Questions</i>	9
<i>Right Sizing Your Evaluation Questions</i>	10
<i>Building Your Evaluation Plan</i>	10
<i>Outcomes and Indicators</i>	11
STEP 4: COLLECT DATA.....	11
<i>Quantitative Methods</i>	11
<i>Qualitative Methods</i>	12
<i>Primary vs. Secondary Data</i>	12
<i>A Reality Check on Sample Size</i>	12
<i>A Note on Participant Burden</i>	12
STEP 5: ANALYZE AND INTERPRET YOUR DATA.....	13
<i>Analyzing Quantitative Data</i>	13
<i>Analyzing Qualitative Data</i>	13
<i>Interpreting Your Findings</i>	13
STEP 6: SHARE AND USE YOUR FINDINGS.....	14
COMMON EVALUATION PITFALLS	14
PART III: EVALUATION WITH LIMITED RESOURCES	15
START WITH THE MOST CRITICAL QUESTIONS.....	15
MATCH YOUR RIGOR TO YOUR PURPOSE	15
MAKE MAXIMUM USE OF EXISTING DATA.....	15
USE FAST, LOW-BURDEN DATA COLLECTION METHODS.....	16
WORKING WITH SMALL NUMBERS	16
COLLABORATE WITH PARTNERS	17
PROTECT WHAT MATTERS MOST	17
PART V: APPENDIX OF TOOLS AND TIPS.....	18

This project is supported by the Centers for Disease Control and Prevention of the U.S. Department of Health and Human Services (HHS) as part of a financial assistance award NUS4CE002305. The contents are those of the author(s) and do not necessarily represent the official views of, nor an endorsement, by CDC/HHS, or the U.S. Government.

How to Use This Toolkit

This toolkit is designed for intimate partner violence (IPV) and sexual violence (SV) practitioners in Tennessee who are interested in strengthening programmatic decision making through gathering evaluation data without hiring an external evaluator. Whether you are new to evaluation or have done some before, this resource will help you think through the key decisions, develop a realistic plan, and collect and use data in ways that are manageable for your organization.

The toolkit is written primarily with primary prevention work in mind — programs and strategies aimed at stopping IPV/SV before it occurs. However, the core concepts and steps apply broadly to any IPV/SV program evaluation, including direct services, community education, and systems change work.

The toolkit is organized into four parts:

- **Part I: Understanding Evaluation** — core concepts and terminology you need before you start planning
- **Part II: Doing Evaluation** — a step-by-step guide to planning and conducting your evaluation
- **Part III: Evaluation with Limited Resources** — practical strategies for doing meaningful evaluation when time, money, and staff are scarce
- **Part IV: Appendix of Tip Sheets and Templates** – tools for carrying out each stage of the evaluation process

Materials draw on evaluation frameworks and resources from the Centers for Disease Control and Prevention (CDC), the W.K. Kellogg Foundation, the National Sexual Violence Resource Center (NSVRC), the Ohio Domestic Violence Network (ODVN), and others. Rather than reprinting any one framework, we have synthesized and adapted these resources into a practical approach suited to the realities of prevention work in Tennessee. A full reference list is included at the end in **Appendix A**.

Part I: Understanding Evaluation

Why Evaluation Matters

Evaluation is not just a funder requirement. Done well, it is one of the most useful tools you have for understanding whether your work is making a difference — and for making it better over time.

Primary prevention of IPV/SV is still a relatively young field. The evidence base is growing, but many prevention strategies — community norm change, bystander intervention, healthy relationship education — are still being studied and refined. That means the evaluations conducted by practitioners like you are genuinely valuable. When you document what you did, who you reached, and what changed, you are contributing to a broader understanding of what works in communities like yours.

At the program level, evaluation can help you:

- Measure the changes your prevention work is helping to create
- Identify what is working well and where adjustments are needed
- Tell the story of your program to funders and community partners with evidence, not just impressions
- Practice accountability to participants, community members, and partners
- Make strategic decisions when you need to shift plans
- Determine whether you are reaching the right people at the right time

Evaluation is most useful when it is treated as an ongoing learning process rather than a one-time report. The goal is not a perfect study — it is useful information that helps you and your organization make better decisions.

What Evaluation Is (and Is Not)

Evaluation is a systematic process of collecting and analyzing evidence to understand the value, quality, and effectiveness of a program or strategy. It is different from informal impressions or anecdotal feedback — though those things have their place too. Evaluation is purposeful, planned, and tied to specific questions you want to answer.

Program evaluation — the type most relevant to IPV/SV prevention workers — asks whether the programs and strategies you implement are meaningful to the people they are meant to serve, and whether they are achieving the outcomes you intend. Knowing that you delivered a curriculum as planned is useful, but evaluation goes further: Did participants actually gain new knowledge or skills? Did the program reach the people who needed it most? Were the outcomes meaningful to them?

A good evaluation is:

- **Useful:** it produces findings that your organization can actually act on

- **Feasible:** it works within your organization’s real capacity
- **Honest:** it reports what the data actually show, including limitations
- **Participatory:** it involves staff, partners, and when possible, participants in shaping what gets asked and how

Types of Evaluation

Not all evaluation questions are the same, and the type of evaluation you conduct depends on where your program is in its development and what you most need to know. There are three main types:

Performance Monitoring

Performance monitoring tracks whether you are doing what you said you would do. It is primarily about accountability — to funders, to leadership, and to your own implementation plan. It answers questions like: How many trainings did we deliver? How many participants did we reach? Are we on schedule? Most grant reporting requires at least this level of evaluation.

Process Evaluation

Process evaluation (sometimes called formative evaluation) examines how a program is being implemented. It goes beyond simply counting activities to ask whether the program is being delivered as intended, whether it is reaching the right people, and what is working or not working about implementation. Process evaluation is especially useful for newer programs or during early implementation when you are still refining your approach. It is well-suited for low-resource settings because it often relies on data you are already collecting.

Outcome Evaluation

Outcome evaluation (sometimes called summative evaluation) examines whether a program achieved the changes it intended — in participants’ knowledge, attitudes, skills, behaviors, or in community conditions. This is the type of evaluation most people imagine when they hear the word “evaluation.” It is also the most resource-intensive, and the strength of the conclusions you can draw depends heavily on your design and sample size. For most community-based prevention programs, a modest outcome evaluation — a pre/post survey with a convenience sample, for example — is realistic and valuable even if it does not meet the bar of rigorous research.

In practice, most prevention programs benefit from conducting all three types, even at different levels of depth. Performance monitoring and process evaluation are often ongoing; outcome evaluation may happen at key points in the program cycle.

A Note on Evaluation Methodology

Methodology refers to the research design and data collection approach you use to answer your evaluation questions. You do not need to be a researcher to understand the basic options — and knowing them helps you make realistic choices about what your evaluation can and cannot claim.

Common approaches include:

- **Pre/post design (quasi-experimental):** Participants complete a survey or assessment before and after a program. This is one of the most practical designs for community-based work. It shows change over time but does not definitively prove your program caused the change.
- **Post-only design:** A survey or assessment administered only after the program. Easier to implement but limited in what it can show — you cannot assess change without a baseline.
- **Case study:** An in-depth look at a single program, site, or community. Useful for understanding how and why something worked in a specific context. Not generalizable to other settings, but rich in detail.
- **Randomized controlled trial (RCT):** Participants are randomly assigned to receive the program or not. Considered the gold standard for establishing causation, but rarely feasible for community-based IPV/SV prevention due to ethical and practical constraints.
- **Outcome mapping:** Focuses on changes in the behavior and actions of the people and organizations your program works with. Particularly useful for community-level or systems change work where outcomes are complex and diffuse.

The right methodology depends on your evaluation questions, your resources, and the stage of your program. You do not have to choose just one approach — many evaluations use a combination. Part II will help you think through what is realistic for your situation.

Building a Culture of Evaluative Thinking

Evaluation is most powerful when it is not just a project you do once a year — but a way of thinking embedded in how your organization works. Evaluative thinking means systematically asking what you expect to happen and why, collecting evidence to check those expectations, and reflecting on what you learn to make better decisions.

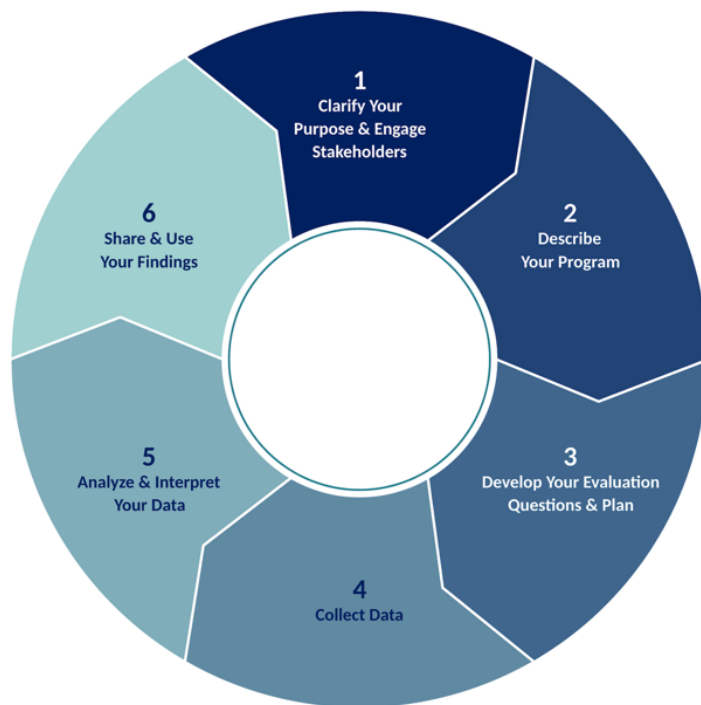
Some practical ways to build evaluative thinking into your organization's culture:

- Set aside time at staff or team meetings to discuss data — not just what the numbers say, but what you think they mean and what you should do differently
- Designate someone (even part-time) to be responsible for tracking and sharing evaluation data on a regular basis
- Make it safe to share findings that show a program is not working as expected — this is how programs improve
- Involve partners and participants in shaping evaluation questions, not just responding to them

Part II: Doing Evaluation

The following steps provide a practical framework for planning and conducting an evaluation. These steps draw on several established evaluation frameworks — including the CDC’s Program Evaluation Framework, the W.K. Kellogg Foundation’s Step-by-Step Guide, and the NSVRC Evaluation Toolkit — and adapt them for the realities of community-based IPV/SV prevention work in Tennessee. You do not need to do every step in a linear way; in practice, evaluation planning is iterative, and you may revisit earlier steps as you learn more.

Figure 1. The Evaluation Process



Step 1: Clarify Your Purpose and Engage Stakeholders

Before collecting a single data point, you need to be clear about why you are doing this evaluation. The purpose of your evaluation will drive every other decision — the questions you ask, the data you collect, and how you use the findings.

Common reasons for evaluating a program include:

- To demonstrate to a funder that you are achieving intended outcomes
- To monitor implementation fidelity and identify problems early
- To improve the program for future participants
- To determine whether you are reaching the right people
- To build the evidence base for a promising approach

The level of rigor your evaluation needs depends in part on its purpose. Data used primarily for internal program learning can be collected with relatively simple methods. Data intended to demonstrate effectiveness to a funder or support a grant renewal will require more systematic and defensible methods. Data intended to contribute to the broader evidence base for prevention will require the most rigorous design. Being clear about your purpose upfront helps you invest your evaluation resources appropriately.

Engaging Stakeholders

A stakeholder is anyone with an interest in your program or in the results of the evaluation. Engaging stakeholders early — not just at the reporting stage — makes your evaluation more useful and more credible. Stakeholders might include program staff and leadership, funders, community partners, program participants, and policymakers. Being intentional about whose voice shapes your evaluation questions and whose interests are reflected in the findings makes for a stronger, more ethical evaluation. See **Appendix B** for stakeholder brainstorming tool.

Step 2: Describe Your Program — Theory of Change and Logic Model

Before you can evaluate a program, you need to clearly articulate what it is, how it works, and what it is supposed to accomplish. Two tools help with this: a theory of change and a logic model.

Theory of Change

A theory of change explains why you believe your program will produce the outcomes you intend. It articulates the assumptions and mechanisms that connect your activities to change. A strong theory of change is grounded in research or accumulated practice wisdom, not just hope.

Logic Model

A logic model is a visual representation of your theory of change. It maps the connections between resources, activities, outputs, and outcomes. Logic models are not just planning tools — they are the foundation of your evaluation. The outputs section informs your process evaluation; the short- and intermediate-term outcomes inform your outcome evaluation.

A logic model typically includes: inputs/resources, activities, outputs, short-term outcomes (1–3 years), intermediate outcomes (3–5 years), long-term outcomes, and context. Developing a logic model is not a one-meeting task — plan for multiple iterations with input from staff and partners. A worksheet for developing your logic model is included in **Appendix C**.

The figure below shows a completed example logic model for a bystander intervention program.

Figure 2. Example Logic Model

Inputs/Resources	Activities	Outputs	Short Term (1-3 years)	Intermediate (4-5 years)	Long Term (5+ years)
DELTA AHEAD Funding	SLT convening and coordination	Quarterly SLT meetings			
TCEDSV Staff		Educational material created and distributed via social media channels, website, trainings, outreach, and professional convenings	Increased state- and community-level networking and collaboration	Increased # of people and organizations within TN exposed to connections between IPV and economic factors and evidence-informed approaches to violence prevention	Decrease in IPV and other forms of violence with common SDoH and targeted risk and protective factors in TN
CPR Evaluation Staff	State-level public education campaign to strengthen work-family supports		Increased implementation of evidence-based PPEs		
State Leadership Team (SLT)	State-level trainings and toolkits on paid parental leave to strengthen work-family supports	Training and toolkit on paid parental leave developed and shared with relevant stakeholders	Increased knowledge of IPV primary prevention approaches and resources	Increase in protective factors in TN including self and collective efficacy, social cohesion, strong workplace IPV policies, increased knowledge and awareness	Increase in health equity and health outcomes through improved SDoH including bridging community dynamics, socioeconomic and political context, and social and physical environment in TN
Community Partners			Increased awareness of IPV as a problem within TN	Decrease in risk factors in TN including weak sanctions against IPV, harmful norms around masculinity and femininity, cultural norms supporting aggression, financial stress, poverty, gender pay inequity	
CDC Staff Support	State- and community-level IPV and the Workplace trainings to improve organizational policies and workplace climate	IPV in the workplace policies developed, trainings conducted	Increased understanding of IPV risk and protective factors and associated SDoH among partners	Increased access to and application of state- and community-level evaluation data	
External TTA Provider support		A Call to Men program implemented and trainings conducted in Hamilton County			
	Community-level A Call to Men trainings and mentorship to engage men and boys as allies in prevention	Evaluation data collected and disseminated	Increased access to and use of data and evaluation in IPV prevention efforts in TN	Increased data monitoring related to IPV primary prevention strategies that promote racial and health equity at state and community levels by partners	
	Process and outcome evaluation at state and community levels	Increased knowledge of and skills related to implementing IPV primary prevention best practices; insights from work in TN shared with peer SDVC's and CDC			
	Participation in national-level dialogue and DELTA CoPs				

Context and Assumptions: 1) There are numerous challenges for violence prevention and equity advancement within the context of TN; 2) Foundational knowledge, infrastructure, and coordination among IPV primary prevention stakeholders exists already as a result of DELTA IMPACT activities; 3) funding for and data relevant to primary prevention are limited in TN; 5) TN is a right-to-work and at-will-employment state with no state-level paid family leave policy.

Step 3: Develop Your Evaluation Questions and Plan

Evaluation questions and your evaluation plan are developed together — you cannot really finalize your questions without knowing what is feasible to measure, and you cannot build a solid plan without knowing what you are trying to learn. This step covers both.

Developing Your Evaluation Questions

Evaluation questions are the high-level questions your evaluation is designed to answer — not the specific survey questions you will ask participants, but the bigger questions that guide what you measure and how. Your evaluation questions should flow directly from your logic model and your purpose.

Process evaluation questions typically ask: Is the program being implemented as planned? Are we reaching our intended audience? Are participants satisfied? What is working well and what needs adjustment?

Outcome evaluation questions typically ask: Did participants gain the intended knowledge, attitudes, or skills? Did the program contribute to the changes we hoped to see? For whom did the program work, and in what ways? Were there unintended outcomes?

Right Sizing Your Evaluation Questions

One of the most common mistakes in prevention evaluation is writing questions that are too broad to answer with available methods and resources. If you cannot name the data source or instrument that would answer a question, it is probably too ambitious for a low-capacity evaluation.

Examples: Right-Sizing Evaluation Questions

Too broad — not answerable without community-level data over many years:

“Are we reducing domestic violence in our county?”

Better — answerable with a pre/post survey:

“Did participants increase their knowledge of warning signs of unhealthy relationships?”

Even better — specific, behavioral, tied to a clear indicator:

“Did participants report greater confidence intervening when they witness concerning relationship behaviors after completing the program?”

Too broad — requires long-term population data:

“Are community norms around relationship violence changing?”

Better — answerable with a partner survey or tracking log:

“Did partner organizations report changes in how they address relationship violence in their policies or programming?”

A good test: if you cannot name the data source or instrument that would answer your question, narrow it until the answer feels within reach.

Prioritize ruthlessly. A small set of well-answered questions is far more valuable than a long list of poorly answered ones. A worksheet for developing and refining evaluation questions is included in **Appendix D**.

Building Your Evaluation Plan

An evaluation plan is your roadmap — translating your evaluation questions into concrete decisions about what you will measure, how, when, and who is responsible. A good plan addresses four core questions:

- **What:** What outcomes and process factors are you evaluating? (Draw from your logic model)
- **How:** What data collection methods and instruments will you use?
- **When:** When will data be collected, analyzed, and used for decision-making?
- **Who:** Who will collect, process, and analyze the data? Who needs to see the findings?

A complete plan also includes program background, sampling approach, considerations for consent and confidentiality, an analysis approach, timeline, communications plan, and a budget. A measurement framework template — which maps each evaluation question to specific indicators, data sources, methods, and frequency — is included in **Appendix F**.

Outcomes and Indicators

A well-written outcome includes a clear direction of change: “Participants will increase their knowledge of healthy relationship characteristics” is measurable; “Participants will benefit from the program” is not. An indicator is how you will know the outcome has been met. Outcomes should be SMART: Specific, Measurable, Relevant, and Time-specific. See **Appendix E**.

Example: Common Outcomes and Indicators in IPV/SV Primary Prevention

These examples illustrate how to move from a broad outcome to a specific, measurable indicator.

Knowledge and awareness

Outcome: Participants increase knowledge of warning signs of unhealthy relationships

Indicator: % of participants who correctly identify 4 of 5 warning signs on a post-training knowledge check (vs. pre-training)

Attitudes and norms

Outcome: Participants decrease acceptance of relationship violence as normal or inevitable

Indicator: Average score on an attitudes scale decreases from pre to post (lower score = less acceptance of violence)

Skills and self-efficacy

Outcome: Participants increase confidence to intervene safely when they witness potentially harmful situations

Indicator: % of participants who report feeling “confident” or “very confident” in their ability to intervene, post-training

Behavior change

Outcome: Participants report using bystander strategies in the 3 months following the program

Indicator: % of participants who report at least one bystander intervention on a 3-month follow-up survey

Community/systems change

Outcome: Partner organizations adopt policies that support healthy relationship norms

Indicator: Number of partner organizations that adopt or update a relevant policy within the program year

Step 4: Collect Data

Data collection methods fall into two broad categories: quantitative and qualitative. Most evaluations benefit from using both. An overview of common methods is included in **Appendix G** and a data collection plan worksheet in **Appendix H**.

Quantitative Methods

Quantitative methods answer questions about what, who, and when — how many people participated, what percentage showed a change in knowledge, whether scores improved from

pre to post. The most common are surveys with closed-ended questions and standardized assessments. Quantitative data are easier to summarize and are generally what funders expect.

Qualitative Methods

Qualitative methods answer questions about how and why — why participants engaged (or didn't), how they applied what they learned, what barriers staff encountered. Common methods include interviews (structured or semi-structured), focus groups, and structured observation. Qualitative data provide the context and texture that numbers alone cannot. For programs serving very small populations — as many rural Tennessee programs do — qualitative methods are especially valuable.

Primary vs. Secondary Data

Primary data is data you collect yourself. Secondary data is collected by someone else and accessed for your evaluation — administrative records, census data, community health assessments, school data, law enforcement data. Before designing a new instrument, always ask: does this data already exist somewhere? Secondary data are particularly valuable for primary prevention evaluation, where community-level outcomes are the goal. See the CDC's Violence Indicators Guide (in **Appendix A**) for a useful database of secondary data sources.

A Reality Check on Sample Size

Many programs work with small numbers of participants. This limits what data can show but does not make evaluation impossible — it means being realistic about what kinds of conclusions can be drawn.

Sample Size and What You Can Reasonably Claim

- **Fewer than 20 participants:** Descriptive findings only. Report what you observed (e.g., “14 of 17 participants reported increased confidence”) but avoid strong conclusions about effectiveness. Qualitative data becomes especially important at this scale.
- **20–50 participants:** Pre/post comparisons are reasonable. Describe change from before to after, with the caveat that you cannot rule out other explanations.
- **50–100+ participants:** More confidence in patterns and subgroup comparisons (e.g., by age, gender, or community).
- **The bottom line:** small numbers call for modest claims, not no evaluation. A thoughtful analysis of 15 surveys is more useful than no data at all.

A Note on Participant Burden

In IPV/SV prevention work, participants are often youth, community members with limited time, or people recruited through partner organizations. Asking too much of them can undermine your evaluation and your relationships.

- Keep surveys as short as possible. Five focused questions answered reliably are worth more than 30 questions half of which are skipped.

- Avoid surveying the same participants more than 2–3 times per program year without a clear reason.
- Be thoughtful about sensitive questions. In IPV/SV prevention, questions about personal experience with violence require careful framing, trauma-informed language, and clear confidentiality information, as included in **Appendix J**. When in doubt, measure attitudes and knowledge rather than personal experience.
- If partners (such as schools) control access to participants, work within their constraints rather than creating separate data collection events.

Where possible, use or adapt existing validated instruments rather than creating new ones. Survey design tips and a pre-use checklist are included in **Appendix I**.

Step 5: Analyze and Interpret Your Data

Analysis turns raw data into findings. Interpretation makes meaning of those findings. Both require keeping your evaluation questions front and center.

Analyzing Quantitative Data

For most community-based evaluations, descriptive statistics are sufficient and are well within reach of staff without a statistics background. Counts, percentages, and averages can be calculated in Excel. Step-by-step guidance for entering and analyzing survey data in Excel is in **Appendix K**.

Basic steps: (1) Clean your data for missing responses and errors. (2) Enter into Excel or statistical software. (3) Calculate frequencies and percentages. (4) Compare pre vs. post or across subgroups where relevant. (5) Draw conclusions by comparing results to expected outcomes.

More advanced techniques (regression, significance testing) require more expertise and are generally not necessary for internal program improvement purposes.

Analyzing Qualitative Data

Qualitative analysis identifies themes and patterns in text — transcripts, open-ended responses, focus group notes. Basic steps: read through all data, label (code) key themes, group coded text by theme, and draw conclusions in relation to your evaluation questions. This can be done in Word or Excel without specialized software. Guidance is included in **Appendix L**.

Interpreting Your Findings

Interpretation means comparing what you found to what you expected. What changes did the data show? Were there surprises? What might explain the results — including factors outside your program’s control? Be honest about what your data can and cannot tell you.

What if my data Show that the program is not working?

Negative or mixed findings are information, not failure. If data suggests a program is not achieving intended outcomes, consider:

- Was the program implemented as planned? (A process problem is different from a design problem)
- Did you reach the intended audience?
- Was the dosage sufficient (sessions, length of engagement)?
- Were there contextual factors — timing, staff turnover, community events — that affected results?
- Are the outcomes you measured appropriate for the stage of the program?

When sharing difficult findings with funders, be factual, contextualize the results, and lead with what you learned and what you will do differently. *Do not suppress findings. Sharing what did not work is how the broader IPV/SV prevention field learns.*

Step 6: Share and Use Your Findings

Findings are only valuable if they are used. Different audiences need findings presented differently: funders want formal reports with data summaries; program staff benefit from reflective conversations; community partners respond to a brief one-page summary; participants appreciate a simple narrative of what you learned and what changes as a result. See **Appendix M** for writing up and interpreting findings.

The most immediate and practical use of evaluation findings is program improvement. Build in deliberate time — even a one-hour team meeting — to ask: What do the findings tell us? What should we do differently? What should we keep doing? Document the decisions you make so that learning is captured for future planning, staff transitions, and funder conversations.

Common Evaluation Pitfalls

Even experienced practitioners run into the same evaluation problems. The following are the most common pitfalls in IPV/SV prevention evaluation — and how to avoid them.

- **Collecting data you never analyze.** Many programs administer pre/post surveys and then never find time to look at the results. Build analysis and reflection time into your evaluation plan as a scheduled activity, not an afterthought.
- **Measuring things that aren't tied to decisions.** Before adding any item to a survey or evaluation plan, ask: what will we do with this information? If you cannot answer that, remove it.
- **Creating surveys without knowing how the data will be used.** Survey design should start with your evaluation questions, not a blank questionnaire. Know what you are trying to learn before you write a single question.

- **Trying to measure too much with too little.** Resist the urge to add outcomes, indicators, and questions beyond what you can realistically collect, analyze, and use.
- **Expecting outcomes too soon.** Knowledge and attitude change can show up quickly. Behavior change takes longer. Community-level change takes longer still. Match your measurement timeline to realistic expectations for when change will occur.
- **Trying to prove causation.** Most community-based evaluations cannot establish that your program caused an observed change. You can document that change occurred and describe a plausible connection without claiming proof.
- **Collecting demographic data that nobody analyzes.** Demographic questions are valuable only if you plan to use them. If you are not going to look at whether outcomes differ by age, race, or gender, do not collect those variables.
- **Surveying participants repeatedly without a reason.** Survey fatigue is real. Every round of data collection should serve a specific purpose. If you cannot articulate why you are collecting data at a particular point, reconsider whether it is necessary.

Part III: Evaluation with Limited Resources

Most IPV/SV prevention programs operate with limited staff time, limited budgets, and competing demands. This section offers concrete strategies for doing meaningful evaluation within those constraints. The goal is not to lower the bar for what counts as evaluation — it is to help you make the most of what you have.

The key principle: don't let resource limitations keep you from doing anything at all. A small, well-executed evaluation that answers one important question is far more useful than an ambitious evaluation plan that never gets implemented.

Start with the Most Critical Questions

Which of your possible evaluation questions would be most actionable — meaning, which answers would clearly lead you to do something differently? Focus your limited resources there first. If you cannot answer the most critical questions with available resources, consider whether existing data sources could approximate an answer, or whether evaluation could be part of a future funding ask.

Match Your Rigor to Your Purpose

Rigor matters, but it is not all-or-nothing. Data for internal program improvement can be collected with simpler methods than data intended to demonstrate effectiveness to a funder or contribute to the evidence base. Just be honest — with yourself and your stakeholders — about what the data can and cannot show (**Appendix G**).

Make Maximum Use of Existing Data

Before designing new data collection, take stock of what already exists: program records (attendance, referral logs), partner organization data (schools, health departments, courts),

community health assessments, state and national data systems, and findings from evaluations of similar programs elsewhere. Secondary data are particularly valuable for primary prevention evaluation, where community-level outcomes are the goal.

Use Fast, Low-Burden Data Collection Methods

- Brief pulse surveys after individual trainings or events (3–5 questions, paper or phone)
- Short staff reflection forms on barriers and facilitators after each implementation cycle
- Attendance and completion tracking built into standard program operations
- Group interviews or focus groups instead of individual interviews (similar depth, lower time cost)
- Partner meeting logs or policy-adoption trackers for systems change work

The CDC’s DELTA FOCUS Data-to-Action framework (**Appendix A**) is a useful model here. It emphasizes collecting “good enough” data to meet your information needs, producing a brief reflective memo, and making decisions based on that data — prioritizing rapid feedback and midcourse correction over large, infrequent studies.

Working with Small Numbers

Many Tennessee programs serve small populations — a single school, a rural county, a community with few partner organizations. Small numbers limit what quantitative data can show, but they do not make evaluation impossible.

- Lean into qualitative methods. Even a handful of interviews can yield rich, actionable information.
- Be transparent about your sample size and explicit about what it does and does not allow you to conclude.
- Document your context. A detailed account of who you reached, what you did, and what barriers you encountered is itself valuable data.
- Consider multi-year aggregation. If annual numbers are too small for meaningful analysis, pooling data across two or three program years may give you more to work with.

A note on political and community context: prevention work in Tennessee does not happen in a vacuum. Some communities have strong support for prevention education; others have skepticism — about discussing relationship violence in schools, about the language of prevention, or about outside organizations working in their community. If community resistance is affecting your ability to reach participants or collect data, that is worth documenting as a contextual factor in your evaluation. Qualitative data — candid notes from community conversations, partner feedback, your own observations — can capture this reality in ways that survey data cannot.

Collaborate with Partners

Consider whether a partner organization is already collecting data you could use, or whether you could pool resources for shared evaluation activities. If you are working in schools, you may be able to add questions to an existing student survey rather than conducting a separate one. Universities and colleges are also potential partners — faculty and graduate students may be interested in your program as a research or practicum site, at low or no cost.

Protect What Matters Most

When resources require cutting back, protect the data collection that directly informs decisions — especially decisions about program continuation, modification, or expansion. A quarterly learning memo that synthesizes available data and drives a reflective team conversation may be more valuable than a formal annual report that arrives too late to change anything.

Part V: Appendix of Tools and Tips

Appendix A | Additional Resources

W.K. Kellogg Foundation Step-by-Step Guide to Evaluation:

https://www.betterevaluation.org/sites/default/files/WKKF_StepByStepGuideToEvaluation_smaller.pdf

NSVRC Evaluation Toolkit: <https://www.nsvrc.org/evaluation-toolkit/>

ODVN Primary Prevention Empowerment Evaluation Toolkit: <https://www.odvn.org/empowerment-evaluation-toolkit/>

Logic model development: <https://ctb.ku.edu/en/table-of-contents/overview/models-for-community-health-and-development/logic-model-development/main>

Developing an evaluation plan: <https://ctb.ku.edu/en/table-of-contents/evaluate/evaluation/evaluation-plan/main>

Key evaluation questions (BetterEvaluation): <https://www.betterevaluation.org/frameworks-guides/rainbow-framework/frame/specify-key-evaluation-questions>

In-depth data management and analysis: http://www.pcar.org/sites/default/files/resource-pdfs/analyzing_evaluation_data.pdf

Violence indicators and secondary data sources: <https://vetoviolence.cdc.gov/apps/violence-indicators-guide-database/database/>

Evaluating policy and systems change: <https://ctb.ku.edu/en/table-of-contents/advocacy/advocacy-research/policy-advocacy-evaluation/main>

Policy, systems, and environmental change approaches — CDC Violence Prevention in Practice:

<https://vetoviolence.cdc.gov/apps/violence-prevention-practice/approach-search#!/>

Community Readiness Model: <https://www.unco.edu/marie/community-readiness.aspx>

The tools in this appendix were adapted or developed using the following resources. We are grateful to these organizations for making their evaluation resources publicly available to prevention practitioners.

Ohio Domestic Violence Network (ODVN)	Empowerment Evaluation Toolkit. Ohio Domestic Violence Network. www.odvn.org
National Sexual Violence Resource Center (NSVRC)	Evaluation Toolkit. NSVRC, 2018. www.nsvrc.org
W.K. Kellogg Foundation	Step-by-Step Guide to Evaluation: How to Become Savvy Evaluation Consumers. www.betterevaluation.org
CDC — VetoViolence EvaluACTION	EvaluACTION: Evaluation Planning Worksheets and Tip Sheets. Violence Prevention in Practice. vetoviolence.cdc.gov

Appendix B | Brainstorming Stakeholders

Who is impacted by the program — and by the evaluation?

Think broadly: direct participants, their families, partner organizations, and the broader community.

Who makes decisions about the program and its implementation?

Who has authority over how the program runs, how it's funded, and whether it continues?

Whose voices need to be included in evaluation planning?

Think about: participants, survivors, funders, staff, community members, partner organizations. Who is missing?

Who are the audiences for this evaluation? Who needs to hear the story of your work?

Different audiences may need different things. Note what each audience cares about most.

What areas of expertise do you need? Who can you draw on?

Consider: data collection, analysis, community knowledge, lived experience, cultural competence.

Appendix C | Logic Model Worksheet

A logic model maps the connection between what your program does and what you expect to happen as a result. Use this worksheet to build or refine your program's logic model. Expect to revise it — that's not a sign something is wrong, it's how the process works.

INPUTS / RESOURCES	ACTIVITIES	OUTPUTS	SHORT-TERM OUTCOMES	INTERMEDIATE OUTCOMES	LONG-TERM OUTCOMES
<i>What staff, funding, materials, partnerships, and time do you bring?</i>	<i>What services, training, or activities does your program deliver?</i>	<i>What countable, direct products result from activities? (e.g., # sessions, # participants)</i>	<i>What will participants know or feel differently right after the program? (0–6 months)</i>	<i>What behavior or attitude changes do you expect over 1–2 years?</i>	<i>What lasting community-level change does this contribute to? (3–5+ years)</i>

CONTEXT *(What community, organizational, or policy factors shape how your program works? Note anything that affects your ability to deliver the program as planned or that might influence your outcomes.)*

Appendix D | Evaluation Question Development Worksheet

This worksheet walks you through developing your evaluation questions step by step. It works best when you've already drafted a logic model (Appendix B). Set aside 45–60 minutes and work through it with at least one colleague — good evaluation questions are developed in conversation, not in isolation.

Step 1: Identify Your Evaluation Purpose

What will you do with the evaluation findings? Who will use them and how?

Step 2: Brainstorm Possible Evaluation Questions from Your Logic Model

Look at each column of your logic model. For each element, ask: 'What would I want to know about this?' Jot down every question that comes to mind — don't filter yet.

Logic Model Element	Possible Questions (brainstorm — don't filter yet)
Activities / Outputs (Was the program implemented as planned?)	
Short-Term Outcomes (What changed right after?)	
Intermediate Outcomes (What changed over time?)	
Context factors (What affected implementation?)	

Step 3: Categorize and Verify Feasibility

For each question on your brainstorm list, **mark it P (process) or O (outcome)**. Process questions ask how the program was implemented. Outcome questions ask whether or what changed.

Then, for each question, **consider feasibility**. A question is feasible if you can answer YES to all three: 1) You can name a specific data source that would answer it; 2) You have (or can realistically get) the data within your timeline and budget; 3) The outcome is something your program can plausibly influence, given its scope and duration

Question	P or O	Feasibility (Y/N)

Step 4: Select Your Final 2–4 Evaluation Questions

From all the questions you've considered, select 2–4 to focus on.

Criteria for Selecting Final Evaluation Questions

- You have a clear data source and method for answering it.
- It addresses something stakeholders actually care about — not just something easy to measure.
- It reflects something your program can plausibly affect, given your scope.
- It's feasible given your staff time, budget, and timeline.

Final Evaluation Question	P or O	Data Source(s)
1.		
2.		
3.		
4.		

An outcome is a **change** — in what people know, believe, do, or experience — that results from your program. Outcomes answer: *What will be different in the world as a result of our efforts?* They should indicate direction using words like *increase, decrease, improve, or maintain*.

Outcome Type	Awareness & Knowledge	Attitudes, Values & Beliefs	Skills & Behavioral Intentions	Behaviors & Actions	Status / Conditions
Individual	Knows warning signs of IPV/SV; aware of local resources	Believes bystanders can make a difference; supports gender equality	Demonstrates healthy communication; plans to intervene if witness harm	Intervened to stop harassment; talked to partner about consent	Connected to services; discloses abuse
Org. / Community	Staff trained on trauma-informed care; community aware IPV occurs locally	Organizational culture shift toward equity and inclusion	Coalition aligned on prevention approach; partners coordinate referrals	Policy implemented; organizations share data across systems	Strong community norms against IPV/SV; policy enacted

Outcome Sequencing: Short- to Long-Term

Short-Term	Intermediate	Long-Term
Knowledge, awareness, attitudes, skills, behavioral intentions	Behavior change, peer norms shift, healthier relationship dynamics	Reduced DV/SV, community conditions change, policy reform

Writing SMART Outcomes

SMART outcomes are **Specific, Measurable, Achievable, Realistic, and Time-Bound**. Use the prompts below to draft outcomes for each time horizon.

Short-Term	Intermediate	Long-Term
Outcome statement:	Outcome statement:	Outcome statement:
Direction of change: ☐ increase ☐ decrease ☐ maintain ☐ improve	Direction of change: ☐ increase ☐ decrease ☐ maintain ☐ improve	Direction of change: ☐ increase ☐ decrease ☐ maintain ☐ improve
By how much?	By how much?	By how much?
By when?	By when?	By when?

Appendix F | Measurement Framework Worksheet

This table connects your evaluation questions directly to **how you will answer them**. For each evaluation question, identify the relevant outcome or process factor, the indicator you'll track, where data comes from, how you'll collect it, and how often. A completed measurement framework is the backbone of your evaluation plan. *Pre-filled row provides an example!*

Evaluation Question	Outcome or Process Factor	Indicator	Data Source	Collection Method	When / Frequency
<i>Did participants increase knowledge of healthy relationship characteristics?</i>	<i>Short-term outcome: knowledge gain</i>	<i>% of participants scoring higher post vs. pre on knowledge items</i>	<i>Pre/post survey</i>	<i>Paper or online survey, matched pairs</i>	<i>Pre: Session 1 Post: Session 6</i>

Outcome = the change you expect to see (e.g., "increased knowledge of healthy relationships").

Indicator = the specific, measurable signal that tells you whether that change happened (e.g., "% of participants scoring higher on the knowledge scale post vs. pre").

Appendix G | Overview of Common Methods

Not every evaluation needs to be a research study. The goal is to match your methods to your purpose — and to be honest about what your evidence can and cannot support.

Purpose	What This Requires	What You Can Reasonably Claim
Internal program learning <i>(improving how you do the work)</i>	Simple methods; small or convenience samples; consistent documentation; honest reflection on what's working	Descriptive findings about implementation; early signals of what participants are responding to; evidence-informed adjustments to delivery
Grant reporting <i>(demonstrating accountability to funders)</i>	Systematic methods; documented process; pre/post data or consistent outcome tracking; evidence that you implemented as designed	Evidence of implementation fidelity and reach; participant-reported short-term outcomes; description of early outcome trends — with appropriate caveats
Building the evidence base <i>(contributing to field knowledge)</i>	Rigorous design (larger samples, comparison groups or quasi-experimental approaches); validated instruments; multi-site or multi-year data	Contribution to causal claims about effectiveness; generalizable findings; evidence strong enough to support adoption or replication

Menu of Methods: Strengths and Fit by Purpose (not exhaustive)

Method	Best Fit	Strengths	Key Limitations
Pre/post surveys (in-person, self-administered)	Internal learning, Grant reporting	Strengths: Anonymous, high response rate at end of session, low cost, repeatable, can be adapted for literacy levels	Limitation: Question wording can bias responses; doesn't capture full context or explain why change occurred
Online surveys (e.g., Qualtrics, SurveyMonkey)	Grant reporting, Internal learning	Strengths: No manual data entry; easy to generate summary reports; good for staff, partners, or other respondents with reliable internet access	Limitation: Not appropriate for participants with limited internet access; lower response rates than in-person
Key informant interviews (staff, partners, stakeholders)	Internal learning, Grant reporting	Strengths: In-depth; flexible; good for exploring attitudes, motivations, implementation issues, and stakeholder perspectives; builds relationships	Limitation: Harder to analyze and compare across respondents; interviewer can bias responses; time-intensive to schedule
Focus groups	Internal learning, Grant reporting	Strengths: Efficient way to gather rich contextual information; useful for exploring perceived effectiveness, satisfaction, or implementation barriers	Limitation: Requires skilled facilitator; recruitment is challenging; analysis is time-intensive and requires care to avoid analyst bias
Administrative records (attendance logs, service data)	All three purposes	Strengths: Data already exists; no additional burden on participants; useful for tracking reach, dosage, and implementation fidelity; free	Limitation: Data quality may be inconsistent; confidentiality considerations for records with identifiers
Observation (structured or unstructured)	Internal learning, Evidence base	Strengths: Captures skills and behaviors that surveys can't; useful for fidelity monitoring and assessing what actually happens in sessions	Limitation: Observers may bias results (especially if staff); requires standardized protocol and trained observers; can influence participant behavior
Document review / content analysis	Grant reporting, Evidence base	Strengths: No new data collection needed; useful for evaluating policies, curricula, media content, MOUs, or legislative change; low cost	Limitation: Limited application; requires a standardized coding protocol to ensure consistency

Appendix H | Data Collection Plan Worksheet

Before designing new data collection tools, take stock of what already exists. Secondary data — records, reports, administrative data, and existing surveys — can answer many evaluation questions at little or no cost. Use Part 1 to inventory what you have, then Part 2 to plan any additional primary data collection needed to fill the gaps.

Part 1 | Secondary Data Inventory — What Do You Already Have?

Examples of secondary data: attendance logs, case management records, program activity reports, hotline call data, census or community-level statistics, previously administered surveys, partner organization data, state surveillance data

Existing Data Source <i>Name/describe the record or data set</i>	What It Contains <i>What information does it include?</i>	Evaluation Question(s) It Can Help Answer <i>Which questions does this address?</i>	Gaps or Limitations <i>What's missing or unreliable?</i>

Part 2 | Primary Data Collection Plan — What New Data Will You Collect?

For each piece of new data you need: identify what you're collecting, how, from whom, and who is responsible.

What <i>Type of information</i>	How <i>Method</i>	From Whom / Source <i>Respondents or data source</i>	When <i>Timing</i>	How Often <i>Frequency</i>	By Whom <i>Data collector</i>

Appendix I | Tips for Writing Good Survey Questions

The quality of your data is only as good as the quality of your questions. Poorly written survey items introduce bias, confusion, and measurement error — and you often don't find out until after you've collected data you can't use. Taking time to review and test your questions before you use them is one of the highest-leverage investments you can make in your evaluation.

Common Pitfalls — and How to Avoid Them

Pitfall	What to Watch For	Weak Example → Better
Leading questions	Questions that suggest a preferred answer. Participants are motivated to please and will follow your lead.	WEAK: "How helpful was this program to you?" BETTER: "How would you describe the program's impact on you?"
Double-barreled questions	A single question that asks about two things at once. Participants can't answer accurately because the two may have different answers.	WEAK: "Did this program increase your knowledge and your confidence?" BETTER: Ask each part as its own item.
Vague or abstract language	Terms like 'often,' 'sometimes,' 'helpful,' or 'significant' mean different things to different people.	WEAK: "I often feel confident talking to my partner." BETTER: Define terms in the question, or anchor to a specific time period.
Response scale problems	Unbalanced scales (more positive options than negative), too many options, or unclear endpoints all introduce error.	WEAK: Strongly Agree / Agree / Somewhat Agree / Disagree BETTER: Strongly Agree / Agree / Disagree / Strongly Disagree (balanced, even scale)
Asking about things people can't accurately recall	Memory is unreliable beyond a few weeks. Asking participants to recall behavior from 6 months ago produces inaccurate data.	WEAK: "In the past year, how many times did you intervene in a situation?" BETTER: "In the past 4 weeks..."
Sensitive questions without appropriate framing	Questions about trauma, violence, or personal behavior need context, safety language, and skip options.	WEAK: [No framing] "Have you ever been hurt by a partner?" BETTER: Add intro text, explain why you're asking, provide a 'prefer not to answer' option, and follow with a resource note.

Appendix J | Consent and Confidentiality in Program Evaluation

1. Key Elements of an Informed Consent Process

Even when formal consent isn't legally required, good practice includes explaining the following to participants before data collection:

- What you are collecting and why: 'We are asking you to complete a brief survey to help us understand whether this program is meeting its goals.'
- Who will see the data: 'Your individual answers will not be shared. We will report only summary results across all participants.'
- Whether participation is voluntary: 'Completing this survey is completely optional and will not affect your participation in this program.'
- What happens to data: How long it's kept, how it's stored, and when it's destroyed.
- How to ask questions or opt out: Provide a contact name and method.

4. Special Considerations for Youth Participants

- Parental consent: For participants under 18, most contexts require written informed consent from a parent or legal guardian. Coordinate with school administrators early in the planning process.
- Assent: In addition to parental consent, youth participants should provide their own assent — a child-friendly agreement to participate. Don't pressure youth to participate even if parental consent is secured.
- Sensitivity: Youth are especially vulnerable to feeling judged or to interpreting questions as evaluating them personally. Frame instruments carefully and consider having a youth advisory review before use.

Sample Consent and Assent Language

ADULT PARTICIPANT NOTICE (adapt as needed)

[Program Name] is asking you to complete a brief survey to help us understand how well this program is working. Your participation is completely voluntary — choosing not to complete the survey will have no effect on any services you receive.

Your individual answers will be kept confidential. We will only share group summary results. Your name will not be attached to your survey responses. We will keep survey data for [X years] and then destroy it.

If you have questions, please contact [Name] at [contact information]. If you experience distress related to any survey questions, support is available at [local resource / hotline].

YOUTH ASSENT LANGUAGE (adapt for age level)

We're asking you to answer a few questions about your experience in this program. You don't have to answer if you don't want to, and it won't matter either way — you'll still be part of the program.

Your answers are private. We won't share them with your parents, teachers, or anyone else. We'll only use the answers from everyone together — not yours individually.

If any question makes you uncomfortable, you can skip it or stop. Do you have any questions before you start?

Appendix K | Quantitative Data Analysis Tipsheet

You don't need specialized software to analyze evaluation data. Excel can handle almost everything a small-to-medium prevention program needs. This guide walks through the basics step by step.

Setting Up Your Data File

- One row per participant. Every participant gets exactly one row.
- One column per variable. Each question or data element gets its own column. Don't combine things.
- Label columns clearly at the top. Use short, descriptive names with no spaces (e.g., ID, School, Grade, Q1_Pre, Q1_Post).
- Use a participant ID, not names. Assign each participant a number (P001, P002...) and keep the ID-to-name link in a separate, secured file.
- Create a codebook tab. On a second tab, document what each column means, the response scale, and any coding decisions.

ID	School	Grade	Q1_Pre	Q2_Pre	Q1_Post
P001	Lincoln HS	9	3	2	4
P002	Lincoln HS	9	2	3	4
P003	Central MS	8	1	2	3

Example: Data structured with one row per participant, one column per variable.

Entering Pre/Post Survey Data

For matched pre/post designs, structure data so both pre and post scores are in the same row.

- Label pre columns: Q1_Pre, Q2_Pre, Q3_Pre, etc.
- Label post columns: Q1_Post, Q2_Post, Q3_Post, etc.
- Match participants using your participant ID. If pre and post data were collected separately, use a VLOOKUP to merge files.
- Enter missing data as blank cells, not zeros. A zero is a real answer; blank means no data.

Calculating Frequencies and Percentages

Step 1: Use COUNTIF to count responses: =COUNTIF(D2:D50, 4) counts how many participants scored 4 on Q1_Pre (data in column D, rows 2–50).

Step 2: Use COUNTA for total valid responses: =COUNTA(D2:D50) counts all non-blank cells. Use as your denominator.

Step 3: Calculate percentage: =COUNTIF(D2:D50,4)/COUNTA(D2:D50) — then format the cell as Percentage.

<i>Response Option</i>	<i>Pre Count</i>	<i>Post Count</i>	<i>Change</i>
<i>Formula used:</i>	<i>=COUNTIF(D2:D50, [value])</i>	<i>=COUNTIF(E2:E50, [value])</i>	<i>Post – Pre</i>
1 = Strongly Disagree	12 (34%)	3 (9%)	–9
2 = Disagree	14 (40%)	8 (23%)	–6
3 = Agree	6 (17%)	16 (46%)	+10
4 = Strongly Agree	3 (9%)	8 (23%)	+5
Total (n)	35	35	—

Calculating a Simple Pre/Post Mean Comparison

Step 1: Calculate the pre-program mean: =AVERAGE(D2:D50) — averages all pre scores for one item.

Step 2: Calculate the post-program mean: =AVERAGE(E2:E50) — same formula for the post column.

Step 3: Calculate the mean change: =E52-D52 (subtract pre mean from post mean). A positive number means scores increased.

Step 4: Repeat for each item or create a composite score by averaging across multiple related items.

Note on statistical significance: For small samples ($n < 50$), be cautious about interpreting mean differences. Describe changes descriptively rather than claiming the program 'caused' the change.

Creating a Basic Bar Chart from Your Results

Step 1: Select your summary data — the labels (response options) and counts or percentages.

Step 2: Go to Insert > Chart > Bar Chart (or Column Chart for vertical bars).

Step 3: Add axis labels: right-click the chart and select 'Select Data' to label axes clearly.

Step 4: Add a descriptive title: double-click the default title and replace it (e.g., 'Knowledge of Healthy Relationships: Pre vs. Post').

Step 5: Simplify the chart: remove gridlines, unnecessary legends, and 3D effects. Simple, clean charts communicate better.

Comparing Subgroups

If you want to know whether outcomes differed by school, grade, or demographic group:

- Manual filtering: Click the filter icon on your column header, select a subgroup (e.g., 'Lincoln HS'), and run your formulas on the filtered data. Use AVERAGEIF for conditional averages: =AVERAGEIF(B2:B50, "Lincoln HS", D2:D50).
- Pivot tables (more powerful): Select your full data range, go to Insert > Pivot Table, and drag variables into Rows (e.g., School) and Values (e.g., Average of Q1_Pre).

Important: Always report subgroup n sizes. A school with 8 participants is too small for reliable subgroup comparisons.

Appendix L | Qualitative Data Analysis Tipsheet

Qualitative data are **words** — responses to open-ended questions, focus group transcripts, interview notes, and written records. Unlike numbers, qualitative data captures *how* and *why*: what participants actually experienced, believed, and meant. Analyzing it well requires a structured process — not just reading and summarizing.

The Analysis Process

Step 1 Organize your data

Assign an ID to each respondent, questionnaire, or transcript. Make copies before marking anything up. Note the source (focus group site, interview date, etc.) so you can trace findings back.

Step 2 Read and absorb

Read all responses before coding anything. Jot initial impressions. What stands out? What surprises you? Note data quality issues — thin responses, unclear wording, missing data.

Step 3 Focus your analysis

Return to your evaluation questions. What do you need to understand? Organize your data around questions, topics, or respondent groups — whichever structure makes most sense for what you're trying to learn.

Step 4 Identify themes and build a coding schema

Read through responses and mark recurring ideas, words, and patterns. Assign short codes (e.g., SAFE = sense of safety; CONN = connection/community). Group coded responses into broader themes. It's okay if themes shift as you go.

Step 5 Look for connections across themes

Do themes cluster together? Do different respondent groups show different patterns? Count how often each theme appears — frequency signals importance. Contradictions and minority views matter too.

Step 6 Interpret and present

Write a summary for each theme. Use direct quotes (exemplars) to anchor your interpretation. Have a colleague review your analysis to check for bias. Address limitations and alternative interpretations when you share findings.

Example: From Raw Responses to Coded Themes

Focus group question: “What has been most valuable about this program?” Each response is read, coded with a short label, and grouped into a theme.

Participant Response	Code	Initial Label	Theme
<i>“I finally felt like I could speak up. Nobody judged me for what I went through.”</i>	SAFE	Psychological safety	Creating a safe, non-judgmental space
<i>“She gets it — she understood our culture and didn’t talk down to us.”</i>	TRUST	Trust in facilitator	Culturally responsive facilitation
<i>“Finding out others had the same experience — I thought I was the only one.”</i>	CONN	Shared experience	Reducing isolation and building community
<i>“I didn’t know what a healthy relationship looked like. Now I can recognize the warning signs.”</i>	SKILL	New knowledge/skill	Building awareness and practical skills
<i>“I wish it were longer. One session isn’t enough to cover everything.”</i>	DUR	Duration/dosage	Desire for more time / deeper engagement

Appendix M | Writing Up and Sharing Evaluation Findings

Good evaluation doesn't end when you collect your data. The work of analysis, interpretation, and communication is where findings become useful. This guide walks through the full arc — from making sense of your data to putting it in front of the people who need it.

Analyze

Organize and classify your data. Look for patterns, count frequencies, compare groups, and run any quantitative tests. Ask: What does each data source show? Where are the differences?

→ Synthesize

Bring findings from multiple sources together. Triangulate — do different methods point to the same conclusion? Note where findings converge, where they diverge, and what you don't yet know.

→ Interpret

Make sense of the findings in context. What do they mean for your program? Judge against your goals, benchmarks, or stakeholder expectations. Involve others — interpretation is strengthened by multiple perspectives.

Establishing a Basis for Judgment

Before you interpret findings, decide what you're measuring them *against*. This should be agreed on early with stakeholders, so everyone is working from the same standard. Common bases include:

Internal standards

Program missions and objectives; your own theory of change; past performance of this program; program protocols and fidelity standards

External standards

Performance of similar programs; national objectives or benchmarks; research literature; judgments by participants, experts, or funders; community values and expectations

Also ask during synthesis: Are your findings consistent with prior research? Do different data sources point in the same direction? Are there alternative explanations for what you found? If you used multiple indicators, did they converge?

Three Levels of Finding Quality

Not all reporting is equally useful. Move beyond describing numbers — aim to interpret what findings mean and connect them to decisions.

Level	What It Looks Like	Example
Describing data (What the numbers say)	Reporting raw results with no interpretation. Accurate but incomplete.	"72% of participants scored higher on the knowledge scale post-program than pre-program."
Interpreting findings (What the numbers mean)	Explaining what the pattern suggests about the program's effect. This is the minimum for useful reporting.	"Most participants demonstrated meaningful knowledge gains after the program, suggesting the curriculum effectively communicates core content about healthy relationships."

Level	What It Looks Like	Example
Making recommendations (What to do about it)	Connecting findings to program decisions. This is what makes evaluation useful.	<i>"Given that 28% of participants did not show knowledge gains, we recommend reviewing Session 3 materials, which address the content assessed on items where gains were smallest."</i>

Common Writing Pitfalls

Pitfall	The Problem	How to Fix It
Overclaiming	Using causal language your data can't support. Avoid: "The program proves that..." or "The program caused participants to..."	Use instead: "Participants reported..." or "Results are consistent with..." or "Findings suggest..." Always match your claim to your evidence.
Underclaiming	Being so hedged that your findings mean nothing. Avoid: "Data seems to potentially suggest that participants may have possibly benefited."	Use instead: Clear, direct language about what you found, followed by honest caveats. State your findings plainly, then note the limitations.
Burying the lead	Presenting all findings at equal weight without signaling what matters most.	Lead with your most important finding. Organize around conclusions, not around the order data were collected.
Ignoring negative findings	Reporting only successes makes findings less credible and less useful for program improvement.	Report what you found, not what you hoped to find. Explain what null or unexpected results might mean for the program.

Know Your Audience: Translating Findings for Use

Different audiences need different things from your findings. Before you write or present, ask: What does this audience already know? What do they need to decide? What format will they actually read? Who do they trust as a credible messenger?

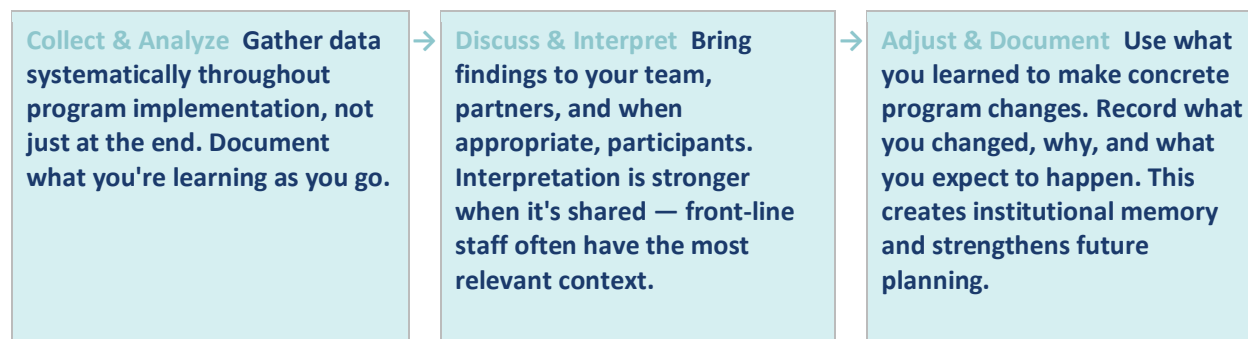
Internal audiences Program staff, supervisors, boards: need to understand what the data mean for daily practice and program direction. Focus on what's working, what isn't, and what to change. Plain language; no jargon.	External audiences Fundors, partners, policymakers, community: need evidence of impact and accountability. Emphasize reach, outcomes, and fit with their priorities. Tailor key messages to what matters to each audience specifically.
--	---

Common Communication Formats

Format	Description and Best Use
Evaluation summary	Overview of evaluation questions, methods, and key findings. Best for internal audiences and funder reports where full documentation is expected.
One-pager	One or two pages with program context and key findings; highly visual. Best for community stakeholders, policymakers, or busy audiences who need high-level takeaways quickly.
Presentation	Highlights high-level results for a specific stakeholder group; opens space for questions and discussion. Useful when you want stakeholders to help interpret findings or make decisions.
Success story	Qualitative or anecdotal evidence of program impact — a participant's experience, a community change. Best embedded within other formats to humanize the numbers.
Newsletter / brief	Short publication highlighting key findings for a specific audience. Useful for keeping partners and stakeholders updated over time as data emerge.
Dashboard	Customized graphical report; typically digital and dynamic, allowing findings to be updated as new data come in. Useful for ongoing monitoring and reporting to funders.

Closing the Loop: Integrating Findings into Ongoing Practice

A report filed and forgotten is an opportunity missed. Evaluation is most valuable when findings are brought back into your organization's planning and decision-making — not as a one-time exercise, but as an ongoing practice. The goal is to build a cycle where what you learn shapes what you do next.



Building this into your organization

Evaluation findings should have a regular home in your organizational calendar — not just in annual reports to funders. Consider:

Practice	What It Looks Like	Why It Matters
Scheduled data review	Set aside time at staff meetings or quarterly check-ins to review emerging data — not just at year-end. What are you seeing? What questions does it raise?	Creates a habit of reflection rather than a one-time exercise. Keeps data from sitting unused between reporting cycles.

Practice	What It Looks Like	Why It Matters
Annual planning integration	Use evaluation findings as a starting point for annual planning. What worked well enough to sustain or expand? What didn't work and should be adjusted or dropped?	Grounds planning decisions in evidence rather than assumption. Helps avoid repeating what didn't work.
Stakeholder feedback loops	Share relevant findings back with partners, participants, and funders — and invite their response. What do they make of the data? What are they seeing on the ground?	Builds trust and shared ownership of the evaluation. Often surfaces interpretations or explanations that internal staff would miss.
Documentation of changes made	Record what you changed, when, and why — including the data that prompted the decision. Keep this accessible to future staff.	Creates organizational memory. Demonstrates responsiveness to funders. Supports continuity when staff turn over.

A question worth asking at every planning meeting: What did our data tell us this year, and what are we going to do differently because of it? If the answer is ‘nothing’ — that's worth examining too.

Evaluation is only useful if it's used. Plan for dissemination from the start — not as an afterthought once analysis is done. Involve stakeholders in interpreting findings, be transparent about limitations, and connect every key finding to a decision someone can actually make. The goal isn't a perfect report. It's a program that learns and improves.