The University of Texas at Austin
Institute on Domestic Violence & Sexual Assault
Steve Hicks School of Social Work

The Bystander Evaluation (BeValued) Toolkit
A Resource to Support Program Growth and Impact

AUGUST 2021
PART 3 IN THE BYSTANDER INTERVENTION SERIES
FOR THE UNIVERSITY OF TEXAS SYSTEM

This page is intentionally left blank.

Authors

Caitlin Sulley, LMSW, Director of Research & Operations
Institute on Domestic Violence & Sexual Assault (IDVSA)
The University of Texas at Austin Steve Hicks School of Social Work

Deirdre Rabideau, LLM, Senior Research Project Coordinator
IDVSA, The University of Texas at Austin Steve Hicks School of Social Work

Lindsay Orchowski, PhD, Associate Professor (Research)
Department of Psychiatry and Human Behavior
Alpert Medical School of Brown University

Adam Jimenez, MSSW, Graduate Research Assistant
IDVSA, The University of Texas at Austin Steve Hicks School of Social Work

Stepha Dragoon, MSSW, MPAff, Graduate Research Assistant
IDVSA, The University of Texas at Austin Steve Hicks School of Social Work

Melanie Susswein, MSW, Director of Translational Communications
IDVSA, The University of Texas at Austin Steve Hicks School of Social Work

Sara Dube, MPH, CHES, Research Project Manager
The University of Texas at Austin

Leila Wood, PhD, MSSW, Research Assistant Professor
The University of Texas at Austin Steve Hicks School of Social Work

Noël Busch-Armendariz, PhD, LMSW, MPA, University Presidential Professor
Director, IDVSA, The University of Texas at Austin Steve Hicks School of Social Work

Contributor

Lara O'Toole, LMSW, Editing and Design

Acknowledgments

The Institute on Domestic Violence & Sexual Assault (IDVSA) researchers express their deep-felt gratitude and appreciation to the community of experts, advisors, and colleagues who participated in creating this toolkit. We are honored to work and learn alongside leaders, scholars, and practitioners who share a vision to increase student safety and work diligently to make that vision a reality. Thank you for your contributions.

We want to acknowledge the bystander intervention program stakeholders from across The University of Texas (UT) academic institutions who have worked with us over the past three years so that we could better understand their bystander intervention programs and ways to support them in improving their programming. They have continuously demonstrated passion and determination as interventionists while working collectively to create programs that support healthy communities. We send a special thank you to the following individuals who provided us with intensive feedback and insights in the Toolkit development process:

The University of Texas at El Paso
- Arely Hernandez, Director, Campus Advocacy, Resources and Education (CARE)
- Jovana Nieto, Outreach Coordinator, CARE
- Charles Gibbens, Assistant Vice President of Student Support, The Office of Dean of Students

We reserve a special shout out for the ongoing consultation and guidance from these partners:

The University of Texas at Austin
- Brittany O'Malley, Assistant Director for Prevention for University Health Services and the Counseling and Mental Health Center
- Sahtiya Hosoda Hammell, Bystander Intervention Programs Coordinator, Longhorn Wellness Center

The University of Texas at Dallas
- Amanda Smith, Associate Vice President for Student Affairs and Dean of Students, Division of Student Affairs

We are particularly thankful for the leadership and vision provided by Chris Brownson, Associate Vice President for Student Affairs and Director of The University of Texas at Austin Counseling and Mental Health Center. He has been instrumental to the success of the numerous initiatives dedicated to reducing interpersonal violence and harm at UT System institutions. We are fortunate to have his leadership, support, and expertise—and that he and his team are ongoing strategic partners.

We remain grateful to the UT System Board of Regents for their dedication and funding allocations to increase student safety. In 2014 and 2018, the UT System Board of Regents funded the bystander intervention initiative to empower students to intervene to prevent multiple types of violence and harm.

Finally, and of equal import, we are grateful to the students of UT institutions. It is a privilege to serve them.

Table of Contents

TABLES

FIGURES

ACRONYMS

CDC	Centers for Disease Control and Prevention
IHEs	Institutions of higher education
IPV	Intimate partner violence (includes dating violence/abuse)
SA/SV	Sexual assault/sexual violence
SAMHSA	Substance Abuse and Mental Health Services Administration

ICONS LEGEND

Section Icons[A]	
	Section I: Learn, Understand, and Remember
	Section II: Analyze and Apply
	Section III: Integrate, Act, and Create
Sidebar Icons	
	Resource available for no-cost download
	Key information and ideas
	Resource may be available through your institution
	Resource available to purchase
	Resource is a video, available to view at no cost.

[A] These three icons were created for The Noun Project. Reprinted with permission. Creators listed below:
- Section I Icon by RAM_DHANI
- Section II Icon by Made x Made
- Section III Icon by Creative Mahira

INTRODUCTION

THE TOOLKIT

This **B**ystand**e**r E**valu**ation **(BeValued) Toolkit** is a comprehensive program evaluation resource for bystander intervention program staff. While researchers from the Institute on Domestic Violence and Sexual Assault (IDVSA) customized it for the eight academic institutions at The University of Texas (UT), the BeValued Toolkit is adaptable and relevant to institutions of higher education (IHEs) across the nation. The BeValued Toolkit provides program staff with the knowledge and tools to critically evaluate their program and improve program growth and impact. This toolkit is based on the highest standard of program evaluation methods endorsed by the Centers for Disease Control and Prevention (CDC).

Toolkit Goals

1. Make program evaluation accessible, flexible, and adaptable for bystander program staff and stakeholders.
2. Facilitate various types of program evaluation, including to assess your current program's effectiveness, with recognition of program development, maturity, and goals.
3. Support program growth and sustainability over time.
4. Address evaluation needs and metrics for the unique multi-harm model.

A Unique Contribution

The BeValued Toolkit is unique among evaluation resources because of its depth, breadth, and adaptability, and because it is tailored to UT System's multi-harm model (see Figure 1).

KEY INFORMATION

Bystander Intervention, Defined

Bystander intervention is evidence- and theory-based violence prevention by engaging community members to actively disrupt and respond to harmful language, acts, and behaviors. By rejecting harmful behaviors, cultural norms, and expectations, there is movement towards a more inclusive, protective environment.

The UT System definition recognizes that bystanders can effectively intervene as a harmful situation occurs and positively influence the outcome for those involved.

Other bystander evaluation programs and IHEs can consider adopting the multi-harm model as one way to address the various harms (e.g., hazing, sexual violence, harassment, suicide) in a campus community. UT's approach aims to create holistically safer and stronger campuses, and this toolkit supports that goal. The multi-harm model also makes program evaluation more complex; the BeValued Toolkit addresses that complexity and offers a series of adaptable questions to fit each unique bystander intervention program.

Figure 1. Multi-Harm Model
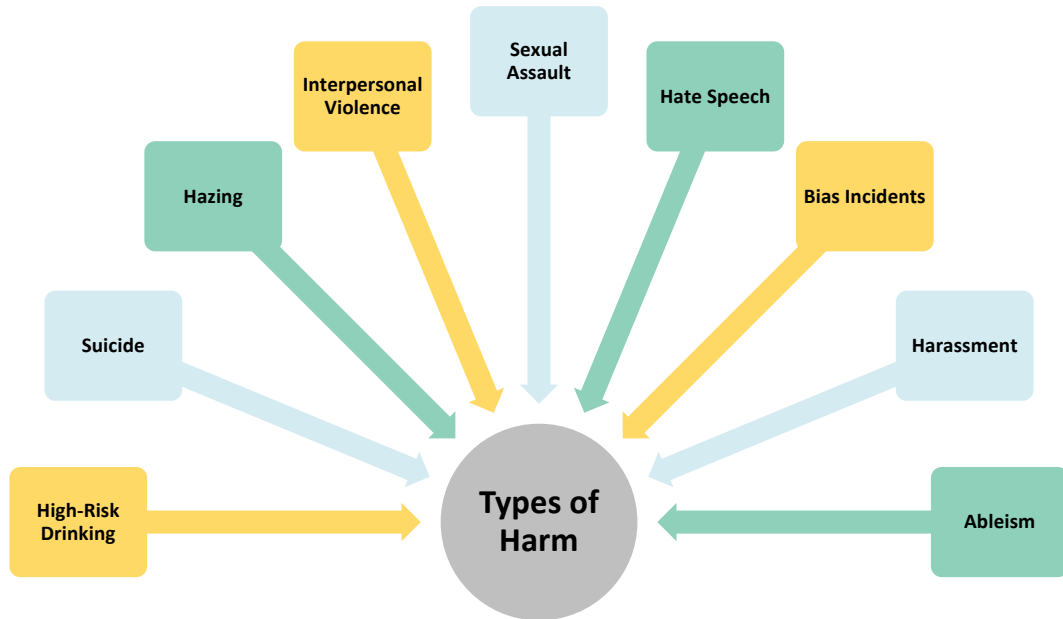
Broad Application

Our goal is to offer a toolkit that will be useful to stakeholders in the UT System and other IHEs who wish to evaluate bystander intervention programs. We also developed this toolkit to be valuable to a broader audience interested in seeing program evaluation best practices applied to a specific type of program. This toolkit not only guides you through assessing the effectiveness of your current program but can assist you in growing your program over time.

This toolkit is:
- Specific to bystander intervention programs.
- Designed for UT academic institutions, but applicable to bystander intervention programs more broadly.
- Highlights intimate partner violence and sexual violence, but applicable to other harms as well.

Supporting Sustainability

This toolkit will help support your program's sustainability by providing the necessary tools and resources for your program evaluation. The ability to continually enhance a program through rigorous evaluation methods is a key predictor of a program's long-term success.[1] We recognize that program evaluation is just one factor among the many that promote program sustainability (see Figure 2).[2] The factors listed in Figure 2 in the green circles with an additional black outline are addressed in this toolkit. The other factors warrant additional exploration and prioritization, and are outside the scope of this toolkit.

Figure 2. Factors Associated With Program Sustainability
Note: The items in green are specifically addressed in this toolkit.

Sulley, C., Rabideau, D., Orchowski, L., Jimenez, A., Dragoon, S., Susswein, M., Dube, S., Wood, L., & Busch-Armendariz, N.B. (2021). *The bystander evaluation (BeValued) toolkit: A resource to support program growth and impact.* Austin, TX: Institute on Domestic Violence & Sexual Assault, The University of Texas at Austin.

HOW TO USE THIS TOOLKIT

While reviewing this toolkit, we encourage you to critically consider your program, capacity, and campus community and adopt the evaluation method and recommendations that fit your setting and situation. The BeValued Toolkit is a multi-functional, flexible resource. Figure 3 lists the recommended steps for maximizing its usability. You can also choose to review specific toolkit sections for a general program evaluation overview or to work through certain evaluation phases at your own pace and in your own desired order.

Adaptability is Key

Regardless of your approach, it is important to note that the evaluation phases in the BeValued Toolkit are designed to be fluid; in other words, you may start at the phase most applicable to your current program needs rather than at the beginning. And you should always feel free to refer to an earlier phase as your program continues to develop and grow. Finally, you should freely adapt the materials provided to best fit your program. For examples on how to tailor this toolkit to your program and institution, see the CDC resource in the sidebar above.

RESOURCES AND TOOLS

Refer to the links, tools, sources, and information provided in these sidebars throughout the toolkit to extend your learning and identify examples and templates helpful to your program evaluation process.

A complete list of the resources and sources cited in this toolkit is also available in the References section.

Example

The CDC's *Using Essential Elements to Select, Adapt, and Evaluate Violence Prevention Approaches*[3]

Figure 3. Recommended Steps for Using the Toolkit
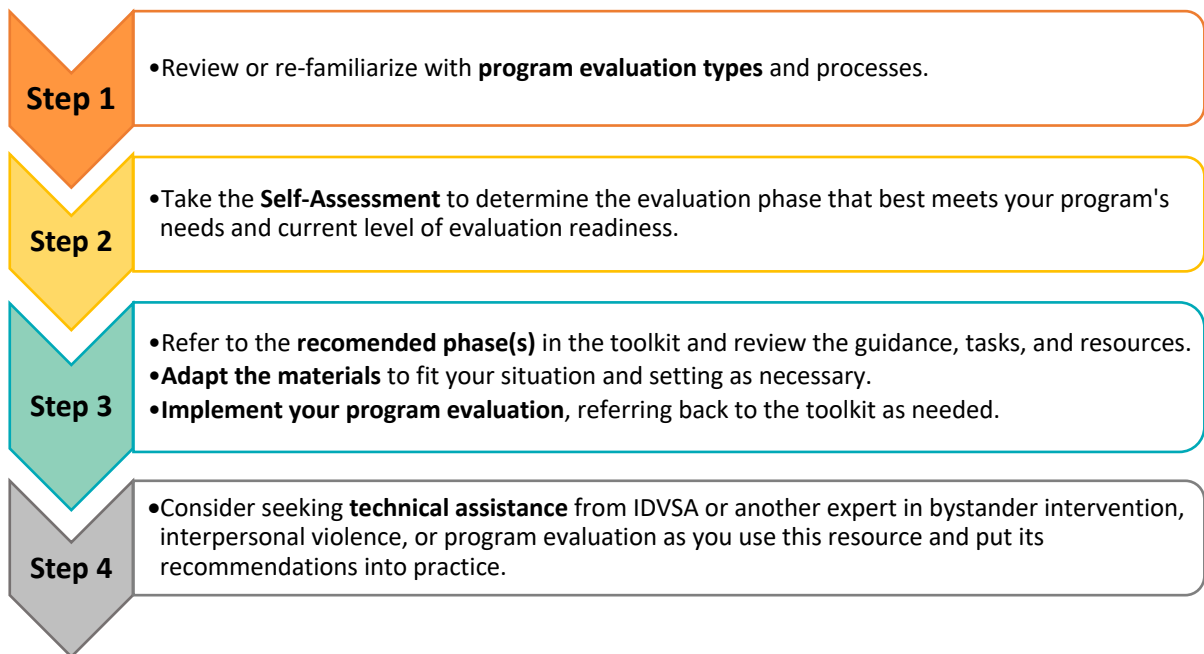
Step 1
- Review or re-familiarize with **program evaluation types** and processes.

Step 2
- Take the **Self-Assessment** to determine the evaluation phase that best meets your program's needs and current level of evaluation readiness.

Step 3
- Refer to the **recomended phase(s)** in the toolkit and review the guidance, tasks, and resources.
- **Adapt the materials** to fit your situation and setting as necessary.
- **Implement your program evaluation**, referring back to the toolkit as needed.

Step 4
- Consider seeking **technical assistance** from IDVSA or another expert in bystander intervention, interpersonal violence, or program evaluation as you use this resource and put its recommendations into practice.

Toolkit Sections

This toolkit is organized into eight main sections. Each section provides a comprehensive overview of the topic, along with guidance, key resources, and examples that can be tailored to fit your program. Table 1 offers a brief overview of each section.

Table 1. Toolkit Sections Overview

Section	Description of Key Elements
Introduction	**Practical context.** This section provides context and background information on the development of the BeValued Toolkit and guidance on how to use it.
Section I: Learn, Understand, and Remember	**Program evaluation basics.** This section includes an overview of program evaluation, program evaluation phases, and the types of program evaluations according to your program's maturity and development. It offers guidance for choosing a type of evaluation and resources to support further learning.
Section II: Analyze and Apply	**Self-assessment.** Our Self-Assessment tool provides a tailored recommendation for your program and directs you to a specific phase and section of the BeValued Toolkit to begin your program evaluation efforts.
Section III: Integrate, Act, and Create	**Phases of program evaluation.** This section provides an overview of the seven unique phases of program evaluation—from program design and planning to sustaining and reporting program outcomes data. The section includes tools, resources, and guidance for moving through the phases.
Common Challenges and Solutions	**Addressing barriers.** This section addresses some of the main challenges of program evaluation and offers potential solutions to improve bystander programs and enhance evaluation efforts.
Tools and Templates	**Tools.** This section provides a set of high-quality program evaluation tools and templates vetted by the IDVSA research team and readily available.
Appendices	**Resources and detail.** This section provides additional evaluation resources, terms and definitions, bystander program outcomes measures in detail, and more about the research methods used to develop the BeValued Toolkit.
References	**Sources.** This section provides a comprehensive list of citations for all content sources and resources referenced in the BeValued Toolkit.

PROJECT BACKGROUND

IDVSA at UT Austin is the only research institute in the nation that approaches research about interpersonal violence with a multi-disciplinary focus. Our expertise includes human trafficking, domestic violence, interpersonal violence, sexual assault, expert witnessing, and resiliency. With many years of cross-campus collaboration and knowledge-built from the *Cultivating Learning and Safe Environments (CLASE)* project to assess, understand, and address sexual misconduct and violence on UT campuses, IDVSA was well-positioned to support the bystander intervention initiative for UT institutions, led by UT Austin's Counseling and Mental Health Center (CMHC). To support bystander efforts, IDVSA focused on building resources and capacity to evaluate bystander intervention programming.

The BeValued Toolkit is one of a set of three resources, which includes:

1. *The Science Behind Bystander Intervention: A Guide to the Literature for UT System Institutions*
   This guide presents a comprehensive summary of the bystander intervention research literature related to a range of harms and high-risk behaviors. It highlights research recommendations on best practices and offers practical guidance to implement and evaluate bystander intervention programming on university campuses.

2. *Bystander Intervention Program Needs Assessment: Findings Across UT Academic Institutions*
   This data brief presents practical and useable findings from a bystander intervention programming needs assessment conducted with UT's eight academic institutions, including an analysis of strengths and challenges.

3. *The Bystander Evaluation (BeValued) Toolkit: A Resource to Support Program Growth and Impact*
   This toolkit presents the basics of program evaluation, a self-assessment to guide programs on getting started with their evaluations, evaluation methods to implement, instructions to adapt for specific program needs, and a host of templates and additional resources.

PROJECT APPROACH

A Thoughtful and Empirical Learning Approach

To organize how readers learn from the BeValued Toolkit, we use Bloom's Taxonomy of Educational Objectives framework.[4] The Taxonomy is an empirical method for education grounded in specific types of learning, and each type of learning provides a way to gain insights and master skills. In the toolkit development process, we thoughtfully identified the educational objectives ("What You Will Learn" in each section) to take an empirical, yet flexible, approach so that the toolkit is valuable and responsive to all readers.

Throughout the toolkit, we use **icons** to identify the type of learning pertinent to specific sections. Figure 4 gives an overview of the icons and the affiliated type of learning based on Bloom's Taxonomy.

Figure 4. Toolkit Educational Framework[B]
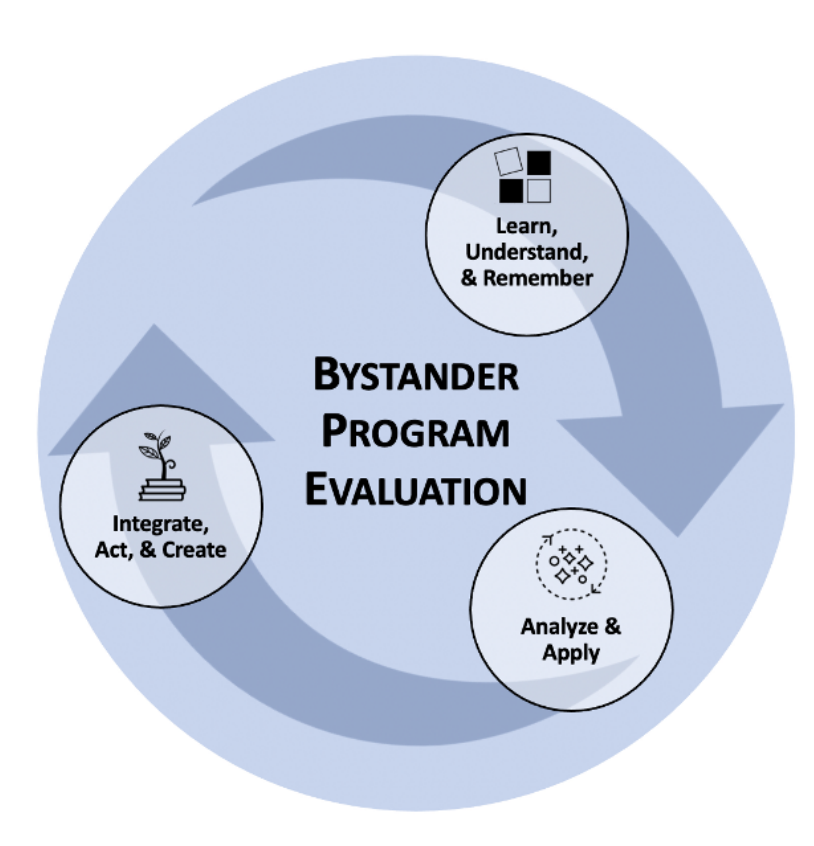
[B] Arrows icon "Refresh" created by Federico Panzano for The Noun Project.

Actively Applying Our Values and Principles

IDVSA's values and principles ground our rigorous, scientific approach to our work. The following guiding principles allowed us to capitalize on our interactions with administrators, staff, faculty, students, and other stakeholders in developing this toolkit.

Values: Leadership and Responsiveness

We embrace our moral, ethical, and legal responsibilities and those of other practitioners working to reduce and prevent violence on college campuses to ultimately realize peace. We are committed to thoughtful, reflective, and restorative dialogue with our partners and recognize the importance of a shared space for aspirational vision.

Principles: Discovery and Intersectional

We understand that influencing institutional change across the UT System and other university settings writ large is predicated on contributing to the existing body of scientific knowledge and providing evidence-based, applicable, and actionable recommendations to campus practitioners so they may build upon and improve their programs and policies. We acknowledge and honor the unique and diverse identities and experiences related to race, gender, ethnicity, sexual orientation, ability, religion/spirituality, nationality, and socioeconomic status. We take active measures to address harms and evaluate bystander programs' efforts to address inequality for our campus communities.

For this project, our hope is that a multidisciplinary approach that prioritizes researcher-practitioner collaboration enhances the quality and relevance of these research and evaluation efforts as they are applied to bystander intervention programming.

A further description of our empirical methods is found in Appendix C. Toolkit Development Methods.

> **KEY INFORMATION**
>
> **Our Mission**
>
> IDVSA's mission is to eliminate abuse and violence with social and economic justice as centering principles. To achieve our aspirational goal, we engage stakeholders in ongoing collaborative decision-making and restorative practices, recognizing that our actions affect their lives.
>
> For more information about IDVSA's work, resources, and publications, visit our website **https://sites.utexas.edu/idvsa/**

SECTION I.
Learn, Understand, and Remember

What You Will Learn

- Define the primary purpose and elements of a program evaluation.
- Identify the types of program evaluations and what type of program evaluation suits programs at different maturity and development levels.
- Distinguish between formative and summative evaluations.
- Recognize this section as a reference for all current and future program evaluation and improvement needs.

BASICS OF PROGRAM EVALUATION

This section provides an overview of program evaluation, including its purpose, types of program evaluations, and standards in the field. This section is valuable to review if you are unfamiliar with program evaluation or considering various evaluation types and how each may benefit your program and current goals.

Purpose of Program Evaluations

Program evaluations are extremely valuable in providing a structured way to gather and report information on a program. They can:

- Monitor program progress and growth.
- Identify program components with the most significant impact.
- Detect issues with program delivery and areas for improvement.
- Assess and refine data collection processes.
- Document program quality.
- Demonstrate impact and effectiveness.
- Quantify changes.
- Highlight innovation and high-quality work.

RESOURCES AND TOOLS
For more guidance with program evaluation basics:
• The CDC's *Introduction to Program Evaluation for Public Health Programs: A Self-Study Guide* [5] • Also see *Appendix A. Key Resources on Program Evaluation* and *Appendix B. Evaluation Terms and Definitions*

The information gathered in a program evaluation can provide evidence to justify program funding, build program support, enable program improvements, and contribute to program growth. By evaluating your program, you can strategically improve and adjust your program to be more effective in achieving its objectives.

The CDC's Continuous Program Improvement Model[6] in Figure 5 depicts how program evaluations are part of an ongoing process to inform program efforts and enhance program effectiveness. Like a continuous feedback loop, the CDC model offers practitioners a way to gather and use real-time information to make necessary adjustments and generate regular reports on program outcomes.

Figure 5. CDC Continuous Improvement Model

Reprinted with permission from: Kidder, D. P., & Chapel, T. J. (2018). CDC's program evaluation journey: 1999 to present. *Public Health Reports, 133*(4), 356–359.

Practitioners who follow this model—design a program, implement it, monitor implementation, measure effects, and prioritize regular evaluations—have more opportunities to use evaluation findings for continuous quality improvement. In other words, by using the CDC approach, you can increase the likelihood that your program will successfully achieve its objectives.

Figure 6 offers a descriptive example of the potential short, intermediate, and long-term benefits of evaluating a bystander intervention program focused on sexual violence and misconduct. The figure also highlights how bystander intervention program evaluations offer one source of information that can be reviewed alongside the institution's other data sources, such as campus climate surveys. By drawing from more robust data, an institution can more comprehensively assess and respond to students' needs around violence and harm to create a campus community that is safer and more supportive overall.[7]

Figure 6. Potential Benefits of Bystander Intervention Program Evaluations
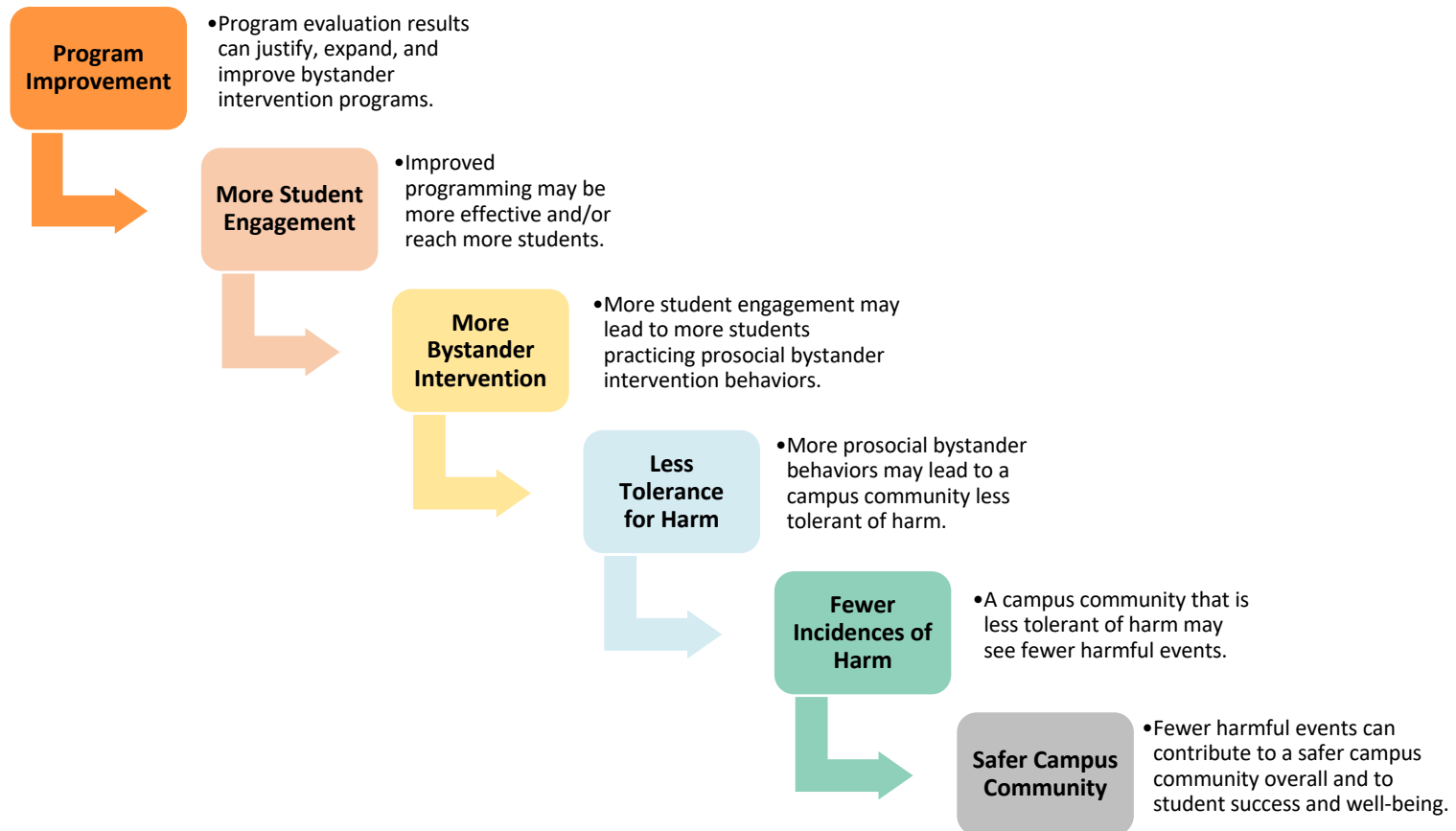
Program Improvement
- Program evaluation results can justify, expand, and improve bystander intervention programs.

More Student Engagement
- Improved programming may be more effective and/or reach more students.

More Bystander Intervention
- More student engagement may lead to more students practicing prosocial bystander intervention behaviors.

Less Tolerance for Harm
- More prosocial bystander behaviors may lead to a campus community less tolerant of harm.

Fewer Incidences of Harm
- A campus community that is less tolerant of harm may see fewer harmful events.

Safer Campus Community
- Fewer harmful events can contribute to a safer campus community overall and to student success and well-being.

Sulley, C., Rabideau, D., Orchowski, L., Jimenez, A., Dragoon, S., Susswein, M., Dube, S., Wood, L., & Busch-Armendariz, N.B. (2021). *The bystander evaluation (BeValued) toolkit: A resource to support program growth and impact.* Austin, TX: Institute on Domestic Violence & Sexual Assault, The University of Texas at Austin.

PHASES OF PROGRAM EVALUATION

This toolkit offers an overview of program evaluation and a deep dive into seven unique phases of the evaluation process (see Figure 7 for an overview of the phases). Each phase represents a different area of consideration for your evaluation, and the BeValued Toolkit offers guidance on the steps involved in each. Links to resources, information, and examples are provided throughout the toolkit. **You do not have to go through the phases in chronological order or start with Phase 0. Your program and evaluation needs will determine your process and starting point.**

Figure 7. Phases of Program Evaluation

Sulley, C., Rabideau, D., Orchowski, L., Jimenez, A., Dragoon, S., Susswein, M., Dube, S., Wood, L., & Busch-Armendariz, N.B. (2021). *The bystander evaluation (BeValued) toolkit: A resource to support program growth and impact.* Austin, TX: Institute on Domestic Violence & Sexual Assault, The University of Texas at Austin.

TYPES OF EVALUATION ACCORDING TO PROGRAM MATURITY

There are different types of program evaluations you can conduct, depending on the information you wish to gather and how mature your program is. Evaluation can be an iterative and ongoing process. Depending on the needs and goals for conducting the evaluation, program staff can employ multiple types of evaluations. For example, when one evaluation leads to findings used to improve programs, other types of evaluation may be used subsequently to review the new program activities. Therefore, we recommend that program staff become familiar with the program and its history before beginning the evaluation planning process. We also recommend carefully considering the purpose of the program evaluation.

Program evaluation methods are broken out by two stages of program development or maturity–**formative** and **summative**. To put it simply, formative evaluations provide information for the program's development and delivery in the early stage of program design and rollout, whereas summative evaluations can show if and how the program achieved its goals and takes place when a program is established and mature. Next, we offer additional detail to describe what we mean.

FORMATIVE EVALUATIONS

Formative evaluations are conducted to inform program development and improve program delivery.

When to do a formative evaluation:

- During the planning or implementation phase of a program in order to provide feedback around program functionality and short-term outcome measures.
- Formative evaluations can help determine whether your program is on track or needs adjustments to reach its intended outcomes.

Next, we describe the two most common types of formative evaluation: needs assessment and process evaluation.

RESOURCES AND TOOLS

For more guidance with different types of evaluations and evaluation activities:

- PA Coalition Against Rape's *Assessing Campus Readiness for Prevention: Supporting Campuses in Creating Safe and Respectful Communities* [8]

- *Using Focus Groups to Evaluate Health Promotion Interventions* [9] (access through your institution)

- *A Stage Model of Behavioral Therapies Research: Getting Started and Moving on From Stage I* [10]

The Needs Assessment That Informed This Toolkit

To prepare for the creation of the BeValued Toolkit, IDVSA administered a survey to learn more about the violence and harm prevention and intervention programming (including bystander intervention) efforts across UT academic institutions. This needs assessment focused on program evaluation efforts as well as needs or gaps in programming and/or program evaluations.

Read our data brief on the needs assessment findings HERE.

- The needs assessment survey was administered to representatives from eight (8) UT academic institutions who served as program staff.
- The survey asked participants to identify strengths, challenges, and needs of their bystander intervention programming as well as their current program evaluation efforts.
- The IDVSA research team conducted a question-by-question analysis to identify trends in the responses and highlight the key findings from each category.
- The survey responses—especially those concerning program evaluation efforts—were a vital resource to us when creating this toolkit. We used them to tailor the toolkit format and contents to address the needs identified by each institution.

Needs Assessment

Needs assessments are conducted to define the problem(s) of interest and can be done at any stage of program evaluation.

The potential goals of a needs assessment include describing:

- The severity of a problem.
- How the problem is distributed. (Does it impact some groups more than others? How?)
- The risk factors associated with the problem.
- The protective factors that can prevent or reduce the risk and/or severity of the problem.

When to do a needs assessment:

- Can be conducted at any stage of program evaluation.
- Typically done during the planning stages of a program—to justify the need for the program and to guide program design.

RESOURCES AND TOOLS

For more guidance with needs assessments and their activities:

- Rutgers' *Understanding and Responding to Campus Sexual Assault: A Guide to Climate Assessment for Colleges and Universities* [11]

- PA Coalition Against Rape's *Assessing Campus Readiness for Prevention: Supporting Campuses in Creating Safe and Respectful Communities* [12]

- When you are looking to expand the program's efforts to address more problems. For example, a bystander intervention program might conduct a needs assessment to find out if they should expand their efforts to include components that focus on hazing in addition to existing program components that focus on intimate partner violence.

Process Evaluation

Process evaluations are typically conducted to determine and describe how a program is being implemented and can serve as a valuable record of historical information. A process evaluation is usually conducted in conjunction with an outcome evaluation to explain and enhance the interpretation of outcomes.

Process evaluations can include an evaluation of:
- The reach of the program
  - Awareness of the program
  - Participation in the program
  - How many doses (or sessions) participants get
  - The dropout rate
  - If there are differences in groups who participate
- Implementation of the program or program fidelity
  - The facilities available to deliver the program
  - The materials and supplies available to deliver the program
  - The funding to deliver the program
  - How many individuals deliver the program
  - What training and supervision individuals delivering the program receive
  - If there are differences in the quality of the program delivered
  - If there are differences in the consistency of the program delivered
  - If the program coordinates with other campus programs
  - The protocols and procedures to deliver the program
  - How satisfied participants are with the program

When to do a process evaluation:
- Process evaluations can start as soon as the program is implemented.
- When identifying or planning program improvements, either on a specific timeline or continuously.
- It is ideal to continue implementing a process evaluation as you move into outcome evaluation. A process evaluation can help you understand why your outcomes were or were not achieved or explain surprising outcome evaluation findings.

RESOURCES AND TOOLS

For an example of a process evaluation tool:

The tool on the next page is an example of how the BeVocal Program at UT Austin tracks its program activities.[13] Campus partners that incorporate BeVocal messaging into their presentations are asked to complete this short survey to provide certain information. Program staff can then collect and tabulate information across trainings provided. That information can be used for a process evaluation, namely to describe program activities.

The BeVocal tool collects basic descriptive information needed to inform a process evaluation, including:

Process Evaluation Component	*Data Source*
Dose delivered	• Date(s) of training(s)
Program Research and Participation	• # of people who received information
	• Audience type
Program Fidelity	• Type of event
	• Type of information provided
Context	• Additional comments or feedback

Process Evaluation Tool: BeVocal Activity Tracker

Facilitator Name:	
Facilitator Email:	
Facilitator Department:	
Issue Addressed:	
Date of Event:	
Name of Event:	
Type of Event	☐ Skills-Based Workshop ☐ Lecture-Style Training ☐ Performance ☐ Tabling ☐ Other: _____ Approximate number of students in attendance: _____ Approximate number of faculty/staff in attendance: _____
Audience	☐ University-wide ☐ Athletes ☐ Sorority & Fraternity Life ☐ Other Student Group: _____ ☐ Student Staff: _____ ☐ Faculty or Staff Group: _____
BeVocal Concepts Covered	☐ Explanation of BeVocal initiative and vision statement ☐ BeVocal definition of Bystander Intervention ☐ BeVocal action steps ☐ Barriers to intervention and motivations to intervene ☐ Resources for intervening
Please check "Yes" here if you are willing to be contacted in the future for a more in-depth follow-up survey or interview about your experience with the content and website	☐ Yes ☐ No
Additional comments or feedback:	

Reprinted with permission from: BeVocal. (n.d.). *BeVocal Activities Tracker*.

https://utexas.qualtrics.com/jfe/form/SV_6l10gn5KfC3vjkV

SUMMATIVE EVALUATIONS

Summative evaluations are conducted to inform program stakeholders about the extent to which a program was able to achieve its objectives once the program is delivered.

When to do a summative evaluation:

- Summative evaluations should be conducted with completed or established programs.
- At the end of a program or at a logical point-in-time (e.g., at the end of an academic year).

Next, we describe the two most common types of summative evaluation: outcome evaluation and impact evaluation.

Outcome Evaluation

Outcome evaluations are conducted to determine if a program has met its objectives or achieved intended effects. Outcome evaluations may assess changes in:

- Short-term program objectives (e.g., participants' psychosocial behavioral determinants, including attitudes, self-efficacy, confidence, or normative beliefs).
- Intermediate program objectives (e.g., participants' behaviors).
- Long-term program objectives (e.g., victimization and perpetration rates campus-wide).

When to do an outcome evaluation:

- Outcome evaluations often occur at the end of a program/program cycle.
- They can also occur throughout a program's duration. As the program continues to change, it is preferable to use outcome evaluations to identify desired results and track progress throughout the program's life span.

RESOURCES AND TOOLS

For more guidance on program evaluations and selecting the appropriate approach:

- The CDC's *Building Our Understanding: Key Concepts of Evaluation, What is it and how do you do it?* [14]

- *Efficacy and Effectiveness Trials (and Other Phases of Research) in the Development of Health Promotion Programs* [15]

- *Multi-College Bystander Intervention Evaluation for Violence Prevention* [16]

- University of New Hampshire's *Sexual Violence Prevention Through Bystander Education: An Experimental Evaluation* [17]

- The CDC's *Intimate Partner Violence and Sexual Violence Victimization Assessment Instruments for Use in Healthcare Settings* [18]

The Importance of Process Evaluation as a Part of Outcome Evaluation

As mentioned previously, process evaluation activities support an outcome evaluation. The direct effects of the program activities on participants only tell some of the story. When used in combination, outcome and process evaluations provide more information about why a program did or did not achieve intended results.

If your outcome evaluation produces disappointing or surprising findings, conducting a process evaluation can help to explain those results. In a process evaluation, you can look at participant training and experiences for additional insights. For example, consider the scenario in which program staff learn that participants did not perceive the training as relevant. When you look more closely at the larger context and process information, you learn that the training in question took place on election day, which could explain participant perceptions.

A Look at Outcome Evaluation Activities

The following strategies and tools are commonly used in outcome evaluations:

- **Measurement tools** – A number of these tools are described in Table 2 and in more detail in Appendix D.
- **Interviews and focus groups** with program participants – A recommended strategy to help you to further understand program outcomes and why they are occurring in a certain way in order to plan your next steps.
- **Exit interviews** with program participants as they complete the program – These are valuable to gather feedback on the specific components of a program that participants found useful or how they utilized program information over time.

RESOURCES AND TOOLS

For more guidance with outcome evaluations:

The resource below (Table 2) lists several existing measures commonly used to assess the short-term, intermediate, and long-term outcomes of bystander intervention programs. The table provides basic information on each measure (e.g., number of questions, construct it measures). You can find more detailed information (e.g., subscales, example questions, answer choices, scoring instructions, source information) in Table 11 in Appendix D. Detailed Outcomes Measures Review. Portions of the measures are reprinted with permission from their authors.

Table 2. Outcomes Measures Matrix

Measurement	Construct	Number of Questions	Audience	Issue(s) or Topic Area(s)	Short-Term[C] Measure	Intermediate[D] Measure	Long-Term[E] Measure	Source
National College Student Bystander Intervention Survey (NCBIS)	Bystander Attitudes, Behaviors, and Barriers	118[F]	College Students	Multi-harm[G]	✔			WITH US Center for Bystander Intervention at Cal Poly, 2020[19]

[C] Short-term outcomes in this resource refer to psychosocial behavioral determinants (e.g., self-efficacy, attitudes, barriers).

[D] Intermediate outcomes in this resource refer to behaviors (i.e., bystander intervention behaviors).

[E] Long-term outcomes in this resource refer to incidence and prevalence of violence.

[F] NCSBIS has 118 possible questions total, which include logic-based items. The survey is split and randomly assigned to sample to limit time to complete (8–10 minutes).

[G] Multi-harm includes SA, IPV, hazing, alcohol and other drug use, and bias.

Measurement	Construct	Number of Questions	Audience	Issue(s) or Topic Area(s)	Short-Term[C] Measure	Intermediate[D] Measure	Long-Term[E] Measure	Source
Perceptions of Peer Helping	Bystander Normative Beliefs	20	College Students	Sexual Assault (SA) Intimate Partner Violence (IPV)	✓			Banyard et al., 2014[20]
Modified Bystander Attitude Scale - Revised (BAS-R)	Bystander Attitudes	11	College Students	SA	✓			McMahon et al., 2014[21]
Bystander Efficacy Scale - Short Form	Bystander Self-Efficacy	8	College Students	SA IPV	✓			Banyard et al., 2014
Readiness to Help Scale - Short Form	Bystander Readiness	18	College Students	SA IPV	✓			Banyard et al., 2014
Brief Intent to Help: Friends and Strangers	Bystander Intentions	17	College Students	SA IPV	✓			Banyard et al., 2014
Barriers to Bystander Action Scale	Bystander Barriers	5	School Personnel	SA IPV	✓			Edwards et al., 2017[22]
Bystander Efficacy Scale	Bystander Self-Efficacy	12	School Personnel	SA IPV	✓			Edwards et al., 2017
Perceptions of School Readiness	Bystander Readiness	12	School Personnel	SA IPV	✓			Edwards et al., 2017
Bystander Intent to Help Questionnaire	Bystander Intentions	7	School Personnel	SA IPV	✓			Edwards et al., 2017

Measurement	Construct	Number of Questions	Audience	Issue(s) or Topic Area(s)	Short-Term[C] Measure	Intermediate[D] Measure	Long-Term[E] Measure	Source
Illinois Rape Myth Acceptance - Updated	Rape Myth Acceptance	22	College Students	SA	✓			McMahon & Farmer, 2011[23]
Bystander Behaviors- Short Form	Bystander Behaviors	20–40	College Students	SA IPV		✓		Banyard et al., 2014
Modified Bystander Behavior Scale - Revised (BBS-R)	Bystander Behaviors	10	College Students	SA		✓		McMahon et al., 2014
Partner Victimization Scale	Violence Victimization	5–49	College Students	SA IPV			✓	Hamby, 2013[24]
Violence Victimization and Perpetration	Violence Victimization/ Perpetration	15–30	College Students	SA IPV			✓	Coker et al., 2015[25]
Sexual Experiences Survey - Short Form Victimization (SES-SFV)/Perpetration (SES- SFP)	Violence Victimization/ Perpetration	54–74	College Students	SA			✓	Koss et al., 2007[26] Johnson et al., 2017[27]
Digital Dating Abuse	Violence Victimization/ Perpetration	18–36	College Students	IPV			✓	Reed et al., 2017[28]

Impact Evaluation

Impact evaluations are conducted to determine if a program met its goals. They are helpful when you want to look at whether innovative programming activities achieved stated aims and when you want to determine whether the program can and should be expanded or replicated in another setting.[33,34]

When to do an impact evaluation:
- When piloting a program or at the beginning of a program in order to evaluate a small sample of results before allocating more time and resources.
- At various intervals during the program.
- At the conclusion of a program.

REVIEW AND SYNTHESIS

Figure 8 depicts which evaluation type is appropriate based on program maturity and lists the main questions asked and answered in each evaluation type.

RESOURCES AND TOOLS

For more guidance with impact evaluations:

- InterAction's *Impact Evaluation Guidance Notes* [29]
- OECD's *Principles of Impact Evaluation* [30]
- USAID's *Technical Note: Impact Evaluation* [31]
- World Bank's *Impact Evaluation in Practice* [32]

Figure 8. Types of Program Evaluations
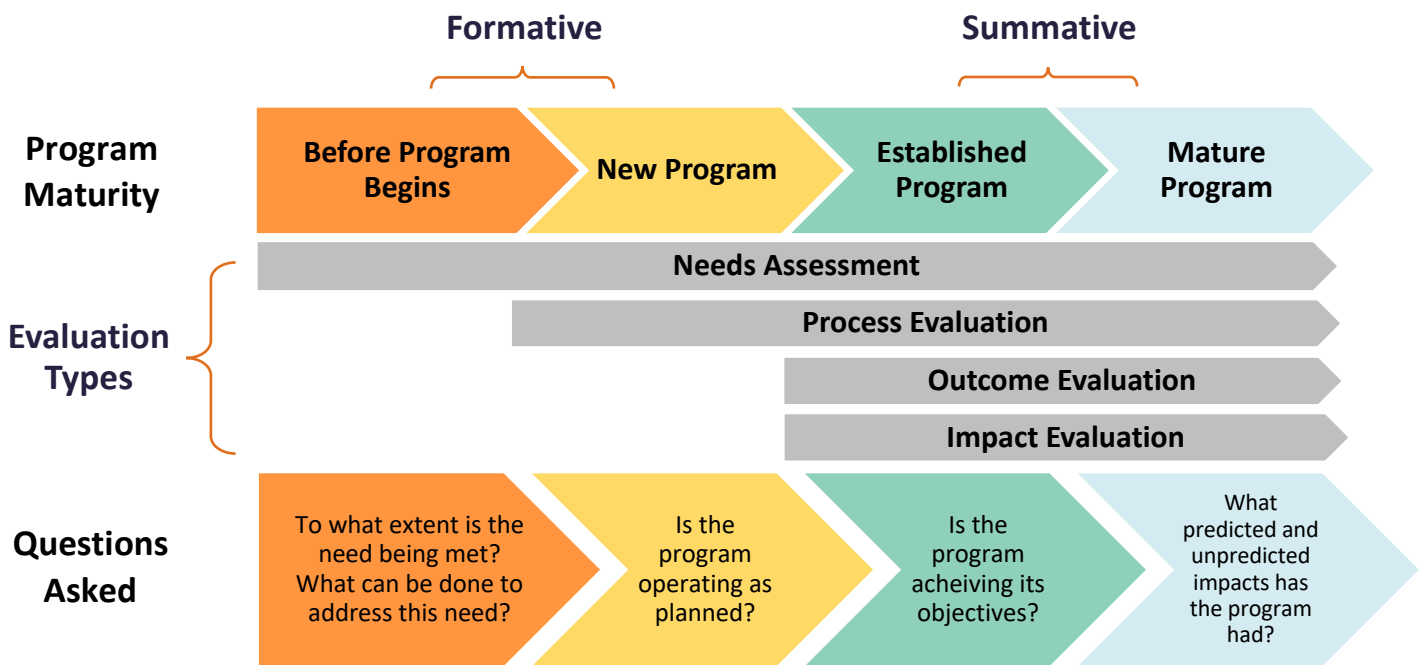
Adapted with permission from:

Centers for Disease Control and Prevention. (2014). *Building our understanding: Key concepts of evaluation: What is it and how do you do it.* https://www.cdc.gov/nccdphp/dch/programs/healthycommunitiesprogram/tools/pdf/eval_planning.pdf

SECTION II.
Analyze and Apply

What You Will Learn

- Utilize the Self-Assessment tool to pinpoint your organizational and program readiness for evaluation.
- Interpret your results on the Self-Assessment tool to discover which BeValued Toolkit sections can support your program's specific needs.
- Connect with other sections in the toolkit for evaluation and program development.

SELF-ASSESSMENT

The BeValued Self-Assessment is an important tool to help you identify your program's and IHE's current program evaluation phase. We developed a Self-Assessment tool with adapted questions from three main resources (see below for source details). The questions are designed to assess evaluation readiness based on the ways your organization and program currently operate.

There are three steps to use the Self-Assessment:

1. Answer the Self-Assessment questions.
2. Review the instructions on how to compute and compare your score to the ranges provided in the scoring matrix.
3. Based on your score, review the recommended evaluation phases in this toolkit to determine the next steps.

The Self-Assessment provides a recommended starting point based on your program's maturity, current capacity for evaluation, and various contextual factors (e.g., resource and staffing availability, support and buy-in, timing, and program objectives). However, you may also find it valuable to review the toolkit in its entirety before diving into the resources associated with your recommended evaluation phase. Since each phase of the toolkit is built upon subject matter in previous sections, reviewing the entire toolkit will allow you to re-familiarize yourself with concepts and develop a complete understanding of the content described in later sections.

Program evaluations are an ongoing effort. By using data, programs can be continually refined. Although you may be at a later phase of program evaluation now, revisiting various phases or sections of the BeValued Toolkit will be helpful as your program grows and changes. Take the Self-Assessment at multiple points throughout your program maturation to ensure that you are utilizing the most applicable tools and resources for your needs.

This Self-Assessment tool was created by adapting the following measures with permission:

- Program Evaluability Assessment Checklist – Penn State University Clearinghouse for Military Family Readiness[35]
- Impact Evaluability Assessment Tool – Zandniapour and JBS International[36]
- Are you ready? Assessing your organization's readiness to conduct evaluation – Manning and Goodman[37,H]

[H] This presentation consulted 7 additional tools in the development of their readiness tool, including:

- A Checklist for Building Organizational Evaluation Capacity (Volkov, B. B. & King, J. A., 2007).
- Capacity and Organizational Readiness for Evaluation (Core) Tool (Innovation Network, 2012).
- Evaluability Assessment: Examining the Readiness of a Program for Evaluation (Justice Research and Statistics Association, 2003).

Important Notes About Terms Used

In this Self-Assessment, we pair the terms "**your program and institution of higher education (IHE)**" to signify that the bystander intervention program fits within the larger context of the IHE and within a specific unit or division. As you review the content in this tool, be sure to consider the unique elements of your IHE and how your program fits into the broader and more complex institutional, local, or state policies, regulations, and mandates.

In addition, this tool asks questions about "**leadership**." You may ask: *Does this mean people who lead the bystander intervention program? The leadership of the division or unit? Or the institution's administration?* Leadership could mean any of all of those. What is important is that you answer the questions about leadership in a way that makes the most sense for your setting.

- Modified Evaluative Thinking Assessment Tool (Bruner Foundation, 2007).
- Monitoring and Evaluation (M&E) Capacity Assessment Tool (Capacity for Health).
- Planning Evaluability Assessments: A Synthesis of the Literature with Recommendations (Davies, R., 2013).
- The Readiness for Organizational Learning and Evaluation Instrument (ROLE) (Preskill, H. & Torres, R. T., 2000).

Scale #1: Organization and Program Maturity Scale

This is the first of two scales to help you determine where your organization and program fall in terms of evaluation readiness. It will point you to the sections of the toolkit that best fit your evaluation needs. This scale focuses on program maturity and development.

1. *Choose the number that best represents your answer to each question.*
2. *Enter that number in SCORE column.*
3. *Calculate your TOTAL SCORE for this scale. (If you are using the electronic toolkit, the score will auto-calculate.)*

Indicate to what extent you agree or disagree with each statement.					Add up this column to calculate.
Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree	SCORE
Organizational & Program Readiness					
Leadership demonstrates commitment to evaluation and evidence-based or data-driven decision making.					
1	2	3	4	5	
Leadership supports staff positions/activities that focus on evaluation, learning, and improvement.					
1	2	3	4	5	
The IHE provides opportunities for and fosters a culture of information sharing, discussion, reflection, learning, and improvement in order to support informed decision-making and practice.					
1	2	3	4	5	

Sulley, C., Rabideau, D., Orchowski, L., Jimenez, A., Dragoon, S., Susswein, M., Dube, S., Wood, L., & Busch-Armendariz, N.B. (2021). *The bystander evaluation (BeValued) toolkit: A resource to support program growth and impact.* Austin, TX: Institute on Domestic Violence & Sexual Assault, The University of Texas at Austin.

	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree	SCORE
Leadership is willing and committed to devoting necessary resources (e.g., staff positions, time, financial, and non-financial resources) to the evaluation.	1	2	3	4	5	
The program has a logic model (visual map that shows the relationships among the program resources/inputs, activities, outputs, outcomes, and impact).	1	2	3	4	5	
The group(s) the program is intended to benefit is(are) clearly identified.	1	2	3	4	5	
The program has an implementation plan (a guide or process for how the program should be delivered and supported).	1	2	3	4	5	
The program has clearly defined the short-term and long-term outcome(s) it aims to address (i.e., the intended benefits participants receive during and after program delivery).	1	2	3	4	5	

	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree	SCORE
Outcomes are defined in quantifiable, measurable terms.	1	2	3	4	5	
Program goals and objectives are related to program activities and are clearly defined, realistic, and attainable.	1	2	3	4	5	
There is a reasonable and shared expectation around the timeframe in which observable/measurable outcomes in the short term, intermediate, and long term will occur.	1	2	3	4	5	
There is staff capacity to operate the program. This means that staff are properly trained to operate the program, and that there are enough qualified staff members to implement planned activities	1	2	3	4	5	

The program is… Which of these options best describes the current state of your campus' bystander intervention program?	…currently being developed or re-worked and has not yet been launched on campus. **1**	…has been in operation for a short length of time and is building its presence on campus. **2**	…has been in operation for a reasonable length of time, but is less known on campus. **3**	…has been in operation for a reasonable length of time and is known on campus. **4**	…mature, stable, and not undergoing any major changes, and is well known on campus. **5**	
Total Organization & Program Score →						

Sulley, C., Rabideau, D., Orchowski, L., Jimenez, A., Dragoon, S., Susswein, M., Dube, S., Wood, L., & Busch-Armendariz, N.B. (2021). *The bystander evaluation (BeValued) toolkit: A resource to support program growth and impact.* Austin, TX: Institute on Domestic Violence & Sexual Assault, The University of Texas at Austin.

Scale #2: Program Structure and Evaluation Foundation

This is the second of two scales to help you determine where your organization and program fall in terms of evaluation readiness. It will point you to the sections of the toolkit that best fit your evaluation needs. This scale focuses on your program's structure and foundation to support an evaluation. 1. *Choose the number that best represents your answer to each question.* 2. *Enter that number in SCORE column.* 3. *Calculate your TOTAL SCORE for this scale. (If you are using the electronic toolkit, the score will auto-calculate.)*	*Indicate to what extent you agree or disagree with each statement.*					*Add up this column to calculate.*
	Strongly Disagree	**Disagree**	**Neutral**	**Agree**	**Strongly Agree**	**SCORE**
Program Structure & Evaluation Foundation						
There is interest and support among stakeholders in advancing the program's level of evidence by conducting an evaluation. Stakeholders see the value of evaluation and have ideas about how the program could benefit.	1	2	3	4	5	
The program has capacity (expertise, skills, staff time, finances) to conduct an evaluation internally or in collaboration with an external partner, or capacity to work with an external evaluator to plan and implement a program evaluation.	1	2	3	4	5	

	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree	SCORE
There is agreement and commitment from all necessary program staff and stakeholders regarding the collection and use of data that is needed for evaluation purposes, including data related to participant/beneficiary satisfaction, outcomes, and impacts.	1	2	3	4	5	
Evaluation questions are clearly stated, and they cover what key stakeholders (including program staff) want to learn about the program.	1	2	3	4	5	
There are structures and systems in place to systematically gather, store, analyze, and use data.	1	2	3	4	5	
The program uses forms to document services provided and activities/participation.	1	2	3	4	5	
The program has developed or adopted measures to assess its progress in achieving goals and objectives.	1	2	3	4	5	
The program has in place—or can develop—procedures to generate data needed for an evaluation.	1	2	3	4	5	
Program staff use (or plan to use) evaluation data outside our organization to increase program funding and share promising practices.	1	2	3	4	5	

Which of these options best describes the current state of your campus' bystander intervention program evaluation?	An evaluation plan has not been developed. 1	An evaluation plan has been developed, but has not yet been implemented. 2	A process evaluation is being conducted to assess the program. 3	An outcome evaluation is being conducted to assess program impact. 4	An outcome evaluation has been conducted to determine program impact. 5	
Total Program Structure & Evaluation Foundation Score						

Sulley, C., Rabideau, D., Orchowski, L., Jimenez, A., Dragoon, S., Susswein, M., Dube, S., Wood, L., & Busch-Armendariz, N.B. (2021). *The bystander evaluation (BeValued) toolkit: A resource to support program growth and impact*. Austin, TX: Institute on Domestic Violence & Sexual Assault, The University of Texas at Austin.

SCORING FOR PERSONALIZED RECOMMENDATIONS

First, take your **Total Organization & Program Maturity Score** from page 31 and compare that number to the scores in blue on the left-hand side of the scoring matrix to match your score with one of the three organization and program score ranges.

Next, take your **Total Program Structure & Evaluation Foundation Score** from page 34 and compare that number to the scores in yellow on top of the scoring matrix to match your score with one of the three evaluation score ranges.

The white box in the matrix where both of your scores meet gives you information on which sections of the BeValued Toolkit and which phases of program evaluation seem to be the most appropriate starting point given your responses on the Self-Assessment tool. Click on the hyperlinked text to be directed to that section of the toolkit or refer to the page numbers.

For example: If you score 26 on the Total Organization & Program Maturity scale and score 39 on the Total Program Structure & Foundation scale, those two scores will intersect in the top right corner of the scoring matrix, indicating your program would be best suited for the content and resources in Phases 3–5

Table 3. Self-Assessment Scoring Matrix

Scoring Matrix		Program Structure & Evaluation Foundation		
		Score: 10–23	**Score: 24–37**	**Score: 38–50**
Organization & Program Maturity	Score: 13–30	See Phases 0–1, starting on page 37.	See Phases 0–3, starting on page 37.	See Phases 3–5, starting on page 56.
	Score: 31–48	See Phases 0–2, starting on page 37.	See Phases 3–5, starting on page 56.	See Phases 5–6, starting on page 72.
	Score: 49–65	See Phases 0–2, starting on page 37.	See Phases 3–5, starting on page 56.	See Phase 6, starting on page 76.

The seven program evaluation phases in the BeValued Toolkit are:

- Phase 0: Prepare for an Evaluation
- Phase 1: Engage Stakeholders
- Phase 2: Describe the Program
- Phase 3: Focus the Evaluation
- Phase 4: Gather Information
- Phase 5: Analyze the Data and Make Conclusions
- Phase 6: Share Results

As a reminder, you do not have to go through the phases in chronological order or start with Phase 0.

SECTION III. Integrate, Act, and Create

What You Will Learn

- Understand the key tasks involved in the phases of evaluation that apply to your program's needs and how to access tools and resources for further support.
- Prioritize your next steps in designing/implementing a program evaluation based on the guidance provided.
- Draft plans and make decisions about implementing your program evaluation.

PHASE 0: PREPARE FOR AN EVALUATION

Before diving into a program evaluation, it is important to prepare and plan specifics and logistics, such as:

- **Evaluation lead and team.** Who is going to conduct the evaluation? Who else will be part of the effort?
- **Budget.** How much is your team able or willing to budget for an evaluation?
- **Timeline.** When do you expect the evaluation to occur? When do you expect to have results?
- **Buy in.** Is there broader support to evaluate this program? Support from leadership?
- **Expectations.** What is your team hoping to achieve by conducting an evaluation?

RESOURCES AND TOOLS

For more guidance in preparing for an evaluation:

- The CDC's *Evaluator Self-Assessment* [38]
- *What Works in Prevention: Principles of Effective Prevention Programs* [39]
- The CDC's *Developing an Effective Evaluation Plan: Setting the Course* [40]

Answering these questions can help set expectations and determine next steps.

Assembling a Team

Before starting a program evaluation, know who will be involved. There may be different levels of engagement among staff and other stakeholders—and you may continue to identify additional team members and stakeholders as you continue your evaluation planning—but having an initial idea of who will be involved is a key first step. Consider the staff, students, faculty, or consultants directly involved in delivering the program. They may be the ones leading the evaluation, collecting, or analyzing the data, and/or reporting on and using the results.

It is also important to pinpoint who will be leading the evaluation effort. Do they have the capacity (relating to workload and "know how") to manage a program evaluation? Is there a possibility that they may introduce bias by making assumptions about the program or drawing conclusions without completing an objective assessment? Sometimes, if the person leading the evaluation works with a consultant, a student, or other faculty/staff member on campus, they can avoid some of the capacity and bias issues. See the CDC's *Evaluator Self-Assessment* in the resource box above to determine if someone has sufficient capacity to

conduct program evaluation.[41]

After reviewing the preparatory steps in this phase, see also information in Phase 1 for guidance on building out roles for your evaluation team and lead.

Preparing a Budget

Identify at least a rough budget for your program evaluation at the start of planning. You don't need to decide what type of incentives to offer or if you want to utilize a specific software for data analysis at this point, but knowing how much you have to spend overall can help you determine if you have the budget to work with an external consultant, other program staff, or a graduate student.

Preparing a Timeline

Start with drafting a "big picture" timeline to set some initial expectations and get you thinking about some initial logistical considerations. Like with preparing a budget, it is not necessary to know all the details of your timeline for evaluation just yet, but it is helpful to start the process. For example, if a graduate student is leading the evaluation and is the only one with the time to analyze the data, it might take them more time to complete program evaluation activities than a dedicated staff member or a consultant. This may be especially true at certain times of the academic year (e.g., midterms and finals).

> **KEY INFORMATION**
>
> **Incentives**
>
> Incentives are an important consideration for program evaluation participation. Incentives can be tailored to fit each IHE. For example, a large commuter school may want to consider drawings for parking passes, whereas a school with a prominent sports team may want to consider athletic ticket giveaways.

Getting Buy-In

Your organization's commitment to evaluating your program is another key element to assess up front. Does your department expect you to report on the progress of your program? Does leadership allow, encourage, and/or expect you to dedicate time and resources to conduct your evaluations? Phase 1 will more directly address how to engage stakeholders, including leadership, but just knowing the answers to these questions now can give you an idea about how easy or difficult it may be to accomplish an evaluation.

Setting Expectations

Taking into consideration your initial ideas for the budget, timeline, and support for program evaluation, now consider what you and your team are expecting to get out of a program evaluation.

Ask yourself:

- What are the general goals of the evaluation?
- Who is the audience or end user for the evaluation findings?
- What resources are available to support the evaluation?
- How deep and thorough can the evaluation be given available funding and resources?
- What are the appropriate expectations to set with your team, with your end user or audience, and with your leadership?

Setting appropriate expectations early on can contribute to a successful program evaluation. For example, while working with a student may be economical and can produce some helpful information for a program, a student's evaluation efforts will not be as robust or informative as a well-funded evaluation conducted by a trained professional. Similarly, if you expect to report the results of an evaluation that used three months of data, your findings will not be as thorough or deep as an evaluation that looks at data from multiple years. While you may only want or need short-term data, depending on the purpose and type of evaluation you are conducting (see Types of Program Evaluations in Section I), it is important to set and discuss evaluation expectations with stakeholders to make sure everyone is on the same page with the plans and goals.

See the tools on the next page and resource box at the start of Phase 0 for additional guidance on preparing for your program evaluation.

TOOLS

Phase 0 Checklist

Checklist for Phase 0: Prepare for an Evaluation	
✓	**TASK**
	Identify and list potential program evaluation team members, including program staff, stakeholders, and/or consultants.
	Review that list and identify key individuals who may qualify as the evaluation lead.
	Engage potential evaluation team lead(s) to assess their interest and capacity to lead this program evaluation.
	Create an initial budget for program evaluation activities.
	Create an initial timeline for program evaluation activities.
	To gain buy-in, provide information about the evaluation and engage in discussion with key stakeholders, like leadership and influential individuals.
	Discuss, clarify, and develop consensus with stakeholders about the evaluation plan, its goals, and eventual outcomes.
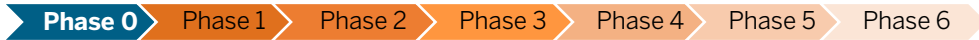
Phase 0 Worksheet: Evaluation Budget Template

This template can help guide you in developing your evaluation budget. Fill in each applicable line item. Feel free to add additional line items specific to your program.

Income & Expense Projections	Amount (Year 1)
Income	
General	
Grants	
Sub-Total (Income)	
Other Income	
Total Income	
Expenses	
Evaluation Expenses	
Staff Salary (percentage of appointment)	
Student Worker	
Consultant	
Total Salaries & Wages Expenses for Evaluation	
Other Expenses	
Evaluation-Related Travel	
Supplies & Materials	
Postage	
Printing & Copying	
Dues & Subscriptions	
Office Supplies	
Staff Training	
Total Operations Expense	
TOTAL EXPENSES	
Balance	

Adapted with permission from: Centers for Disease Control and Prevention. *WISEWOMAN evaluation toolkit*. Atlanta, GA: Centers for Disease Control and Prevention, US Dept of Health and Human Services; 2015. https://www.cdc.gov/wisewoman/docs/WW_Evaluation_Toolkit.pdf

PHASE 1: ENGAGE STAKEHOLDERS

Stakeholders can move a program and program evaluation forward in several ways. They can help:

- Obtain resources for the program/evaluation.
- Plan and implement the evaluation.
- Develop and prioritize evaluation questions.
- Identify data sources and assist in participant recruitment.
- Interpret and use evaluation results.
- Increase awareness and support for the program/evaluation.
- Increase credibility for the program/evaluation and reduce distrust.

Stakeholder Roles and Types

The ways a stakeholder can help a program and/or program evaluation often depend on what kind of stakeholder they are. Different types of stakeholders may also interpret and use program evaluation results differently. Table 4 lists examples of different types of stakeholders that may be involved in bystander intervention programs. For example, **students** may assist in the recruitment for the program/evaluation and may reference evaluation results incorporated into program marketing and messaging. **Department directors** may help prioritize evaluation questions and secure funding for a program evaluation, and they might use evaluation results to promote the program to campus leaders to secure more resources for the future. **Community practitioners** or other community partners can broaden the reach of the program or evaluation, enhance a needs assessment, or offer other contextual data, and can help to align IHE programming with community efforts to prevent harm and violence.

When engaging stakeholders, carefully consider your goals in doing so. It is best to strategically engage individuals rather than engaging everyone right away. First, work with a smaller group to establish a general foundation and understanding of your evaluation plan before broadening your stakeholder group. Wait to engage additional people until you are ready to do so. Also, consider the influence needed to achieve your evaluation goals and identify potential stakeholders best suited for your goals (e.g., aspirational thinkers, content specialists, strategic planners). A mix of skills and abilities to influence and inspire will create a dynamic group with the capacity to push the program evaluation goals forward.

RESOURCES AND TOOLS

For more guidance in preparing for an evaluation:

- RWJ Foundation's *A Practical Guide for Engaging Stakeholders in Developing Evaluation Questions* [42]

- *Grassroots Action and Learning for Social Change: Evaluating Community Organizing* [43]

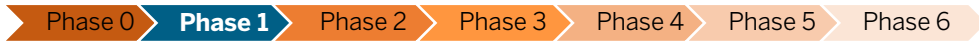- *Researchers, Practitioners, and Funders: Using the Framework to Get Us On the Same Page*[44] (access through your institution)

Table 4. Types of Stakeholders

Type	Description	Examples
Implementer	Program staff and others directly involved in the delivery and operations of a program	• Peer educators • Other student employees, interns, or volunteers • Program Coordinator/Specialist/ Manager • Program Director
Decision Maker	Those in a position to do or decide something about the program	• Program Director • Department Director • Associate Vice President • Associate Dean of Students
Partner	Those who support or are actively invested in the program	• Title IX Coordinator • Faculty/Staff • Funder • Community practitioners
Participant	Individuals who are served by the program	• Undergraduate Students • Graduate Students • Faculty/Staff
Supporter	Individuals interested in the program or its evaluation results	• Professionals working in sexual violence prevention or another relevant harm prevention field • General community members or practitioners who wish to learn from the IHE's work • Policy makers and committee or task force members working on issues related to your program

Adapted with permission from: Centers for Disease Control and Prevention. *WISEWOMAN evaluation toolkit.* Atlanta, GA: Centers for Disease Control and Prevention, US Dept of Health and Human Services; 2015. https://www.cdc.gov/wisewoman/docs/WW_Evaluation_Toolkit.pdf

Engaging a diverse group of stakeholders in evaluation planning and implementation can result in all the practical benefits discussed, strengthening your program and program evaluation approach.

| Phase 0 | Phase 1 | Phase 2 | Phase 3 | Phase 4 | Phase 5 | Phase 6 |

The Importance of Engaging Community Practitioners

It is valuable to include community partners as stakeholders. Depending on your IHE and its surrounding community, you may have already established community partners. These partnerships are important so that various prevention programs can be on the same page and learn from one another. In addition, data sharing with community agencies can help enhance evaluation activities (i.e., needs assessments).

How Many Stakeholders to Engage

As noted earlier, it is important to consider how many stakeholders you engage. Stakeholder group size may vary depending on your goals and capacity. Having a small group of five (5) individuals that are new or less familiar with your program may help you develop trust and rapport. A large group of 20 may bring varying perspectives and campus-wide reach. Preparation and good communication are vital to high-quality stakeholder engagement. Remember to practice facilitation before meetings, and communicate with individual stakeholders one-on-one when building relationships and buy-in or when vetting sensitive information.

Building Buy-In

As discussed in Phase 0, engaging stakeholders also means working to build their commitment or "buy-in" for the program and/or the program evaluation. This may mean having repeated discussions, using persuasive communication, and raising stakeholder awareness on the value of bystander intervention programming and the benefits of conducting a program evaluation. In all these interactions, it is helpful to use what you know about each stakeholder so that your pitch makes sense (to them). If you thoughtfully meet stakeholders where they are (e.g., appreciate their experiences and language; intentionally work to be inclusive and considerate), you can help them understand what you are doing, what you need, and how they can help. You also avoid a disconnect in expectations about your program or evaluation.

KEY INFORMATION

Also Engage Critics

The CDC recommends engaging with potential critics of the evaluation as well as with supporters. Critics may identify issues related to program or evaluation activities and offer an opportunity for you to strengthen your process.

Engaging with individuals who may critique or be apathetic to your efforts can also highlight areas of dissonance or problems with assumptions being made.

For example, someone may have a different idea of what the bystander program should be trying to accomplish and reporting on than you or other program implementers do. Having a conversation and engaging with them may clear up confusion and open up new opportunities for program growth and partnerships.

Another strategy to enhance buy-in with stakeholders is to utilize a "data give-back" approach. This refers to your ability to collect, analyze, and share your program's evaluation findings to establish a positive rapport with professional stakeholders, students, and the broader campus community. See Phase 6: Share Results for more information.

Engaging Faculty and Staff

In our work with UT academic institutions, program staff across institutions noted a mutual goal to engage faculty and staff, both as bystander program stakeholders and as participants. It is beneficial to involve faculty/staff in strategic ways so that they help extend the reach of the program beyond student audiences, act as the first step in expanding impact, and enhance the "buy-in" needed for program sustainability.

Example in Practice

UT Austin has started doing this in a number of ways, including presenting at a symposium held by the Faculty Innovation Center. A faculty member and a staff member who are both trained in bystander intervention programming through BeVocal presented to faculty on the BeVocal model, and then worked with the audience through three bystander-related scenarios (seeing harm, causing harm, and receiving harm). As a result, program staff were able to assess the nuances between faculty, staff, and student training needs. The symposium was a tipping point for staff and faculty members to become more engaged participants and stakeholders in BeVocal, furthering program support and partnerships.

Source: Russell, M., & Kester, B. (2019, January). *Understanding and engaging faculty: UT Austin's BeVocal sharable project.* The University of Texas System Bystander Intervention Conference.

See the tools that follow and the resource boxes earlier in this section for additional guidance on engaging stakeholders.

TOOLS
Phase 1 Checklist

Checklist for Phase 1: Engage Stakeholders	
✓	**TASK**
	Identify and list potential program evaluation stakeholders. These stakeholders should represent various departments and offices across campus.
	Review list of stakeholders and identify key individuals who may improve credibility, program implementation, advocacy, or funding/authorization decisions. Also, consider who has the influence to help you achieve your evaluation goals.
	Reach out to individual stakeholders and/or 1–2 representatives from key stakeholder organizations to assess their interest and capacity to participate/contribute
	Create an engagement plan for stakeholders. This should include your goals for engagement and your target stakeholder group size (a small group of 5 individuals or a larger group of up to 20 or more).
	Implement the stakeholder engagement plan.

Phase 1 Worksheet: Stakeholder Engagement Plan

Stakeholder Name	Title or Department	Evaluation Interest	Level of Participation	Role	Preferred Mode of Communication & Frequency	Other Notes
EXAMPLE: Malika Smith	IDVSA	Short-term & intermediate outcomes of UT bystander programs	High	Provide consultation, technical assistance, & training for UT bystander program staff	E-mail & phone calls as needed Quarterly Zoom or phone call update/check-in meetings	IDVSA may also be able to provide support for data analysis

Adapted with permission from: Centers for Disease Control and Prevention. *WISEWOMAN evaluation toolkit.* Atlanta, GA: Centers for Disease Control and Prevention, US Dept of Health and Human Services; 2015:

https://www.cdc.gov/wisewoman/docs/WW_Evaluation_Toolkit.pdf

PHASE 2: DESCRIBE THE PROGRAM

Before fully planning and implementing a program evaluation, all key stakeholders—especially the evaluation lead—must have a solid understanding of your bystander intervention program. By following the guidance here in Phase 2, you have an opportunity to get all stakeholders on the same page about the program, its purpose and goals, its scope, and its expected outcomes.

Narrative

A narrative description of the program can provide context for the program and evaluation. The program narrative can include information like:

- The need for the program
- Program mission
- Program objectives
- Program audience(s)
- The program's maturity and development
- The program's setting
- Resources available

A one-page narrative summarizing some of the bullet points above is an excellent way to concisely summarize your program and evaluation efforts and reach a larger audience. This one-pager can be sent via email or printed and passed out on your campus to increase program awareness and establish expectations.

Intervention Mapping for Program Development

A theoretical and research-informed understanding of the risk and protective factors that underlie a given health problem is central to developing an intervention program, such as a bystander intervention program. This includes identifying and grasping the **theory or theories of change** that ground your program approach (i.e., bystander intervention theories of change). One design and planning approach to intervention development is known as "intervention mapping," which breaks program design and development into several steps.

Those steps are:

- Develop a logic model of the problem based on a needs assessment.
- State program outcomes and objectives.
- Develop the program implementation plan, including scope and sequencing.
- Produce the intervention content, including the program manual, materials, and messages.
- Plan program use, including program acceptance, implementation, and maintenance.

Most of the intervention mapping steps are "heavy lifts" and require a large amount of planning and resources. However, intervention mapping provides a valuable roadmap for program planning and development. *Clearly planned interventions prepare you for a clearly planned evaluation of the program.* More detail about this process is provided in the two resources about intervention mapping in the sidebar above.

Note that if you are using or adapting a pre-packaged, research-based program, you can still follow the steps of intervention mapping. Start by reviewing all the information provided in the package and checking to be sure all the information is present for intervention mapping. Work with the company or organization that the package was purchased from if you need additional materials or explanations. Fill in any missing information as needed based on your analysis.

RESOURCES AND TOOLS

For more guidance on intervention mapping and theory of change:

- Harvard Family Research Project's *Introduction to Theory of Change* [45]

- *Planning Health Promotion Programs: An Intervention Mapping Approach* [46]

- *Intervention Mapping: Theory- and Evidence-Based Health Promotion Program Planning* [47]

Logic Model

A program logic model is a critical tool to help stakeholders thoroughly think through and describe all the pieces of a program and link these to specific outputs and outcomes that you expect will occur as a result of the program. Logic models are *different from strategic planning* because they: include performance metrics or benchmarks to ensure a program is on track. They include both outputs and longer-term outcomes with information on how to achieve those. Logic models detail the program's specific expected:

- **Inputs** – the funding, staffing, and other resources used by a program.
- **Activities** – what exactly will be done during the program.
- **Outputs** – the tangible results of the activities.
- **Outcomes** – the expected results of a program, which can be short-term, intermediate, and long-term.
  - o **Short-Term Outcomes** – changes in psychosocial behavioral determinants.
  - o **Intermediate Outcomes** – changes in behaviors or policies.
  - o **Long-Term Outcomes** – changes in health indicators and/or prevalence rates.

> **RESOURCES AND TOOLS**
>
> For more guidance on logic models:
>
> - The CDC's *Learning to Love Your Logic Model* [video][48]
> - University of Wisconsin's *Enhancing Program Performance with Logic Models* [virtual course][49]
> - *Logic Models as a Tool for Sexual Violence Prevention Program Development* [50]
> - W. K. Kellogg Foundation's *Using Logic Models to Bring Together Planning, Evaluation, and Action: Logic Model Development Guide* [51]
>
> Also see the templates and examples in the **Tools section** at the end of Phase 2.

Logic models can be extremely valuable in helping to guide program evaluations too. For example, process evaluations can assess what is laid out in a program's inputs, activities, and outputs columns to see if the program was implemented as intended. Outcome evaluations can look at the short-term, intermediate, and long-term outcomes columns to assess program effectiveness in those areas. In addition, logic models can help to focus an evaluation by using your specific program details to determine what aspects/factors/elements of the bystander intervention program you want to evaluate. There are specific indicators that help determine this, such as SMART goals, (specific, measurable, achievable, relevant, time-

bound) which serve as statements for program activities and outcomes and help guide the development of a logic model.

See Figure 9 for an example logic model for a bystander intervention program.

Logic Model vs. Theory of Change

There is often some confusion between the terms "logic model" and "theory of change." A logic model should be a graphical illustration that helps all program stakeholders clearly identify outcomes, inputs, and activities. The theory of change links the outcomes and activities to explain how and why the desired changes are expected to come about. A theory of change requires that you justify why something—an attitude change, a skill built—is expected to cause a reduction in a health risk or harmful behavior. For example, a theory of change would suggest what needs to occur for sexual violence to be prevented on the campus. Table 5 summarizes the difference between logic models and theories of change.

Bystander intervention programs generally already follow a specified theory of change. Other harm prevention programs on your campus may also follow a similar theory of change, and a partnership could be synergistic and beneficial.

> **KEY INFORMATION**
>
> **SMART Goals:**
> - Specific
> - Measurable
> - Achievable
> - Relevant
> - Time-bound
>
> By creating goals in this way, you are able to monitor progress toward program goals clearly.
>
> An example of a SMART goal:
> *Train 5 freshman facilitators in bystander intervention by Spring 2022.*
>
> The CDC's *Writing SMART Objectives* [52]

Table 5. Logic Model vs. Theory of Change

Logic Model	Theory of Change
• Graphic Representation/Visuals	• Explanatory
• Lists of components	• Hypotheses of how change will occur
• Descriptive	

See the tools and sidebars in this section for additional guidance on describing your program(s), including on creating logic models.

Multiple Programs and Logic Models

If you are using separate programs that, together, create a comprehensive prevention package, there may need to be a separate logic model for each program specific to the goals and activities of that program, and then a larger logic model that shows how each of the programs fit together into a larger prevention package and approach.

In the article in the callout box below, the authors discuss the integration of sexual assault resistance, bystander, and men's social norms programs. Each program has its own logic model and the authors show how the programs fit together synergistically.

RESOURCES AND TOOLS

For an example of integrating multiple harm prevention programs:

- *Integrating Sexual Assault Resistance, Bystander, and Men's Social Norms Strategies to Prevent Sexual Violence on College Campuses: A Call to Action* [53]
  (access through your institution)

TOOLS

Figure 9. IDVSA's Example of a Bystander Intervention Program Logic Model

Inputs	Activities	Outputs	Short-Term Outcomes	Intermediate Outcomes	Long-Term Outcomes
Funding	Create bystander intervention workshop presentations	- # of topics addressed - # of workshop presentations created	Participants have increased knowledge of the prevalence of violence, the spectrum of harm, available resources, and the importance of community-level prevention	Participants speak up and intervene more in potentially harmful situations or interactions	Fewer individuals engage in potentially harmful behaviors, or use potentially harmful language
Personnel (staff/students/volunteers)	Deliver bystander workshops	- # of workshop requests - # of workshops delivered - # of workshop participants (individuals and groups) - # of incentives distributed	Participants have increased knowledge of bystander intervention, potential barriers, potential motivators, and tactics	More individuals speak up and intervene in potentially harmful situations or interactions	Fewer reports of hate speech, harassment, and biased incidents
Facilities	Create bystander initiative messages	- # of topics addressed - # of social media posts created - # of posters created - # of flyers/cards created	Participants have increased self-efficacy (or confidence) to speak up or intervene in potentially harmful situations or interactions	Participants provide more support to survivors of violence or harm	Fewer reports of interpersonal violence and sexual assault
Materials & Supplies	Deliver bystander initiative messages	- # of social media posts posted - # of posters posted - # of flyers/cards handed out - # of other groups delivering initiative messages - # of initiative messages other groups deliver	Participants think that their peers would speak up or intervene in potentially harmful situations or interactions	More individuals provide support to survivors of violence or harm	Fewer reports of hazing, high-risk drinking, academic dishonesty, and suicide
Devices & Software	Train bystander initiative peer educators/facilitators to deliver workshops	- # of peer educators/facilitators recruited - # of peer educators/facilitators trained	Participants have more positive attitudes about intervening in potentially harmful situations or interactions	Participants engage in more proactive prevention efforts (e.g., attending rallies, joining advocacy groups, changing policies)	A dedicated social media platform for bystander intervention programs run by peer educators
Incentives (money, tickets, t-shirts, other swag)	Launch peer educators/facilitators	- # of peer educators/facilitators delivering workshops - # of workshops each peer educator/facilitator delivers	Participants don't feel they have as many barriers, or feel they can overcome barriers, to intervene in potentially harmful situations or interactions	More individuals engage in more proactive prevention efforts (e.g., attending rallies, joining advocacy groups, changing policies)	All trained participants feel confident and responsible to intervene in potentially harmful situations
Leadership & Community Support	Build cross-campus partnerships	- # of meetings with potential partners - # of partners	Participants have more intentions to intervene if they encounter potentially harmful situations or interactions	A community of committed bystanders	A culture of campus safety with fewer incidents of violence on campus

Sulley, C., Rabideau, D., Orchowski, L., Jimenez, A., Dragoon, S., Susswein, M., Dube, S., Wood, L., & Busch-Armendariz, N.B. (2021). *The bystander evaluation (BeValued) toolkit: A resource to support program growth and impact.* Austin, TX: Institute on Domestic Violence & Sexual Assault, The University of Texas at Austin.

Phase 2 Checklist

| ✓	TASK
Checklist for Phase 2: Describe the Program	
	Write a comprehensive program description, including need, mission, objectives, audience(s), program maturity/development, context, resources, and activities.
	Complete the steps of intervention mapping.
	Ensure that program staff members have a foundational understanding of the program's theory of change.
	Create a program logic model, including inputs, activities, outputs, and outcomes.

Phase 2 Worksheet: Logic Model Checklist

Fill out this checklist to assess whether you have completed the necessary tasks to create a complete logic model that will be an effective tool for program planning, implementation, and evaluation.

☐   Have you stipulated the specific needs or desired outcomes for the bystander intervention program or programs?

☐   Have you included the inputs/resources that go into the program? For example:
- Staff, peer educators, consultants/speakers
- Materials, curriculum, brochures, trainings needed
- Funding (budget, grants)
- Participants (peer educators, staff, faculty, students, administrators)
- Program setting (campus or community settings, online forums)

☐   Have you identified the various locations where the program(s) will be conducted?

☐   Have you identified the number and length of program activities?

☐   Have you described an expected unit of service for each activity in your program?
- Number of students served
- Type of training provided
- Training audience (students, staff, faculty, other)
- Hours of service provided

☐   Does each output describe an expected number of participants served?
- For example: 2 staff trained, 15 peer educators trained, 144 staff/faculty trained
- Do you have a plan in place to count these targets over time?

☐   Does the dosage and intensity of the programming that is planned seem appropriate?

☐   Are the planned activities well matched to your desired outcomes? Examine both.

☐   Are program outcomes described using specific, observable terms that relate to changes in knowledge, attitudes, beliefs, skills, or behaviors?

☐   What follow-up efforts for groups needing more assistance will you include to address gaps in the program?

☐   When you look at the completed draft logic model, do the activities fit well together? What is missing? Do they build upon one another in a staged way to advance participant knowledge over time? Is the program likely to be synergistic? What can be done to knit together the different program activities into a more comprehensive package?

Phase 0 ⟩ Phase 1 ⟩ **Phase 2** ⟩ Phase 3 ⟩ Phase 4 ⟩ Phase 5 ⟩ Phase 6
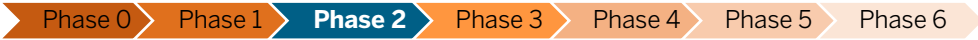
Phase 2 Worksheet: Logic Model Template[54]

Name of Program:

Problem Statement:

Inputs/Resources (What you have now)	**Needed Capacity** (What you need to build)	**Strategies/ Activities**	**Outputs**	**Outcomes**		
				Short-term or immediate *changes…*	Intermediate *changes…*	Long-term *changes…*

Impact

Contextual and Influential Factors

Phase 0 〉 Phase 1 〉 Phase 2 〉 **Phase 3** 〉 Phase 4 〉 Phase 5 〉 Phase 6

PHASE 3: FOCUS THE EVALUATION

As you begin to plan your program evaluation, it is essential to figure out what you and your team want the focus of the evaluation to be. Answering these questions can help you narrow down your evaluation so it is more manageable, matches your expectations, and can better support your program:

- What is the **purpose** of the evaluation?
- What **information** are you hoping to get out of the evaluation?
- What **existing data**/information is already available to you?
- Do you have an idea for your **evaluation plan** (or design)?

RESOURCES AND TOOLS

For more guidance on focusing your evaluation:

- University of Wisconsin's *Planning a Program Evaluation* [55]
- *Getting the Evaluation Focus Clear* [56]
- *Handbook of Practical Program Evaluation* [57]

Purpose

One of the first things to do is to determine the purpose of this program evaluation and how results will be used.

- Will results be shared with campus leaders to show program impact?
- Will results only be used internally to improve program activities?
- Will results be used for any reporting requirements?

The evaluation's purpose can also be determined by stakeholder needs and expectations, the context of the program and evaluation (e.g., the program's stated mission and objectives), what resources you have available to you, and the maturity of the program.

Knowing the purpose of the evaluation can also help narrow down which type of evaluation is most appropriate for your current goals and expectations. The types of program evaluations include:

- Needs assessments
- Process evaluations
- Outcome evaluations
- Impact evaluations

To learn more about each of these types of program evaluations—as well as the difference between formative and summative evaluations—go back and review Section I, particularly the Types of Program Evaluations information. It provides additional details, examples, and guidance on when to use each type of evaluation.

Evaluation Questions

Developing evaluation questions is an important next step after narrowing down the focus of your program evaluation. Stakeholders can help you brainstorm an initial list of potential questions you want to ask during your evaluation, then help prioritize and refine those questions.

When you develop and prioritize evaluation questions, it is also important to make sure they can be logically traced back to your actual program activities. For example, an evaluation question that asks about interactive theater performances won't help a program that only has workshops.

Below are a few **example evaluation questions** developed using the example logic model from Phase 2 (Figure 9). As you can see, the evaluation questions cover a mix of outputs and outcomes.

- How many requested workshops were delivered per semester?
- How many participants were in each workshop?
- What was the program's effect on participants' self-efficacy?
    - What was the program's effect from pre-test to post-test?
    - What was the program's effect two months after program exposure?
    - What was the program's effect six months after program exposure?

Choosing the right **timeframe** is also important to be able to assess program effects. To document change created as a result of bystander intervention programs, participants must have a chance to see a harmful situation and intervene. If short-term follow-ups occur, say a week or one month after a workshop or other intervention, that may not allow for the opportunity for a bystander intervention to occur. For this reason, *it is important to be thoughtful as well as pragmatic about the best follow-up window*. If there is not enough time to survey behavior change among participants, then consider other measures of intentions and confidence, such as self-perceptions or peer perceptions.

Existing Sources of Data

A program may not need to collect new data to answer some evaluation questions but instead can use sources of available institutional data. For example, if you are looking at the impact of bystander intervention on medical emergencies, you could ask other departments for information, such as the number of calls to the emergency medical service on campus. If you are trying to assess program effects on high-risk drinking, you might request data on the number of Resident Advisors (RAs) involved in alcohol-related concerns at various points in time. Consider what sources of data might be available to you to support your program evaluation efforts.

Indicators

We recommend defining indicators for program achievements when designing your program evaluation and determining program success. Indicators are typically statements that follow the SMART template; they are: specific, measurable, achievable, relevant, and time-bound (i.e., there should be a timeline). Depending on the maturity of your program and the model behind it, you may develop indicators by looking at what your program was able to accomplish in previous years, what similar programs and similar institutions are

able to accomplish, and/or what the literature says we should expect.

> **KEY INFORMATION**
>
> **Campus Climate Survey Data**
>
> The UT CLASE Survey of Prevalence & Perceptions used the ARC3 instrument as its foundation. The ARC3 Campus Climate Survey includes questions related to bystander intervention perceptions and attitudes, baseline knowledge about the presence of a bystander intervention program on campus, and how students obtain bystander related information.
>
> If your campus has done a campus climate survey, inquire about the protocols, data, and results. The information can be helpful for planning, measuring knowledge about the program's availability, offering ideas to better reach students, and identifying content to work on with the campus community.

Outputs vs. Indicators

To explain further, indicators are different from outputs in the logic model. Think of indicators as an *indicator of progress toward the desired goal*. Outputs are what you did, while indicators are how well you did it. In other words, a program might have certain outputs but does not meet its indicators of success. For example, let's say you completed 10 program sessions, but the facilitators only showed 70% consistency in measurement. The planned outputs might be 10 programs administered, and the indicator of progress is that across these 10 programs, 75% of program participants showed an increase in score. An indicator allows you to show the goal you want to hit with program outputs.

Sulley, C., Rabideau, D., Orchowski, L., Jimenez, A., Dragoon, S., Susswein, M., Dube, S., Wood, L., & Busch-Armendariz, N.B. (2021). *The bystander evaluation (BeValued) toolkit: A resource to support program growth and impact.* Austin, TX: Institute on Domestic Violence & Sexual Assault, The University of Texas at Austin.

Below are some example indicators developed for a bystander intervention program:

- The program will deliver three open workshops on campus each semester during the academic year.
- Over the next six months, five freshman facilitators will be trained to deliver the program to freshman peers.
- Facilitators will achieve 90% scores on assessments about program knowledge to ensure consistency in message delivery.
- From pre-test to post-test, in trainings provided during the academic year, 75% of program participants will increase their score for intention to intervene.

Evaluation Design and Methods

Focusing your evaluation also includes figuring out what your evaluation will look like: your evaluation design. Evaluation design is the way in which you structure your evaluation activities and the data collection/research methods you use in order to answer the evaluation questions. Your evaluation design will, in part, depend on the resources you have available, the training/expertise of the evaluation team, your capacity to conduct the evaluation, how rigorous you want the evaluation to be, the evaluation type you have selected (e.g., impact evaluation), and what kind of information you want to get from the evaluation.

Table 6 includes a description of common types of evaluation designs.

Table 6. Evaluation Designs

Design	Description	Advantages	Disadvantages
Experimental	Two groups with random sampling/ random assignment	The gold-standard for evaluation due to a rigorous research design and ability to infer cause and effect. If you have the resources and time, strongly consider this design.	It may not be feasible to randomly assign students to groups. There also may not be an alternative program available as a comparison or not enough participants in a program to make a reliable comparison between groups. Statisticians call this being "underpowered" to detect a difference between groups. It also takes more resources to run two different program groups at the same time.

Table 6. Evaluation Designs (continued)

Design	Description	Advantages	Disadvantages
Quasi-experimental	Two groups; nonrandom	Helpful in detecting program effect. If there are two existing programs, this design takes advantage of having a natural comparison group for each program.	Many of the disadvantages of experimental designs apply here as well (i.e., it takes more resources to run two program groups; there may not be enough participants to compare in two groups). Also, because the design does not include random assignment, if a difference is detected between groups, it may not be possible to say that the effect was due to the program itself.
Observational	Cross-sectional survey, case studies	Provides a snapshot of information	Introduces bias from the observer's perception or from the perspective of the participants themselves (if using data from a case study).
Pre-test / Post-test	Pre-tests tell you where participants are starting from, and post-tests give information on how much participants change after a program intervention. To report short-term change and program effect, include a pre-test and post-test.	This is a basic program evaluation design that can be used with almost any program. Pre- and post-test surveys can readily be administered to any audience immediately before and after a program by asking participants to complete the same survey items at each point in time.	Does not provide information about behavior change. If a post-test is administered immediately after a program, there is a possibility for "demand characteristics" (i.e., participants rate the program in an overly positive manner). Also, post-test immediately after a program may tell you more about how participants liked a program, rather than whether attitudes have truly changed behaviors—time must pass to determine that.

Next, see the tools and the resource boxes in this section for additional guidance on focusing your evaluation.

Phase 0 > Phase 1 > Phase 2 > **Phase 3** > Phase 4 > Phase 5 > Phase 6

TOOLS

Phase 3 Checklist

✓	Checklist for Phase 3: Focus the Evaluation
	TASK
	Describe the purpose of the evaluation, including potential users and uses.
	Select the parts of your program logic model that should be part of the focus of the evaluation.
	Decide what type of evaluation to conduct.
	Write evaluation questions and review them with stakeholders.
	Create or adapt indicators and review them with stakeholders.
	Develop the evaluation design and make sure it answers the evaluation questions.

Phase 3 Worksheet: Logic of Evaluation Questions or Indicators

Logic Model Component	Indicators (I) and Evaluation Questions (Q)	Potential Data Sources
EXAMPLES: Workshops Delivered	I: Three (3) open workshops will be delivered on campus each semester. Q: How many requested workshops will be delivered per semester? Q: How many participants will there be in each workshop?	Program activity logs Workshop request forms Sign-in sheets
Quality of Workshops Delivered	I: Workshops will be delivered with 85% fidelity to the planned script. Q: Is the anticipated content of the workshop consistently provided to participants?	External ratings of the extent to which a facilitator addressed each item in a script. The facilitators' own rating of the content covered in the program
Utility of the Workshops	I: 85% of participants will report that they are satisfied with the content of the workshop. Q: Do workshop participants find the content of the program useful?	Post program rating forms completed by the participants

Adapted with permission from: U.S. Department of Health and Human Services Centers for Disease Control and Prevention, Office of the Director, Office of Strategy and Innovation. (2011). *Introduction to program evaluation for public health programs: A self-study guide.* https://www.cdc.gov/eval/guide/index.htm

Phase 0 | Phase 1 | Phase 2 | Phase 3 | **Phase 4** | Phase 5 | Phase 6

PHASE 4: GATHER INFORMATION

After you hone in on the focus and design of your program evaluation, further specify what information you are hoping to gather. As discussed earlier, consider your stakeholders' needs and the importance of the context in which you are gathering information. In addition to talking with stakeholders, you can also consider these key elements that relate to the program, its ability to meet its goals, and the larger context it operates within:

- The participants' needs
- The program's stated mission and objectives
- The program's protocols and procedures
- Similar programs and/or control or comparison groups
- The values, goals, and norms of your institution and your community
- Any relevant institutional, local, or state policies, regulations, and mandates
- If you will need to go through your university's institutional review board (IRB)
- How you will obtain consent from participants to participate in the evaluation
- Potential ethical implications of your program evaluation and data collection

> **RESOURCES AND TOOLS**
>
> For more guidance on collecting data:
>
> $ • _Data Collection: Planning for and Collecting All Types of Data_ [58]
>
> ☁   • _Likert-Type Scale Response Samples_ [59]
>
>   • The CDC's _Developing an Effective Evaluation Plan: Setting the Course_ [60]

Types and Sources of Data

When you and your evaluation team consider what information you'd like to collect, also consider the types of data and the quality of these data. There are two main types of data: qualitative and quantitative. Table 7 has a description, advantages and disadvantages, and examples of each of these two types of data.

- Quantitative methods (like surveys) are helpful in getting information from more people.
- Qualitative methods (like interviews and focus groups) dive deeper into the issue.

Qualitative data and quantitative data have different strengths and often answer different questions, but _one type is not better than the other._ For example, a poorly formed survey will not provide you with as much reliable information as an expertly conducted in-depth interview. The quality of the data and data collection methods are what is important here.

Data can come from many sources. Table 8 describes the most common sources of data for program evaluations.

Table 7. Types of Data

Type	Description	Advantages	Disadvantages	Examples
Quantitative	Quantitative data is numerical or quantifiable, meaning that they can be used for mathematical calculations or statistical analysis to infer findings.	Can be less time-consuming to collect. Can use web-based survey platforms (like Qualtrics, Survey Monkey, etc.) which can generate descriptive results from a survey.	May only be able to give you a shallow view into your activities (e.g., doesn't give you a meaningful understanding of people's experiences).	• Close-ended survey questions • Ratings made in observation reports • Statistics in secondary data analysis
Qualitative	Qualitative data are non-numerical and categorical by nature. Qualitative methods are helpful in categorizing data based on attributes and properties of a thing or phenomenon, themes, or recurring ideas.	Tells you more about experience in a participant's own words	Can be time-consuming to collect and analyze	• Interviews • Focus groups • Observation reports • Secondary data analysis • Open-ended survey questions observation

Table 8. Sources of Data

| Data Source	Description
Survey	Structured tools for data collection with standard questions, and typically conducted over the phone, by mail, or online. Surveys can be self-paced or can be led by the person conducting the survey. Researchers or evaluators conducting a survey should strive for a sample that is representative of the population of interest.[61]
Interview	One-on-one discussions with professionals, victims, or other stakeholders. Interviews can be structured (proceed from a set script of questions), semi-structured (include some questions but can progress naturally), or unstructured (open-ended and free form).
Focus Group	Structured group discussions led by one or two facilitators.
Observation	A data collection approach in which people are examined in a natural setting or a naturally occurring situation in order to document their responses or reactions to various choices, decisions, or opinions. Observations are both qualitative and quantitative in nature.
Document Review	Refers to the organization, analysis, and interpretation of data by reviewing existing documents, such as reports, program logs, performance ratings, funding proposals, meeting minutes, or newsletters.
Secondary Data	Refers to data that have already been collected in the past through primary sources and therefore are readily available for analysis.

Data can also be collected from program activities. For example, the number of clicks, "likes," and shares from a social media post can give you some information on the reach of your program's messages. A push-pin map (where individuals indicate locations on campus where they saw a harm and where they intervened; see Figure 10) can give you information about participants' behaviors and about the potential community impact of your bystander intervention program. Being creative about how and where you collect data can give you information about how well your program is doing.

Figure 10. Example of a Push-Pin Map [62]
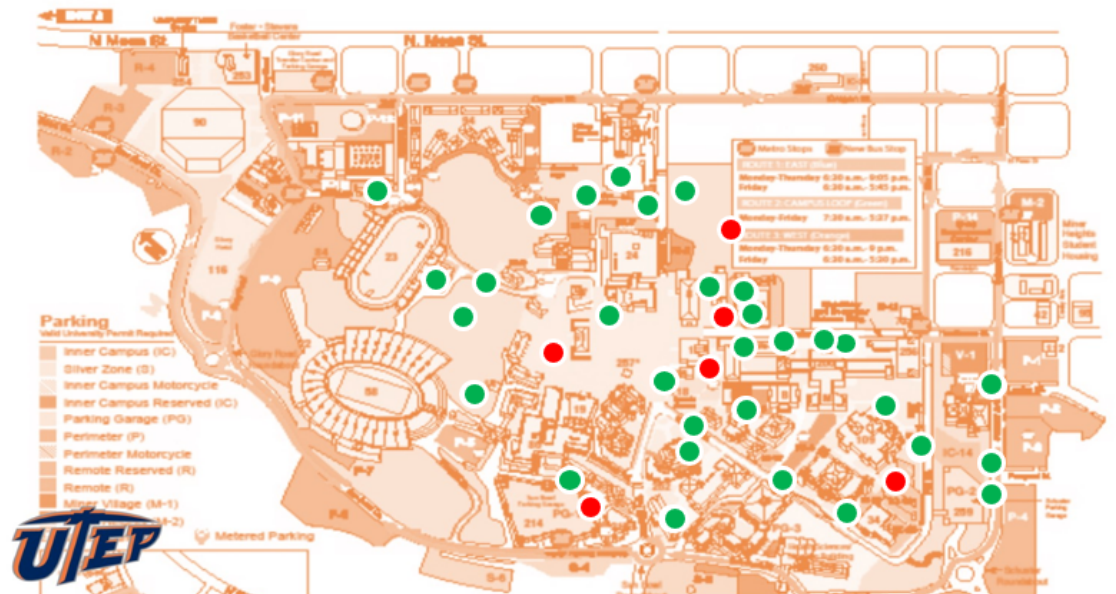
Note: This map is for illustration purposes only and does not represent any actual incidents on the UTEP campus. Reprinted with permission.

Questions to Consider Related to Data Types and Data Sources:

As you consider your sources of data, review the questions below. Answering these questions with your evaluation team will help you to further design your program evaluation and detail what you expect to learn from it. In these questions, a sample refers to the group of people who participate in your evaluation, which is a subset of the larger group of participants or people in the group of interest.

1) How much data do you want/need?

Typically, data gathered from larger groups of participants or "samples" are better for detecting small changes. If you have a large total sample, are you going to select a sub-sample for your evaluation?

If a large sample size is not feasible, consider your options and capacity. If your total sample of data (e.g., workshop post-tests) is small, can you use the whole sample (all post-tests)?

2) How will you choose your evaluation participants?

You can choose a sample of participants in various ways, including random selection, by group, by date, or by recruiting from student organizations, to name a few.

3) Will the data reflect responses from a diverse sample?

The data from a bystander intervention program or program evaluation are vital to understanding unique campus needs, specifically meeting students' needs, and drawing conclusions from the data to refine your program accordingly and make it more effective.

4) Is institutional review board (IRB) approval needed for these evaluation activities?

Your university's IRB provides oversight for research involving human subjects. Most program evaluations are deemed quality improvement efforts given that the original data collected (or raw data) is expected to be kept internal to the program staff in order to continually revise the program. IRB approval is necessary when there is a specific research question that is going to be evaluated. IRB approval should be obtained if you intend to publish quality improvement data in a scientific journal or publication. You will need to work with a Principal Investigator (generally a PhD-level researcher) who can provide methodological and human subjects guidance.

5) Are the data collection strategies culturally sensitive to race, gender, and other aspects of cultural identity?

Consider using student engagement strategies to enhance culturally specific bystander program scenarios by asking questions during evaluation activities such as, "Do you feel represented by these scenarios?" and "Does this workshop reflect issues that may arise in your own social life?"

6) What potential types of harm may evaluation participants experience as you collect data?

Participants may experience survey fatigue or overburden during an evaluation (survey respondents become bored or apathetic and have low-quality responses). It is important that your evaluation plan takes this into consideration and that you implement strategies to avoid or counteract any harms or threats to quality data collection. Read more on this next as well as in the Common Challenges and Solutions section later in the Toolkit.

Data Collection Plan

To efficiently collect data in a way that upholds the quality of the data and evaluation and avoids introducing bias or added burden, spend time thinking through your data collection plan. Map out the general timing, the procedures by which data will be collected, and how the data will be managed and stored once collected.

It is crucial to consider the potential harm that may arise during an evaluation. As noted, some common types of harm in program evaluations include survey fatigue and overburden.

The challenge will be to ensure that the selected data collection method is appropriate for the target population, is easy to administer, has a minimal burden, and is not too complex.[63] Thinking through these processes and their potential impacts will help bring the details of your program evaluation into focus.

Data Collection Logistics

Evaluation logistics are an essential component of your overall plan and require thoughtful attention.

To address logistics in your plan, describe:
- Who the ideal participants are.
- How they will be recruited.
- What sample size you are hoping to get.
- When you are hoping to recruit participants.
- When they will actually participate in the evaluation.

Think also about data collection, including:
- Who the data collectors will be.
- How they will be trained.
- How they will collect the data.
- How they will enter and store the data.
- Who else will have access to the data that is being collected.

> **KEY INFORMATION**
>
> **The Value of Pre-Testing Tools**
>
> You can improve the quality of the data you collect by pre-testing your **measurement tools** (e.g., survey questions or interview guides) to help iron out any issues (e.g., wording). Pre-testing can also involve testing your **data collection procedures and training** for the individuals who will be collecting data for your evaluation. Pre-testing might help you address any issues that arise from your selected data sources, how you manage the data, or how you code the data.
>
> This process can help you develop quality control measures to ensure the data you collect is high quality before officially launching your program evaluation.

Consider any security or confidentiality needs with your data and determine if there should be any specific protocols put into place to protect the data:
- Do (or should) participants need to provide their consent before their data is collected and used for your evaluation?
- Do schedules and checklists need to be created for data collectors?
- Should a manual of procedures (MOP), protocol, or standard operating procedure (SOP) be produced?

Determining Measures to Use

When designing an evaluation, there are costs and benefits to designing your own survey versus using an established questionnaire. While you may be tempted to draft a survey or other data collection tools on your own to measure precisely what you want to capture,— and nothing more—we advise caution. Evaluation instruments that use valid and reliable measures (e.g., survey questions that have been thoroughly tested and capture exactly what is intended when used repeatedly) produce more informative data with a lower likelihood of error. This ensures the data credibility and supports any conclusions you draw from those data. Credible information can also be collected from (trustworthy) secondary/existing data sources. Relying on an established questionnaire can provide you with the benefit of knowing that the questions have been vetted by other researchers and practitioners, are readily understood by a specific target audience, and hang together to measure a larger construct.

Established measurement tools:

- Have been tested and piloted with various audiences and assessed for reliability and validity. (See Appendix D. Detailed Outcomes Measures Review for a comprehensive look at established measurement tools for bystander intervention programs.)
- Allow you to compare your evaluation with other evaluations that used the same questions or tool.

That said, existing questionnaires can be long and may not perfectly fit your program outcomes. As such, you may need to select a smaller subset of questions from a survey (i.e., a subscale, if one exists) or create your own questions. Given that there are many surveys available to assess bystander intervention and related constructs, *creating your own items should be a last resort.*

If you do decide to create your own tool, keep in mind:

- Only ask one thing at a time. Avoid double-barreled questions.
- Be concise, not wordy.
- Be clear.
- Use a scale with a number of answer choices (more than two or three choices) for nuance and variability in your data. See Likert-Type Scale Response Samples HERE.

See the tools and sidebars in this section for additional guidance on gathering information for your program evaluation.

Phase 0 〉 Phase 1 〉 Phase 2 〉 Phase 3 〉 **Phase 4** 〉 Phase 5 〉 Phase 6

TOOLS

IDVSA's Bystander Evaluation Question Generator

For more bystander-specific evaluation questions and measures, see IDVSA's Bystander Evaluation Question Generator. This evaluation tool can help you create and combine the most relevant bystander evaluation questions based on harm types addressed by bystander intervention (i.e., sexual assault, intimate partner violence, hazing, etc.) and various behavioral determinants that your program intends to measure (i.e., willingness to intervene, perceived barriers, self-efficacy etc.).

This evaluation tool is accessible via a Qualtrics survey link **HERE** and will ultimately provide a unique link to the recommended questionnaires based on your responses.

Phase 4 Checklist

Checklist for Phase 4: Gather Information	
✓	TASK
	Review available and potential data sources and choose the best fit.
	Create or adapt data collection instruments (e.g., surveys or focus group guides).
	Create a detailed data collection and management plan and protocols, and review them with stakeholders.
	Pre-test data collection protocols with data collectors and individuals from the target audience(s) to work out the logistics.
	Pre-test new instruments with individuals from the target audience to refine them.

Phase 0 〉 Phase 1 〉 Phase 2 〉 Phase 3 〉 **Phase 4** 〉 Phase 5 〉 Phase 6

Phase 4 Worksheet: Data Collection Planning and Logistics

Use the template below to draft your data collection plan and logistics, filling in as many details as possible and communicating them to your evaluation team members. The team should collectively review and refine the plan.

Evaluation Questions	Data Sources & Methods	Timing of Data Collection	Responsible	Quality Assurance	Timing of Data Analysis	Notes
EXAMPLE: **How many workshops will be delivered per semester?**	Program activity logs Workshop request forms	Weekly review of requests each semester Monthly review of activity logs each semester	John E. Song D.	John E. will collect & analyze data. Song D. will review each semester.	End of each semester & academic year	Song D. will train John E. to collect/analyze & provide ongoing training as needed.
EXAMPLE: **Is there a program effect on participant bystander efficacy in the short- or long-term?**	Workshop pre-tests & post-tests Online surveys	Pre-/post-test for each workshop Participant online survey (sent via email) 2 & 6 months afterwards	John E. Song D.	John E. will collect data. Song D. will review monthly & conduct analysis.	End of each semester & academic year	Song D. will train John E. to collect & enter data. Jackson W. will give incentives to participants of online surveys ($5 gift cards).

Adapted with permission from: Centers for Disease Control and Prevention. *WISEWOMAN evaluation toolkit.* Atlanta, GA: Centers for Disease Control and Prevention, US Dept of Health and Human Services; 2015. https://www.cdc.gov/wisewoman/docs/WW_Evaluation_Toolkit.pdf

PHASE 5: ANALYZE THE DATA AND MAKE CONCLUSIONS

This section provides you with high-level information on data analysis. To guide your learning, we offer various analysis strategies, information on industry standards, and a list of resources where you can learn more. However, this section is not intended as a standalone resource on how to conduct data analysis. IDVSA provided technical assistance to UT institutions through June 2021 and can continue to be a resource for data analysis direction and expertise.

As you begin to look at the data you have collected through your program evaluation, it is important to analyze those data, interpret the results, and draw conclusions that are directly linked to the evidence the data offer.

Data Analysis Steps

Analysis of quantitative and qualitative data can look very different. Table 9 provides basic steps to help you prepare and guide your analysis efforts.

Confirm Values and Expectations

As you are working to analyze the data and interpret the evaluation findings, we recommend that you continue to meet with your stakeholders to revisit the evaluation aim, goals, and objectives to ensure everyone is on the same page regarding standards for program success. These ongoing conversations can help to remind everyone involved in the evaluation and program of what is important, and help focus your efforts and interpretations.

RESOURCES AND TOOLS

For more guidance on data analysis:

- *Toward a Definition of Mixed Methods Research* [64] (access through your institution)

- *Toward a Conceptual Framework for Mixed-Method Evaluation Designs* [65] (access through your institution)

- The CDC's *Developing an Effective Evaluation Report*

- McGill University's *Codebook Cookbook* [66]

Table 9. Steps for Data Analysis

Steps	Quantitative Data	Qualitative Data
1. Transfer or transcribe the data.	Transfer the data to a new document (e.g., an answer sheet) to make it easier to enter the information into a database.	Transcribe focus group discussions, recordings, or field notes into a format that you can use (i.e., Microsoft Word).
2. Code the data, if necessary.	Create a codebook if one does not exist. Code the information so it can be entered into a database. You may need to revise your codebook to increase the accuracy of coding. During this phase, program staff should also consider if there are any biases or other influences that could affect how the data are interpreted and if they impact any of the conclusions being made.	Create a codebook and apply the codes you developed to the segments of text that align with your theme(s) of interest. You may need to revise your codebook to increase the accuracy of coding.
3. Use computer software to assist with data analysis.	Quantitative data entry and analysis can be done using any of the following software: • Microsoft Excel or Access • SPSS • SAS • Stata Using basic Excel functions, you can report on findings like: • % of participants who agreed or disagreed in a survey. • Average scores for certain questions.	Qualitative data can be organized using Microsoft Word, while data analysis can be conducted using any of the following software: • ATLAS.ti • MAXQDA • CDC EZ-Text (free) Instead of using software, you can also analyze data manually with another person. For example, two people can each read the data independently, write down themes they identify, come back together to compare and refine, and then make decisions about the final themes.
4. Review the data for completeness and accuracy.	You may need to "clean" your data to ensure it is complete and ready to be analyzed. This includes: • Verifying that the data file has the correct and expected number of participants. • Checking for any errors or inconsistent responses in the file. • Regularly reviewing data to ensure data quality. • Checking for missing data and determine how that will be handled.	You may need to "clean" your data to ensure it is complete and ready to be analyzed. This includes: • Assessing legibility of the text and recordings. • Assessing the quality of open-ended responses to interview/focus groups questions. • Having a colleague review the information you have collected for accuracy. • Regularly reviewing the data to ensure data quality over time.
5. Review your data management system.	Be sure to review the data analysis steps from your evaluation plan prior to implementation (consider having a colleague review them too) to identify any potential problems.	Be sure to review the data analysis steps from your evaluation plan prior to implementation (consider having a colleague review them too) to identify any potential problems.

Adapted with permission from: Centers for Disease Control and Prevention. *WISEWOMAN evaluation toolkit.* Atlanta, GA: Centers for Disease Control and Prevention, US Dept of Health and Human Services; 2015.

https://www.cdc.gov/wisewoman/docs/WW_Evaluation_Toolkit.pdf

Phase 0 〉 Phase 1 〉 Phase 2 〉 Phase 3 〉 Phase 4 〉 **Phase 5** 〉 Phase 6

Interpret Findings

When you begin to interpret your evaluation results, continue to keep your program and its goals in mind. Also, determine whether your evaluation efforts could have been influenced or there are any known limitations to the evaluation (missing data, alternative explanations, potential bias, the validity and reliability of results, etc.). You can also consider how the results of similar programs compare to your results, and if your results are consistent with what you might expect from your program's theory of change or other applicable behavioral change theories. During this phase, it is important for program staff and others closely involved in the evaluation to consider if there are any biases, or other influences, that could affect how the data is interpreted and if they impact any of the conclusions being made. Finally, if you are preparing to share your findings and interpretations, think about your audience and how they might interpret and use what you share.

Utilize Findings to Improve Your Program

Throughout your evaluation activities, and certainly as you formally analyze your results, you will identify ways to improve your program. We encourage you to discuss needed changes with staff and other stakeholders, document next steps, and implement those steps to strengthen your program. And then, once you make changes to your program, plan for another evaluation to assess their impact.

See the tool and the sidebars in this section for additional guidance on analyzing and interpreting your data.

KEY INFORMATION

Consider Evaluation Limitations

You may be wondering how to understand the limitations of your evaluation. In other words, how reliable and valid are my evaluation results? We encourage you to consider whether the evaluation could have missing data, alternative explanations, or potential bias. Other than missing data, you may need another level of evaluation expertise to know if certain limitations exist.

It is valuable to make a statement with your results and to your audience about how limitations may exist within your findings, but it is outside the scope of the program evaluation to explore them in-depth.

TOOLS

Phase 5 Checklist

Checklist for Phase 5: Analyze the Data and Make Conclusions	
✓	TASK
	Check data for errors.
	Determine ways to analyze your data with other staff involved in the evaluation effort.
	Consult with IDVSA researchers or other evaluators on specific data analysis plans and protocols needed.
	Document any notes on context to help guide interpretation (e.g., as you review the data and begin to make sense of it, write down your questions that arise, ways to explain responses seen, or various factors that help to explain the data).
	Assess results against findings from the literature, similar programs, and results from previous program cycles/years (if applicable).
	Compare actual outcomes to the intended outcomes from your logic model.
	Document your analysis/interpretations/findings, including a description of potential biases, alternative explanations for your findings, and other limitations.

Phase 0 › Phase 1 › Phase 2 › Phase 3 › Phase 4 › Phase 5 › **Phase 6**

PHASE 6: SHARE RESULTS

The final phase in a program evaluation is to share your results. This process may differ depending on your audience(s) and how results might be used, but the important thing is to make sure that results of your hard work do get used.

When you are preparing to share your program evaluation results, also think about:

- **Timing** – when to share your findings.
- **Format** – how you will share and/or present your findings.
- **Strategy –** how you might prioritize any recommendations you have.

> **RESOURCES AND TOOLS**
>
> For more guidance on sharing results:
>
> - The CDC's *Developing an Effective Evaluation Report* [67]

Using the Evaluation Results

To ensure that your results reach stakeholders and are used, try some of these strategies:

- Prepare tailored recommendations for different audiences.
- Prepare audience(s) to use findings by creating a series of questions to help them explore what the findings mean for the program and what changes could be made for program improvement.
- Elicit feedback from audience(s) by asking them about what surprised them, if anything, about the results.
- Follow up with audience(s) by providing a copy of your results and encourage any additional questions or comments.
- Share your findings in an evaluation report as well as in shorter, visual, and/or creative formats.

The Value of an Evaluation Report

An evaluation report can provide a comprehensive look at a program and the results of a program evaluation. A standard evaluation report should include the following sections.

- An executive summary
- Program background and the purpose of the evaluation
  - This section can also include the program stakeholders and program description.
- Evaluation methods, including:
  - Evaluation questions
  - Evaluation design
  - Measures used
  - Data collection methods
  - Data analysis methods
  - Any limitations identified

- Evaluation results
- Discussion and any recommendations for the future

Evaluation reports can be very useful to both document your evaluation process and findings in a comprehensive way for record keeping as well as to share the results of your program evaluation with stakeholders involved in making program-related decisions.

Sharing Your Findings With the Community

If you'd like to share your evaluation findings with stakeholders who are not as closely involved with your program—or want to share your results in other engaging formats—consider these options:

- PowerPoint presentations
- Program website feature
- Publications, briefs, or white papers
- Fact sheets or one-pagers
- Success stories and/or testimonials
- Podcasts or media features (interviews, press release)
- Posters and/or infographics
- Group-specific discussion meetings (which may be especially important for historically marginalized groups)
- Town hall meetings
- Social media campaigns

Figure 11 is an example of a chart developed from the *Bystander Intervention Program Needs Assessment: Findings Across UT Academic Institutions,* which is an example of a "two-pager" or "data brief" with a number of visual elements.

> **KEY INFORMATION**
>
> **The Importance of Working with University Communications and Administrators**
>
> We encourage you to work directly with university communications or your department's communications staff and administrators on sharing findings with the community. They can support your efforts and ensure that proper administrators are aware of and sign off on any promotion that goes beyond internal stakeholders.
>
> For example, your program leadership should obtain buy-in and support from administrators before doing a town hall meeting. Communications staff can help to support this process and provide guidance

Figure 11. Example Chart for Communicating Evaluation Findings
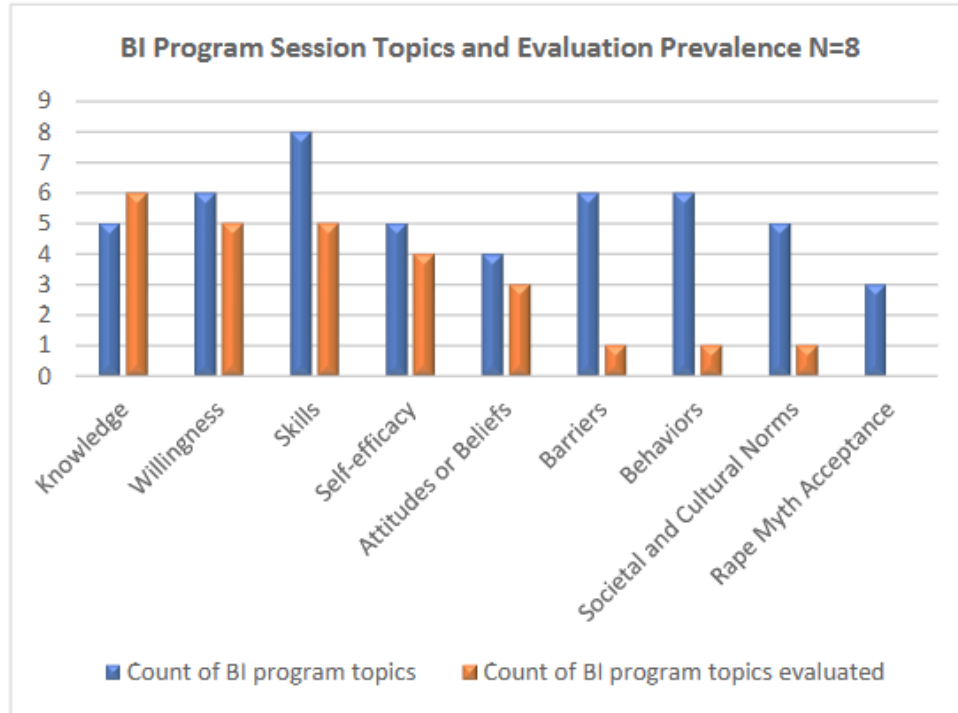
Communications items and activities are ways to raise awareness about your program and show the community how it is working. They also offer an opportunity to reinforce messages related to bystander intervention that are specific to your program. Sharing your data with the community after a program evaluation also fosters accountability. And, finally, sharing the data lets evaluation participants know that the input they provided was meaningful and valued.

See the tools next and the sidebars in this section for additional guidance.

TOOLS

Phase 6 Checklist

Checklist for Phase 6: Share Results	
✓	**TASK**
	Prepare stakeholders to use evaluation findings by identifying your program's training and technical assistance needs.
	Consider ways to use evaluation findings to support long-term program planning, program promotion, and public knowledge about the benefits of the program. Come back to this task repeatedly over time as you work with stakeholders.
	Schedule meetings to present and discuss program evaluation findings, conclusions, and next steps – possibly for interim and final results.
	Create tailored and streamlined evaluation reports or presentations for various internal and external stakeholders.
	Share evaluation findings through multiple avenues, clearly and succinctly, and in a timely manner.

COMMON CHALLENGES AND SOLUTIONS

In developing the BeValued Toolkit, IDVSA worked closely with UT academic institutions to understand their programs, their evaluation-related needs, needs as well as the challenges they face when designing, implementing, and evaluating their bystander intervention programs. In this section, we address some key challenges that they shared associated with program evaluation—and offer ideas for solutions. Although we recognize that additional challenges exist for bystander intervention programming, the challenges included here are the ones most closely related to program evaluation.

COMMON EVALUATION CHALLENGES[68]

- Lack of evaluation training
- How to avoid over-burdening students during an evaluation process
- Staff turnover
- How to implement an evaluation to assess long-term outcomes/impact
- Lack of capacity to do a program evaluation
- Addressing a disconnect in expectations/priorities between leadership and program implementers

SOLUTION-ORIENTED STRATEGIES

How to Obtain Additional Training

Training to conduct evaluations is a valid concern. This toolkit has the information needed to get you started. You can also search for specific information in areas where you feel you need the most help. Reading a few articles on the subject can go a long way. You can also utilize no-cost trainings offered by your university, which often are provided by the university's research office. You can also connect with your university/department/office's in-house evaluator. These positions often rotate on annual schedules to support evaluations across the university. IDVSA provided technical support to UT institutions until June 2021 and can continue to be a resource for training and consultation.

How to Avoid Survey Fatigue

Survey request fatigue occurs when people have been asked for feedback so many times that they do not want to respond anymore. If individuals never see the results of a survey, they may also begin to wonder if providing their feedback is even worthwhile. There is also survey-taking fatigue, which occurs when someone is in the middle of a survey and becomes tired of the activity. Survey-taking fatigue can result in poor data because individuals may skip questions or answer questions quickly, giving little thought to their answers.

To mitigate the effects of survey fatigue:

- Communicate the value of the survey.
  - For example: "We are going to use the results from this survey to make this program better in the future."
- Keep the survey short.
- Only include questions that are meaningful.
- Only include items on the survey that you plan to analyze.
- Take the survey yourself.
- Have a colleague who is unfamiliar with the program take the survey.
- Use open-ended questions sparingly. These questions require people to write in a response. If there are too many, people may not give meaningful responses to each item.
- Host a program evaluation data "give-back" at the end of the year. The goal of a give-back is to share the information you learned in the program evaluation with your community. Data can be shared in a variety of ways, including a poster presentation of your findings, distributing flyers, or providing your key takeaways via social media outlets. People may be more willing to complete a survey in the future when they know that their feedback is discussed and valued by the program.

RESOURCES AND TOOLS

For more guidance on avoiding survey fatigue:

- *Multiple Surveys of Students and Survey Fatigue* [69]
- *Finding a Cure for Survey Fatigue* [70] (access through your institution)
- *Tired of Survey Fatigue? Insufficient Effort Responding Due to Survey Fatigue* [71]
- *The Survey Fatigue Challenge: Understanding Young People's Motivation to Participate in Survey Research Studies* [72]

For more guidance on staff training:

- The CDC's *Training Professionals in the Primary Prevention of Sexual and Intimate Partner Violence: A Planning Guide* [73]

Sulley, C., Rabideau, D., Orchowski, L., Jimenez, A., Dragoon, S., Susswein, M., Dube, S., Wood, L., & Busch-Armendariz, N.B. (2021). *The bystander evaluation (BeValued) toolkit: A resource to support program growth and impact.* Austin, TX: Institute on Domestic Violence & Sexual Assault, The University of Texas at Austin.

Transferring and Preserving Knowledge

Given the nature of college campuses, turnover in staff is expected, which puts programs at risk of losing key historical knowledge. Students trained as peer educators will be continually graduating, and individuals within higher education may also be transferring to new positions or advancing in their roles via promotions. Developing a well-established training for peer educators and staff can ensure that the knowledge, skills, and abilities fostered by your office can be readily transferred to new staff over time.

In addition, *a commonly overlooked component of process evaluation is documenting the training activities conducted with staff or student educators*. Staff training and supervision is an essential component to understanding how prevention activities are implemented and may provide insight into why—after the completion of a prevention effort—the program did or did not have the desired results.

> **RESOURCES AND TOOLS**
>
> For more guidance on ensuring high-quality programming:
>
> - *Treatment Fidelity in Outcome Studies* [74] (access through your institution)
> - *Fidelity Criteria: Development, Measurement, and Validation* [75]
> - *Adoption, Adaptation, and Fidelity of Implementation of Sexual Violence Prevention Programs* [76]

Ensuring High-Quality Programming

To make the most of your resources, it is important to ensure that peer educators and staff are running the program "as intended"—in other words, if there is **program fidelity**. There are many factors that influence whether a program is implemented according to the original model. Utilizing program fidelity assessments can help you to evaluate the consistency and competency of program delivery.

It is also important to create a system for documenting why specific deviations from protocol take place. For example, was key content left out because there was not enough time for discussion? Did participants engage in lengthy discussion of some topics, but avoid discussion of others? Was key content left out because attendees needed to leave the building for a fire drill? Ensuring that there is ample space to annotate why deviations from protocol occur will allow for a narrative description of *how* a program is being implemented in your community.

To assess program fidelity, the program's key constructs (broad concepts or topics) must be clearly defined, along with the measures and response options for each (for examples, see Appendix D. Detailed Outcomes Measures Review). To obtain objective data using fidelity ratings in program evaluation, it is most effective to involve a third-party "expert" to rate the adherence and competency of the individuals delivering the program. This could be a model developer, a program evaluator, or a researcher. It can also be useful to have program

facilitators assess their own adherence and competency in implementing the program using a checklist of key intervention content elements and style. Supervisors can also use fidelity assessment forms to structure their observations of programs and provide constructive feedback to program facilitators.

MINDSET AND EXPECTATIONS

Keep the right mindset.

Remember, the evidence base is continuing to develop regarding the causes of health risk behaviors, which are complex and multifaceted. As a result, risk prevention and bystander intervention efforts are continually evolving to reflect new knowledge regarding risk and protective factors. Thus, program evaluation will be ongoing, and you can expect to be continually changing your program in response to participant feedback and the latest research over time.

Have realistic expectations.

Given the complex nature of human behavior and health, a single intervention is unlikely to reduce incidence rates of a health risk behavior. If you are planning, considering, or operating a single-component program, evaluate its immediate effects *and* implement it in combination with other efforts and partners. And just because a program is not showing decreases in the rates of a target health risk behavior, it does not mean that it isn't working to shift other positive outcomes.

It is difficult to show that a prevention program reduces rates of a behavior.

When an event—such as a harmful behavior—occurs relatively infrequently, it is difficult (statistically) to show that a prevention program had an effect on reducing the event's occurrence. It is important to ensure that your efforts to evaluate your bystander intervention programs assess a range of intervention effects and take into account what is needed to make immediate, short- and long-term impacts on reducing rates of health risk behaviors.

It takes time to document the effect of a prevention effort.

Documenting that a prevention program has worked requires that participants be followed over a long enough period of time to show that effects are maintained. What appears to be an immediate benefit of a program may wear off. Programs may also have long-term benefits that are only captured after multiple evaluations over time. *Program evaluation should be considered an ongoing process*, where evidence for the outcomes is accumulated over time.

TOOLS AND TEMPLATES

Below, there is a complete list of the worksheets, tools, and templates we included after information on certain evaluation types and after each evaluation phase in the toolkit. The items listed are available for download, and some are editable as well.

There are many evaluation resources available, and we list and recommend a number of them throughout the BeValued Toolkit in **sidebars**. The IDVSA team has vetted and selected the following items as particularly valuable for specific tasks and considerations that are part of planning and implementing a bystander intervention program evaluation in a higher education setting.

SUMMARY LIST OF WORKSHEETS, TOOLS, AND TEMPLATES

Note: There are additional resources provided in sidebars throughout the document.

Each item in the following list is linked to an editable or interactive version of the tool:

- Process Evaluation Tool: BeVocal Activity Tracker
- BeValued Toolkit Self-Assessment
- Phase 0 Worksheet: Evaluation Budget Template
- Phase 1 Worksheet: Stakeholder Engagement Plan
- Phase 2 Worksheet: Logic Model Checklist
- Phase 2 Worksheet: Logic Model Template
- Phase 3 Worksheet: Logic of Evaluation Questions and Indicators
- Phase 4 Worksheet: Data Collection Planning and Logistics
- IDVSA's Bystander Evaluation Question Generator
- IDVSA's Bystander Evaluation Questionnaires

EVALUATION PLAN TEMPLATE

This template shows all of the necessary components for an evaluation plan and provides you with a checklist to track tasks and text boxes to insert key information on each component. An editable version of this template is available **HERE**.

✓ OR ✗	PLAN COMPONENT	PLAN INFORMATION
	PROGRAM DESCRIPTION	
	Program need *(national/state statistics; campus statistics; risk/protective factors)*	
	Program mission & objective(s)	
	Program logic model *(put on separate page)*	
	EVALUATION DESCRIPTION	
	Evaluation focus	
	Evaluation questions	
	Evaluation methods	
	Analysis plan	
	Dissemination plan	
	Evaluation budget	

APPENDICES

APPENDIX A. KEY RESOURCES ON PROGRAM EVALUATION

If you wish to delve further into evaluation resources and guidance, or explore other toolkits, Table 10 below offers a list of quality program evaluation resources vetted by the IDVSA research team. These are the resources that we found most relevant and useful during our systematic review of resources about program evaluation. *For a list of all resources and sources noted in the BeValued Toolkit, see the References section following the Appendices.*

Table 10. Key Program Evaluation Resources

Name	Developer	Description	Website
Program Performance and Evaluation Office (PPEO)	CDC	Provides a great deal of guidance for program staff, from basics to more complex evaluations. Includes many links to other CDC and non-CDC resources (examples).	https://www.cdc.gov/eval/
Framework for Program Evaluation in Public Health	CDC	1999 CDC program evaluation guide from Morbidity and Mortality Weekly Report.	https://www.cdc.gov/mmwr/PDF/rr/rr4811.pdf
CDC Evaluation Framework	CDC	Video about the CDC framework for program evaluation.	https://www.youtube.com/watch?v=tOjieBh1ce0
CDC Learning to Love Your Logic Model	CDC	Video about how to improve program logic models.	https://www.cdc.gov/eval/videos/logic-model/LearningToLoveModel-LowRes.mp4
Evaluator Self-Assessment	CDC	Self-assessment for program staff to determine their own capacity to conduct program evaluations.	https://www.cdc.gov/eval/tools/self_assessment/EvaluatorSelfAssessment.pdf
Evaluation Planning: What is it and how do you do it?	CDC	Provides an overview of types of program evaluations and the purpose of each kind. Also has a worksheet with guiding questions for each step of program evaluation.	https://www.cdc.gov/healthcommunication/pdf/evaluationplanning.pdf

Name	Developer	Description	Website
Introduction to Program Evaluation for Public Health Programs: A Self-Study Guide	CDC	Toolkit provides guidance, recommendations, examples, checklists, and worksheets for program evaluations.	https://www.cdc.gov/eval/guide/CDCEvalManual.pdf
Developing an Effective Evaluation Plan	CDC	Workbook provides step-by-step guidance on the evaluation process and how to effectively develop an evaluation plan. Worksheets, templates, and examples are provided throughout.	https://www.cdc.gov/obesity/downloads/CDC-Evaluation-Workbook-508.pdf
Developing an Effective Evaluation Report	CDC	Workbook provides guidance and examples on the evaluation process and how to effectively communicate results from program evaluations throughout.	https://www.cdc.gov/eval/materials/Developing-An-Effective-Evaluation-Report_TAG508.pdf
WISEWOMAN Evaluation Toolkit	CDC	Program evaluation toolkit for WISEWOMAN programs/grantees, a public health stroke and heart disease prevention effort. Provides a guided process of evaluation, templates, and examples.	https://www.cdc.gov/wisewoman/docs/WW_Evaluation_Toolkit.pdf
Community Toolbox	University of Kansas	Toolbox provides practitioners with a step-by-step guide on how to plan, develop, implement, evaluate, and sustain community-based initiatives; each chapter focuses on a particular area of this process (46 chapters). Chapters 36–39 introduce and discuss evaluation, evaluation methods, and the benefits of program evaluation. Each chapter has multiple sections and information, links, tools, slides, checklists, and more to give practitioners step-by-step guidance.	https://ctb.ku.edu/en/evaluating-community-programs-and-initiatives

Name	Developer	Description	Website
The Evaluation Center: Evaluation Checklists	Western Michigan University	Evaluation checklists for multiple evaluation planning and implementation task areas.	https://wmich.edu/evaluation/checklists
Enhancing Program Performance with Logic Models	University of Wisconsin-Extension	Free online course on improving programs with logic models.	https://lmcourse.ces.uwex.edu/#
A Practical Guide for Engaging Stakeholders in Developing Evaluation Questions	Robert Wood Johnson Foundation	Workbook provides practical guidance on identifying stakeholders and methods to engage them in evaluation planning as well as worksheets to use in the process.	http://www.pointk.org/resources/files/rwj.stakeholders.final.1.pdf
The 2010 User-Friendly Handbook for Project Evaluation	National Science Foundation	Handbook explains the details of program evaluation and methodologies.	https://www.purdue.edu/research/docs/pdf/2010NSFuser-friendlyhandbookforprojectevaluation.pdf
Evaluation Toolkit	Cottage Center for Population Health	Online toolkit with step-by-step guidance to plan, implement, and share program evaluations. Provides information, links to other resources, tools (including worksheets), and case examples.	https://www.cottagehealth.org/population-health/learning-lab/toolkit/

APPENDIX B. EVALUATION TERMS AND DEFINITIONS

The following definitions are offered to facilitate a common understanding of terms.

Behavioral Determinants: Constructs or conceptual elements used as part of theories and models to concretely describe and explain what influences individuals' behaviors.[77]

Codebook: A list of category labels and descriptions that relate to the themes identified while reviewing qualitative data.[78]

Coding Data: Identifying units of information in qualitative data and assigning them to categories from an evaluation codebook.[79]

Data: Information that can be analyzed.[80] Data can be qualitative or quantitative, and can be collected through a variety of sources including interviews, focus groups, and surveys.

Data Analysis: Reviewing and interpreting data to identify themes, findings, and conclusions.

Data Cleaning: The process of identifying inaccurate, incomplete, duplicate, or improbable data, and then correcting it when possible. Data cleaning is a two-step process that includes detection and correction.[81]

Data Sources: The individuals, existing data sets, or other sources of information that will provide data for an evaluation.

De-Identifying: Removing identifying information, such as names and addresses, from a data set. This process protects participants' privacy and confidentiality.

Document Review: Refers to the organization, analysis and interpretation of data by reviewing existing documents such as, reports, program logs, performance ratings, funding proposals, meeting minutes, newsletters, and more.

Evaluation: Systematic investigation of the process or outcomes of a particular program or endeavor.[82]

Focus Groups: Structured group discussions led by one or two facilitators.

Formative Evaluations: Conducted to inform program development and improve program delivery.

Frequency: In statistics, how often a response occurs.

Institutional Review Board (IRB): An administrative body, typically at an institution of higher education, tasked with overseeing research involving human subjects and protecting the rights of participants.[83]

Interviews: One-on-one discussions with professionals, survivors, or other stakeholders. Interviews can be structured (proceed from a set script of questions), semi-structured (include some questions but can progress naturally), or unstructured (open-ended and free form).

Logic Model: An outline of the program's resources/inputs, activities, outputs, and outcomes. The CDC defines logic model as "graphic depictions of the relationship between a program's activities and its intended outcomes." [84]

Mean: The average of the data. The sum of a set of numbers divided by the quantity. In this toolkit, the sum of responses divided by the number of responses.

Median: The middle number of a data set when numbers are arranged from smallest to largest.

Mode: The most frequently occurring response in a data set.

Needs Assessment: Explores the extent of a problem, considers barriers, prioritizes needs, and identifies available resources and interventions to address those needs.[85]

Normative Beliefs: "Individuals' beliefs about the extent to which other people who are important to them think they should or should not perform particular behaviors. In general, researchers who measure normative beliefs also measure motivations to comply—how much individuals wish to behave consistently with the prescriptions of important others. Each normative belief about an important other is multiplied by the person's motivation to comply with that important other and the products are summed across all of the person's important others to result in a general measure that predicts subjective norms. Subjective norm is a predictor of intention to behave which, in turn, is a predictor of actual behaviors." [86]

Observation: Researchers attend during routine activities of the group they are studying to watch and gather information without direct interaction.[87]

Open-Ended Questions: Questions that do not have yes or no answers and inspire critical thinking and layered responses.[88]

Operationalize: Determine how to define and measure a variable or outcome for research or evaluation.[89]

Outcome Evaluation: A method to determine whether the program met its stated outcomes.[90]

Process Evaluation: A method to assess how a program functions to determine whether activities and outputs occurred as intended.[91]

Program Evaluation: A method to determine the impact and effectiveness of a program, initiative, or services, and implement empirically based improvements.[92] The CDC defines program evaluation as, "the systematic collection of information about the activities, characteristics, and outcomes of programs to make judgments about the program, improve program effectiveness, and/or inform decisions about future program development."[93]

Qualitative Data: Words that come from interviews, focus groups, observation reports, secondary data analysis, and possibly some items on surveys. This type of data is used to highlight themes or recurring ideas.

Quantitative Data: The resulting numbers from closed-ended survey questions, ratings made in observation reports, or statistics in secondary data sources that can be analyzed via statistical methods.

Range: The difference between the highest occurring data point and the lowest occurring data point for one question.

Sampling: The process of selecting a subset of individuals from the population of interest for evaluation or other investigative inquiry (intervention, surveys, interviewing, etc.). The goal of sampling is to obtain a group that is representative of the larger population from which it is drawn so that any findings can be generalized back to that population. When sampling, one must always be mindful of potential bias that impacts the ability to infer characteristics about the larger population from the sample used.

Secondary Data: Refers to data that has already been collected in the past through primary sources and is therefore readily available for analysis.

Secondary Data Analysis: Reviewing and analyzing existing data.

Summative Evaluations: Conducted to inform program stakeholders about the extent to which a program was able to achieve its objectives once the program is delivered.

Surveys: Structured tools for data collection with standard questions that are typically conducted over the phone, by mail, or online. Surveys can be self-paced or can be led by the person conducting the survey. Researchers or evaluators conducting a survey should strive for a sample that is representative of the population of interest.[94]

## APPENDIX C. TOOLKIT DEVELOPMENT METHODS

The research process and overall development process for the BeValued Toolkit was tailored to respond to the specific needs expressed by the representatives from each of the UT System academic institutions (see Figure 12) while at the same time, remaining inclusive enough to cover the more typical elements of a general program evaluation so that this resource is applicable to additional bystander intervention practitioners and stakeholders within and beyond the UT System. IDVSA took active measures to ensure no overburden or harm in the data collection process and also ensured the recommendations included in the toolkit do not cause harm to students or campus stakeholders, nor interfere with the primary educational mission of the institutions.

Research Questions

The questions guiding the development of this toolkit were:

1. What are best practices and recommendations regarding program evaluation?
   a. Specifically, what are evaluation best practices and recommendations for bystander intervention programs?
2. How can practitioners apply best practices and recommendations for program evaluation to best meet the needs of their campus?

Figure 12. UT System Institutions Participating in IDVSA Bystander Intervention Project
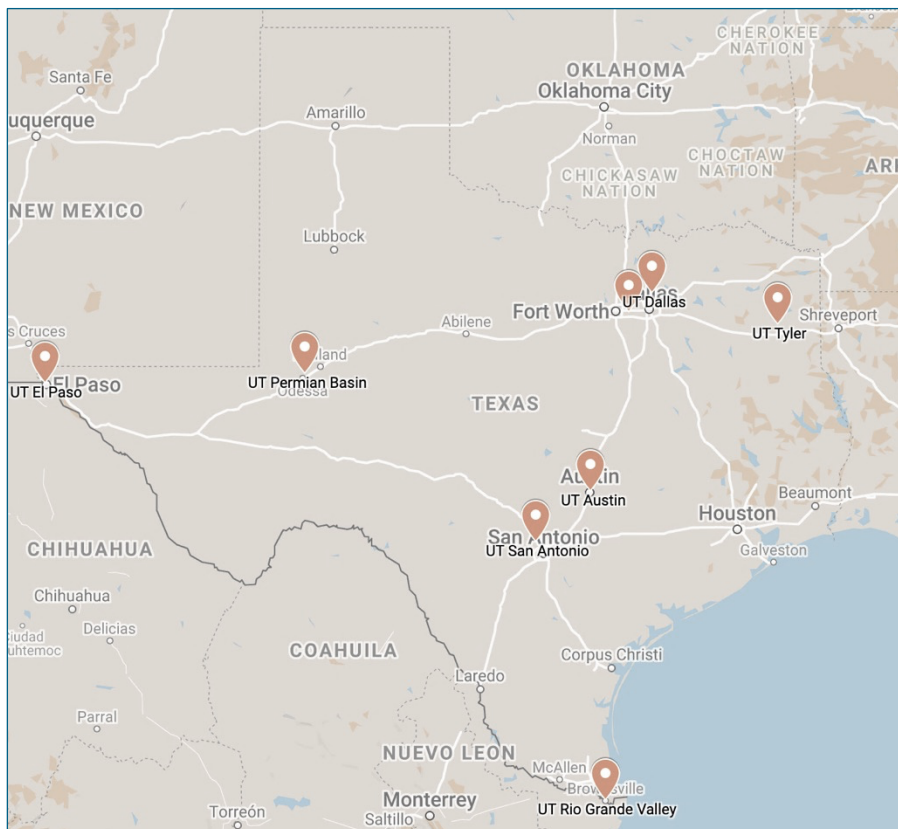
Sulley, C., Rabideau, D., Orchowski, L., Jimenez, A., Dragoon, S., Susswein, M., Dube, S., Wood, L., & Busch-Armendariz, N.B. (2021). *The bystander evaluation (BeValued) toolkit: A resource to support program growth and impact.* Austin, TX: Institute on Domestic Violence & Sexual Assault, The University of Texas at Austin.

Data Collection

Several sources of data contributed to the development of this toolkit.

Literature Review

The literature regarding bystander intervention provided insight into current practices in the field, including evaluation practices and measures. Search terms included (but were not limited to): "bystander intervention," "bystander," "sexual assault prevention," "alcohol," "drinking," "college," "university." Articles were selected if they focused on colleges or universities in the United States, bystander intervention initiatives, factors that influence individuals' bystander intervention behaviors, and were published in peer-reviewed academic journals in the last five years. A comprehensive review of the literature on bystander intervention is available **HERE**.

Internet Search

In addition to the literature, the IDVSA team searched through reputable sources to build a preliminary inventory of program evaluation resources. First, IDVSA researchers reviewed the website for the Centers for Disease Control and Prevention (CDC) Program Performance and Evaluation Office (PPEO).[95] Researchers also reviewed resources linked from the PPEO's website for relevance, credibility, and utility.

Needs Assessment

IDVSA also conducted original research to determine the needs of practitioners implementing violence prevention initiatives within the UT System academic institutions, specifically bystander intervention programs. This assessment was specific to their programming and evaluation needs and challenges. These data informed the content and organization of this toolkit. The findings are available **HERE.**

Consultation With National Experts and Stakeholders

As this toolkit was being planned and developed, the IDVSA team consulted with some key individuals. First, stakeholders for this project, including key staff who led the project for the UT System and other representatives from UT bystander intervention initiatives, provided valuable insight on the format and structure of this toolkit during initial planning discussions all the way through final development. These conversations ground the toolkit and assure its utility to practitioners. Additionally, national experts on bystander intervention and their program evaluations contributed. This included Lindsey Orchowski, who ultimately became a co-author for this toolkit, and Sarah McMahon of Rutgers University, who provided early feedback and insights to ensure this toolkit was consistent with best practices.

A Focus on Trauma-Informed Methods

The Substance Abuse and Mental Health Services Administration (SAMHSA) describes a trauma-informed approach as a program, organization, or system that can recognize the impact of trauma through signs and symptoms and identify various paths of recovery while

also actively resisting the potential of re-traumatization. There are four fundamental principles to this approach:

- Realize that trauma exists.
- Recognize the impact of trauma and signs/symptoms.
- Respond.
- Resist re-traumatization.[96,97]

In each research project, IDVSA utilizes a universal research design that addresses trauma conceptually, procedurally, and interpersonally. The design acknowledges that researchers and interview subjects may live with behavioral health issues or trauma in their past and present lives.

Program Evaluation Standards

As noted above, the CDC was a primary source consulted in the development of this toolkit. The CDC is a highly reputable center with rigorous scientific information about program evaluation that is available to the public. The CDC's framework for program evaluation provides practitioners with practical guidance for conducting program evaluations.[98] The CDC outlines four standards of evaluation that should guide the overall work of practitioners conducting program evaluations.

The CDC standards or guiding principles of a program evaluation:

- **Utility** - Make sure the evaluation serves the needs of the intended users.
- **Feasibility** - Make sure the program evaluation is realistic given the context and setting.
- **Propriety** - Make sure the evaluation is legal, ethical, and takes into consideration the welfare of those involved or who will be affected by results.
- **Accuracy** - Make sure the evaluation will discover and conveys intended information, and that it is grounded in logical or factual cogency.

Program Evaluation Phases

The CDC also provides practitioners with guidance on the steps a "good" program evaluation should take to meet the standards and expectations listed above. The CDC's six program evaluation phases (see Figure 13) provide practitioners with specific guidance on basic program evaluation tasks and activities.

These steps guide the layout of this toolkit and shaped how we identified our seven phases. However, we adapted them by adding a phase and slightly adapted the wording for clarity and applicability with our audience. We added the preparation phase (Phase 0: Prepare for an Evaluation) to increase the applicability of this toolkit and provide foundational information on the activities and learning necessary before an evaluation can begin.

Figure 13. CDC Program Evaluation Framework
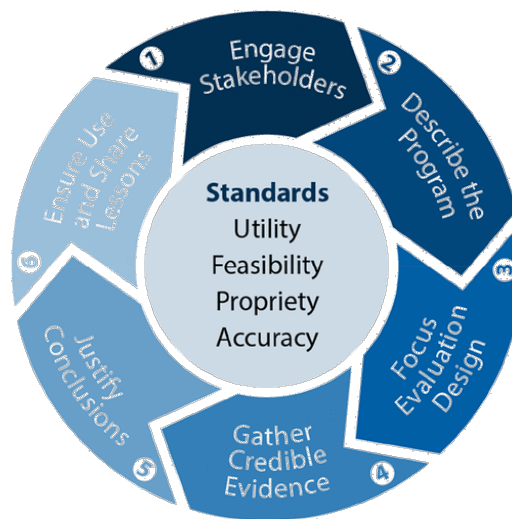
Reprinted with permission from: Centers for Disease Control. (1999). Framework for program evaluation in public health. *MMWR: Recommendations and Reports, 48*, 1–40.

Considerations When Adapting This Toolkit

While our hope is that this toolkit is generally applicable to assist all program evaluation efforts for programs working to prevent violence, there are some considerations that were noted throughout this process for those looking to adapt it.

Designed for UT, but applicable to bystander intervention programs more broadly. First, the toolkit was designed primarily for bystander intervention programs at UT System academic institutions. UT System health institutions and organizations that are not part of UT System may need to adapt this toolkit when using it.

Specific to bystander intervention programs. Second, while much of the information and resources provided in this toolkit may be applicable to other types of prevention approaches, the primary focus was bystander intervention programming. Practitioners looking to evaluate other types of programs may need to adapt this toolkit to better fit their needs.

Highlights intimate partner violence and sexual violence, but applicable to other harms. Third, most of the example measures provided in this toolkit focus on intimate partner violence and sexual violence. While the UT System bystander intervention initiative extends to other harms, such as hazing, high-risk drinking, and suicide prevention, *the bystander intervention approach was originally developed to prevent sexual violence.* Therefore, validated and reliable measures for bystander intervention tend to focus on this area. Practitioners looking to evaluate bystander intervention programs that address other harms may need to adapt this toolkit to better fit their needs, or look further to the literature to find other relevant evaluation measures.

APPENDIX D. DETAILED OUTCOMES MEASURES REVIEW

Table 11. Outcomes Measures in Detail for Bystander Intervention Programs

All measures were reprinted with permission from their authors. This Appendix presents Pros/Cons identified by the IDVSA research team. While the table provides considerable detail on the measures, it is not necessarily exhaustive of all possible pros/cons or information on for each measure.

Measurement	Scale Instructions		
NATIONAL COLLEGE STUDENT BYSTANDER INTERVENTION SURVEY (NCSBIS)	Instructions vary.		
	Subscales	**Example Items**	**Answer Choices**
	Bystander Attitudes	I believe [behavior] is a problem at my college/university.	Vary across subscales.
		I believe students should say or do something when they witness [behavior].	
	Bystander Behaviors	When you witnessed [behavior], what did you do?	
	Bystander Barriers (below)	When you witnessed [behavior], why didn't you help? (below)	
	• Bystander Apathy	I didn't think it was a serious or dangerous situation.	
	• Failure to Notice Situation	I was busy with my own activities and didn't notice what was going on.	
	• Failure to Diagnose the Situation as Intervention Worthy	I was uncertain about what was going on and whether my help was needed.	
	• Failure to Take Intervention Responsibility	I felt it was someone else's responsibility to take action.	
	• Lack of Skills	I was afraid that my intervention might make things worse.	
	• Audience Inhibition & Intervention Costs	I thought my friends/peers wouldn't approve of my saying or doing something.	
	Scoring Instructions	**Pros/Cons**	**Other Notes**
	Contact WITH US for scoring information.	**Pros:** • WITH US can tailor the measure to the unique needs of various programs. • Measure addresses a range of harms at IHEs. • Measure includes questions about behaviors witnessed by students, which may add depth to knowledge if IHE has conducted a campus climate survey (prevalence and perceptions).	For access to the full measure and other resources, visit the WITH US Center for Bystander Intervention at Cal Poly at www.withus.org.

Sulley, C., Rabideau, D., Orchowski, L., Jimenez, A., Dragoon, S., Susswein, M., Dube, S., Wood, L., & Busch-Armendariz, N.B. (2021). *The bystander evaluation (BeValued) toolkit: A resource to support program growth and impact*. Austin, TX: Institute on Domestic Violence & Sexual Assault, The University of Texas at Austin.

NATIONAL COLLEGE STUDENT BYSTANDER INTERVENTION SURVEY (NCSBIS) (continued)		• Measure provides information to tailor programs to address barriers to helping and students' motivations to intervene. **Cons:** • Measure does not capture extent of behaviors witnessed, or extent of students' intervention (captures incidence).	WITH US offers fee-based services, which include NCSBIS survey administration and analysis.
	Citation		
	For more information, see: WITH US Center for Bystander Intervention at Cal Poly. (2020, July). *National college student bystander intervention survey: NCBIS data summary report, national aggregate report.* WITH US. https://www.withus.org/bystanderdatareports		

Measurement	Scale Instructions		
PERCEPTIONS OF PEER HELPING	Please use the following scale to rate how likely YOUR FRIENDS are to do each of the following behaviors.		
	Subscales	**Example Items**	**Answer Choices**
	None indicated	• Ask a stranger if they need to be walked home from a party or get their friends to do so. • Tell a campus or community authority if they see a person who has had too much to drink and is passed out. • Accompany a friend to the police department or other community resource if they needed help for an abusive relationship. • Speak up to someone who is calling his/her partner names or swearing at them. • Go to a community resource (crisis center, counseling center, police, professor, supervisor, etc.) if they saw someone grabbing or pushing their partner.	Not at all likely (1) Unlikely (2) Neither likely nor unlikely (3) Likely (4) Extremely likely (5)
	Scoring Instructions	**Pros/Cons**	**Other Notes**
	Take the average across questions.	**Pros:** • Measure was developed by a team of researchers with extensive experience in bystander intervention. **Cons:** • Measure was tested with mostly White, socioeconomically privileged students. • Measure was not tested across different cultural contexts.	Higher scores indicate participants believe their peers are more likely to engage in prosocial bystander behaviors.
	Citation		
	For more information and the full measure, see: Banyard, V. L., Moynihan, M. M., Cares, A. C., & Warner, R. (2014). How do we know if it works? Measuring outcomes in bystander-focused abuse prevention on campuses. *Psychology of Violence, 4*(1), 101–115. https://www.researchgate.net/publication/263916119_How_Do_We_Know_If_It_Works_Measuring_Outcomes_in_Bystander-Focused_Abuse_Prevention_on_Campuses		

Measurement	Scale Instructions		
MODIFIED BYSTANDER ATTITUDE SCALE-REVISED (BAS-R)	Read the following statements. How likely are you to…		

	Subscales	Example Items	Answer Choices
	High-risk situations	Confront a friend who plans to give someone alcohol to get sex.	Not at all likely (1)
	Post-assault support for victims	Go with a female friend to the police department if she says she was raped.	Unlikely (2) Neither likely nor unlikely (3)
	Post-assault reporting of perpetrators	Tell an RA or other campus authority about information I might have about a rape case even if pressured by my peers to stay silent.	Likely (4) Extremely likely (5)
	Proactive opportunities	Visit a website to learn more about sexual violence.	

Scoring Instructions	Pros/Cons	Other Notes
Sum all responses to create a total score.	**Pros:** • Measure was developed and updated by a team of researchers with extensive experience in bystander intervention. • Measure focuses on empirically reliable indicators. **Cons:** • Measure was tested with students on one college campus. • Measure does not focus on personal actions (e.g., sexist language or pornography use).	Higher scores indicate participants report being more willing to engage in prosocial bystander behaviors and have more positive attitudes about bystander intervention.

Citation
For more information and the full measure, see: McMahon, S., Allen, C. T., Postmus, J. L., McMahon, S. M., Peterson, N. A., & Lowe Hoffman, M. (2014). Measuring bystander attitudes and behavior to prevent sexual violence. *Journal of American College Health, 62*(1), 58–66. https://doi.org/10.1080/07448481.2013.849258

Measurement	Scale Instructions		
BYSTANDER EFFICACY SCALE - SHORT FORM	Please read the following behaviors. Indicate how confident you are that you could do them. Rate your degree of confidence by recording a whole number from 0 to 100 using the scale given below:		
	Subscales	**Example Items**	**Answer Choices**
	Sexual Abuse Short Form	• Get help and resources for a friend who tells me they have been raped. • Do something to help a very drunk person who is being brought upstairs to a bedroom by a group of people at a party. • Do something if I see a woman surrounded by a group of men at a party who looks very uncomfortable. • Speak up to someone who is making excuses for forcing someone to have sex with them.	0% – Can't do 10% – Quite uncertain 20% 30% 40% 50% – Moderately certain 60%
	Intimate Partner Abuse Short Form	• Talk to a friend who I suspect is in an abusive relationship. • Get help if I hear of an abusive relationship in my dorm or apartment. • Speak up to someone who is making excuses for using physical force in a relationship. • Speak up to someone who is calling their partner names or swearing at them.	70% 80% 90% 100% – Very certain
	Scoring Instructions	**Pros/Cons**	**Other Notes**
	Take the average across questions.	**Pros:** • Measure was developed and updated by a team of researchers with extensive experience in bystander intervention for practitioners to use. **Cons:** • None identified.	High scores indicate participants have greater self-efficacy (confidence) to intervene as a bystander.
Citation			
For more information, see: Prevention Innovations Research Center. (2015). *Evidence-based measures of bystander action to prevent sexual abuse and intimate partner violence: Resources for practitioners (short measures).* Durham, NH: University of New Hampshire. https://www.unh.edu/research/sites/default/files/media/2019/09/bystander program evaluation measures - short version compiled.pdf For the full length measure, see: Banyard, V. L. (2008). Measurement and correlates of pro-social bystander behavior: The case of interpersonal violence. *Violence and Victims, 23,* 83–97.			

Measurement	Scale Instructions
READINESS TO HELP SCALE - SHORT FORM	For the next set of questions, please keep in mind the following definitions: <u>Sexual abuse</u> refers to a range of behaviors that are unwanted by the recipient and include remarks about physical appearance, persistent sexual advances that are undesired by the recipient, as well as unwanted touching and unwanted oral, anal, or vaginal penetration. These behaviors could be initiated by someone known or unknown to the recipient, including someone they are in a relationship with. <u>Intimate partner abuse</u> refers to a range of behaviors experienced in the context of any type of intimate relationship or friendship. These behaviors include use of physical force or threats of force against a partner including slapping, punching, throwing objects, threatening with weapons or threatening any kind of physical harm. It can also include extreme emotional abuse such as intimidation, blaming, putting down, making fun of, and name calling. <u>Stalking</u> refers to a range of behaviors that are unwanted by the recipient and that cause fear including repeatedly (two or more times) maintaining unwanted visual or physical proximity to a person, repeatedly conveying oral or written threats, or other activities that are intended to make someone afraid. Examples of stalking include unwelcome communication, including face-to-face, telephone, voice message, electronic mail, written letter, and/or contact; unwelcome gifts or flowers, etc.; threatening or obscene gestures and/or pursuing or following; surveillance; trespassing; or vandalism. Please read each of the following statements and indicate how true each is of you using the following scale.

Subscales	Example Items	Answer Choices
Action - Sexual Abuse Short Form	• Actively involved in projects to deal with sexual abuse on campus. • Recently volunteered on projects focused on ending sexual abuse on campus. • Have been/currently involved in efforts to end sexual abuse on campus.	Not at all true (1) Somewhat untrue (2) Neither true nor untrue (3) Somewhat true (4) Very much true (5)
Taking Responsibility - Sexual Abuse Short Form	• Sometimes I think I should learn more about sexual abuse. • I think I can do something about sexual abuse. • Planning to learn more about the problem of sexual abuse on campus.	
No Awareness - Sexual Abuse Short Form	• Don't think sexual abuse is a problem on campus. • Don't think there is much I can do about sexual abuse on campus. • Not much need for me to think about sexual abuse on campus.	

		Subscales	Example Items	
READINESS TO HELP SCALE-SHORT FORM **(continued)**		Action - Intimate Partner Abuse Short Form	• Actively involved in projects to deal with intimate partner abuse on campus. • Recently volunteered on projects focused on ending intimate partner abuse on campus. • Have been/currently involved in efforts to end intimate partner abuse on campus.	
		Taking Responsibility - Intimate Partner Abuse Short Form	• Sometimes I think I should learn more about intimate partner abuse. • I think I can do something about intimate partner abuse. • Planning to learn more about the problem of intimate partner abuse on campus.	
		No Awareness - Intimate Partner Abuse Short Form	• Don't think intimate partner abuse is a problem on campus. • Don't think there is much I can do about intimate partner abuse on campus. • Not much need for me to think about intimate partner abuse on campus.	**Other Notes**
		Scoring Instructions	**Pros/Cons**	High scores for *Action* and *Taking Responsibility* subscales, and low scores for the *No Awareness* subscale, indicate participants are more ready to engage in violence prevention efforts.
		Take the average of subscale questions.	**Pros:** • Measure was developed and updated by a team of researchers with extensive experience in bystander intervention for practitioners to use. **Cons:** • None identified.	
		Citation		
		For more information, see: Prevention Innovations Research Center. (2015). *Evidence-based measures of bystander action to prevent sexual abuse and intimate partner violence: Resources for practitioners (short measures)*. Durham, NH: University of New Hampshire. https://www.unh.edu/research/sites/default/files/media/2019/09/bystander_program_evaluation_measures_-_short_version_compiled.pdf For the full length measure, see: Banyard, V. L., Moynihan, M. M., Cares, A. C., & Warner, R. (2014). How do we know if it works? Measuring outcomes in bystander-focused abuse prevention on campuses. *Psychology of Violence, 4*(1), 101–115. https://www.researchgate.net/publication/263916119_How_Do_We_Know_If_It_Works_Measuring_Outcomes_in_Bystander-Focused_Abuse_Prevention_on_Campuses		

Measurement	Scale Instructions		
BRIEF INTENT TO HELP: FRIENDS AND STRANGERS	Please read the following list of behaviors and check how likely YOU ARE to engage in these behaviors using the following scale:		
	Subscales	**Example Items**	**Answer Choices**
	Intent to Help Friends: Brief Version	• I approach someone I know if I thought they were in an abusive relationship and let them know I'm here to help. • I let someone who I suspect has been sexually assaulted know I'm available for help and support. • I ask someone who seems upset if they are okay or need help. • If someone said they had an unwanted sexual experience but don't call it rape, I express concern or offer to help. • I express concern to someone I know who has unexplained bruises that may be signs of abuse in a relationship. • I stop and check in on someone who looks intoxicated when they are being taken upstairs at party. • I see a guy talking to a woman I know. He is sitting close to her and by look on her face I can see she is uncomfortable. I ask her if she is okay or try to start a conversation with her. • I see someone and their partner. Partner has fist clenched around the arm of person and person looks upset. I ask if everything is okay. • Ask someone who is being shoved or yelled at by their partner if they need help. • Tell someone if I think their drink was spiked with a drug.	Not at all likely (1) Unlikely (2) Neither likely nor unlikely (3) Likely (4) Extremely likely (5)

Sulley, C., Rabideau, D., Orchowski, L., Jimenez, A., Dragoon, S., Susswein, M., Dube, S., Wood, L., & Busch-Armendariz, N.B. (2021). *The bystander evaluation (BeValued) toolkit: A resource to support program growth and impact*. Austin, TX: Institute on Domestic Violence & Sexual Assault, The University of Texas at Austin.

	Subscales	Example Items	
BRIEF INTENT TO HELP: FRIENDS AND STRANGERS **(continued)**	Intent to Help Strangers: Brief Version	• I talk with people I don't know about sexual abuse and intimate partner abuse as issues for our community. • I talk with people I don't know about going to parties together and staying together and leaving together. • I talk with people I don't know about what makes a relationship abusive and what warning signs might be. • I express concern to someone I don't know if I see their partner exhibiting very jealous behavior and trying to control them. • I share information or resources about sexual assault and/or intimate partner abuse with someone I don't know. • I approach someone I don't know if I think they are in an abusive relationship and let them know that I'm here to help. • I let someone I don't know who I suspect has been sexually assaulted know that I am available for help and support.	
	Scoring Instructions	**Pros/Cons**	**Other Notes**
	Take the average across questions.	**Pros:** • Measure was developed and updated by a team of researchers with extensive experience in bystander intervention for practitioners to use. **Cons:** • None identified.	High scores indicate participants have greater intentions to engage in prosocial bystander behaviors.
	Citation		
	For more information, see the following: Prevention Innovations Research Center. (2015). *Evidence-based measures of bystander action to prevent sexual abuse and intimate partner violence: Resources for practitioners (short measures)*. Durham, NH: University of New Hampshire. https://www.unh.edu/research/sites/default/files/media/2019/09/bystander_program_evaluation_measures_-_short_version_compiled.pdf Banyard, V. L., Moynihan, M. M., Cares, A. C., & Warner, R. (2014). How do we know if it works? Measuring outcomes in bystander-focused abuse prevention on campuses. *Psychology of Violence, 4*(1), 101–115. https://www.researchgate.net/publication/263916119_How_Do_We_Know_If_It_Works_Measuring_Outcomes_in_Bystander-Focused_Abuse_Prevention_on_Campuses		

Measurement	Scale Instructions		
BARRIERS TO BYSTANDER ACTION SCALE - SCHOOL PERSONNEL (BBAS-SP)	Please read the following statements and indicate how much you agree or disagree with each.		
	Subscales	**Example Items**	**Answer Choices**
	Intervening Worsens Situation	Intervening could have negative impacts.	Strongly Disagree (1) Disagree (2)
	Cannot Intervene	I am too busy to intervene.	Agree (3) Strongly Agree (4)
	Scoring Instructions	**Pros/Cons**	**Other Notes**
	Take the average across questions.	**Pros:** • Measure was developed and updated by a team of researchers with extensive experience in bystander intervention. • Measure assesses bystander intervention in school faculty/staff. • Measure can be used for general bystander intervention harm types. **Cons:** • Measure was developed and tested with high school personnel. • Measure was tested in a sample that lacked racial and ethnic diversity. • Measure has lower reliability ratings (i.e., Cronbach's $\alpha =$ .63 - .65).	Higher scores indicate participants have more perceived barriers for intervening.
	Citation		
	For more information and the full measure, see: Edwards, K. M., Sessarego, S. N., Stanley, L. R., Mitchell, K. J., Eckstein, R. P., Rodenhizer, K. A. E., Leyva, P. C., & Banyard, V. L. (2017). Development and psychometrics of instruments to assess school personnel's bystander action in situations of teen relationship abuse and sexual assault. *Journal of Interpersonal Violence, 36*(3–4), NP1586-1606NP. https://doi.org/10.1177/0886260517746946		

Measurement	Scale Instructions		
BYSTANDER EFFICACY SCALE - SCHOOL PERSONNEL (BES-SP)	Please read the following behaviors. Indicate how confident you are that you could do them. Rate your degree of confidence by recording a whole number from 0 to 100 using the scale given below:		
	Subscales	**Example Items**	**Answer Choices**
	Verbal Disagreement	Express concern/discomfort if a student makes a joke about a women's body.	0% – Can't do 10% – Quite uncertain 20%
	Providing Help	Get help for a student who tells me they are in an abusive relationship.	30% 40% 50% – Moderately certain
	Speak Out Against Excuses for Assault	Speak up to a student making excuses for forcing someone to have sex.	60% 70% 80% 90% 100% – Very certain
	Scoring Instructions	**Pros/Cons**	**Other Notes**
	Take the average across questions.	**Pros:** • Measure was developed and updated by a team of researchers with extensive experience in bystander intervention. • Measure assesses bystander intervention in school faculty/staff. **Cons:** • Measure was developed and tested with high school personnel. • Measure was tested in a sample that lacked racial and ethnic diversity.	High scores indicate participants have greater self-efficacy (confidence) to intervene with students.
	Citation		
	For more information and the full measure, see: Edwards, K. M., Sessarego, S. N., Stanley, L. R., Mitchell, K. J., Eckstein, R. P., Rodenhizer, K. A. E., Leyva, P. C., & Banyard, V. L. (2017). Development and psychometrics of instruments to assess school personnel's bystander action in situations of teen relationship abuse and sexual assault. *Journal of Interpersonal Violence, 36*(3–4), NP1586-1606NP. https://doi.org/10.1177/0886260517746946		

Measurement	Scale Instructions		
PERCEPTIONS OF SCHOOL READINESS	Please read the following statements and indicate how much you agree or disagree with each.		
	Subscales	**Example Items**	**Answer Choices**
	Efforts	There are programs and policies to try to prevent sexual assault and abuse.	Strongly Disagree (1) Disagree (2) Agree (3) Strongly Agree (4)
	Knowledge of Efforts	Students know about the programs, rules, and services for sexual assault and abuse.	
	Knowledge of Issue	Students know a lot about relationship abuse and sexual assault.	
	Resources	There is a lot of time/money spent on programs to prevent abuse and sexual assault.	
	Climate Related to School Personnel	Staff at this school care about preventing abuse and sexual assault among students.	
	Scoring Instructions	**Pros/Cons**	**Other Notes**
	Take the average across questions.	**Pros:** • Measure was developed and updated by a team of researchers with extensive experience in bystander intervention. • Measure assesses bystander intervention in school faculty/staff. **Cons:** • Measure was developed and tested with high school personnel. • Measure was tested in a sample that lacked racial and ethnic diversity.	High scores indicate participants believe their school is more ready to address and prevent violence.
	Citation		
	For more information and the full measure, see: Edwards, K. M., Sessarego, S. N., Stanley, L. R., Mitchell, K. J., Eckstein, R. P., Rodenhizer, K. A. E., Leyva, P. C., & Banyard, V. L. (2017). Development and psychometrics of instruments to assess school personnel's bystander action in situations of teen relationship abuse and sexual assault. *Journal of Interpersonal Violence, 36*(3–4), NP1586-1606NP. https://doi.org/10.1177/0886260517746946		

Measurement	Scale Instructions		
BYSTANDER INTENT TO HELP QUESTIONNAIRE - SCHOOL PERSONNEL (BIHQ-SP)	How likely would you be to…		
	Subscales	**Example Items**	**Answer Choices**
	Proactive Bystander Intentions	Tell a group of boys calling a girl a "slut" to stop.	Very Unlikely (1)
			Unlikely (2)
	Reactive Bystander Intentions	Comfort a teen who is a victim of abuse.	Likely (3)
			Very Likely (4)
	Scoring Instructions	**Pros/Cons**	**Other Notes**
	Take the average across questions.	**Pros:** • Measure was developed and updated by a team of researchers with extensive experience in bystander intervention. • Measure assesses bystander intervention in school faculty/staff. **Cons:** • Measure was developed and tested with high school personnel. • Measure was tested in a sample that lacked racial and ethnic diversity.	High scores indicate participants are more willing to intervene with students.
	Citation		
	For more information and the full measure, see: Edwards, K. M., Sessarego, S. N., Stanley, L. R., Mitchell, K. J., Eckstein, R. P., Rodenhizer, K. A. E., Leyva, P. C., & Banyard, V. L. (2017). Development and psychometrics of instruments to assess school personnel's bystander action in situations of teen relationship abuse and sexual assault. *Journal of Interpersonal Violence.* https://doi.org/10.1177/0886260517746946		

Measurement	Scale Instructions		
ILLINOIS RAPE MYTH ACCEPTANCE-UPDATED	Please read the following statements and indicate how much you agree or disagree with each.		
	Subscales	**Example Items**	**Answer Choices**
	She asked for it	If a girl is raped while she is drunk, she is at least somewhat responsible for letting things get out of control.	Strongly Disagree (5) Disagree (4)
	He didn't mean to	When guys rape, it is because of their strong desire for sex.	Neither agree nor disagree (3)
	It wasn't really rape	If a girl doesn't physically resist sex - even when protesting verbally - it can't be considered rape.	Agree (2) Strongly Agree (1)
	She lied	A lot of times, girls who say they were rape agreed to have sex and then regret it.	
	Scoring Instructions	**Pros/Cons**	**Other Notes**
	Take the average across questions.	**Pros:** • Measure was developed and updated by researchers with extensive experience in bystander intervention. • Measure uses terms and language potentially more relatable to college students than the original IRMA scale. **Cons:** • Measure focuses on heteronormative concepts. • Measure language may need to be updated again to be culturally competent and relevant to college students. • Measure doesn't distinguish between acquaintances or strangers.	High scores indicate participants are more likely to reject rape myths.
	Citation		
	For more information and the full measure, see: McMahon, S., & Farmer, G. L. (2011). An updated measure for assessing subtle rape myths. *Social Work Research, 35*(2), 71–81. https://www.jstor.org/stable/42659785		

Measurement	Scale Instructions		
BYSTANDER BEHAVIORS - SHORT FORM	Now, please read the list below and indicate <u>Yes</u> or <u>No</u> for all the items indicating <u>behaviors</u> you have actually engaged in <u>IN THE LAST 2 MONTHS</u>. If you have not been in a situation like that in the past two months, select "no opportunity." **OR** How many times have you done each behavior in the past two months?		
	Subscales	**Example Items**	**Answer Choices**
	Risky Situation	• I let a friend I suspect has been sexually assaulted know that I was available for help and support. • I approached a friend if I thought they were in an abusive relationship and let them know that I was there to help. • If I saw a friend grabbing or pushing their partner, I said something to them. • If a friend said they had an unwanted sexual experience, but they don't call it "rape," I expressed concern and/or offered to help. • I heard a friend talking about using physical force with their partner, spoke up against it, and expressed concern for their partner. • I confronted a friend who made excuses for abusive behavior by others. • If I saw a friend taking a very intoxicated person up to their room, I said something and asked what the friend was doing. • I supported a friend who wanted to report sexual assault or intimate partner abuse that happened to them even if others could get in trouble. • I heard a friend talking about forcing someone to have sex with them, spoke up against it, and expressed concern for the person who was forced. • I expressed disagreement with a friend who said having sex with someone who is passed out or very intoxicated is okay. • If I heard a friend insulting their partner, I said something to them.	Yes (1) No (0) No opportunity

Sulley, C., Rabideau, D., Orchowski, L., Jimenez, A., Dragoon, S., Susswein, M., Dube, S., Wood, L., & Busch-Armendariz, N.B. (2021). *The bystander evaluation (BeValued) toolkit: A resource to support program growth and impact.* Austin, TX: Institute on Domestic Violence & Sexual Assault, The University of Texas at Austin.

	Subscales	Example Items	
BYSTANDER BEHAVIORS - SHORT FORM **(continued)**	Access Resources	• I called a crisis center or community resource for help when a friend told me they experienced sexual or intimate partner abuse. • I called 911 or authorities when a friend needed help because of being hurt sexually or physically. • When I heard that a friend was accused of sexual abuse or intimate partner abuse, I came forward with what I knew rather than keeping silent. • I went with a friend to talk with someone (community resource, police, crisis center, etc.) about an unwanted sexual experience or intimate partner abuse.	
	Proactive Behavior	• I encouraged others to learn more and get involved in preventing sexual or intimate partner abuse. • I talked with a friend about sexual and/or intimate partner abuse as an issue for our community. • I talked with a friend about what makes a relationship abusive and what warning signs might be.	
	Party Safety	• I made sure a friend didn't leave an intoxicated friend behind at a party. • I walked a friend home from a party when they had too much to drink.	
	Scoring Instructions	**Pros/Cons**	**Other Notes**
	Sum all responses to create a total score **OR** skip the 'No opportunity' responses and take the average of the other questions.	**Pros:** • Measure was developed and updated by a team of researchers with extensive experience in bystander intervention for practitioners to use. **Cons:** • None identified.	Higher scores indicate participants report more prosocial bystander behaviors. "Friend" language can be changed to "Stranger" to assess bystander intervention behaviors for strangers.

Sulley, C., Rabideau, D., Orchowski, L., Jimenez, A., Dragoon, S., Susswein, M., Dube, S., Wood, L., & Busch-Armendariz, N.B. (2021). *The bystander evaluation (BeValued) toolkit: A resource to support program growth and impact*. Austin, TX: Institute on Domestic Violence & Sexual Assault, The University of Texas at Austin.

Citation	
BYSTANDER BEHAVIORS - SHORT FORM **(continued)**	For more information, see: Prevention Innovations Research Center. (2015). *Evidence-based measures of bystander action to prevent sexual abuse and intimate partner violence: Resources for practitioners (short measures)*. Durham, NH: University of New Hampshire. https://www.unh.edu/research/sites/default/files/media/2019/09/bystander_program_evaluation_measures_-_short_version_compiled.pdf For the full measure, see: Banyard, V. L., Moynihan, M. M., Cares, A. C., & Warner, R. (2014). How do we know if it works? Measuring outcomes in bystander-focused abuse prevention on campuses. *Psychology of Violence, 4*(1), 101–115. https://www.researchgate.net/publication/263916119_How_Do_We_Know_If_It_Works_Measuring_Outcomes_in_Bystander-Focused_Abuse_Prevention_on_Campuses

Measurement	Scale Instructions		
MODIFIED BYSTANDER BEHAVIOR SCALE – REVISED (BBS-R)	Have you actually participated in the following behaviors in the past 12 months?		
	Subscales	**Example Items**	**Answer Choices**
	Intervention opportunities before, during, and after an assault	Confront a friend if I hear rumors that they had forced someone to have sex.	Yes (1) No (-1) Wasn't in the situation (0)
	Proactive opportunities	Visit a website to learn more about sexual violence.	
	Scoring Instructions	**Pros/Cons**	**Other Notes**
	Sum all responses to create a total score.	**Pros:** • Measure was developed and updated by a team of researchers with extensive experience in bystander intervention. • Measure focuses on empirically reliable indicators. **Cons:** • Measure was tested with students on one college campus. • Measure does not focus on personal actions (e.g., sexist language or pornography use).	Higher scores indicate participants report more prosocial bystander behaviors.
Citation			
For more information and the full measure, see: McMahon, S., Allen, C. T., Postmus, J. L., McMahon, S. M., Peterson, N. A., & Lowe Hoffman, M. (2014). Measuring bystander attitudes and behavior to prevent sexual violence. *Journal of American College Health, 62*(1), 58–66. https://doi.org/10.1080/07448481.2013.849258			

Measurement	Scale Instructions		
PARTNER VICTIMIZATION SCALE	Questions for participants who have had at least one romantic partner: Answer the next questions about any boyfriend, girlfriend, husband, or wife you have had, including exes.		
	Subscales	**Example Items**	**Answer Choices**
	Primary questions	Not including horseplay or joking around, my partner threatened to hurt me, and I thought I might really get hurt.	Yes (1) No (0)
	Follow-up questions	How old were you when this happened?	Early Childhood (birth to 5) (1) Childhood (6–12) (2) Adolescence (13–18) (3) Early Adulthood (19–25) (4) Adulthood (26 or older) (5)
	Scoring Instructions	**Pros/Cons**	**Other Notes**
	Sum all primary responses to create a total victimization score. Follow-up responses are not included in the total victimization score or summed together.	**Pros:** • Measure correlates with other indicators for victimization and adversity. **Cons:** • Measure was tested with a rural, low-income sample. • Measure focuses on intimate partners only.	High victimization scores indicate participants report more violence victimization. Participants only need to answer follow-up questions if they answered 'yes' to a primary question.
	Citation		
	For more information and the full measure, see: Hamby, S. (2013). *Partner Victimization Scale.* Sewanee, TN: Life Paths Research Program. https://www.lifepathsresearch.org/wp-content/uploads/Partner-Victimization-Scale1.pdf		

Measurement	Scale Instructions		
VIOLENCE VICTIMIZATION AND PERPETRATION	[FOR UNWANTED SEX AND HARASSMENT/STALKING QUESTIONS] Since fall of [YEAR], the following things happened to me (or I did the following things to someone else). [FOR DATING VIOLENCE QUESTIONS] Since fall of [YEAR], my partner did the following (or I did the following to my partner). By partner, we mean any current or former spouse, boyfriend, girlfriend, or dating partner or any person with whom you have ever been romantically or sexually involved.		
	Subscales	**Example Items**	**Answer Choices**
	Unwanted Sex	You had unwanted sexual activities because someone threatened to end your relationship if you did not, or you felt pressured by the other person.	0 times (0) 1–2 times (1) 3–5 times (2) 6 or more times (3)
	Sexual harassment and stalking	Someone made gestures, rude remarks, or used sexual body language to embarrass or upset you.	
	Dating Violence - Physical	My Partner pushed or shoved me.	
	Dating Violence - Psychological	My Partner shouted, yelled, insulted, or swore at me.	
	Scoring Instructions	**Pros/Cons**	**Other Notes**
	Take the sum for each subscale to get a frequency measure. For dating violence items, students not in a dating relationship in the past year were included in the analyses with a frequency of violence of 0 times.	**Pros:** • Measure draws items from other commonly used surveys. **Cons:** • The measure's reliability varies across subscales (i.e., Cronbach's $\alpha =$ .54 - .80). • Measure groups sexual harassment and stalking together.	High scores indicate participants report more violence victimization or perpetration. Language of victimization items can be altered to assess perpetration. Questions have been compiled from other measurements (i.e., National Intimate Partner and Sexual Violence Survey, Sexual Experiences Questionnaire, National Violence Against Women Survey, Revised Conflicts Tactic Scale).
	Citation		
	For more information and the full measure, see: Coker, A. L., Fisher, B. S., Bush, H. M., Swan, S. C., Williams, C. M., Clear, E. R., & DeGue, S. (2015). Evaluation of the Green Dot bystander intervention to reduce interpersonal violence among college students across three campuses. *Violence Against Women, 21*(12), 1507–1527. https://doi.org/10.1177/1077801214545284		

Measurement	Scale Instructions
SEXUAL EXPERIENCES SURVEY - SHORT FORM VICTIMIZATION (SES-SFV)/ PERPETRATION (SES-SFP)	The following questions concern sexual experiences that you may have had that were unwanted. We know that these are personal questions, so we do not ask your name or other identifying information. Your information is completely confidential. We hope that this helps you to feel comfortable answering each question honestly. Place a check mark in the box () showing the number of times each experience has happened to you. If several experiences occurred on the same occasion—for example, if one night someone told you some lies and had sex with you when you were drunk, you would check both boxes a and c. "The past 12 months" refers to the past year going back from today. "Since age 14" refers to your life starting on your 14th birthday and stopping one year ago from today.

Subscales	Example Items	Answer Choices
None indicated	Someone fondled, kissed, or rubbed up against the private areas of my body (lips, breast/chest, crotch or butt) or removed some of my clothes without my consent (but did not attempt sexual penetration) by: • Telling lies, threatening to end the relationship, threatening to spread rumors about me, making promises I knew were untrue, or continually verbally pressuring me after I said I didn't want to.	[How many times in the past 12 months?] 0 1 2 3+ [How many times since age 14?] 0 1 2 3+

Scoring Instructions	Pros/Cons	Other Notes
Take the percentage of respondents who responded "yes" to each type of unwanted sex act and/or tactic to compel the unwanted sex act.	**Pros:** • Measure is widely used. • Measure has high reliability ratings (i.e., Cronbach's α = .92 - .99). **Cons:** • Measure may need to update or clarify language to maintain relevance with college students.	High scores indicate participants report more violence victimization or perpetration.

Citation
For more information and the full SES-SFV measure, see: Koss, M. P., Abbey, A., Campbell, R., Cook, S., Norris, J., Testa, M., Ulman, S., West, C. & White, J. (2007). Revising the SES: A collaborative process to improve assessment of sexual aggression and victimization. *Psychology of Women Quarterly, 31*(4), 357–370. https://doi.org/10.1111/j.1471-6402.2007.00385.x

Measurement	Scale Instructions		
DIGITAL DATING ABUSE	Think about your current or most recent dating relationship when answering the following questions: Using the Internet or a cell phone, my dating partner … OR Using the Internet or a cell phone, I…		
	Subscales	**Example Items**	**Answer Choices**
	Digital Direct Aggression	Shared an embarrassing photo or video with others without permission.	Never (1) Sometimes (2)
	Digital Monitoring/Control	Pressured [*me/someone*] to respond quickly to calls, texts, or other messages.	Often (3) Very Often (4)
	Digital Sexual Coercion	Pressured [*me/someone*] to "sext."	
	Follow-up questions: Digital Dating Abuse Victimization Distress	Thinking about the LAST TIME this happened, how much did this upset you?	Not at all (1) A little (2) Some (3) A lot (4)
	Scoring Instructions	**Pros/Cons**	**Other Notes**
	Take the average for each subscale for a frequency score.	**Pros:** • Measure assesses digital behaviors and victimization distress. **Cons:** • Measure only has one item to assess victimization distress. • Measure focuses on digital abuse, which may be associated with offline behaviors that should also be measured.	High scores indicate participants report more digital dating abuse victimization or perpetration. Language of victimization items can be altered to assess perpetration.
	Citation		
	For more information and the full measure, see: Reed, L. A., Tolman, R. M., & Ward, L. M. (2017). Gender matters: Experiences and consequences of digital dating abuse victimization in adolescent dating relationships. *Journal of Adolescence, 59*, 79–89. https://doi.org/10.1016/j.adolescence.2017.05.015		

References

[1] W.K. Kellogg Foundation. (2004). *W.K. Kellogg Foundation evaluation handbook.*
https://www.wkkf.org/resource-directory/resources/2004/01/logic-model-development-guide

[2] W.K. Kellogg Foundation. (2004).

[3] Perkinson, L., Freire, K. & Stocking, M. (2017). *Using essential elements to select, adapt, and evaluate violence prevention approaches.* Centers for Disease Control and Prevention.
https://www.cdc.gov/violenceprevention/pdf/adaptationguidance.pdf

[4] Bloom, B.S. (1984.) *Taxonomy of educational objectives.* Boston, MA: Allyn and Bacon.

[5] Centers for Disease Control and Prevention (CDC). (2011). *Introduction to program evaluation for public health programs: A self-study guide.* Atlanta, GA. CDC.
https://www.cdc.gov/eval/guide/CDCEvalManual.pdf

[6] Kidder, D. P., & Chapel, T. J. (2018). CDC's program evaluation journey: 1999 to present. *Public Health Reports, 133*(4), 356–359.

[7] Driver-Linn, E., & Svensen, L. (2017). *Moving toward a "data ecosystem" to assess campus responses to sexual assault and misconduct: A resource for college and university decision-makers.*
https://www.aau.edu/sites/default/files/AAU-Files/Key-Issues/Campus-Safety/Data-Ecosystem-Approach-Resource-Guide.pdf

[8] Wasco, S. M. & Zadnik, L. (2013). *Assessing campus readiness for prevention: Supporting campuses in creating safe and respectful communities.* Enola, PA: Pennsylvania Coalition Against Rape.
https://www.pcar.org/sites/default/files/resource-pdfs/campus_readiness_manual_2-12_final.pdf

[9] Mitchell, K., & Branigan, P. (2000). Using focus groups to evaluate health promotion interventions. *Health Education, 100*(6), 261–268.

[10] Rounsaville, B. J., Carroll, K. M., & Onken, L. S. (2001). A stage model of behavioral therapies research: Getting started and moving on from stage I. *Clinical Psychology: Science and Practice, 8*(2), 133–142.

[11] McMahon, S., Moran, J., Cusano, J., O'Connor, J., & Chaladoff, N. (2016). Chapter 6: Action planning and dissemination. In *Understanding and responding to campus sexual assault: A guide to climate assessment for colleges and universities* (pp. 32–35). Center on Violence Against Women and Children, School of Social Work, Rutgers, The State University of New Jersey: New Brunswick, NJ. https://socialwork.rutgers.edu/file/1730/download

[12] Wasco, S. M., & Zadnik, L. (2013).

[13] BeVocal. (n.d.). *BeVocal Activities Tracker*. https://utexas.qualtrics.com/jfe/form/SV_6l10gn5KfC3vjkV

[14] CDC Healthy Communities Program. (2014). *Building our understanding: Key concepts of evaluation - What is it and how do you do it?* https://www.cdc.gov/nccdphp/dch/programs/healthycommunitiesprogram/tools/pdf/eval_planning.pdf

[15] Flay, B. R. (1986). Efficacy and effectiveness trials (and other phases of research) in the development of health promotion programs. *Preventive Medicine, 15*(5), 451–474..

[16] Coker, A. L., Bush, H. M., Fisher, B. S., Swan, S. C., Williams, C. M., Clear, E. R., DeGue, S. (2016). Multi-college bystander intervention evaluation for violence prevention. *American Journal of Preventive Medicine, 50*(3), 295–302.

[17] Banyard, V., Moynihan, M. M., & Plante, E. G. (2007). Sexual violence prevention through bystander education: An experimental evaluation. *Journal of Community Psychology, 35*, 463–481.

[18] Basile, K. C., Hertz, M. F., Back, S. E. (2007). *Intimate partner violence and sexual violence victimization assessment instruments for use in healthcare settings: Version 1*. Atlanta (GA): CDC, National Center for Injury Prevention and Control. https://www.cdc.gov/violenceprevention/pdf/ipv/ipvandsvscreening.pdf

[19] WITH US Center for Bystander Intervention at Cal Poly. (2020, July). *National college student bystander intervention survey: NCBIS data summary report, national aggregate report*. WITH US. https://global-uploads.webflow.com/5ba1264a5a1f9d4953de4829/6035cb76afd6d7141fa25df2_NCBIS%20Survey%20Summary%20Report%202020.pdf

[20] Banyard, V. L., Moynihan, M. M., Cares, A. C., & Warner, R. (2014). How do we know if it works? Measuring outcomes in bystander-focused abuse prevention on campuses. *Psychology of Violence, 4*(1), 101–115. https://www.researchgate.net/publication/263916119_How_Do_We_Know_If_It_Works_Measuring_Outcomes_in_Bystander-Focused_Abuse_Prevention_on_Campuses

[21] McMahon, S., Allen, C. T., Postmus, J. L., McMahon, S. M., Peterson, N. A., & Lowe Hoffman, M. (2014). Measuring bystander attitudes and behavior to prevent sexual violence. *Journal of American College Health, 62*(1), 58–66. https://doi.org/10.1080/07448481.2013.849258

[22] Edwards, K. M., Sessarego, S. N., Stanley, L. R., Mitchell, K. J., Eckstein, R. P., Rodenhizer, K. A. E., Leyva, C. & Banyard, V. L. (2017). Development and psychometrics of instruments to assess school personnel's bystander action in situations of teen relationship abuse and sexual assault. *Journal of Interpersonal Violence, 36(3–4), NP1586-1606NP.*. https://doi.org/10.1177/0886260517746946

[23] McMahon, S., & Farmer, G. L. (2011). An updated measure for assessing subtle rape myths. *Social Work Research, 35*(2), 71–81. https://www.jstor.org/stable/42659785

[24] Hamby, S. (2013). *The Partner Victimization Scale.* Sewanee, TN: Life Paths Research Program. https://www.lifepathsresearch.org/wp-content/uploads/Partner-Victimization-Scale1.pdf

[25] Coker, A. L., Fisher, B. S., Bush, H. M., Swan, S. C., Williams, C. M., Clear, E. R., & DeGue, S. (2015). Evaluation of the Green Dot bystander intervention to reduce interpersonal violence among college students across three campuses. *Violence Against Women, 21*(12), 1507–1527. https://doi.org/10.1177/1077801214545284

[26] Koss, M. P., Abbey, A., Campbell, R., Cook, S., Norris, J., Testa, M., Ulman, S., West, C. & White, J. (2007). Revising the SES: A collaborative process to improve assessment of sexual aggression and victimization. *Psychology of Women Quarterly, 31*(4), 357–370. https://doi.org/10.1111/j.1471-6402.2007.00385.x

[27] Johnson, S. M., Murphy, M. J., & Gidycz, C. A. (2017). Reliability and validity of the sexual experiences survey–short forms victimization and perpetration. *Violence and Victims, 32*(1), 78–92.

[28] Reed, L. A., Tolman, R. M., & Ward, L. M. (2017). Gender matters: Experiences and consequences of digital dating abuse victimization in adolescent dating relationships. *Journal of Adolescence, 59*, 79–89. https://doi.org/10.1016/j.adolescence.2017.05.015

[29] InterAction. (2019). *Impact evaluation guidance note and webinar series.* https://www.interaction.org/blog/impact-evaluation-guidance-note-and-webinar-series/

[30] Organisation for Economic Co-operation and Development. (n.d.). *Outline of principles of impact evaluation.* http://www.oecd.org/dac/evaluation/dcdndep/37671602.pdf

[31] USAID. (2013, September). *Technical note: Impact evaluations.* https://www.usaid.gov/sites/default/files/documents/1870/IE_Technical_Note_2013_0903_Final.pdf

[32] Gertler, Paul J., Martinez, S., Premand, P., Rawlings, L. B., & Vermeersch, C. M. J. (2016). *Impact evaluation in practice* (2nd edition). Washington, DC: Inter-American Development Bank and World Bank. https://openknowledge.worldbank.org/handle/10986/25030

[33] CDC. (n.d.). *Types of evaluation.* https://www.cdc.gov/std/program/pupestd/types%20of%20evaluation.pdf

[34] American University. (2020). *What is impact evaluation?* American University, Washington, DC – Online Programs. https://programs.online.american.edu/online-graduate-certificates/project-monitoring/resources/what-is-impact-evaluation

[35] Clearinghouse for Military Family Readiness. (2017). *Program Evaluability Assessment Checklist.* Penn State University. https://militaryfamilies.psu.edu/wp-content/uploads/2018/08/Program_Evaluability_Assessment_Checklist-9-9-17.pdf

[36] Zandniapour, L., and JBS International. (2014). *Impact Evaluability Assessment Tool.* Washington, DC: Corporation for National Community Service. https://americorps.gov/sites/default/files/document/2015_09_03_ImpactEvaluabilityAssessmentTool_ORE.pdf

[37] Manning, C.F. & Goodman, I.F. (2016). *Are you ready? Assessing your organization's readiness to conduct evaluation.* [PowerPoint slides]. Goodman Research Group, Inc. https://cdn.ymaws.com/scholarshipproviders.site-ym.com/resource/resmgr/Files/Conference/2015_Conference/Conference_Materials/12_11_PPT_AreYouReady.pdf

[38] CDC. (2018). *Evaluator self-assessment.* https://www.cdc.gov/eval/tools/self_assessment/EvaluatorSelfAssessment.pdf

[39] Nation, M., Crusto, C., Wandersman, A., Kumpfer, K., Seybolt, D., Morrissey-Kane, E., & Davino, K. (2003). What works in prevention: principles of effective prevention programs. *American Psychologist, 58*(6/7). 449–456.

[40] CDC. (2013). *Developing an effective evaluation report: Setting the course for effective program evaluation.* Atlanta, Georgia: CDC. https://www.cdc.gov/eval/materials/Developing-An-Effective-Evaluation-Report_TAG508.pdf

[41] CDC. (2018). *Evaluator self-assessment.*

[42] Preskill, H., & Jones, N. (2009). *A practical guide for engaging stakeholders in developing evaluation questions.* http://www.pointk.org/resources/files/rwj.stakeholders.final.1.pdf

[43] Foster, C. C., & Louie, J. (2010). *Grassroots action and learning for social change: Evaluating community organizing.* http://www.innonet.org/client_docs/File/center_pubs/evaluating_community_organizing.pdf

[44] Emshoff, J.G. (2008). Researchers, practitioners, and funders: Using the Framework to get us on the same page. *American Journal of Community Psychology, 41*, 393–403.

[45] Anderson, A. (2005). Introduction to theory of change. *The Evaluation Exchange, XI*(2). https://archive.globalfrp.org/evaluation/the-evaluation-exchange/issue-archive/evaluation-methodology/an-introduction-to-theory-of-change

[46] Eldredge, B. K. L., Markham, C. M., Ruiter, R. A. C., Fernàndez, M. E., Kok, G., & Parcel, G. S. (2016). *Planning health promotion programs: An intervention mapping approach* (4th Edition). Wiley: San Francisco, CA.

[47] Fernandez, M. E., Ruiter, R. A. C., Markham, C. M., & Kok, G. (2019). Intervention mapping: Theory- and evidence-based health promotion program planning: Perspectives and examples. *Frontiers in Public Health*, *7*, 209. https://doi.org/10.3389/fpubh.2019.00209

[48] Chapel, T. J. (n.d.). *Learning to love your logic model.* [video]. Atlanta, Georgia: CDC. https://www.cdc.gov/eval/videos/logic-model/LearningToLoveModel-LowRes.mp4

[49] Taylor-Powell, E., Jones, L., & Henert, E. (2002). *Enhancing program performance with logic models.* University of Wisconsin-Extension. http://lmcourse.ces.uwex.edu/

[50] Hawkins, S. R., Clinton-Sherrod, A. M., Irvin, N., Hart, L., & Russell, S. J. (2009). Logic models as a tool for sexual violence prevention program development. *Health Promotion Practice, 10*, 29–37. https://www.jstor.org/stable/26736795

[51] W. K. Kellogg Foundation. (2004). *Using logic models to bring together planning, evaluation, and action: Logic model development guide.* https://www.wkkf.org/resource-directory/resources/2004/01/logic-model-development-guide

[52] CDC. (2018). Evaluation briefs: Writing SMART objectives. https://www.cdc.gov/healthyyouth/evaluation/pdf/brief3b.pdf

[53] Orchowski, L. M., Edwards, K. M., Hollander, J. A., Banyard, V. L., Senn, C. Y., & Gidycz, C. A. (2020). Integrating sexual assault resistance, bystander, and men's social norms strategies to prevent sexual violence on college campuses: A call to action. *Trauma, violence & abuse, 21*(4), 811–827. https://doi.org/10.1177/1524838018789153

[54] McCawley, P. F. (2013). *The logic model for program planning and evaluation.* University of Idaho Extension. https://www.extension.uidaho.edu/publishing/pdf/CIS/CIS1097.pdf

[55] Taylor-Powell, E., Steele, S., & Douglah, M. (1996). *Planning a program evaluation.* Madison, WI: University of Wisconsin Cooperative Extension. https://core.ac.uk/download/pdf/5220766.pdf

[56] Riddle, J. M., Halverson, A., & Barnes, M. (2020). Getting the evaluation focus clear: A shared understanding of what is being evaluated. *Journal of Graduate Medical Education*, *12*(4), 499–500. https://www.jgme.org/doi/pdf/10.4300/JGME-D-20-00701.1

[57] Newcomer, K. E., Hatry, H. P., & Wholey, J. S. (2015). *Handbook of practical program evaluation.* John Wiley & Sons.

[58] Phillips, P. P., & Stawarski, C. A. (2008). *Data collection: Planning for and collecting all types of data* (Vol. 175). John Wiley & Sons.

[59] Vagias, W. M. (2006). *Likert-type scale response anchors.* Clemson International Institute for Tourism & Research Development, Department of Parks, Recreation and Tourism Management. Clemson University.

[60] CDC. (2013). *Developing an effective evaluation report: Setting the course for effective program evaluation.*

[61] Royse, D., Thyer, B.A., & Padgett, D.K. (2009). *Program evaluation: An introduction* (5th ed.). Belmont, CA: Cengage Learning.

[62] Cano, M. & Chanoi, Z. (2019). *Do one thing: Shareable project.* [PowerPoint slides]. The University of Texas at El Paso.

[63] Ontario Centre of Excellence for Child and Youth Mental Health. (n.d.). *Program evaluation toolkit: Tools for planning, doing and using evaluation.* Ontario Centre of Excellence for Child and Youth Mental Health. https://www.cymh.ca/modules/ResourceHub/?id=6D1CDF70-8A99-4432-ABA6-E19862DA6857

[64] Johnson, R. B., Onwuegbuzie, A. J., & Turner, L. A. (2007). Toward a definition of mixed methods research. *Journal of Mixed Methods Research,1*(2), 112–133.

[65] Greene, J. C., Caracelli, V. J., & Graham, W. F. (1989). Toward a conceptual framework for mixed-method evaluation designs. *Educational Evaluation and Policy Analysis, 11*(3), 255–274.

[66] Bélisle, P., & Joseph, L. (2020). *Codebook cookbook: A guide to writing a good codebook for data analysis projects in medicine.* McGill University. http://www.medicine.mcgill.ca/epidemiology/joseph/pbelisle/CodebookCookbook/CodebookCookbook.pdf

[67] CDC. (2013). *Developing an effective evaluation report: Setting the course for effective program evaluation.*

[68] Busch-Armendariz, N. B., Rabideau, D., Sulley, C., Jimenez, A; & Duran, J. (2020). *Bystander intervention program needs assessment: Findings across UT academic institutions.* Austin, TX: Institute on Domestic Violence & Sexual Assault, The University of Texas at Austin.

[69] Porter, S. R., Whitcomb, M. E., & Weitzer, W. H. (2004). Multiple surveys of students and survey fatigue. *New Directions for Institutional Research, 121*, 63–73. http://oia.unm.edu/surveys/survey-fatigue.pdf

[70] Sinickas, A. (2007). Finding a cure for survey fatigue. *Strategic Communication Management, 11*(2), 11. https://search.proquest.com/openview/2d5769d2c0348b8238763c1b46debe2c/1?pq-origsite=gscholar&cbl=44514

[71] Nguyen, H. L. T. (2017, May). *Tired of survey fatigue? Insufficient effort responding due to survey fatigue.* [Master's thesis, Middle Tennessee State University]. https://jewlscholar.mtsu.edu/handle/mtsu/5302

[72] Karlberg, C. (2015, June). *The survey fatigue challenge: Understanding young people's motivation to participate in survey research studies.* [Master's thesis, Lunds Universitet]. http://lup.lub.lu.se/luur/download?func=downloadFile&recordOId=7792773&fileOId=7793183

[73] Fisher, D., Lang, K. S., & Wheaton J. (2010). *Training professionals in the primary prevention of sexual and intimate partner violence: A planning guide.* CDC.

http://www.ncdsv.org/images/CDC_TrainingProfessionalsInThePrimaryPreventionOfSexualAndIPVaPlanningGuide_2010.pdf

[74] Moncher, F. J., & Prinz, R. J. (1991). Treatment fidelity in outcome studies. *Clinical Psychology Review, 11*(3), 247–266.

[75] Mowbray, C. T., Holter, M. C., Teague, G. B., & Bybee, D. (2003). Fidelity criteria: Development, measurement, and validation. *American Journal of Evaluation, 24*(3), 315–340.

[76] Noonan, R. K., Emshoff, J. G., Mooss, A., Armstrong, M., Weinberg, J., & Ball, B. (2009). Adoption, adaptation, and fidelity of implementation of sexual violence prevention programs. *Health Promotion Practice, 10*(1), 59S–70S. https://www.jstor.org/stable/26736798

[77] Katz, J., & Moore, J. (2013). Bystander education training for campus sexual assault prevention: An initial meta-analysis. *Violence and Victims, 28*(6), 1054-1067.

[78] Olney, C. A. & Barnes, S. J. (2013). *Collecting and analyzing evaluation data (2nd edition).* Outreach Evaluation Resource Center. National Network of Libraries of Medicine. https://nnlm.gov/sites/default/files/neo/files/booklet-three.pdf

[79] Olney, C. A. & Barnes, S. J. (2013).

[80] Data (n.d.). In *Collins Dictionary*. https://www.collinsdictionary.com/us/dictionary/english/data

[81] CDC.(2013). *Managing data.* https://www.cdc.gov/globalhealth/healthprotection/fetp/training_modules/10/managing-data_pw_final_09252013.pdf

[82] Royse, D., Thyer, B.A., & Padgett, D.K. (2009).

[83] Oregon State University (n.d.). *What is the institutional review board (IRB)?* http://research.oregonstate.edu/irb/frequently-asked-questions/what-institutional-review-board-irb

[84] CDC. (2011). *Developing an effective evaluation plan.* Atlanta, Georgia: CDC, National Center for Chronic Disease Prevention and Health Promotion, Office on Smoking and Health; Division of Nutrition, Physical Activity, and Obesity. https://www.cdc.gov/obesity/downloads/CDC-Evaluation-Workbook-508.pdf

[85] Royse, D., Thyer, B.A., & Padgett, D.K. (2009).

[86] National Cancer Institute. (2020, September). *Normative beliefs.* National Cancer Institute, Division of Cancer Control & Population Sciences – Behavioral Research Program. https://cancercontrol.cancer.gov/brp/research/constructs/normative-beliefs

[87] Royse, D., Thyer, B.A., & Padgett, D.K. (2009).

[88] Royse, D., Thyer, B.A., & Padgett, D.K. (2009).

[89] Royse, D., Thyer, B.A., & Padgett, D.K. (2009).

[90] W.K. Kellogg Foundation (2004).

[91] Royse, D., Thyer, B.A., & Padgett, D.K. (2009).

[92] Royse, D., Thyer, B.A., & Padgett, D.K. (2009).

[93] CDC. (2011). *Developing an effective evaluation plan.*

[94] Royse, D., Thyer, B.A., & Padgett, D.K. (2009).

[95] CDC. (2018). *Program performance and evaluation office (PPEO).*
https://www.cdc.gov/program/index.htm

[96] Substance Abuse and Mental Health Services Administration (SAMHSA). (2014). *SAMHSA's concept of trauma and guidance for a trauma-informed approach.* HHS Publication No. (SMA) 14-4884. Rockville, MD: Substance Abuse and Mental Health Services Administration.
https://s3.amazonaws.com/static.nicic.gov/Library/028436.pdf

[97] SAMHSA. (2015). *A guide to GPRA data collection using trauma-informed interviewing skills.*

[98] CDC. (1999). *Framework for program evaluation in public health.* Tech. rep., U.S. Department of Health and Human Services, Atlanta.
https://www.cdc.gov/mmwr/preview/mmwrhtml/rr4811a1.htm