Planting the Seeds for High-Quality Program Evaluation in Public Health

National Center for Environmental Health

Division of Environmental Health Science and Practice

Table of Contents

FOREWORD

It is my pleasure to welcome you to *Planting the Seeds for High-Quality Program Evaluation in Public Health.* This textbook is the culmination of a lengthy and extensive journey undertaken by the National Asthma Control Program (NACP). NACP aims to build evaluation capacity within our program and among our partners, including state and local health departments, federal agencies, non-governmental organizations, professional associations, and academic communities.

NACP is a program within Centers for Disease Control and Prevention (CDC), and has long demonstrated its commitment to program evaluation. In 2009, we adopted a formal plan to provide comprehensive evaluation support to the asthma programs we fund. Part of this plan included developing a guidance manual that would help asthma program staff members learn as they were doing—learn about the principles and practices of evaluation as they worked alongside stakeholders to plan, implement, and learn from their evaluations.

We called the manual *Learning and Growing through Evaluation* because we wanted to emphasize that evaluation is a tool for systematically learning how to improve programs—to make them more equitable, effective, and efficient—and, by growing stronger in this evaluative capacity, we further CDC's mission to save lives and protect people.

Over time, our asthma program partners requested additional resources. We worked with them to transform our manual into a set of modules. The modules maintain the learning and growing theme.

We heard from our partners in the public health field that a gap exists between what practitioners need to know when engaging in evaluation as part of their jobs and how evaluation is typically taught in academic settings. We saw—and seized—the opportunity to adapt our Learning and Growing modules for a broader audience. We are pleased to share this textbook with you. Our hope is that this text will plant the seeds for high-quality evaluation throughout the public health field and help all programs share in the success that asthma programs have achieved through high-quality evaluative efforts.

Josephine Malilay, Ph.D., M.P.H.
Chief, Asthma and Community Health Branch
Division of Environmental Health Science and Practice
National Center for Environmental Health, CDC

ACKNOWLEDGEMENTS

The authors gratefully acknowledge the following individuals and organizations who contributed to the development of *Planting the Seeds for High-Quality Program Evaluation in Public Health.* We would like to thank Chief Josephine Malilay and Deputy Chief Pamela Collins of the Asthma and Community Health Branch, and Director Erik Svendsen of the Division of Environmental Health Science and Practice for their direction, encouragement, and support of this effort.

We thank Adrianne Peschard, Javier San Miguel, Nia Hand, and Rachel Ament at Sensis for the production and design that transformed our content into this e-text. We would also like to thank Samuel Dunklin for his assistance in reviewing drafts of this text.

Planting the Seeds for High-Quality Evaluation in Public Health is based on the *Learning and Growing through Evaluation* series. We acknowledge Elizabeth Herman, Kari Cruz, Meredith Stocking, Joanne Abed, Linda Winges, and Shyanika Wijesinha Rose for their contributions to that series. We would also like to acknowledge the asthma program staff members who served as reviewers. For a complete list of these reviewers, visit http://www.cdc.gov/asthma/program_eval/guide.htm.

This text was developed under contract HHSD2002015M8163B, with Darrin L. Brown, CDC/ACHB, acting as contract representative.

The contents of this document do not necessarily reflect the official position of the U.S. Department of Health and Human Services or Centers for Disease Control and Prevention.

Suggested reference:

Wilce, M., Fierro, L.A., Gill, S., Perkins, A., Kuwahara, R., Barrera-Disler, S., Orians, C., Codd, H., Castleman, A.M., Nurmagambetov, T., Anand, M. *Planting the Seeds for High-Quality Program Evaluation in Public Health.* Atlanta, GA: Centers for Disease Control and Prevention, National Center for Environmental Health, Division of Environmental Health Science and Practice, Asthma and Community Health Branch. June 2021.

PREFACE

Evaluation is an important—in fact, central—component of public health practice. Evaluation is recognized as an essential service of public health (CDC, n.d.) and included within several public health professional competency sets (ASPPH, n.d.; CDC & CSTE, 2008; NCHEC, 2020). Additionally, the demand for evaluation is quite high, especially within federal public health agencies such as the Department of Health and Human Services (Lemire et al., 2018). Despite the importance of evaluation in public health, research studies suggest that many public health professionals who enter the field through formal academic routes (e.g., a Master's in Public Health from accredited institutions) may graduate without completing a course in program evaluation (Fierro & Christie, 2011). For students who are fortunate enough to gain exposure to evaluation in their academic pursuits, research suggests that this training often omits topics that are recognized, within the formal discipline of evaluation, as essential to effective evaluation practice. Omitted topics include evaluability assessments, meta-evaluations, and negotiation and conflict resolution skills (Hobson et al., 2019).

As a result of gaps in their training, public health practitioners can find themselves in a situation where they are planning and conducting evaluations or commissioning and monitoring evaluation contracts without the skills and knowledge they need to fulfill the job. These circumstances result in practitioners seeking out professional development activities in evaluation to round out their skillset. We created this text to help shift this dynamic in public health evaluation. To this end, we intentionally created an online evaluation textbook that is free of charge, making it available to anyone who wishes to learn about evaluation, irrespective of their circumstances. Second, we made a conscious effort to integrate knowledge from the formal discipline of evaluation into this text. There was much knowledge to draw upon from 144 professional associations for evaluators across the globe (IOCE, n.d.), a rich and growing empirical and applied knowledge base documented in several professional journals (e.g., the *American Journal of Evaluation*, *New Directions for Evaluation, the Canadian Journal of Program Evaluation*), professional competency sets (King & Stevahn, 2020), and guidelines for practice (AEA, 2011, 2018).

We designed the content so that each chapter will be informative for many audiences, from novice to advanced. After reading the text, individuals who are brand new to evaluation will have a solid foundation for practice. Those who are more advanced will acquire skills and knowledge about evaluation approaches and techniques that they likely have not encountered. Individuals who lead evaluations will benefit from the step-by-step process we present for engaging in evaluation planning and implementation efforts, the introduction we provide to the different approaches that can be used to tailor these steps in practice, and the features we include to call the reader's attention to ways they can reflect on their own actions and context, in real time, to improve their practice. Other professionals who commission or fund evaluations, who supervise staff members or contractors who are responsible for evaluation, or who are participants in evaluation will also benefit from the content included in this text. After engaging with this text, these individuals will have a better understanding of what is entailed in carrying out high-quality evaluations, what they can do to support and sustain high-quality evaluation practices in their organizations, and what to consider in selecting the right external evaluator for their organization.

Individuals working in public health departments, community-based organizations, non-profit organizations, and other contexts where public health evaluation practice is prominent may elect to consult this text regularly as part of their ongoing professional development activities. However, instructors may also find this text to be useful. Instructors delivering a course in program evaluation within public health schools or programs may wish to adopt this text. Others who are delivering workshops on evaluation for public health audiences may find it helpful to list this text as a reference.

This text is organized into three sections. The first provides a general framing for evaluation and includes an introduction to the professional practice of evaluation, the role of evaluation in public health, and various approaches for practicing evaluation. As part of this, you are introduced to the "how" of practice, which includes

critical reflection, **situational awareness**, **cultural responsiveness**, and **interpersonal competence**. These approaches are highlighted at various points throughout the text to emphasize their importance to evaluation practice. In the second section we cover the details of evaluation planning—at both a macro and micro level—as well as what to expect when the rubber meets the road during implementation. The second section also discusses how to support planning and implementation through institutionalizing evaluation capacity in public health organizations. Evaluation planning at the macro level is an advanced exercise. Readers who are just beginning their journey into evaluation (i.e., novice evaluators) may wish to skip Chapter 3 and revisit it after reading the other chapters. The final section of the text focuses on evaluating common infrastructural elements of public health programming (i.e., partnerships, surveillance) and public health interventions.

Several learning tools have been incorporated into the text to deepen your understanding of evaluation and develop your skills. Each chapter includes review and skill-building questions which provide you with an opportunity to assess and apply various evaluation concepts. Sample answers to the review questions are provided in **Appendix A**, so you can check your understanding of key concepts. The skill-building exercises are meant to give you some initial practice, recognizing that in real-world settings similar activities will usually be conducted with stakeholders rather than alone. There is a glossary of key terms for you to reference throughout your reading. We also have included a rich set of appendices to help you deeper into subjects that may have immediate relevance to your work.

We hope that this text provides you with an opportunity to dig into the details of the craft of evaluation at your convenience and that it provides the insights you need to feel increasingly confident in engaging in evaluation practice. It has been our pleasure building our own evaluation capacity since we first put pen to paper in developing the precursor to this text— *Learning and Growing through Evaluation*—and our hope is that sharing the lessons we have learned on our journey will benefit your practice as well.

SECTION I

Framing Evaluation in the Public Health Context

SECTION I: Framing Evaluation in the Public Health Context

Evaluation and evidence-informed decision making are central to public health practice. In recent decades, there has been tremendous growth in the professional discipline of public health evaluation. In this section, we introduce evaluation as a discipline and professional practice, position evaluation within the broader public health context, and offer ideas for expanding existing public health evaluation frameworks with recent advances from the broader field of evaluation.

Whether planning and implementing evaluations within a public health organization or delivering evaluation services to a public health organization, evaluators need a strong foundation on which to build and strengthen their practice. Embedded in this foundation is knowledge of evaluation's role in society, the discipline's core principles and theories, and the concepts that drive high-quality practice. **Chapter 1** introduces you to evaluation and describes its path to becoming a professional discipline and practice. This includes discussions regarding key events in the development of the field and the concepts fundamental to its practice.

In **Chapter 2**, you will become familiar with the CDC's *Framework for Program Evaluation in Public Health* (1999) and the steps considered by many as essential to conducting evaluations. This framework is supplemented by additional dimensions that are essential to high-quality and ethical practice. The chapter also introduces you to prescriptive approaches for evaluation practice (i.e., evaluation theory). There are many evaluation theories. In chapter 2 we highlight several and then describe the connection between the theories, the CDC Framework, and evaluation practice in general.

CHAPTER ONE

Introduction to Evaluation

CHAPTER ONE: Introduction to Evaluation

In everyday communication, the term evaluation is used regularly. Just think about the number of times you have heard, "We will evaluate our options and get back with you on that." When we think about evaluation as a formalized practice, what exactly are we talking about? In this chapter, we define evaluation. We also describe some important milestones from the history of the evaluation field and progress made toward professionalization. Ultimately, this text is about evaluation within public health. Consequently, we describe several ways that evaluation is recognized formally within the field of public health and point to indications of its importance within public health practice.

By the end of Chapter One, the reader will be able to

- Define evaluation.
- Explain some key advancements in the field of evaluation.
- Describe the evolution of evaluation practice.
- Identify two or more ways that evaluation contributes to public health practice.

What is Evaluation?

Evaluation[1, 2] has been defined as the "process of determining the merit, worth, or value of something, or the product of that process" (Scriven, 1991, p. 139). Although we will see that evaluation is well on its way to being established as a formal professional practice, it is important to recognize that evaluation occurs informally, as a routine part of daily life. Humans are inclined to make sense of the world around them. As part of this process, they regularly weigh the pros and cons of different options (Mark et al., 2000). People engage in evaluations to make decisions from the moment they wake (e.g., determining what they will eat for breakfast and what they will wear that day) to the moment they go to sleep (e.g., deciding which pajamas to put on and what time to actually curl up in bed). When making these decisions, evaluation often occurs informally and naturally, sometimes with limited intention.

One might wonder why society needs a formal evaluation discipline given that we get so much practice evaluating every day. As described by Mark, Henry, and Julnes (2000), the natural sense we make of the world around us, though helpful, can also be fraught with bias. In fact, large bodies of research have established recognition of these biases (Mark et al., 2000, p. 5). Mark et al. (2000) argued that the "field of evaluation has been developed to assist, support, and extend natural human abilities to observe, understand, and make judgments about policies and programs" (p. 5).

[1] Though there are several types of evaluation, in this text we focus our discussion on program evaluation rather than other types such as personnel or product evaluation.

[2] The Government Accountability Office's *Key Terms and Concepts* provides additional helpful insights on program evaluation, its role in organizations, and history.

The Professional Practice of Evaluation

Evaluation is a discipline and professional practice, subject to its own standards and principles. It is rooted in evaluation theory and a broad array of methods. Over the years, many different descriptions for evaluation have been presented, but it is generally understood as "the **systematic** collection of information about the activities, characteristics, and results of programs, to make **judgments** about the program, improve or further develop program effectiveness, inform decisions about future programming, or increase understanding" (Patton, 2008, p. 39). The words *systematic* and *judgments* are bolded because they represent two important aspects of evaluation. Systematic collection is an empirical element, which involves the use of scientific methodology to collect and analyze accurate and credible evidence. Judgments are a normative element. Judgments involve deciding about the value, merit, or worth of something (Fourier, 2005, p. 140). This normative aspect very much distinguishes evaluation from other types of inquiry such as research and epidemiologic investigations (Fourier, 2005). For instance, an epidemiologist may conduct an investigation to understand what caused a foodborne outbreak at a recent workplace gathering. This study is systematic in nature; for example, the epidemiologist will gather data about potential sources of exposure from individuals in attendance. To detect similarities and differences in exposure, evidence will include confirmed cases, suspected cases, and individuals who did not get the illness. However, this study would not include a *normative* aspect; we know that foodborne illness is not good, and the epidemiologist is solely interested in uncovering the cause. An evaluation, on the other hand, might examine what *options are best* for preventing foodborne illness during workplace gatherings.

A Brief History of Professional Evaluation Practice

The act of evaluation as a routine process of our daily lives predates professional evaluation. We evaluate our options regularly; for example, selecting a restaurant for dinner, deciding which movie to watch, and opting for one outfit over another. However, by many accounts, evaluation as a profession is a relatively young field that is still taking shape. Most historical accounts of the field suggest that evaluation started in the 1960s with the passage of the Great Society Legislation—a social reform agenda initiated by President Kennedy and expanded under Presidents Johnson and Nixon (Shadish et al., 1991). The Great Society Legislation involved a series of legislation and initiatives (e.g., the Elementary and Secondary Education Act of 1965) which required that an evaluation amendment be attached to every education bill, to improve the social and economic conditions of Americans (House, 1993; Shadish et al., 1991). Prior to this time, evaluation as a professional practice was not commonly performed.

Following the passage of the Great Society Legislation, the demand for program evaluation services and funding to support this work increased substantially. Because there was not an established field of evaluators, professionals from social science research disciplines (e.g., sociology, political science, economics, psychology) filled the void by providing evaluation services (Alkin & King, 2016; Shadish et al., 1991). As such, the field grew organically, leveraging methodologies and the philosophies about how knowledge is constructed from a variety of disciplines. Thus, evaluation, even in the current day, is open to a variety of designs (e.g., case studies, randomized controlled trials), data collection methods (qualitative, quantitative, mixed), and philosophies (e.g., post positivist, constructivist).

Between 1965 and present day, the field has undergone numerous changes. Evaluation has witnessed an expansion in the number of approaches and ideas about what it means to conduct high-quality evaluation; many options are available. Well-recognized textbooks in the field such as Alkin (2013) and Mertens and Wilson (2019) document and describe these approaches in detail. In addition, evaluation has made extensive progress toward becoming a formal profession with the creation of guidelines, the adoption of professional competencies (King & Stevahn, 2020), the establishment of a variety of peer-reviewed journals, and the growth of professional associations worldwide.

The Journey Toward High-Quality Evaluation. Early in the field's development, evaluators were primarily focused on examining program effectiveness to support policy makers in determining which programs

to continue to support and which to defund. Thus, prevailing evaluation approaches of this time tended to focus on causal attribution—answering whether the program produced its intended outcomes. Evaluators of the time also operated under the impression that if they used their expertise to design and implement a high-quality evaluation, those who commissioned the work would naturally make use of it in their decision making. As a result, there was no real need to engage these intended end-users in planning for evaluations. In fact, the prevailing opinion was that doing so could introduce bias (Shadish et al., 1991). Though an accepted practice today, engaging *program participants* (i.e., intended beneficiaries of the program) in designing and perhaps even implementing an evaluation would have been doubly unthinkable.

The Importance of Utility. As more evaluations were performed, it became evident that the end users were, sadly, not using evaluation findings in their decision making (Shadish et al. 1991). In fact, it appeared that policy makers would often make decisions that were in opposition to the empirical evidence (Alkin & King, 2016). Frustrated with these prevailing norms, evaluation scholars began to conduct research studies, what we now recognize as **research on evaluation**, that examined the conditions under which evaluation was and was not used for decision making. Highlights from the research suggested that use of the findings from empirical studies, including evaluation studies, could take many forms. For instance, end users might take near-term actions that were directly in response to the findings. We call near-term actions **instrumental use** (Alkin & King, 2016). Researchers also identified instances in which findings from evaluations contributed to decision makers' *thinking*. The insights from these studies, along with insights from other sources over time, affected future decisions. When insights inform decisions, they are called **conceptual** or **enlightenment use** (Alkin & King, 2016).

On the journey to developing better evaluations, evaluation scholars recognized *utility* as an important characteristic of high-quality evaluation. Observations about evaluation practice suggested that several changes to practice might increase the likelihood of instrumental use. First, as previously noted, evaluations conducted at the inception of the field tended to answer "yes or no" questions about program effectiveness. Such evaluations might provide insights valuable for making decisions about whether to keep or defund a specific program (i.e., **summative evaluation**), but they provide very little information if this is not the decision of interest. For instance, evaluation scholars working in the context of government programs noted that once government programs are put into place, they tend to remain. As a result, what is often needed in these contexts is information that will help program planners and decision makers alter the program (i.e., **formative evaluation**). A simple "yes or no" response from an evaluation does not provide insights about *what* worked well and *what* did not work well nor about *how* change occurred. It is important to know how change occurred so the strategy used or mechanism underlying the change can be leveraged in the future.

New evaluation approaches emerged with the intention of opening up the black box of programs to answer evaluation questions that focused on whether a program was effective (yes or no), *why* it was or was not effective, and *how* change occurred. This type of evaluation was termed **Theory Driven Evaluation** (Chen, 1990; Weiss, 2000a, 2000b). Theory driven evaluation incorporates the idea that the causal mechanisms underlying programs are important to articulate and test. Tools such as theories of change (Weiss, 1998), action models, change models (Chen, 2015), and logic models (McDavid et al., 2019) were developed to describe these causal pathways. Evaluation scholars noted several information sources that could be used to design such models, including practical experience and wisdom garnered from reviews of program documents, conversations with stakeholders, and findings from social science research and tested theories (Chen, 2015; Donaldson, 2007).

Additional insights about how to enhance evaluation use emerged through specific studies on the topic of use. Patton (1977) conducted a multi-site case study of programs operating under what is now the Department of Health and Human Services. As part of this study, Patton made several observations about evaluation use that fundamentally changed the practice of evaluation for decades to come. First, he noticed that when instrumental use occurred, there were almost always individuals involved from the program who were interested in the evaluation and had clear ideas about what they would use the findings to inform. This phenomenon was coined as the **personal factor** by Patton (2008). Later, personal factors played a central role in the creation of an evaluation approach known as **Utilization-Focused Evaluation**.

The personal factor highlighted the central importance of engaging individuals who have an interest or stake in an evaluation. Evaluators previously were concerned that such engagement would lead to bias in evaluations, but the findings from the scholarship, performed by Patton (2005) and others, consistently found that engaging stakeholders in meaningful ways, from the planning phase of an evaluation to the end, was an important factor in facilitating use (Johnson et al., 2009). As such, stakeholder engagement plays an important role in evaluation practice. As you will see in subsequent chapters, it comprises the first step in the CDC's (1999) *Framework for Program Evaluation in Public Health.*

Emphasizing Cultural Responsiveness and Equity. The practice of evaluation continues to evolve. For instance, one aspect of evaluation practice that has taken shape, and continues to grow, relates to cultural responsiveness and the role of evaluation in fostering social justice. **Culturally responsive evaluation**, as described by Hood, Hopson, and Kirkhart (2015),

> is a holistic framework for centering evaluation in culture (Frierson, Hood, Hughes, and Thomas, 2010). It rejects culture-free evaluation and recognizes that culturally defined values and beliefs lie at the heart of any evaluative effort. Evaluation must be designed and carried out in a way that is culturally responsive to these values and beliefs, many of which may be context-specific. (p. 283)

Culturally responsive evaluation is recognized as starting toward the end of the 1990s (though instances of evaluations using principles that are part of this approach occurred as early as the 1930s), and stemmed from work conducted in education, which focused on cultural responsiveness in assessment and pedagogy (Hood, Hopson, & Kirkhart, 2015). More recently, explicit attention has been given to the role of evaluation in promoting social justice in society (Mertens & Wilson, 2019; Neubauer et al., 2020; Symonette, Miller, & Barela, 2020). Several evaluation approaches, including **Deliberative Democratic Evaluation** (House & Howe, 2004), **Transformative Participatory Evaluation** (Mertens & Wilson, 2019), and **Values-Engaged Educative Evaluation** (Greene et al., 2011) explicitly seek to promote social justice and human rights through conducting evaluation in a manner that is attuned and responsive to power differentials and systemic inequities (Mertens & Wilson, 2019). These approaches might be considered a correction to practices early in the field's development when evaluators did not "challenge the underlying assumptions of program designers (social theory), or listen carefully to the experiences of stakeholders" (Hall, 2020, p. 17).

As global concerns (e.g., climate change, novel viruses, forced migration) continue to evolve, the topics addressed in evaluation practice and the logistic and political pressures faced by evaluators will also change. As it has throughout its 60-year history, the field of evaluation will certainly respond to the new challenges by developing more approaches and methodologies. These approaches will be influenced by the growing diversity among evaluation practitioners and the rise in prominence of formerly marginalized voices in the field (Hall, 2020).

Evaluation as a Professional Activity. As noted in the previous section, the field of evaluation has undergone several changes between its start in the mid-1960s and present day. As evaluation started to emerge as a practice distinct from the act of research, professional societies representing evaluators formed. The first evaluation societies included the Eastern Evaluation Research Society in 1978 (EERS, n.d.), followed by the Canadian Evaluation Society in 1981 (CES, n.d.) and the American Evaluation Association in 1986, which formed from a merger between The Evaluation Network and The Evaluation Research Society (Alkin & King, 2016, p. 570). As of December 2020, there were 144 recognized evaluation associations across the globe, known as Voluntary Organizations for Professional Evaluation (VOPEs) (IOCE, n.d.). These organizations support the continued growth of the discipline by providing forums for networking, knowledge exchange, and the advancement of issues of topical importance.

Many VOPEs have disseminated statements to guide evaluators in their practice. For instance, the American Evaluation Association has established *Guiding Principles for Evaluators* (AEA, 2018), which offers guidelines

for ethical evaluation practice. They also provided a *Statement on Cultural Competence in Evaluation* (AEA, 2011), which describes the importance of cultural competence and responsiveness in evaluation and offers several considerations for evaluators in improving their cultural competency. *Program Evaluation Standards* have also been established by the Joint Committee on Standards for Educational Evaluation. The standards provide benchmarks for high-quality evaluation practice both within and beyond educational contexts (Yarbrough et al., 2010). Most recently, in 2018, the AEA Board approved a set of 49 evaluator competencies (King & Stevahn, 2020), organized under five domains. According to AEA (n.d.), these competencies provide "a common language and set of criteria to clarify what it means to be included in the definition of evaluator." The competencies paint a portrait of the many dimensions of evaluation practice and the knowledge, skills, and abilities that individuals or teams conducting evaluation should have. The competencies also provide an opportunity for evaluators to self-reflect on areas of strength in their practice and identify areas for further development (AEA, n.d.; Stevahn et al., 2020).

An important component of establishing a field as a profession includes the presence of academic training institutions (Ayoo et al., 2020). Currently, most evaluators come to the profession as "accidental evaluators" (King & Stevahn, 2013). They often have a full-time job and begin practicing evaluation because of a limited internal evaluation capacity to respond to requests, often from funders, to do evaluation. As a result, many individuals choose to participate in professional development opportunities available within the field (Christie et al., 2013) to enhance their competency in evaluation. For individuals seeking formalized academic training in evaluation, several universities and colleges offer certificates in evaluation and master's or doctoral degrees that include coursework in evaluation, both in the U.S. and globally (LaVelle, 2019).

Evaluation as a Routine Component of Social Programming. Over time, funding for evaluation services has ebbed and flowed. However, research studies performed in the past few years suggest that in recent decades there has been a consistent and strong demand for evaluation services from both the federal government and philanthropic sectors (Kinarsky, 2018; Lemire et al., 2018). For instance, federal agencies increased their spending on contracts supporting evaluation-related activities between Fiscal Year 2010 and Fiscal Year 2017, from $394 million to $651 million (Lemire et al., 2018). In an analysis of data from the Foundation Center, Kinarsky (2018) found that the annual spending from Fiscal Year 2010 and Fiscal Year 2014 for evaluation by foundations was relatively constant, ranging from $237 million to $334 million. Federal **evaluation policies** likely contributed to this strong and consistent demand for evaluation.

In 1993, Congress passed into law the Government Performance and Results Act (GPRA) (S. 20, 1993). Under GPRA, most federal agencies were required to furnish Congress and the Office of Management and Budget (OMB) with a five-year strategic plan, an annual performance plan that clearly stated annual performance goals and measures, and an annual program performance report. The annual performance report was to document the performance attained during the previous year and any actions the agency would take to improve performance when a goal was not attained. GPRA focused heavily on program *results* and was designed to "improve federal program effectiveness and public accountability by promoting a new focus on results, service quality, and customer satisfaction" (S. 20, 1993, Sec 2b3). In 2010, GPRA was expanded slightly with the Government Performance and Results Modernization Act (GPRAMA), which remains in place to the current day (Lemire et al., 2018, p. 67).

In early 2019, the Foundations for Evidence-Based Policymaking Act of 2018 (Evidence Act), which complements and expands upon GPRAMA, was signed into law. As part of the Evidence Act, the same federal agencies covered under GPRAMA are now required to conduct a **capacity assessment**, develop a **learning agenda**, establish an agency-wide **evaluation policy**, designate a chief evaluation officer, and perform several other activities that are intended to strengthen existing data and evaluation efforts (Evidence Act, 2019). Since 2019, the Office of Management and Budget has written several memos to provide guidance to agencies on how to implement the requirements outlined in the Evidence Act (OMB, 2019; OMB, 2020). Most recently, the Presidential Memorandum on Restoring Trust in Government through Scientific Integrity and Evidence-Based

Policy (January 27, 2021) emphasized the importance of evaluation in evidence-based policy making and ensuring the scientific integrity and use of a broad array of methods reflected in the social and behavioral sciences in its design and implementation.

Given the stability of funding for evaluation within the public and philanthropic sectors in recent decades and the consistency with which the federal government has passed evaluation-related statutes over the past three decades, it seems reasonable to suggest that evaluation is here to stay as part of social-service programming. What was once viewed as a "sideline" activity (House, 1993, p. 15) now appears to be mainstream.

Evaluation in Public Health

Evaluation is a ubiquitous practice in many sectors, and public health is no exception. The practice of public health is steeped in evidence-based decision making, and evaluation is acknowledged as an important competency for both individuals and public health organizations. The Certified Public Health Examination (Foster et al., 2018) tests public health practitioners' competency in specific aspects of evaluation practice. At the organizational level, evaluation is part of the criteria for accrediting public health schools, programs, and public health departments. In this section we provide details about some of the many ways evaluation is woven into the field of public health.

Evidence-informed Practice and Evaluation

An important aspect of keeping the public safe from myriad threats includes understanding whether the approaches we use to address critical needs, promote equity, reduce exposure to risk factors, and increase exposure to protective factors, ultimately lead to improvements in health—that is, are they **effective**? Furthermore, ensuring that funding for public health goes as far as possible requires that we understand the extent to which the processes public health practitioners use are **efficient** and result in **cost-effective** solutions. When conducted well, evaluation can provide insights on all these facets of work.

Given the importance of evidence in public health practice, it may come as no surprise that evaluation is viewed as an essential aspect of public health practice. Importantly, both the *conduct* of evaluations and *use* of the findings from evaluations are key to successful public health practice. This cycle—planning → implementation → action—is reflected in the Essential Public Health Services (PHNCI, 2020) as well as in the efforts taken within the public health community to promote the use of evidence-based practices.

Essential Services. Since 1994, the field of public health has recognized ten essential services that should be provided by public health practitioners to support healthy and safe communities. In 2020, the Public Health National Center for Innovations and the de Beaumont Foundation updated these essential services to align with recent advancements in public health practice (CDC, n.d.). Two observations are readily apparent in viewing the list of essential services (**Figure 1.1**) and the details of each (PHNCI, 2020). First, using data and evidence is central to public health practice. In order to "assess and monitor population health status" with an eye toward greater equity, public health professionals require the collection and careful analysis of public health surveillance data to "determine the root causes of health disparities and inequities" (PHNCI, 2020, p. 2) and promote cultural responsiveness. Data are at the heart of the work public health practitioners perform to "investigate, diagnose, and address health problems and hazards affecting the population" (PHNCI, 2020, p. 3). To do so, they must analyze data in real-time to identify and respond

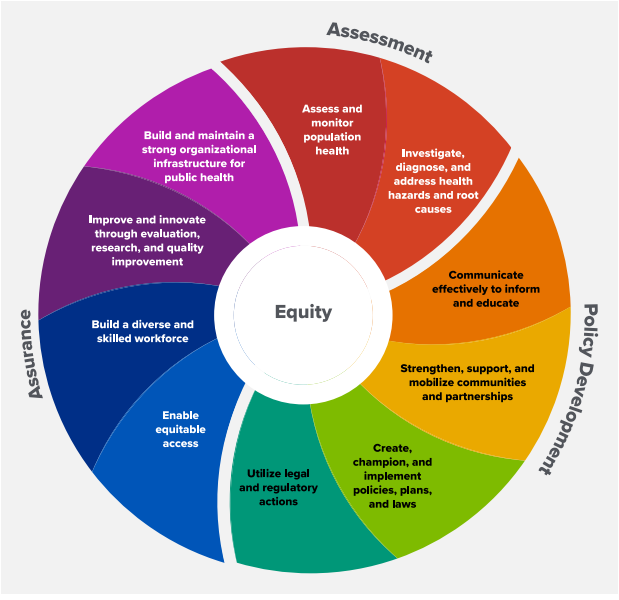


Figure 1.1. 10 Essential Public Health Services
Source: Public Health National Center for Innovations [PHNCI]. (2020). 10 Essential Public Health Services.

to outbreaks, examine surveillance data to address patterns in chronic disease incidence, and make use of findings from analyses of big data. Thus, public health practitioners need evidence to carry out the ten essential services.

A second observation is that evaluation features prominently on the list of essential services, comprising one of the ten: "Improve and innovate public health functions through ongoing evaluation, research, and continuous quality improvement" (PHNCI, 2020, p. 10). Evaluation helps ensure that public health activities are facilitating healthier populations and not resulting in harm. Under this essential public health service, public health practitioners are to use evaluative insights to inform decision making, including actions such as deciding whether to improve or expand existing programming, identifying areas in need of programming, monitoring for unintended consequences, and understanding how many resources should be allocated for future programming. Additionally, as more evaluations are conducted and disseminated, they can be used to inform an important evidence base that can be used for collective decision making in public health about what works in practice.

Compiling the Evidence Base. Evidence of the effectiveness of a public health intervention stems from findings produced through evaluations, The evidence is often a combination of information from observation, theory, and quantitative and qualitative data (Brownson et al., 2009). Evidence-informed public health refers to "the process of distilling and disseminating the best available evidence from research, context and experience to inform and improve public health practice and policy" (National Collaborating Centre for Methods and Tools (NCCMT), n.d.). Evidence-informed public health considers not only the research available on a specific public health topic, but also local community knowledge, needs, and preferences for the acceptability of a particular intervention or policy as well as the political climate and public health resources available for implementation (Ciliska et al., 2008; NCCMT, 2009). Evaluation contributes to building the evidence base by testing the effectiveness of new public health service models and identifying promising practices that can then be scaled (GOA, 2016).

You may also encounter the term "evidence-based practice." Evidence-based practice often refers to public health interventions that are identified as being effective. Effectiveness, in this case, means that the intervention in question appears to consistently produce the intended outcomes (i.e., effects, typically improvements in a specific health outcome) across peer-reviewed studies. Several clearinghouses exist to support evidence-based practice. These entities perform systematic reviews of the existing literature and make recommendations about which interventions public health practitioners should adopt based upon multiple criteria about the strength of evidence for an intervention's effectiveness. These clearinghouses include, but are not necessarily limited to Cochrane, the Campbell Collaboration, the Community Guide, and Health Evidence. The source of data underlying these evidence syntheses are results from evaluations. The methods for scoring or ranking the existing evidence varies depending upon the clearinghouse. Most will include different types of evaluation study designs, such as **randomized controlled trials**, **quasi-experimental designs**, and **non-experimental designs**. However, several assign more weight to evidence from more tightly controlled study designs (e.g., experimental) than other study designs (e.g., non-experimental). We discuss types of study designs in later chapters and provide details for several in **Appendix E**. Though each of these entities uses slightly different categories for classifying levels of evidence, they are similar to the categories in **Table 1.1**.

Table 1.1. Categories of the Effectiveness of Interventions

Proven	Likely effective	Promising	Emerging
Evidence of effectiveness (not efficacy) from an authoritative group (e.g., recommended by the Community Guide, Clinical Guide, Cochrane Collaboration)	Sufficient evidence of effectiveness based on adequately designed peer-reviewed research studies (e.g., results of high-quality studies or evaluations, systematic reviews, or health impact assessments)	Evidence of effectiveness limited based on non-peer reviewed studies, published reports, books, or expert consensus	Evidence of effectiveness absent; theoretical basis or practice-based experience only; plausible
Benefits clearly exceed harms	Benefits exceed harms	Benefits may exceed harms	Benefits may not exceed harms
Used in public health practice: implemented in multiple settings and with different populations, suitable reach, feasible, evaluable, sustainable	Used in public health practice: implemented in multiple settings, suitable reach, feasible, evaluable, sustainable	Used in public health practice: implemented in some settings, but suitability, reach, feasibility, sustainability are uncertain; may have evidence from process evaluation	Use in public health practice limited
Cost effectiveness reasonable	Cost effectiveness likely reasonable	Cost effectiveness uncertain	Cost effectiveness uncertain
Some level of external validity (applicability in multiple settings and different populations) has been demonstrated	Suitable for evidence-based review	Needs further evaluation in controlled studies and community practice	Needs systematic evaluation and practical experience

Source: Secretary's Advisory Committee on National Health Promotion and Disease Prevention Objectives for 2020. (2010, July 26). *Evidence-based clinical and public health: Generating and applying the evidence.* U.S. Department of Health and Human Services.

Public health practitioners work in a variety of settings, including academia, foundations, think tanks, research organizations, and federal, tribal, state, local, and territorial health departments. They play an important role in contributing to the evidence base. As such, public health practitioners, and the institutions in which they work, should be knowledgeable about the act of evaluation, including how to support, plan, and conduct it, as well as how to disseminate and use its findings.

Supporting Evidence-informed Practice.[3] Fortunately, there are several efforts embedded within the public health system that support the conduct and use of evaluation. For instance, future public health practitioners who are engaged in formalized coursework to obtain a Master of Public Health (M.P.H.) or Doctor of Public Health (Dr.P.H.) at an academic institution accredited by CEPH are required to have grounding "in foundational public health knowledge" (CEPH, 2016, p.15). As seen in **Table 1.2**, the ability to identify and appropriately select methods to evaluate programs and to carry out evaluations of public health policies are included in this foundational knowledge.

[3] The terms evidence-informed practice and evidence-based practice are discussed in the text. These concepts are not used interchangeably, rather they are treated as distinct, yet related concepts. As the name suggests, evidence-informed practice refers to practice that is enriched by scientific evidence, but not necessarily restricted to it (Epstein, 2009; Kumah et al., 2019). Evidence-based practice, on the other hand, is the application of scientific evidence, whereby practice directly aligns with research (Kumah et al., 2019).

Public health practitioners sitting for the exam to attain Certified in Public Health (CPH) status also are expected to demonstrate knowledge of evaluation. Based upon a recent job task analysis, the National Board of Public Health Examiners (NBPHE) has structured the CPH exam questions into ten sections, each of which comprises 10% of the total score (Foster et al., 2018; NBPHE, n.d.). One section focuses on program planning and evaluation; however, as seen in **Table 1.2**, evaluation appears within several of the exam sections. Based upon this breakdown, it is evident that NBPHE anticipates that individuals attaining the CPH will have a broad range of knowledge regarding evaluation in public health.

In addition to individual-level knowledge in evaluation, various accrediting bodies also anticipate that public health organizations will engage in evaluative work. For instance, in accrediting public health schools and programs, CEPH expects these academic institutions will regularly engage in evaluating their program and acting on the findings to make improvements (CEPH, 2016). Additionally, the Public Health Accreditation Board (PHAB), which accredits tribal, state, local, and territorial health departments, expects that these departments will engage in continuous quality improvement and that they will share and use insights that come from evaluations to improve the health of populations (see **Table 1.2**). Evaluation, performance management, and quality improvement are related concepts. CDC provides a description of each of these in Evaluation, Performance Management, and Quality Improvement: Understanding the Role They Play to Improve Public Health Webinar Slides.

Table 1.2. Integration of Evaluation in Public Health Competencies and Criteria*

CEPH – MPH Competencies	Foundational Competencies ✓ Planning & Management to Promote Health (p. 17) • Select methods to evaluate public health programs ✓ Policy in Public Health (p. 17) • Evaluate policies for their impact on public health and health equity
National Board of Public Health Examiners (NBPHE)	Evidence-based approaches to public health ✓ Apply evidence-based theories, concepts, and models from a range of social and behavioral disciplines in the development and evaluation of health programs, policies, and interventions. Communication ✓ Provide a rationale for program proposals and evaluations to lay, professional, and policy audiences ✓ Communicate results of evaluation efforts Leadership ✓ Develop strategies to motivate others for collaborative problem solving, decision making, and evaluation ✓ Implement a continuous quality improvement plan ✓ Develop a continuous quality improvement plan ✓ Evaluate organizational performance in relation to strategic and defined goals ✓ Maximize efficiency of programs Collaboration and partnership ✓ Implement methods of shared accountability and performance measurement* with multiple organizations ✓ Identify critical stakeholders for the planning, implementation. and evaluation of health programs, policies, and interventions Program planning and evaluation ✓ Develop and conduct formative evaluation plans ✓ Develop and conduct outcome evaluation plans ✓ Develop process evaluation plans ✓ Apply qualitative evaluation methods ✓ Apply quantitative evaluation methods ✓ Evaluate the benefits of qualitative or quantitative methods for use in evaluation ✓ Assess evaluation reports in relation to their quality, utility, and impact ✓ Assess program performance ✓ Utilize evaluation results to strengthen and enhance activities and programs ✓ Apply evidence-based practices to program planning, implementation, and evaluation ✓ Identify challenges to program implementation ✓ Ensure that program implementation occurs as intended ✓ Plan and communicate steps and procedures for the planning, implementation, and evaluation of health programs, policies, and interventions ✓ Apply evaluation frameworks to measure the performance and impact of health programs, policies, and systems Program Management ✓ Develop monitoring and evaluation frameworks to assess programs
Public Health Accreditation Board (PHAB)	Domain 9: Evaluate and continuously improve processes, programs, and interventions (p. 203) ✓ Use a performance management system to monitor achievement of organizational objectives ✓ Develop and implement quality-improvement processes integrated into organizational practice, programs, processes, and interventions. Domain 10: Contribute to and apply the evidence base of public health (p. 219) ✓ Promote understanding and use of the current body of research results, evaluations, and evidence-based practices with appropriate audiences

* Note that items within this table are extracted verbatim from the listed sources. Page numbers indicate reference page, where available.

Chapter Summary

In this chapter we provided an overview of evaluation as a discipline, its history, and its evolution as a profession. We also examined the role of evaluation in public health, from its inclusion in the Essential Public Health Services to its contribution to evidence-informed and evidence-based decision making in public health contexts. What is clear from this chapter is that evaluation plays a critical role in public health practice and that integrating the lessons learned over the past 60 years, from the formal discipline of evaluation, can only strengthen existing practice. In **Chapter 2**, we more formally integrate public health evaluation practice with advances from the evaluation field.

Review Questions

1. What is the professional practice of evaluation, and why is it important? How is it distinct from research and epidemiologic investigations?
2. Why is it important to conduct systematic evaluations of public health programs? What risks are present when program evaluations are not systematic?
3. Why is use important to evaluation? How might each kind of evaluation use (i.e., instrumental, conceptual, or enlightenment) improve health programs?
4. In what ways is evaluation and the use of evidence for decision making already integrated into public health practice?

CHAPTER TWO

Approaches to Evaluation Practice

CHAPTER TWO: Approaches to Evaluation Practice

There are many ways to plan and conduct evaluations (Alkin, 2013). We begin this chapter by describing a well-established, and well-recognized overarching framework for evaluation practice: CDC's (1999) *Framework for Program Evaluation in Public Health*. The framework is comprised of six common steps that evaluators can use to plan and implement evaluations. First, we provide a description of the CDC Framework. Then, we modify the framework to incorporate some advancements that have occurred in the evaluation field since the Framework's inception. Finally, we review several approaches to evaluation (i.e., evaluation theories) and connect their prescriptions to the steps articulated in CDC's framework.

By the end of Chapter Two, the reader will be able to

- Describe the general content of the six steps in CDC's Framework.
- Explain at least two additional dimensions of evaluation practice beyond those reflected in CDC's Framework.
- Provide a definition of evaluation theory.
- Name at least two evaluation theories and discuss how their application would affect evaluation implementation.

CDC's Evaluation Framework: A General Approach to Evaluation in Public Health

In 1999, CDC offered support for public health practitioners in conducting high-quality evaluation with the publication of their *Framework for Program Evaluation in Public Health*. The CDC Framework provides guidance on developing evaluation strategies that are appropriate to the public health field and consists of six steps and four evaluation standards to guide strategic choices in developing an evaluation (**Figure 2.1**).

Though arranged in numeric order (Step 1–Step 6), the Framework is not intended to be used in a linear manner. Rather, it is important to consider the dynamic interplay among the steps while planning and implementing an evaluation. When initiating an evaluation, it is essential to keep the end-goal in mind—namely ensuring the use of the evaluation findings among **stakeholders** to take actions for improving the public's health. As such, in *Step 1—Engage Stakeholders*, evaluators should consider the world of possible intended end users of the evaluation, from those who may take immediate actions on the findings for programmatic improvement to those who may be unaffected by the evaluation but interested in using

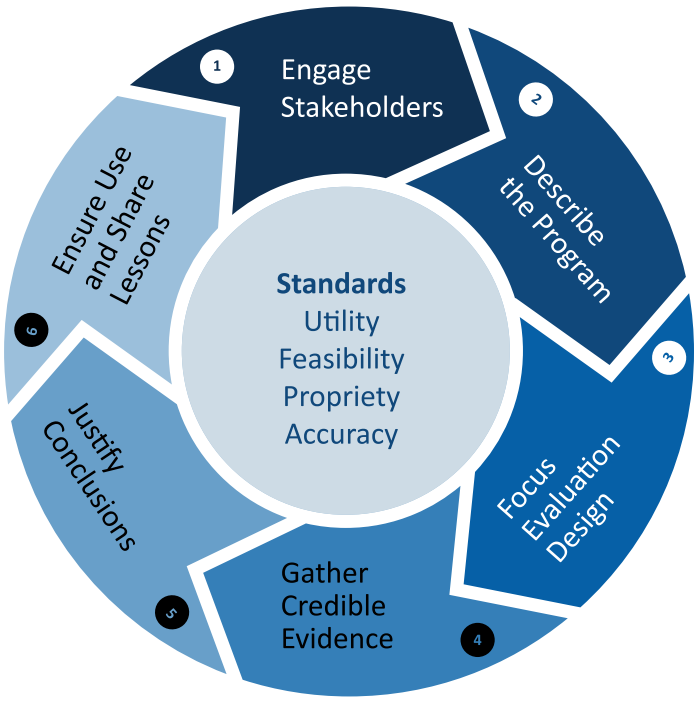


Figure 2.1. CDC Framework for Program Evaluation in Public Health (1999)

the findings to consider how they may approach future public health practice. In *Step 2—Describe the Program*, evaluators should engage with stakeholders and leverage information from other sources (e.g., prior evaluations, research findings, social science, other relevant theories) to develop a visual image that describes what actions the program will take to improve public health. The visual should also describe the general pathways through which the actions lead to the intended improvements. Such pictorial representations of the program may be depicted using several tools, well known to the field of evaluation, such as **logic models** or **theories of change**.

Once the program logic is articulated clearly, evaluators can engage stakeholders in the process of establishing the boundaries of the evaluation study and the methods that will be used for the inquiry. In *Step 3—Focus the Evaluation Design*, evaluators help stakeholders identify a range of possible questions that could be examined through the evaluation and, recognizing that not all aspects of a program can be evaluated in any given study, facilitate the prioritization and refinement of the questions that will ultimately be answered. Using the **key evaluation questions** as a guide, evaluators can work with stakeholders to identify one or more evaluation study designs such as experimental, quasi-experimental, and non-experimental, that will provide the most accurate insights, while also being feasible to implement given logistical and resource constraints. In *Step 4—Gather Credible Evidence*, it is time to decide what specific data will be accessed or collected to formulate responses to the key evaluation questions.

As described in **Chapter 1**, a key difference between research and evaluation is that in evaluation, a *judgment* is formulated about the object that is the focus of the evaluation—what we call the **evaluand**. This is the topic of *Step 5—Justify Conclusions*, in which the findings from an evaluation are compared against stakeholder values. As noted earlier, planning for *Step 6—Ensuring Use and Sharing Lessons Learned*, starts at the beginning of an evaluation with consideration of which stakeholders to engage and the anticipated end-uses of the findings. As you will see in future chapters, sharing lessons learned from an evaluation is a process that occurs throughout the entirety of an evaluation, not just the end.

The standards located at the center of the CDC Framework—utility, feasibility, propriety, and accuracy—were originally developed by the Joint Commission on Standards for Educational Evaluation (JCSEE). The standards provide guideposts for judging the quality of evaluations (Yarbrough et al., 2010). Evaluators and their teams can consult these standards when they are trying to make decisions about how to structure the evaluation during the design phase (e.g., which questions to prioritize, which of several evaluation designs might be best to use). They can also consult these standards during the evaluation to reflect upon how the implementation process is progressing and where there may be opportunities to improve the quality of the evaluation. In **Table 2.1** and **2.2**, we summarize the six steps of CDC's framework and the four standards developed by the JCSEE.

Table 2.1. Six Steps in the CDC Framework for Program Evaluation

Step	Description
Step 1 **Engage Stakeholders**	Evaluation stakeholders are people or organizations that are invested or interested in the results of the evaluation or have a stake in what will be done with evaluation results. Representing their needs and interests throughout the process is fundamental to good program evaluation. A checklist to assist with the implementation of Step 1 is available at https://www.cdc.gov/eval/steps/step1/index.htm
Step 2 **Describe the Program**	A comprehensive program description clarifies the need for your program, the activities you are undertaking to address this need, and the program's intended outcomes. This can help you when it is time to focus your evaluation on a limited set of questions of central importance. Note that in this step, you are describing the program and not the evaluation. Various tools such as logic models, Program Impact Models, and theories of change will be introduced to help you depict your program and the anticipated outcomes. Such models can help stakeholders reach a shared understanding of the program. A checklist to assist with the implementation of Step 2 is available at https://www.cdc.gov/eval/steps/step2/index.htm
Step 3 **Focus the Evaluation Design**	Focusing the evaluation involves determining the most important evaluation questions and the most appropriate design for an evaluation, given time and resource constraints. An entire program does not need to be evaluated all at once. Rather, the right focus for an evaluation will depend on such items as the length of time the program has been in place, what questions are being asked, who is asking them, and what will be done with the resulting information. A checklist to assist with the implementation of Step 3 is available at https://www.cdc.gov/eval/steps/step3/index.htm
Step 4 **Gather Credible Evidence**	Once you have described the program and focused the evaluation, the next task is to gather data to answer the evaluation questions. Evidence gathering should include consideration of each of the following: indicators, sources of evidence or methods of data collection, quality, quantity, and logistics.
Step 5 **Justify Conclusions**	When agencies, communities, and other stakeholders agree that evaluation findings are justified, they will be more inclined to take action on the evaluation results. As stated in the CDC Framework, "Conclusions become justified when analyzed and synthesized evidence is interpreted through the 'prism' of values that stakeholders bring, and then judged accordingly." This step encompasses analyzing the data you have collected, making observations or recommendations about the program based on the analysis, and justifying the evaluation findings by comparing the evidence against stakeholder values.
Step 6 **Ensure Use and Share Lessons Learned**	The purpose(s) you identified early in the evaluation process should guide the use of evaluation results (e.g., demonstrating effectiveness of the program, modifying program planning, accountability). To ensure that evaluation results are used by key stakeholders, consider the timing, format, and key audiences for sharing information about the evaluation process and findings.

Table 2.2. Standards * included in the CDC Framework for Program Evaluation

Standard	Critical Questions
Utility	Who needs the evaluation results? For what purpose do they need the evaluation results and why are they interested in the evaluation? Will the evaluation provide relevant information in a timely manner for them?
Feasibility	Are the planned evaluation activities realistic given the time, resources, and expertise? How can planned evaluation activities be implemented with minimal program disruption?
Propriety	Does the evaluation protect the rights of individuals and the welfare of those involved? Does it engage those most directly affected by the program and changes in the program, such as participants or the surrounding community?
Accuracy	Will the evaluation produce findings that are valid and reliable, given the needs of those who will use the results?

*These standards were originally developed by the Joint Committee on Standards for Educational Evaluation. An updated version was published in 2010, which includes a fifth standard: evaluation accountability (Yarbrough et al., 2010).

Since the publication of the CDC Framework in 1999, the JCSEE has updated the program evaluation standards. One key change included the addition of a fifth standard: **evaluation accountability** (Yarborough et al., 2010). This standard emphasizes the importance of documenting evaluations appropriately and putting into place methods to evaluate the evaluation itself—what the evaluation field calls **meta-evaluation**. Ensuring that appropriate documentation is retained throughout an evaluation, regarding the planned procedures and any changes that took place during implementation, is important for upholding the accountability of the work. In addition, periodically examining the quality of the evaluation implementation and outcomes by using the JCSEE standards can also foster an environment of accountability while providing insights to improve the evaluation in real-time and enhance future practice. Such meta-evaluation can be performed by the evaluation team carrying out the evaluation or by a third party (or both), during or after, an evaluation.

Enhancing the CDC Evaluation Framework

The steps presented in the CDC Framework describe the general process most evaluators would agree needs to be considered in any evaluation. Several advancements have transpired in the evaluation field since its publication. Given this, we present an expanded conceptualization of the CDC Framework in **Figure 2.2**, in which we integrate several aspects of evaluation practice that are now recognized as mainstream in most evaluation circles. In this section, we describe each of these additions.

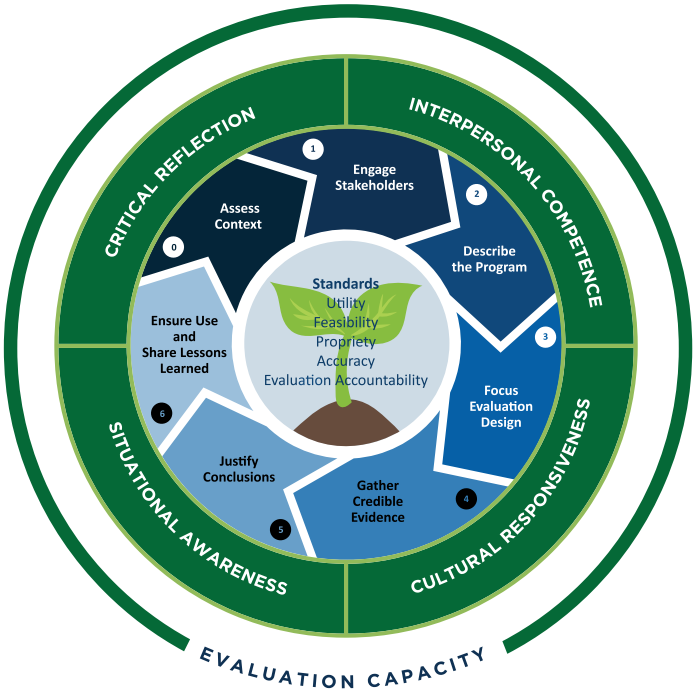


Figure 2.2. Enhanced CDC Framework

Getting Started—Assessing Context

Prior to embarking on the steps described in the CDC Framework, one must first gather a rich understanding of the context within which the evaluand (the object of focus for the evaluation) is situated and integrally intertwined (Greene at al., 2011). For instance, as noted in Public Health 3.0,[4] "today a

[4] Public Health 3.0 is a new public health model in which local governments partner with communities and local leaders in different sectors to address the social, economic, and environmental determinants of health (DeSalvo et al., 2017).

person's zip code may be a stronger determinant of health than is his or her genetic code" (DeSalvo et al., 2017, p. 1). In advocating for an upgrade to public health practice, DeSalvo et al. (2016, 2017) call our attention to the central role that social determinants of health play in morbidity and mortality. Furthermore, these authors point out many systemic biases that are giving rise to the large disparities witnessed in health outcomes. Place, community, history, power, and privilege are all part of the constellation of factors that drive these inequities in health. For public health evaluators to produce an accurate evaluation that provides rich insights that are actionable in context, it is critical the evaluators understand the context within which the evaluand and its participants reside. Public health evaluators should also understand the dynamics that affect the design and implementation of the evaluand. Understanding the context in which an evaluand is situated is so important in evaluation practice that it is now included as a full domain in the AEA Evaluator Competencies (King & Stevahn, 2020). As explained by King and Stevahn (2020),

> The context domain includes competencies relevant to the various features of settings in which evaluation occurs, such as place and context, people and their values or beliefs, cultural and historical circumstances, ways that power and privilege are brought to bear, and other pertinent characteristics. (p. 51)

The "How" of Evaluation Practice

As shown in **Figure 2.2**, the CDC Framework is now located within another circle. The factors included inside of the external circle—critical reflection, cultural responsiveness, situational awareness, and interpersonal competence—describe characteristics that are important for evaluators to embody as they carry out the steps described in the CDC Framework. As evaluators, our ability to be aware of ourselves, others, and the broader environments in which we are working—the *how* of our practice—is equally as important as the technical steps we carry out.

First, who we are as people is central to who we are as professionals and how we practice. As much as we may see our personal and professional selves as separate entities, the distinction is likely minimal, if it exists at all. Each of us holds a worldview that shapes how we perceive and respond to the events of daily life, regardless of the context (Mertens & Wilson, 2019; Mezirow, 1990). The beliefs, values, and assumptions that inform our worldview shape how we make sense of and participate in the world around us. However, worldviews are also problematic in that they, often unbeknownst to us, prescribe what we do and do not pay attention to. Worldviews also prescribe how we perceive and judge information. This can present both quality and ethical limitations in evaluation (Mezirow, 1990, 2009).

The methodologist within us may argue that such assumptions can be controlled through the application of scientific rules and procedures. Yet, that view alone is rooted in an empirical assumption. A different approach is to engage in ongoing **critical reflection** throughout evaluation processes. Critical reflection enables us to unearth and scrutinize the values, beliefs, and assumptions underlying our worldviews so they can be understood, challenged, and modified accordingly (Mezirow, 1990, 2009). Doing so allows us to gain a more truthful and valid interpretation

The "How" of Practice
Critical Reflection

Approaches to critical reflection can, and should, take many forms. Some people are more comfortable with solitary pursuits, while others appreciate a little help from a critical friend(s). Think about your passions and creative outlets and how they can foster your critical reflection. At the same time, don't be afraid to try something new—it may stimulate new or different thinking.

Like many things, it's content over form. If you're new to critical reflection, a good starting point is examining your biases and their practice implications. Not sure what your biases are? Try taking the Implicit Association Test, developed by Project Implicit. The test was designed to help people identify their hidden biases. Other strategies such as Brookfield's (2017) Four Lenses of Critical Reflection and the DEAL model (Ash & Clayton, 2009) can also help to push your critical reflection.

of our experiences, both in the present and in the future. We can also gain greater awareness of ourselves and our practice.

This sense of self-awareness is also key in our ability as evaluators to be **culturally responsive**. As noted by DHHS (2014),

> When we conduct an evaluation, everything we do reflects our own cultural values and perspectives—from the evaluation purpose, the questions we develop, and the methodologies we select to our interpretation of the findings and the recommendations we make based on those findings. Because culture is influenced by many characteristics (i.e., race, ethnicity, language, gender, age, religion, sexual orientation, education, and experience), it is important that we stop and reflect on our own culture before embarking on an evaluation. To conduct culturally competent evaluations, we must learn and appreciate each program's cultural context and acknowledge that we may view and interpret the world differently from many evaluation stakeholders. (p.3)

Thus, we must know the intersectionality of the factors that comprise our own culture and how this may positively or negatively impact our work as evaluators in other contexts. Such reflection will help as we navigate the complexities of designing and implementing a culturally responsive evaluation—where each step in the process has direct and indirect ties to culture (Hood, Hopson, & Kirkhart, 2015).

Evaluations are conducted in a complex world. Although you may have a clear plan **(Chapter 4)** to carry out the evaluation, the best laid plans do not always come to fruition **(Chapter 5)**. As a result, evaluators need to be acutely **aware of the situation** within which they are operating. They should be willing to adapt and change as needed to the ever evolving context in which staff members come and go, priority needs are refined, budgets are adjusted up or down, and new questions surface that need immediate attention. As described by Patton (2008), to implement a high-quality evaluation, evaluators need to be "active, reactive, interactive, and adaptive" (p. 207). Evaluators can acknowledge the reality of an ever-changing context by building flexibility into work plans and budgets as much as possible at the design phase of an evaluation. Funders of evaluation can do the same by adopting internal policies that include the ability to refine and revise budgets and workplans without creating undue inefficiencies.

Additionally, as seen in CDC's framework, a significant portion of evaluation work includes engaging stakeholders. As described by King and Stevahn (2013), and as we discuss later in this chapter, the extent to which evaluators engage with stakeholders varies based on several factors. Some of these factors include evaluator preference, what the evaluator envisions as their role (e.g., teacher, coach, critical friend, unbiased methodologist), and even what they envision the role of evaluation to be in society (e.g., social betterment, social justice, knowledge development). Even when evaluators minimally engage with stakeholders, there is *still* engagement. As such, embracing skills in **interpersonal competence** is important for any evaluator. Similar to our discussion on context earlier in this section, the 2018 AEA Evaluator Competencies include an entire domain dedicated to interpersonal competence. The interpersonal competence domain highlights the importance of establishing trust, understanding power and privilege in context, understanding how power and privilege affect the evaluation, being able to address conflicts that may arise, and many other aspects of practice (King & Stevahn, 2020).

Evaluation Capacity

Finally, high-quality evaluation practice that aligns with the elements reflected in **Figure 2.2** requires that adequate evaluation capacity exists for individuals, teams, and public health organizations. Since this evaluation capacity supports an effective evaluative function, we have placed it at the bottom of the figure. For evaluation planning and implementation to work well, program stakeholders need to value evaluation and have a strong foundational knowledge of evaluation with satisfactory evaluation skills. In addition, public health organizations need to have an infrastructure that provides what is necessary to carry out evaluations and to *use* evaluative

insights. As you will see in subsequent chapters, it is important to integrate activities for **evaluation capacity building (ECB)** into evaluation plans. Here we provide some insights about ways to build evaluation capacity and what this capacity looks like at the individual and organizational level. Prior to creating an evaluation capacity building plan, be sure to assess the current capacity. Often individuals, teams, and organizations already have capacity that simply needs to be strengthened or expanded upon.

ECB is "…the intentional work to continuously create and sustain overall organizational processes that make quality evaluation and its uses routine" (Stockdill, Baizerman & Compton, 2002, p. 14). This "intentional work" includes building capacity in evaluation at both an individual and organizational level. ECB develops individual knowledge, skills, and attitudes to do or use evaluation and helps to create the practices, processes, policies, and resources needed to sustain evaluation practice within a particular organization (Preskill & Boyle, 2008).

Some Components of Evaluation Capacity. Evaluation capacity can be strengthened within an individual by building knowledge of how to do or use evaluation (e.g., understanding how evaluation can contribute to decision making or what qualities to look for when hiring an external evaluator) (Preskill & Boyle, 2008). Skills needed for evaluation capacity include practical abilities or behaviors such as the ability to write an evaluation plan, collect data, or communicate findings to different audiences. Individual attitudes toward evaluation, such as the belief that evaluation is a worthwhile investment or that evaluation should be incorporated into the program design process, are examples of how attitudes can affect the individual and organizational capacity to conduct evaluation. Furthermore, evaluation capacity efforts can begin sensitizing stakeholders to important nuances of evaluation practice such as the "hows" articulated in the previous section, in particular the importance of cultural responsiveness and being situationally aware.

To understand which areas need ECB, it may be helpful to think of the stakeholders involved in the evaluation process and the evaluation activities that will be carried out when planning and implementing the evaluation. Are there certain staff members who will be implementing the evaluation that may need training or support? Is the leadership supportive of evaluation or is there a need to garner their support as the evaluation progresses? Is this the first time some stakeholders have been involved in evaluation? Will they need some upfront training, or ongoing coaching or technical support to fully engage in creating a strategic evaluation plan **(Chapter 3)**, developing an individual evaluation plan **(Chapter 4)**, or assisting with implementing the evaluation **(Chapter 5)**?

Many types of organizational-level evaluation capacities are important for supporting an environment where high-quality evaluation is practiced, used, and sustained. For example, having leaders who support evaluation by ensuring adequate resources are available and staff members know the importance of integrating evaluation findings into their work is essential (Lopez, 2018). Additionally, having an organizational culture that values and champions not only doing but also learning from evaluation is an enabling factor (Preskill & Boyle, 2008). Other organizational evaluation capacities may seem intuitively obvious, but in our experience, they are often overlooked during the planning phase. For example, do focus groups have audio recorders so you can generate transcripts for qualitative analysis? Have you arranged a subscription to an online survey platform? Does the survey platform include the flexibility to ask and respond to survey questions in a manner that aligns with what is outlined in the evaluation plan? Is appropriate data analysis software available to perform the quantitative and qualitative analyses outlined in the evaluation plan? If it is the plan to share findings with community members in person, will you need to rent a physical space or do you have adequate space within your organization? Can community members easily access the meeting space?

Some ECB Strategies. Many strategies exist for building evaluation capacity (Bourgeois et al., 2021). Prior to selecting specific strategies, consider what types of capacities need to be built, in which people or groups, and why. For instance, it may be that your aim is to develop organizational evaluation capacity to support high-quality evaluation. More specifically, perhaps you would like to include a simple statement in the staff member onboarding packet that articulates the importance of evaluation and how staff are expected to contribute to or use evaluation as part of their regular work. To make this a reality, the individuals in positions of authority within the

organization must recognize the importance of evaluation, value evaluation, and ultimately give permission to staff members within the human resources department to include this statement in the onboarding materials. Thus, in this case, and most others, building evaluation capacity starts with *individuals*.

At an individual level, knowledge, skills, and attitudes about evaluation can be cultivated by using several types of ECB strategies. Strategies may include providing written materials, training, and workshops about evaluation; convening communities of practice to share evaluation information and experiences; receiving mentoring, coaching, or technical assistance from an evaluator; or participating in an evaluation internship or fellowship (Preskill & Boyle, 2008). Holding data dialogues to discuss and interpret findings, participating in an evaluation working group, and building a network of people interested in evaluation are other strategies (Preskill & Boyle, 2008; Bourgeois et al., 2021).

Perhaps most surprising to newcomers is the idea that individual-level evaluation capacity can also be built by *engaging in the act of evaluation*. Many people are unfamiliar with the process of evaluation—they often come to the evaluation table with many assumptions about what evaluation is, what questions it can answer, and how it is done. By participating in an evaluation, people learn about evaluation and can subsequently adjust their existing mental models about the value, utility, and practice of evaluation. Learning about evaluation by engaging in the process and leveraging insights from the process itself to make change is known as process use. Patton and Horton (2009) describe **process use** in the following way:

> Process use occurs when those involved in the evaluation learn from the evaluation process itself or make programme changes based on the evaluation process rather than the evaluation's findings. Process use, then, includes cognitive, attitudinal, and behaviour changes in individuals, and programme or organizational changes resulting, either directly or indirectly, from engagement in the evaluation process and learning to think evaluatively (for example, goal clarification, conceptualizing the programme's theory of action, identifying evaluation priorities, struggling with measurement issues, participation in design and interpretation). Process use is reflected in statements like this: "During the evaluation, we realized some ways to improve our work with partners, and we began implementing them even before the evaluation was done and the report was written." Process use includes the effects of evaluation procedures and operations. Such uses of evaluation processes can affect programmes as much as, or even more than, the use of evaluation findings disseminated in evaluation reports. (p. 4)

Applying Evaluation Theory - Variations on CDC Framework Implementation

Not all evaluators will carry out each step in the CDC Framework in the same manner. One factor that guides practitioners in making their decisions is the **evaluation theory** to which they subscribe. Evaluation theories are prescriptive approaches to evaluation practice (Alkin, 2013). They were primarily developed within the field of evaluation by practitioners with extensive experience and are intended to impart wisdom about what an evaluator ought to do in practice.

The steps comprising the CDC Framework are general in nature. After reviewing it, we may understand that we need to engage stakeholders, but should we involve them in all subsequent steps outlined in the Framework? How deeply should we engage them? Should stakeholders ever take the lead or a co-leadership role in designing and implementing data collection efforts? Evaluation theories help us understand the many options that are available to us. In this section, we provide a snapshot of some evaluation theories, demonstrating the types of prescriptions that they provide and summarizing them in **Table 2.3**. We extend this thinking in **Chapter 4**, where we highlight how evaluation theories can help in designing elements of an evaluation plan that correspond with each step of the CDC Framework.

Assess Context (Step 0)

What does it mean to assess the context and what are some options for doing so? Several evaluation theorists provide direction on what is important to examine about the context and ways of making these observations. For instance, as part of Values-Engaged Educative Evaluation, Greene et al. (2011) suggest it is important to understand the past and current context within which the evaluand resides. Specifically, Greene et al. (2011) highlight the importance of understanding characteristics such as the political dynamics, population demographics, cultures embedded in the setting, and physical and economic features of a setting (p. 24–25). To help you become familiar with these features of context, Greene et al. (2011) suggest many activities, including volunteering within the community over a short period of time, informally exploring the community by taking a walk around the area, and reading or viewing local media (p. 25).

Excited about evaluation theory and want to learn more?

- Alkin, M.C. (2013). *Evaluation roots (2nd ed.).* Thousand Oaks, CA: Sage.
- Mertens, D.M. & Wilson, A.T. (2019). *Program evaluation theory and practice (2nd ed.).* New York, NY: Guilford Press.
- Shadish, W. R. Jr., Cook, T. D., & Leviton, L. C. (1991). *Foundations of program evaluation: Theories of practice*. Newbury Park, CA: Sage.
- Hood, S., Hopson, R.K., & Frierson, H. (Eds.) (2015). *Continuing the Journey to Reposition Culture and Cultural Context in Evaluation Theory and Practice.* Charlotte, NC: Information Age Publishing, Inc.

Engage Stakeholders (Step 1)

What is the appropriate breadth and depth of stakeholder engagement in evaluations? Evaluation theories range in their prescriptions regarding how many stakeholders to engage and how deeply to engage them in each step of the process. Some theories prescribe engaging a wide range of stakeholders in the evaluation process (Fetterman, 2013, 2017); others suggest engaging a more limited set of stakeholders, often those who have the direct authority to make program modifications (Wholey, 2013)

King & Stevahn (2013) outline the range of stakeholder involvement that is possible in evaluation with their *Interpersonal Participation Quotient*, which classifies evaluation practice on a continuum. Points on the continuum include evaluator-directed evaluation, collaborative evaluation, and participant-directed evaluation (p. 27). Participant-directed evaluation has the highest level of stakeholder engagement. An example of this type of evaluation includes one in which stakeholders participate in (and may even lead) all aspects of the design and implementation. For instance, stakeholders may generate and prioritize the evaluation questions and could be at the helm when designing data collection instruments, collecting data, analyzing data, and making meaning of the findings from these analyses. Collaborative evaluation has equal involvement from the stakeholders and the evaluator; in this case the stakeholders might work hand-in-hand with the evaluator in identifying relevant evaluation questions, reviewing data collection instruments created by the evaluator for readability and cultural responsiveness, and interpreting the findings from the evaluation alongside the evaluator. Last, evaluator-directed evaluation includes little to no engagement from stakeholders. In this case, the evaluator may simply take guidance from a general scope of work and early conversations with the stakeholders who requested the evaluation to design and implement the evaluation, circling back to the stakeholders only at the end of the evaluation to share recommendations.

Describe the Program (Step 2)

Most theories discuss the use of models to describe the evaluand (Alkin & Christie, 2005). However, theory-driven evaluation emphasizes the central importance of understanding the theory underlying how a program is intended to operate (Chen & Rossi, 1992; Chen, 2015). Though there are several features to theory-driven evaluation, one feature involves developing action and change models (Chen, 2015). **Action models** describe the inner workings of a program (e.g., the desired characteristics or qualifications of program implementers) and the protocols used for program implementation. **Change models** depict how the effective implementation of a program translates into the intended programmatic effects, documenting **mediators**, and often **moderators**, of

the presumed mechanism of change. We describe each of these models in more detail in **Chapter 4** and provide examples to articulate their potential use in a public health setting in **Chapter 7**.

Focus the Evaluation (Step 3)

In **Chapters 3** and **4** we discuss the importance of identifying a small set of questions to focus the evaluation. Out of the many evaluation questions that exist, how do we know which ones should have priority in any given evaluation? Some evaluation scholars and theorists have provided suggestions for criteria to consider when prioritizing evaluation questions (Chen, 2015; Cronbach, 1982) such as how long a program has been in operation. More mature programs might focus on effects of the program because sufficient time has elapsed. Newer programs might focus on identifying areas to improve the process of implementing a program. Additionally, these theories provide insights about how to structure the process of engaging stakeholders to create, revise, and prioritize evaluation questions (Cronbach, 1982) and the types of questions that could be asked in evaluations to promote equity (Greene et al., 2011).

Gather Credible Evidence (Step 4)

All theories incorporate advice for how to gather credible evidence (Alkin, 2013). Some evaluation theories, such as Values-Engaged Educative Evaluation (Greene et al., 2011) and Transformative Evaluation (Mertens, 2009), place stronger emphasis on the use of mixed methods—the collection, analysis, and *integration* of qualitative and quantitative data—than other evaluation theories. Most evaluation theories leverage traditional social science methods in the collection and analysis of data. However, some evaluation theorists introduce other inquiry approaches into the process. For example, Scriven (2015) intricately describes the logic of evaluation, including identification of **criteria of merit**, standards, and synthesis. In contrast, Eisner (2013) leverages the act of connoisseurship and criticism from the arts.

Justify Conclusions (Step 5)

Prescriptions for identifying criteria and standards (i.e., benchmarks), that are reflective of values about what constitutes good and poor performance, and that synthesize findings to render judgements about an evaluand are sparse in evaluation theory. However, some evaluation theorists shine a light on how to develop criteria and standards. For some, this is a purely objective process guided by the discipline of logic (Scriven, 2013); for others, development is an act of articulating, deliberating, and representing what various stakeholders value about the evaluand (Greene, 2013; House, 2013). Nevertheless, evaluation theories help us to better understand what it means to make values explicit in the act of evaluation and some of the ways this can take shape when articulating criteria and associated standards.

Ensure Use and Share Lessons Learned (Step 6)

Sharing lessons learned from an evaluation can aid in fostering use, however as discussed earlier, to truly foster use evaluators must be thinking about it from the very beginning of the evaluation. Patton (2008) is probably the most well-known authority on fostering use and has developed multiple evaluation theories including Utilization-Focused Evaluation (U-FE), and several other theories that fall under the umbrella of U-FE (Developmental Evaluation, Patton, 2011; Principles-Focused Evaluation, Patton, 2018). An important component of sharing lessons learned is developing means for communicating about the progress and findings from evaluations. Russ-Eft and Preskill (2009) leverage their organizational development approach to evaluation and share numerous considerations in communication and reporting, including, communication plan formats and multiple modes of reporting. Though not recognized as formal evaluation theorists, recent advances in data visualization within evaluation stem from the work of Evergreen (2020) and Hutchinson (2017).

Table 2.3. Sample of Evaluation Theories by Evaluation Activity or Characteristic

Criteria	Sample of Evaluation Theories with Related Suggestions
Assessing the Context	✓ Values Engaged Educative Evaluation (Greene et al., 2011) ✓ Utilization-Focused Evaluation (Patton, 2008, 2013) ✓ Transformative Participatory Evaluation (Mertens & Wilson, 2019)
Engaging Stakeholders	✓ Interactive Evaluation Practice (King & Stevahn, 2013) ✓ Utilization-Focused Evaluation (Patton, 2008, 2013) ✓ Empowerment Evaluation (Fetterman, 2017)
Describe the Program	✓ Theory Driven Evaluation (Chen, 2015) ✓ Program Theory Driven Evaluation Science (Donaldson, 2007
Focus the Evaluation	✓ Theory Driven Evaluation (Chen, 2015) ✓ Values Engaged Educative Evaluation (Greene et al., 2011) ✓ Prescriptive insights from Cronbach (1982)
Gather Credible Evidence	✓ All evaluation theories ✓ Emphasis on mixed methods: Greene et al. (2011), Mertens (2009) ✓ Alternatives to social science methodology: Eisner (2013), Scriven (2015)
Justify Conclusions	✓ Logic of evaluation (Scriven, 2015) ✓ Deliberative Democratic Evaluation (House & Howe, 1999, 2000) ✓ Values Engaged Educative Evaluation (Greene et al., 2011)
Ensure Use and Share Lessons Learned	✓ Utilization-Focused Evaluation (Patton, 2008) ✓ Evaluation and organizational development (Russ-Eft & Preskill, 2009) ✓ Data visualization (Evergreen, 2020; Hutchinson, 2017
"How" of Practice	✓ Critical reflection: Transformative Participatory Evaluation (Brookfield, 2017; Mezirow, 2009; Mertens, 2009) ✓ Cultural responsiveness: Culturally responsive evaluation (Hood, Hopson, & Kirkhart, 2015; Hood, Hopson, & Frierson, 2015); Culturally Responsive Indigenous Evaluation (Waapalaneexkweew & Dodge-Francis, 2018) ✓ Situational awareness: Utilization-Focused Evaluation (Patton, 2008) and Developmental Evaluation (Patton, 2011) ✓ Interpersonal competence: Interactive Evaluation Practice (King & Stevahn, 2013)

Chapter Summary

In this chapter we introduced the general steps involved in planning and conducting an evaluation as represented in CDC's (1999) *Framework for Program Evaluation in Public Health*. We supplemented the CDC Framework with advances in the evaluation field. We also acknowledged the importance of the evaluand's context and the effect evaluators' actions and reactions can have on implementing the Framework steps and upholding the standards. Additionally, we described the variations that exist in evaluation and we leveraged insights from evaluation theory to portray some of this variation and point to specific theories that include prescriptions relevant to each step of the CDC Framework. In the next section of the text, we turn attention to the practical aspects of evaluation practice—designing evaluations, implementing evaluations, and supporting these processes through ongoing efforts to create, maintain, and grow evaluation capacity.

Review Questions and Skill Building Exercise

Review Questions

1. This chapter introduced you to the Program Evaluation Standards developed by the Joint Committee on Standards for Educational Evaluation (JCSEE) (i.e., utility, feasibility, propriety, accuracy, and evaluation accountability). Select two standards. How might these standards (a) complement and (b) conflict with one another in an evaluation? You do not need to use the same two standards for a and b.
2. The chapter presented four additional dimensions of evaluation practice: critical reflection, situational awareness, interpersonal competence, and cultural responsiveness. Out of these four dimensions, which would you like to improve upon and why? What steps might you take to further develop your abilities in this area over the next year?
3. What is evaluation theory and how does it help evaluators carry out the steps of the CDC Framework?

Skill Building Exercise

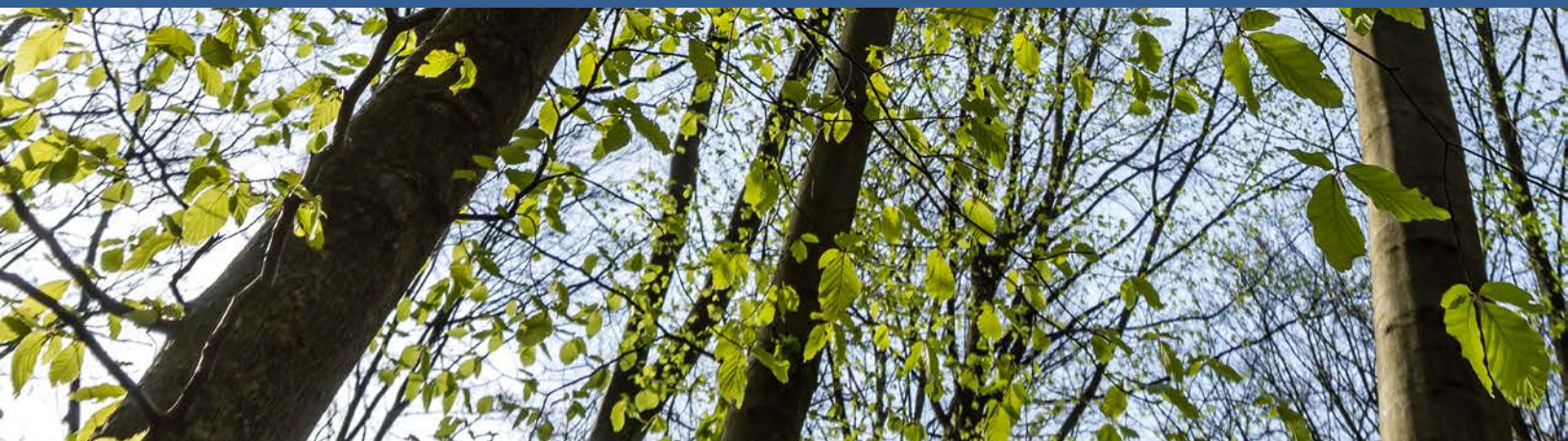
This chapter included a discussion about the importance of context. Context includes the nature and implementation of programs, how the programs are evaluated, and whom they are evaluated by. As such, gaining a sound understanding of context is a critical first step when preparing for an evaluation. This exercise will help you to develop skills to assess context.

Think of a program or intervention in which you are currently involved. If you are not currently involved with a program or intervention, think of one you have been involved with in the past or are familiar with. You have been asked to prepare a context description of the program or intervention as part of an evaluation. The description should include both current and historical information on the program's or intervention's context as well as the organization and community in which it is situated. The description must also include information pertaining to the evaluation context.

1. What contextual factors do you feel are most important to understand for the evaluation? Try to think of aspects that relate to the program or intervention, organization, community, and evaluation.
    a. For each contextual factor, write down reasons why that factor is important to understand for the evaluation.
    b. Re-examine the program or intervention's context, but this time, through the eyes of those (1) served and (2) underserved by it. What contextual factors would they most likely view as important and why?
    c. Compare your lists from step 'a' and 'b. Where are the differences and overlaps? How might the differences be reconciled?
    d. What activities would you engage in to develop an understanding of each of these contextual factors?
2. In what ways could or should each of the contextual factors influence the evaluation?
3. What constraints may exist and how could these limit your ability to gain an understanding of the context?

SECTION II

Planning and Implementing High-Quality Evaluation



SECTION II: Planning and Implementing High-Quality Evaluation

In Section I, we introduced you to the topic of evaluation by defining it, describing its history, and orienting you to the professional practice and its role in social service programming as it stands today. We connected evaluation to the discipline of public health by identifying several ways that evaluation exists or is valued within public health. First, we introduced you to the six steps of the CDC *Framework for Program Evaluation in Public Health* and considered some ways the CDC Framework can be enhanced given recent advancements in the evaluation field. We then introduced you to evaluation theories which are different approaches to evaluation that provide prescriptions about *how* evaluation should be performed. We also provided a brief overview of ways evaluation theories could affect the implementation of each of the six steps articulated in the CDC Framework.

Now that you have a good sense of what evaluation is, how it may apply to your current or future work in public health, and the general process involved in conducting evaluation, we will delve into more details about how to actually *do* evaluation. In this section, we focus on how to plan and implement evaluations drawing upon the steps in the CDC Framework. We start with evaluation planning because, when done in a thoughtful way, planning sets the stage for conducting high-quality evaluations. In **Chapter 3**, we describe *strategic evaluation planning*, a process through which a suite of evaluations is prioritized for implementation to meet the information needs of various stakeholder groups over an extended period of time. This type of planning can be thought of as occurring at a macro level within an organization or program. In **Chapter 4**, we zoom in on the process of developing a plan for a single evaluation (i.e., the micro-level) represented in the overall suite of evaluations—what we have chosen to call *individual evaluation planning.*

Of course, another aspect of evaluation practice that is important is implementing the well thought out plans. No matter how well you plan an evaluation, surprises will emerge as the plan is implemented. Though changes are inevitable during implementation, most issues should not fully derail an evaluation. When evaluators are aware of the common issues that can arise during implementation, they can more readily detect these issues and generate some basic contingency plans. Knowing what to look for and recognizing the importance of being flexible and adaptable during the implementation phase are equally as important as planning is to ensure a high-quality evaluation. In **Chapter 5**, we share some common challenges that evaluators encounter during the implementation phase and provide suggestions for managing such risks.

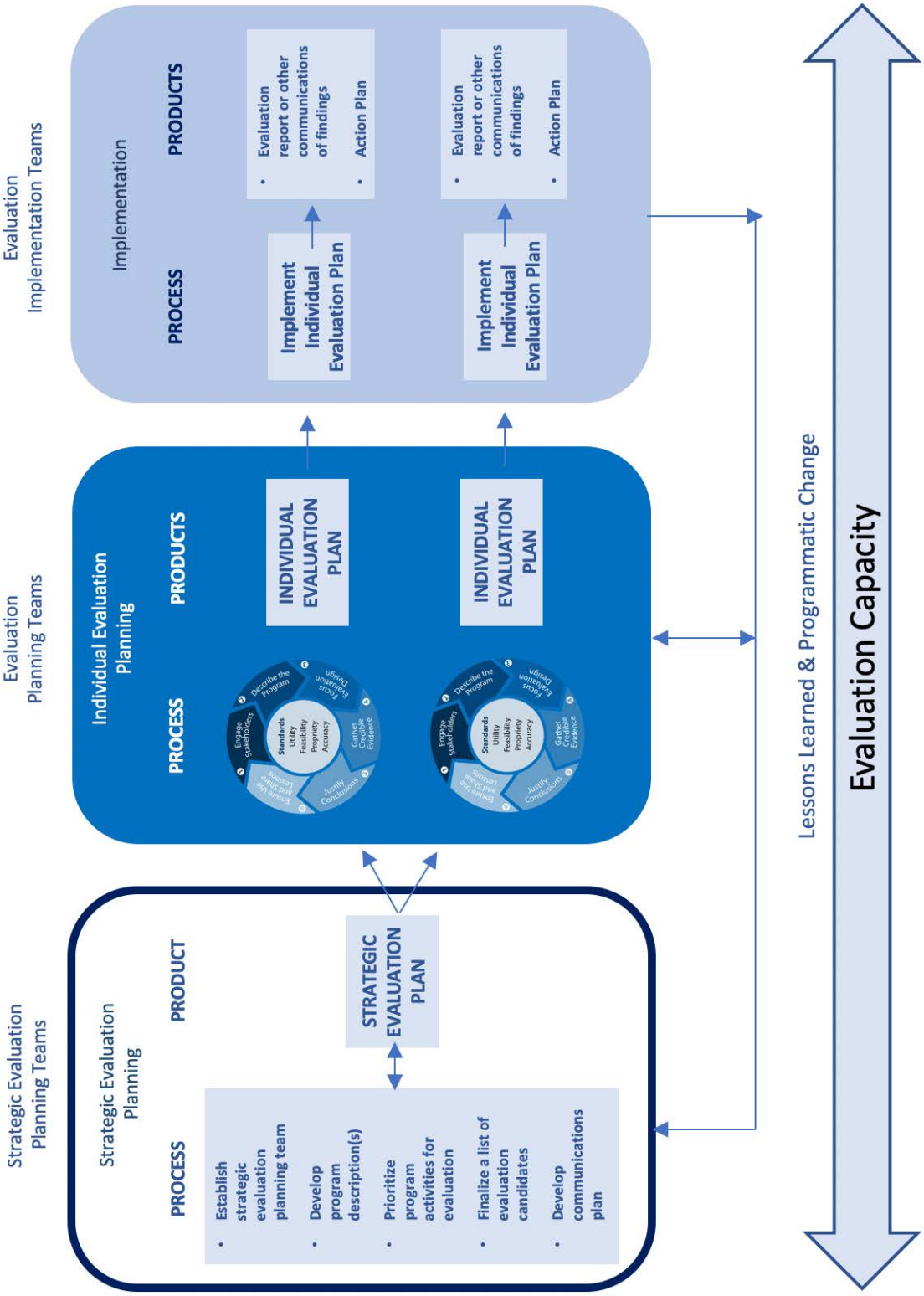
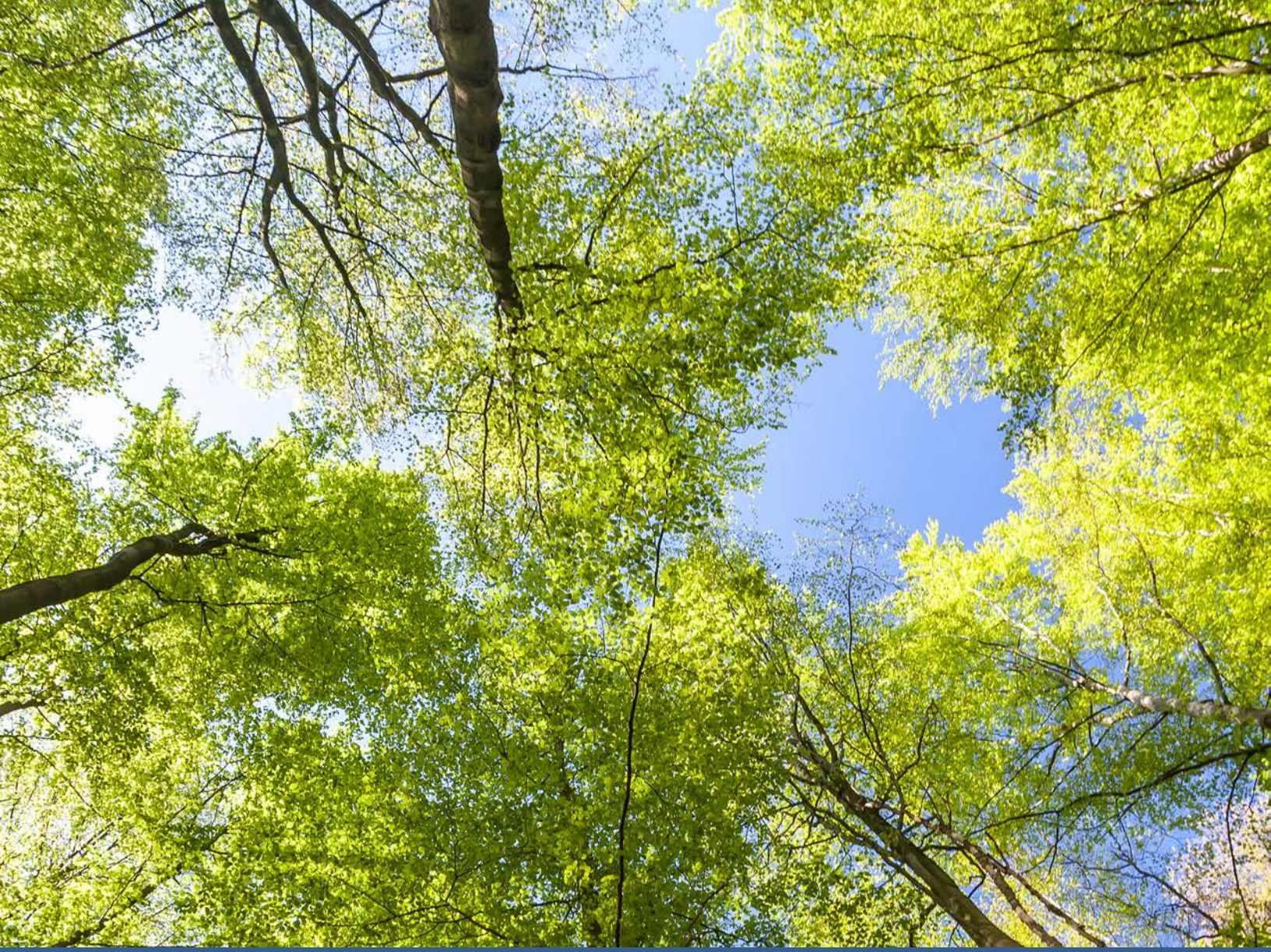


Figure 3.0 Connections Among Planning, Implementation, and Evaluation Capacity



CHAPTER THREE

Thinking Strategically: The Strategic Evaluation Plan

CHAPTER THREE: The Strategic Evaluation Plan

Evaluating all aspects of a public health program at once is simply not possible; the monetary costs alone would be prohibitive. As a result, public health practitioners must select which evaluations they will perform among several options. All too often, individuals identify these evaluation topics without using a systematic approach. This does not need to be the case. Most public health programming is structured and planned out for a foreseeable time frame. In this type of situation it is possible to identify in advance what evaluation needs exist in the near and long term, so that adequate time is available to thoughtfully plan each evaluation and ensure that the resources available for evaluation activities can be allocated well.

In this chapter, we introduce you to an approach for systematically prioritizing evaluations: strategic evaluation planning.[5] Using the CDC Framework as our general guide, we provide step-by-step instructions for engaging in a strategic evaluation planning process. At the end of the chapter, we share a detailed outline of the final product of this process, what we call a **strategic evaluation plan**. Finally, to give you a better sense of what the strategic evaluation planning process and final product looks like in practice, we present a fictional scenario that models the process and associated plan in **Appendices C and D**.

As mentioned in the preface, strategic evaluation planning is an advanced concept in evaluation and could be hard to digest, especially for novice evaluators. Readers who are new to evaluation may wish to just skim this chapter or read it after the other chapters.

By the end of Chapter Three, the reader will be able to

- Explain the purpose of strategic evaluation planning.
- Articulate how strategic evaluation plans and individual evaluation plans differ.
- Lead a strategic evaluation planning process.
- Develop a strategic evaluation plan.

Introduction

Experiences from CDC's National Asthma Control Program, which has conducted strategic evaluation planning since 2009, indicate that this planning process can be very helpful in allocating resources for evaluation and in facilitating more useful evaluation results than unstructured processes. Strategic evaluation planning increases the likelihood that the highest-priority evaluation needs are addressed over a specific time period (e.g., lifetime of a grant, contract, or cooperative agreement) and that evaluation activities responding to these needs are conducted in an appropriate sequence, on a reasonable timeline, and within existing budget constraints. Furthermore, such plans make it likely that several components of a public health program receive attention over time, while also permitting evaluation of emerging issues as they arise.

[5] Learning agendas are another way to prioritize evaluations and is a very similar process to strategic evaluation planning. Learning agendas are used to prioritize a set of questions for filling in knowledge gaps (OES, n.d.). They help an organization plan its activities within a particular budget or timeframe as well as promote organizational learning and improvement. One major difference between a learning agenda and a strategic evaluation plan is that the former uses many types of data and information to answer the learning questions—these sources include evaluations but also research studies, environmental scans, peer-to-peer discussions, and more (USAID, 2017). The questions presented in a learning agenda, therefore, are not necessarily evaluative in nature (i.e., understanding value and improving performance).

Findings from a recent research study by Doll (2020) on strategic evaluation initiatives, including strategic evaluation planning, suggested many benefits. Benefits include enhancing the alignment of work across an organization, improving stakeholder understanding of how a given effort relates to other efforts within a complex system, and improving resource allocation for evaluation. The study demonstrated that additional benefits appear to stem from simply participating in the strategic evaluation planning process (e.g., increased understanding and valuing of evaluation among stakeholders, a tendency among stakeholders to ask better questions about programming in the future, and better integration of evaluation into existing workstreams).

The Strategic Evaluation Planning Process

Given all the potential benefits of this process you may be asking what is entailed in developing a strategic evaluation plan? The overarching process is depicted in **Figure 3.1** and includes seven steps, each of which aligns with one or more steps of CDC's framework. In general, the process is participatory in nature and starts with identifying and inviting a core group of individuals to engage in the strategic evaluation planning process. This group forms the Strategic Evaluation Planning Team. The team then sets out on a journey to describe the overarching public health program, outline how they will determine what aspects of this program should be evaluated, prioritize what will be evaluated including the key evaluation questions to answer, explore and document potential methods for responding to the key evaluation questions, and compare the proposed priority evaluations against available resources and adjust the plan accordingly. Ultimately, the team documents the process and results of strategic evaluation planning in a written strategic evaluation plan. In the following sections, we describe each step of this process in detail.

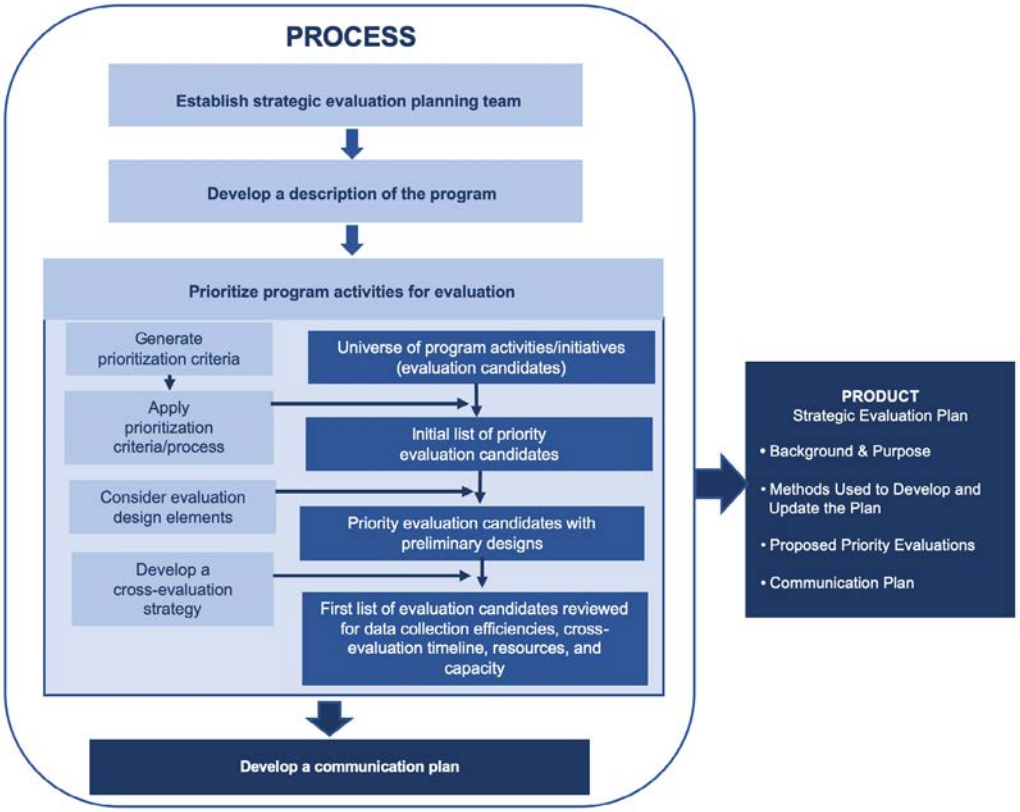


Figure 3.1 Strategic Evaluation Planning Process and Product

Establish a Strategic Evaluation Planning Team

In alignment with Step 1 of the CDC Framework—*Engage Stakeholders*—we suggest beginning the strategic evaluation planning process by forming a small **Strategic Evaluation Planning Team**, of about four to six individuals, to develop the plan. Ideally, the Strategic Evaluation Planning Team will serve as champions for evaluation on an ongoing basis. This team should also monitor progress in implementing the plan and be actively involved in annual reviews and updates.

A program evaluator should lead or co-lead this team. This evaluator may be external or internal to the program.[6] Deciding who the right evaluator is will depend on what the program decision makers desire in terms of the mix of technical skills, familiarity with the program or context, and personal characteristics. At a high level, it is important to take the following into consideration:

The "How" of Practice
Cultural Responsiveness

Addressing "aspects of culture in planning and managing evaluations" (AEA, n.d., p.3) is an essential characteristic of competent evaluation practice. Ideally, all members of the Strategic Evaluation Planning Team should understand the cultural context of the program and evaluation, or at least be committed to learning about and honoring it.

- Experience with program evaluation.
- Recognition of the importance of cultural responsiveness and their ability to describe how they will foster culturally responsive evaluation in this context
- Ability to communicate effectively
- Basic knowledge of similar public health programs
- Experience with the range of data collection strategies and evaluation designs that will best serve the program
- Good references from trusted sources

See **Appendix B** for additional insights should you have a need to hire an evaluator external to your organization.

Other members of the Strategic Evaluation Planning Team should include stakeholders knowledgeable about the program, its history, its goals and objectives, the role of evaluation in program improvement, and the resources available for evaluation. The team should also reflect the diversity of the community served by the program.

Numerous resources are available to assist programs in hiring and working with external evaluators. Here are a few you may find helpful:

- International Development Research Centre. (2004). *Selecting and managing an evaluation consultant or team.*
- Bruner Foundation. *Commissioning evaluation. Tips for grantmakers and grant seekers.*
- CDC's National Asthma Control Program. *Finding the Right People for your Program Evaluation Team: Evaluator and Planning Team Job Descriptions.*

It is likely that the program has many more stakeholders than those who comprise the Strategic Evaluation Planning Team. Therefore, consider how best to communicate with this larger group of stakeholders about the activities of the Strategic Evaluation Planning Team. Some of these stakeholders may become involved when specific plans are crafted for each evaluation that will be implemented (i.e., individual evaluation plans, the subject of **Chapter 4**). However, prior to that time, the Strategic Evaluation Planning Team may find it helpful to consult briefly with the broader group of stakeholders to better understand what type of information they would find helpful to have about the program's performance and how they might use these insights.

At the first meeting of the Strategic Evaluation Planning Team, it is important to establish some ground rules and expectations. The team should plan to discuss roles and responsibilities, a schedule for meetings, and a

[6] For the sake of brevity, we refer to *programs*, but a strategic evaluation plan may also be created for an entire organization or a portfolio of programs.

timeline to complete the team's activities. The team may also find it helpful to preview their intended end product. A preview can be accomplished by sharing the outline of the strategic evaluation plan provided at the end of this chapter, or **Appendix D**, with the team. Reviewing this outline can be part of the evaluation capacity building processes you undertake.

Describe the Program

In alignment with Step 2 of the CDC Framework—*Describe the Program*—the next step in creating a strategic evaluation plan is to develop a description of the program and its major components. To do so, we recommend that the evaluator engage in three preliminary activities:

1. Review program documents.
2. Share a detailed list of program activities with the Strategic Evaluation Planning Team.
3. Depict the connections between overarching program activities and intended outcomes.

Review Program Documents. The following documents often contain a wealth of information about planned activities and anticipated program outcomes: previous or current program plans, progress reports to funders, performance monitoring summaries or reports, surveillance products (e.g., reports, fact sheets, maps, web tables, briefs, newsletters), prior evaluation plans and the products resulting from these evaluations, and program funding applications and associated work plans (GAO, 2005). If the evaluator is new to the program, conducting a review of program documents is a good way to become familiar with the program.

Share a Detailed List of Program Activities. The evaluator's next step is to summarize what they have learned for the Strategic Evaluation Planning Team. Preparing a series of program activity profiles (see **Table 3.1** for an example) may be helpful prior to convening the first team meeting. Activity profiles document the important features about key programmatic activities that each member of the Strategic Evaluation Planning Team will need to understand in order to make decisions at a later stage about what to prioritize for evaluation in the upcoming years. Individuals on the Strategic Evaluation Planning Team have likely played a role in designing or implementing these activities and, therefore, will be able to help finalize the information in the profiles. The team can then reference these profiles as they discuss which program activities are most important to evaluate over the next several years.

As mentioned previously, the Strategic Evaluation Planning Team will want to consider how the broader group of stakeholders may be able to contribute to this process, especially those who were engaged in planning for the public health program (e.g., participated in strategic planning sessions for the program itself; helped develop a plan for the program; or assisted with applying for original funds that started, continued, or expanded the program). The team may find it helpful to share the activity profiles, or a list of the profiles, with a broader group and invite them to identify additional activities that should be profiled. Sharing will help

- Fill knowledge gaps regarding ongoing activities
- Foster a sense of ownership among partners for the strategic evaluation plan and evaluations to follow
- Familiarize partners with aspects of the program other than those they are directly working on

Conversations with stakeholders during this step also afford an opportunity to engage in Step 0 of the enhanced CDC Framework we presented in Chapter 2—*Assess the Context*. The Strategic Evaluation Planning Team and other stakeholders may find it helpful to discuss the context within which each of the program activities depicted in the profiles are implemented. These insights could be summarized and added to the template provided in **Table 3.1**, as they will likely be helpful later when the Strategic Evaluation Planning Team engages in discussions about who to invite to design and implement individual evaluation plans (more on this in **Chapter 4**) and what evaluation designs and methods may be most feasible.

Table 3.1 Program Activity Profile Template

Program Component	(Some programs can be thought of as having categories of activities, for instance those relating to the partnership or surveillance functions, or related to a specific intervention or suite of interventions.)
Title of Activity	(Name of activity that is generally recognized by key audiences)
Description of Activity	(Describe the activity)
Duration of Activity	(Start date and end date or ongoing)
Partner Involvement	(Describe whether partners of the program are involved in the activity and, if so, specify the major partners and their roles)
Cost of Activity	(Provide a rough or ballpark estimate of what the activity costs overall or annually, including funds from all sources; specify what portion, if any, comes from partner contributions)
Contribution to Intended Program Outcomes	(Describe what results or outcomes are anticipated based on conducting this activity)
Known Challenges in Conducting the Activity	(List any known challenges in conducting the activity, this may stem from prior evaluations, analyses of data regularly collected to monitor program performance, or be anecdotal in nature)
Prior Evaluation	(List any prior evaluations conducted of this activity and when they occurred)

Depict the Connections Among Overarching Program Activities and Intended Outcomes. Program descriptions, such as **logic models**, are an important starting point to generate a common understanding of how a program's activities are expected to work together to lead to one or more long-term programmatic results. The W.K. Kellogg Foundation Evaluation Handbook (1998), defines a program logic model as a

> picture of how your program works—the theory and assumptions underlying the program…[The logic model] provides a roadmap of your program, highlighting how it is expected to work, what activities need to come before others, and how desired outcomes are achieved. (p. 35)

Logic models show the expected sequence of activities and consequences that ultimately lead to public health results. A typical logic model depicts what goes into a program (inputs or resources), what the program does (activities), the immediate products of these activities (outputs), and the *changes* we anticipate will result when the activities are implemented well (a series of programmatic outcomes ranging from short- to long-term; DHHS, 2005). Often, the program processes (i.e., inputs or resources, activities, outputs) are depicted on the left-hand side of the model, while the anticipated changes or outcomes are on the right-hand side. It may be helpful to think of the left-hand side of the model as the "sphere of control" and the right-hand side as the "sphere of influence."[7]

Figure 3.2 presents a logic model for a state public health program focused on reducing the burden of asthma through improved asthma management practices. By detailing the pathways between program inputs or resources, activities, and a series of outcomes, **Figure 3.2** helps us see how different program components contribute to achieving the envisioned long-term program outcomes. This overhead view of the entire program is valuable because it allows members of the Strategic Evaluation Planning Team to see, at a glance, the many aspects of the program that could be topics for an evaluation.

[7] This use of "sphere of control" and "sphere of influence" were coined by CDC's former Chief Evaluation Officer, Thomas Chapel.

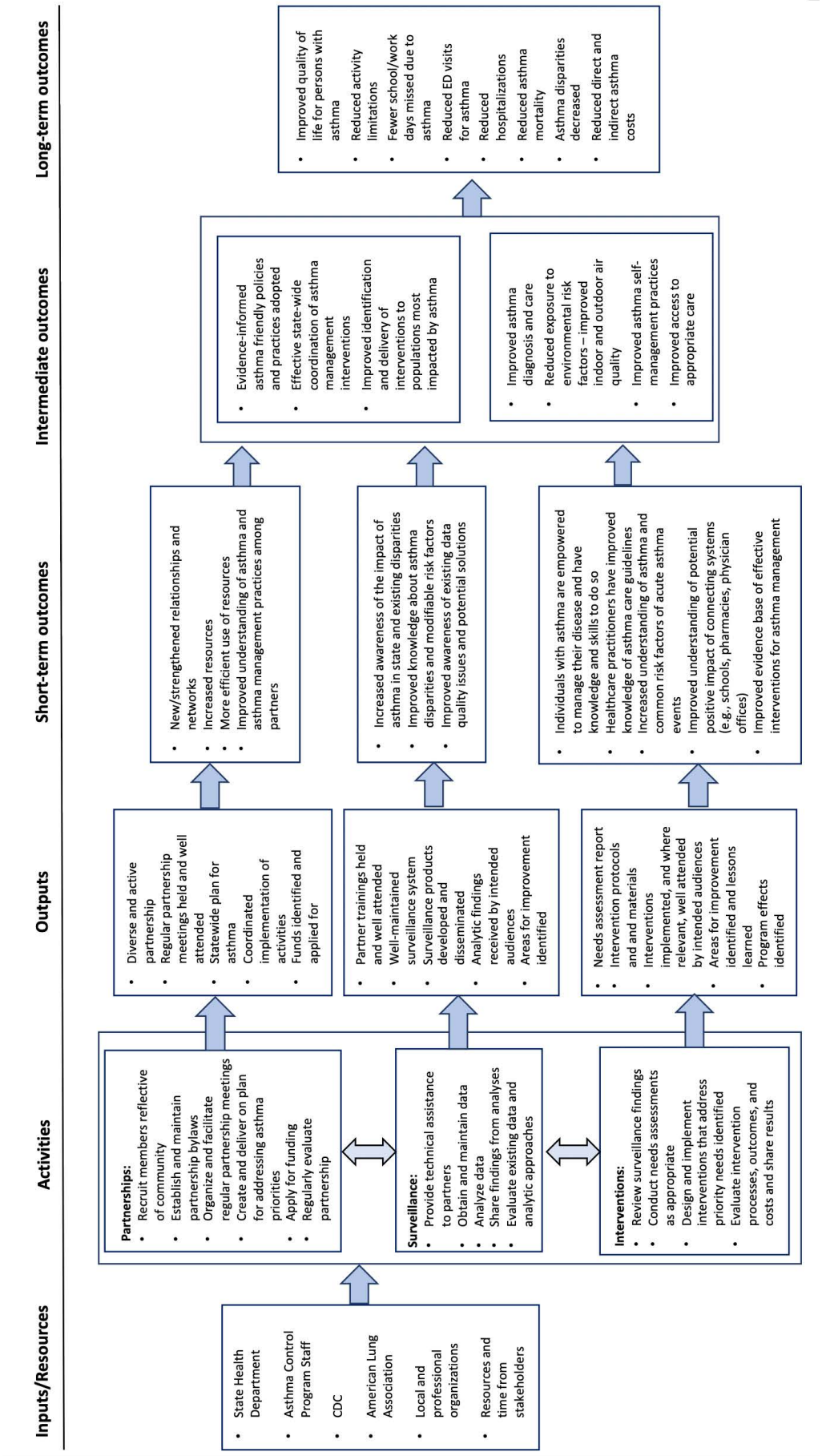


Figure 3.2 Example State Asthma Control Program Logic Model

The Strategic Evaluation Planning Team should consider creating a narrative description to accompany the overarching program logic model. Such a description provides an opportunity to reflect on and document important details about the context within which the program operates and may affect the program's successful implementation (e.g., the existing need for this program; culture and history; important disparities in health outcomes, access to care, and access to other important resources). The team can also articulate key assumptions that underly the successful implementation of what is depicted in the logic model (e.g., for the program in **Figure 3.2** this might include consistent or dependable allocation of program funds, partner interest and availability, adequate dose of interventions to affect change). Such assumptions may also form the basis for evaluations.

Prioritize Program Activities for Evaluation

Once the Strategic Evaluation Planning Team has described the program, they are ready to start thinking about what to evaluate (i.e., *Focus the Evaluation Design, Step 3* of the CDC Framework). As you may imagine when looking at **Figure 3.2**, resources will not be available to evaluate every activity that comprises a public health program. Therefore, it is important to engage in a systematic process to prioritize what will be evaluated in the upcoming years. It is also important to document the process used to establish these evaluation priorities so that other stakeholders can understand the process.

There are many methods for prioritizing what will be evaluated, including the Nominal Group Technique, the Simplex Method, or Criteria Weighting. These techniques vary in terms of how stakeholders are engaged and how criteria are applied. It is best to examine various prioritization techniques to determine which most suits your program. Regardless of the method selected, the Strategic Evaluation Planning Team will need to

- Develop clear prioritization criteria
- Apply the criteria to a list of potential **evaluation candidates**
- Generate a rank-ordered list of priority evaluation candidates

To get a well-rounded set of evaluations for the program, first consider the major buckets of activities that comprise the program (often noted the in first row of the Activity Profile; see **Table 3.1**). Next, the team should make sure one or more aspects of programming within each bucket is evaluated at some point during the timeframe covered by the strategic evaluation plan.

In **Table 3.2**, we list several possible criteria to consider using in the prioritization process. Both objective criteria (e.g., prior evaluation, cost) and subjective criteria (e.g., stakeholder interest, sustainability) are important to consider. This list is not intended to be comprehensive, nor does the order imply that one criterion is more important than another. The Strategic Evaluation Planning Team may also identify criteria not on this list. It is most important that the ultimate criteria selected resonate with the Strategic Evaluation Planning Team and are easy to apply consistently across evaluation candidates.

Table 3.2 Potential Criteria for Evaluation Prioritization

Criterion	Information Required for Prioritization
Cost	What financial resources have we invested in this activity?
Labor or time intensive	How much staff members' time have we invested in this activity?
Prior evaluation	Have we evaluated this activity before? When?
Performance	Does information from our performance measurement system indicate a need for in-depth examination of this activity?
Maturity	What is the stage of development or implementation for this activity?
Stakeholder interest	How interested are program stakeholders in this activity?
Sustainability	How much does this activity contribute to the sustainability of the program?
Centrality	How connected is this activity to our partners across the jurisdiction?
Plan alignment	How closely aligned is this activity with our jurisdiction's public health plan?
Plausible outcomes	Can this activity reasonably be expected to lead to relevant outcomes?
Disparities	Will this activity reduce health disparities?
Equity	Does this activity promote health equity?
Focus	Does this activity affect those most burdened by the health condition?
Reach	How many people in our jurisdiction are (or could be) affected by this activity?
Challenges	Are we (or do we anticipate) struggling with this activity?
Pilot	Do we plan to expand this activity?
Information need	How critical is the evaluation information for making near-term decisions?
Improvements	Would evaluating this activity likely result in recommendations for programmatic improvement?
Use	Is it likely that results or recommendations from this evaluation will be used by the intended audiences?

In the event the Strategic Evaluation Planning Team opts to use a qualitative rating system to identify priorities for evaluation, they might produce a table like that shown in **Table 3.3**. In this example, the Strategic Evaluation Planning Team is rating several activities that pertain to the state asthma program depicted in **Figure 3.2**. The team applied qualitative ratings (low, medium, high) to each activity based upon a subset of criteria in **Table 3.2**. After scoring the activities based upon the selected criteria, the Strategic Evaluation Planning Team can identify which activities rise to the top for evaluation (indicated by shaded rows in **Table 3.3**). The activities that rise to the top are their *priority evaluation candidates*.

In looking over the information contained in **Table 3.3**, you may find that it is not immediately obvious which activity rises to the top. Should an activity scored high-medium-medium be ranked higher than one scored low-high-high? If the Strategic Evaluation Planning Team had used quantitative ratings to do the scoring where high = 3, medium = 2, and low = 1, both activities would have a total of 7 points. In some cases, having a tie may be perfectly fine—there may be the option to move both forward as priority evaluation candidates. However, there are other ways to structure the prioritization process to provide finer details, and perhaps, reflect the values of the Strategic Evaluation Planning Team more accurately.

For instance, as the team develops the prioritization criteria, they may want to consider whether some criteria are more important than others or whether they want to establish a threshold for one

The "How" of Practice
Interpersonal Competence

A systematic process is an effective approach to prioritizing future evaluation activities. However, team members are likely to have different perspectives on what is and is not a priority, regardless of the process. When conflict arises, as it often does, evaluators need to rely on interpersonal competence to guide the team. Mediation, negotiation, and communication are critical skills in leading teams out of conflict and toward a resolution. Importantly, conflict is not always explicit, which emphasizes the importance of 'reading the room' to detect signs of disagreement or consternation.

or more criteria. Using **Table 3.3** as a referent, the team could decide that to be considered a priority candidate, an activity must score at least medium on the criterion Information Need. Given this, the activities "provision of technical assistance on surveillance to partners" and "identification of priority needs to address in interventions" would automatically be eliminated as priority evaluation candidates. Establishing some ground rules for scoring ahead of time with the Strategic Evaluation Planning Team will help the team more readily come to agreement as they rank the activities, and ultimately leave the team in a good position to document the decisions made.

Table 3.3 Activities Rank Ordered by Criteria

Activity	Criteria		
	Equity	Information Need	Challenges
Partnerships			
Partnership member recruitment	Medium	Medium	Low
Implement statewide asthma plan	Medium	High	High
Identification and acquisition of funds	Low	Medium	Low
Surveillance			
Provision of technical assistance on surveillance to partners	Medium	Low	Low
Maintenance of asthma surveillance	High	Medium	Medium
Dissemination of asthma surveillance products	Low	High	Low
Interventions*			
Identification of priority needs to address in interventions	High	Low	Low
School-based asthma self-management training	Medium	Medium	Low
Intervention to improve indoor air quality in multi-unit housing	High	High	High
Healthcare practitioner education series for improved diagnosis and treatment of asthma in metro clinics	High	High	Low

*You may notice that specific interventions are not mentioned in **Figure 3.2**. Programs may or may not be at the point in development where they can identify specific interventions that will be or are being implemented. Here to make the example clearer, we have selected some possible interventions that the program in **Figure 3.2** might choose to implement.

Once the Strategic Evaluation Planning Team has generated a priority list of evaluation candidates, it may be helpful to look back to the logic model developed in the previous step. In examining the list of priority candidates, the team should ask themselves "What types of activities are we including? What outcomes are represented by those activities? Which pathways are we considering?" Viewing the list of evaluation candidates through this lens can help the team focus on how well the list captures the bigger picture of the public health program.

In the next two steps, the Strategic Evaluation Planning Team will review and modify the list of evaluation candidates. First, they will consider the types of questions they might like to answer about each evaluation candidate, subsequently considering the evaluation designs that could be used to answer the questions and resource requirements for each potential evaluation to determine what is feasible (Issel, 2009; DHHS, 2003; GAO, 1991). Then, they will look across the list of candidate evaluations to ensure a strategy is in place that appropriately sequences the proposed evaluations. At the end of this process, the goal is to have an evaluation strategy that yields the most comprehensive and useful information possible while using evaluation resources wisely.

Consider Evaluation Design Elements

Now that the Strategic Evaluation Planning Team has a list of priority evaluation candidates, it is time to think about what the evaluation itself may look like. In alignment with Step 3—*Focus the Evaluation Design* and Step 4—*Gather Credible Evidence* of the CDC Framework, the Strategic Evaluation Planning Team should do

each of the following for every priority evaluation candidate:

1. Generate evaluation questions
2. Sketch out possible evaluation designs and data collection methods
3. Estimate the resource requirements and feasibility of conducting the evaluation
4. State how evaluation information can be used

Keep in mind that later in this process, individual evaluation plans will be created for each evaluation proposed in the strategic evaluation plan. As a result, there is no need to provide *detailed* information about study designs for the priority evaluation candidates at this stage. Rather, all that is needed is a broad strategy so that the Strategic Evaluation Planning Team can calculate ballpark estimates of the resources required and assess the general feasibility of conducting the evaluation. This information will help the team decide how many evaluations can be conducted in a year and if all the priority evaluation candidates can actually be performed during the period of time of interest. For the evaluations that can be implemented, these insights will help to inform when it is most appropriate to conduct the evaluation.

Crafting Good Evaluation Questions

CDC's good evaluation questions checklist provides additional information on developing evaluation questions. Consider using this resource while developing and refining your evaluation questions.

Generate Evaluation Questions. An important first step is to identify what the Strategic Evaluation Planning Team, and perhaps other stakeholders, feel is most important to know about each priority evaluation candidate. It can sometimes be very tempting to begin a conversation about design by discussing what data are readily available. You will identify data sources in Step 4, so try to stay away from that conversation now; considering data availability before addressing information needs runs the real risk of answering questions that stakeholders do not have a use for.

As stakeholders generate questions, consider the entire continuum of the logic model developed earlier. For example, the stakeholders may want to know whether the activity is conducted in the manner intended (a **process evaluation** question), or to what extent it is contributing to programmatic outcomes (an **outcome evaluation** question). The following are some examples of evaluation questions that stakeholders might consider.

- ***Process.*** In what ways was the activity implemented as intended? How did implementation differ from the original plan? What were the barriers or facilitators to implementation? How can implementation of the activity be improved? To what extent are there adequate resources (e.g., financial, personnel, expertise, partner relations) in place to implement the activity? What is the quality of the product?

- ***Outcome.*** To what extent did this activity lead to successfully achieving the stated program goals? What types of participant outcomes have been achieved? Who benefited the most? The least? What types of long-term outcomes can be attributed to this activity? What unintended outcomes (positive or negative) occurred? What did the activity cost in relation to the benefit observed?

At this stage, we recommend aiming for no more than five questions per evaluation candidate. If you have difficulty narrowing down the list of potential questions, consider the following questions:

- How would a sound answer to this question help the program?
- How important is this question to program staff members and stakeholders?
- How likely is it that answering this question would lead to improvement?

It may be helpful to organize the questions into an evaluation question worksheet. In **Table 3.4**, we provide an example of what a completed evaluation question worksheet would look like for one priority evaluation candidate, a healthcare practitioner education series for improved diagnosis and treatment of asthma in metro clinics.

The "How" of Practice
Situational Awareness

When considering various evaluation questions, evaluators should remain sensitive to the political context of the evaluation. If it is likely that the answer to a question will not be acted on for political reasons, it is probably best to exclude it from immediate evaluation activities and focus on questions that will be acted on.

Table 3.4 Example Evaluation Question Worksheet

Evaluation Candidate	Question Type	Questions	Question Priority (Low, Medium, High)
Interventions			
Healthcare practitioner education series for improved diagnosis and treatment of asthma in metro clinics	Process	How frequently do participants complete the entire training sequence? Why do those who complete remain? Who do those who do not complete leave early?	Low
	Process	To what extent are we reaching and effectively enrolling intended participants?	Medium
	Outcome	To what extent has this intervention led to improvements in knowledge regarding evidence-based treatment guidelines?	High
	Outcome	To what extent are improvements in the use of evidence-based guidelines among practitioners translating into health improvements among their patients?	High

Narrowing the scope of an evaluation may be challenging, but this is a very important task that should not be omitted. Public health programs often have limited resources to support evaluation work; tackling scope issues as a group early on will help to focus these evaluation resources on the items that have the most utility and importance at the given time.

Define Evaluation Designs, Data Collection Methods, and Timeline. The next step in developing an evaluation strategy is to sketch out possible methods that the team conducting the evaluation can use to answer the potential evaluation questions. Remember, this is rough, preliminary planning at this stage to help the Strategic Evaluation Planning Team develop an overall strategy. Once the team finalizes the evaluation strategy, additional work will need to be performed to develop much more precise and detailed designs for each individual evaluation. At this stage, the Strategic Evaluation Planning Team may find it helpful to consider the following:

1. **Evaluation designs.** Many evaluation designs are possible for an evaluation. Sometimes it is appropriate to use more than one. Examples of evaluation designs include **experimental designs** (e.g., randomized controlled trials), quasi-experimental designs (e.g., **pre-post-test design** with a comparison group, interrupted **time series**, **regression discontinuity**), and non-experimental designs (e.g., **case study**, **post-only design**) (Campbell & Stanley, 1966; DHHS, 2003; DHHS, 2005; EPA, 2007; Johnson &

Christensen, 2008; Trochim, 2020; Yin, 2018; Salabarría-Peña et al., 2007). Which designs are best suited to answering the evaluation questions given the context within which the evaluand resides? (See **Appendix E** for additional details on evaluation designs.)

2. **Data collection methods.** In addition to an overarching evaluation design, it is important to consider what data are needed to respond to the evaluation questions. In addition to using existing data, there are numerous data collection strategies to consider, including document reviews, surveys, interviews, observations, and focus groups. As the Strategic Evaluation Planning Team members discuss the many possible evaluation designs and data collection methods to use, keep in mind what the intended users of the evaluation will view as credible evidence. For example, some audiences may view quantitative data as more accurate and valid than qualitative data, others may place greater weight on stories that come from intensive and focused case studies employing qualitative data collection (DHHS, 2005; Salabarría-Peña et al., 2007). Mixed-methods designs that combine quantitative and qualitative data collection methods are also an option (Creswell & Clark, 2018). (See Chapter 4 for additional details about data collection methods.)
3. **Timelines.** An additional consideration is when data collection should occur. The optimal time to collect data will be driven by several factors:
    - Information need. Are there any programmatic decisions pending, for the program or its partners, that the evaluation could help to inform?
    - Design. Some designs require baseline data and follow-up data. In these cases, the data collection schedule will be determined in large part by the program's delivery schedule (Chappelle, 2014).
    - Maturity. If the Strategic Evaluation Planning Team suggests that evaluation questions pertaining to outcomes are a priority, consider when outcomes are most likely to occur relative to the activities that give rise to them.
    - Logistical constraints. Not all data are easy to come by. Therefore, for the team needs to consider whether there are times when it will be easier or more challenging to gain access to or collect data.

Consider Resource Requirements and Feasibility of Data Collection. After the Strategic Evaluation Planning Team has identified potential evaluation designs and data collection methods, it will be important to step back and consider the resource requirements and feasibility of implementing these proposed methods. Specifically, consider the following:

- **Resource requirements.** What are the resource requirements (personnel and funding) for the methods proposed? Detailed budget data are not needed at this stage, but categorizing the methods as requiring a low, medium, or high level of resources will be helpful.
- **Feasibility.** How feasible are the proposed methods? Is it likely that support will be available to ensure a high-quality evaluation that meets the standards of utility, feasibility, propriety, accuracy, and evaluation accountability?
- **Available expertise.** What level of expertise exists within the program or among the partners to carry out the proposed methods? Will assistance be needed from an external evaluator? Will existing staff members need to garner new skills or knowledge through professional development activities? Can either of these be supported financially (i.e., contributes to resource requirements)?
- **Instrumentation.** If new data collection is proposed, will the evaluator need to develop data collection instruments or are there existing instruments that can be used? If instrumentation development is necessary, what resources will be needed, including time to pilot and refine the instruments?

- **Information technology.** Is the existing technological infrastructure sufficient to carry out the evaluation? Will the purchase of data collection software or services be required to successfully complete the evaluation?

The Strategic Evaluation Planning Team may find it helpful to complete **Table 3.5** to summarize the proposed methods for each prioritized evaluation. Doing so may be helpful for organizing discussions around possible designs, methods, timelines, and resources. More specifically, it will be helpful to review this table when developing a cross-evaluation strategy during the next step of the process.

Table 3.5 Example Evaluation Design and Data Collection Summary (partially completed)

Question	Possible Evaluation Design(s)	Potential Data Collection Methods	Possible Data Sources	Data Collection Begins	Final Results Due	Resources Required
To what extent are we reaching and effectively enrolling intended participants?	Non-experimental	Secondary data analysis	Logs of invites and attendance	Year 1	Year 2	Minimal
To what extent has this intervention led to improvements in knowledge regarding evidence-based treatment guidelines?	Randomized design	Knowledge assessment tests	Healthcare practitioners who do or do not attend Phase I trainings	Year 2	Year 4	Modest

Develop a Cross-Evaluation Strategy

At this point, the Strategic Evaluation Planning Team has prioritized evaluation candidates; identified potential evaluation questions, designs, and data collection methods for each candidate; and discussed potential resource needs and feasibility considerations. Now, it is time to package all the information into a coherent evaluation strategy and document it in the strategic evaluation plan. To do so, we leverage ideas and concepts from Step 3—F*ocus the Evaluation Design*, Step 4—*Gather Credible Evidence*, and Step 5—*Justify Conclusions* of the CDC Framework.

To create this coherent strategy, the Strategic Evaluation Planning Team should look across the proposed evaluations and examine them for

1. **A good mix of evaluation activities and questions.** This is an excellent time to double check that the mix of evaluations proposed is a good representation of the important elements of the program. Will the proposed evaluations provide the information that primary stakeholders need to improve the program or identify successes?
2. **Opportunities for data collection efficiencies.** Look across the proposed evaluations to identify areas where the proposed methods for priority evaluation candidates can be integrated. Is it possible to modify activities to collect data that support more than one evaluation question?
3. **Timing across evaluations.** The Strategic Evaluation Planning Team has already considered the optimal timing of data collection activities for each priority evaluation candidate. Now it is important to revisit the timeline in light of the proposed evaluations. At this stage, it is helpful to develop a timeline indicating the duration of each proposed evaluation along with key milestones for each. When all the proposed evaluations are placed together on one timeline, it will be easier for the Strategic Evaluation Planning Team to assess the feasibility of the suite of evaluations proposed.

Table 3.6 summarizes additional considerations that the Strategic Evaluation Planning Team may find helpful to improve the coherence and efficiencies in the plan.

Table 3.6 Issues to Consider When Looking Across Proposed Evaluation Strategies

Area	Overarching Question	
Evaluation Design	What evaluation designs are proposed?	• Will a proposed evaluation design be suitable for answering multiple evaluation questions? • What, if any, unintended consequences may result from implementing the proposed evaluation design?
Data Collection: Respondent Population	From whom is data being collected?	• If several data collection strategies have the same respondent population, can you collect information for more than one purpose using a single data collection tool? • Are data collection activities concentrated too heavily on one respondent population? • Can the burden be shared more equitably? • What are respondents' previous experience with evaluation? How will these experiences shape engagement with respondents? • To what extent do these data collection methods align with the values and interests of respondents? Will data collection methods allow for authentic input from respondents?
Data Collection: Timeline	When are data being collected?	• How can evaluation data collection needs be integrated into the program timeline? For example, if baseline data need to be collected, program activities may need to be delayed. • If data about different evaluation activities need to be collected at the same time, do you have the resources to conduct multiple evaluation activities simultaneously? • What contextual factors need to be accounted for when considering the timing of data collection (e.g., school breaks, holidays, busy periods for respondents)?
Data Collection: Source	From where are data being collected?	• How can evaluation data collection needs be integrated into the program timeline? For example, if baseline data need to be collected, program activities may need to be delayed. • If data about different evaluation activities need to be collected at the same time, do you have the resources to conduct multiple evaluation activities simultaneously? • What contextual factors need to be accounted for when considering the timing of data collection (e.g., school breaks, holidays, busy periods for respondents)?

Who	Who will conduct the evaluation activity?	• Do you have the personnel and resources to conduct the prioritized evaluations given the strategy proposed? • Do they have the necessary skills and expertise? If not, how could they obtain these skills? • Can you leverage additional evaluation assistance from partners?
Analysis	How will the data be analyzed?	• Who will do the analysis? • Do they have the necessary skills and expertise? If not, how could they obtain these skills? • Can you leverage additional analytic capability from partners? • How will the results of the analysis be validated?
Use	How will the information from the evaluation likely be used?	• Will the information be provided in time to inform decisions? • Who will use the information provided? • In what ways, if any, will these findings benefit communities? • Are there capacity-building activities that need to be conducted with intended users to increase the likelihood that results will be used? • What is the potential for the misuse of findings and how will this be mitigated?

Promote Use through Communication

In alignment with Step 6 of the CDC Framework, it is important to consider how to Ensure Use and *Share Lessons Learned* with respect to the strategic evaluation plan. Communicating with key audiences about the progress on the activities articulated in the strategic evaluation plan, or about upcoming activities, for which their involvement may be needed is important. Communication also entails some additional resource expenditure (e.g., personnel time), so accounting for communication as part of the strategic evaluation plan will increase the likelihood that the communication will take place.

The overarching communication strategy should focus on sharing high-level information about the strategic evaluation plan itself. For instance, progress on developing, modifying, and implementing the plan. In addition, the plan should include disseminating a summary of the findings across all of the evaluations conducted over the time period covered in the plan (e.g., five years). Developing a communication plan like **Table 3.7** and including it as part of the written strategic evaluation plan, can help to keep communication organized and ensure accountability. It is possible that several audiences will be interested in receiving updates about the progress made on implementing the strategic evaluation plan and, later, the lessons learned from conducting all the evaluations. These audiences include funding partners, the Strategic Evaluation Planning Team, sister programs within the organization, and organizational leadership. For each activity and product, consider who the audience might be, and think outside the box. Who has a need to know? What level of information do they need? Who might be able to act on the results? For each audience, consider the best format for sharing information. Will they respond best to detailed results or high-level overviews? Is a written, oral, or visual format better?

Table 3.7 Example Communication Plan * (partially completed)

Audience 1 (e.g., Strategic Evaluation Planning Team)			
Purpose	**Possible Formats**	**Timing**	**Notes**
Inform about specific upcoming evaluation planning activities	Email	Bi-weekly	
Keep informed about progress of developing the strategic evaluation plan	Email	Monthly	For those unable to attend meetings
Present final strategic evaluation plan	PowerPoint Presentation	End-of-year meeting	Consider receiving general formative feedback on process
Communicate need to update strategic evaluation plan	Email	As need arises	
Provide general update on status of evaluations as proposed in the strategic evaluation plan	Email	Semi-annually	
Document and share synthesis of findings and lessons learned during lifecycle of program	Final report Formal presentation Working sessions	End of strategic evaluation plan implementation period	Use working sessions to generate ideas for specific use of findings in future program plans
Acknowledge contributions	Formal thank you letter	End of strategic evaluation plan implementation period	From program leadership
Audience 2 (e.g., Program Staff Members)			
Purpose	**Possible Formats**	**Timing**	**Notes**
Inform about specific upcoming evaluation planning activities	Email	Monthly	

*Adapted from Russ-Eft & Preskill, 2009, p. 407-411

Write and Revise the Strategic Evaluation Plan

The strategic evaluation plan should be considered a living document. As you may have noticed, considerable guesswork and uncertainty are involved in creating its content. There will naturally be things that the Strategic Evaluation Planning Team did not consider or have on their radar during the planning phase that will require adjustments. Furthermore, evaluation is a dynamic process. New information and unanticipated events are normal. The plans in place need to be flexible enough to adjust in response. Committing to an annual review of the strategic evaluation plan, with the Strategic Evaluation Planning Team, will ensure there is a built-in process to revisit the assumptions made, and adjust in relation to the context. It will be important to document the changes made to the strategic evaluation plan as part of these revision processes.

Chapter Summary

In this chapter we introduced you to the concept of a strategic evaluation plan, the process for developing it, and some of the potential benefits of engaging in the process. We drew upon the CDC Framework to structure this macro-level planning process and shared insights for

- Identifying members of the Strategic Evaluation Planning Team
- Identifying, designing, and prioritizing evaluations for implementation
- Documenting a strategy for communicating with various stakeholders

In the next chapter, we turn our attention to the process of engaging stakeholders in crafting a detailed plan for each priority evaluation articulated in the strategic evaluation plan.

Review Questions and Skill Building Exercise

Review questions

1. What are some benefits of strategic evaluation planning? How does a strategic evaluation plan differ from an evaluation plan?
2. What are some qualities to look for when seeking members of a Strategic Evaluation Planning Team?
3. What types of efficiencies might a Strategic Evaluation Planning Team identify when developing a cross-evaluation strategy?

Skill-building Exercise

Consider a public health program or organization in which you work or with which you are familiar. Practice some of the following procedures for developing a strategic evaluation plan with this program or organization in mind:

1. Who would you invite to participate as a member of the Strategic Evaluation Planning Team and why?
2. What are the key activities of this program or organization? List each, and then develop a Program Activity Profile (**Table 3.1**) for at least one
3. What prioritization criteria might you use to identify candidates for inclusion in the strategic evaluation plan? Why these instead of others?
4. Draft a plan for evaluating one of the activities by completing the details of an evaluation profile as seen in **Table SEP.5** (in the Strategic Evaluation Plan Outline). What are some possible strengths of this proposed evaluation? What are some limitations?
5. Try developing a communication plan for the strategic evaluation plan. Complete just a few rows in **Table SEP.6** (in the Strategic Evaluation Plan Outline) to get a sense of who you might communicate with in the context you are thinking about, how you might communicate with them, and why

Strategic Evaluation Plan Outline

1. Program Background and Purpose of Strategic Evaluation Plan

This section provides background information on the public health program and explains how a strategic approach to evaluation, as documented in this plan, will assist the program in meeting its aims.

Program Background

- Provide an overview of the program and the primary goals for a specific timeframe (e.g., the program funding cycle).
- Provide an overarching logic model for the program with narrative text describing it.

Purpose of Plan

- What is the role of evaluation in achieving the program's purpose?
- How will evaluation help tell the program's story?
- What are the expectations for how program staff members and stakeholders will use this plan?

2. Methods for Developing and Updating the Strategic Evaluation Plan

This section provides information about the methods the Strategic Evaluation Planning Team used to develop the plan, who was involved, how decisions were made, and how the plan will be kept up to date.

Stakeholders

- Who is the program's evaluation lead?
- Who are the stakeholders involved in developing the strategic evaluation plan?
- What role did they play in developing the strategic evaluation plan?
- What role will these stakeholders play in implementing the evaluations?
- How will participation from stakeholders, especially from individuals whose perspectives are often excluded from similar planning processes, be supported?

Table SEP.1 Strategic Evaluation Planning Team – Contributions, Roles, and Future Involvement

Stakeholder Name	Title and Affiliation	Contribution to Evaluation Planning	Role in Implementing Evaluations	Considerations to Support Participation

Methods Used to Develop the Strategic Evaluation Plan

- What process did the Strategic Evaluation Planning Team use to identify candidates for evaluation? (Narrative description)
- How are diverse stakeholder perspectives represented in the criteria?
- How were the criteria applied to establish priority evaluation candidates?
- What information sources did the team use to support assessment of criteria?

Table SEP.2 Prioritization Criteria

Criteria Used	How Criteria Were Applied	Information Supporting Criteria Determination
E.g., Cost	Higher cost activities supported by existing funds were rated as higher priority for evaluation.	Program budgets
E.g., Performance	Activities (and associated outcomes) where questions for further investigation were raised from trends or patterns in associated performance metrics were rated as higher priority for evaluation.	Performance metrics
E.g., Equity	Activities with potential to diminish structural supports for inequities were given a higher priority.	Situational analysis based on stakeholder discussions.

Proposed Methods for Reviewing and Updating the Strategic Evaluation Plan

- How will the team reflect on or assess its work?
- How often will the team review and update the strategic evaluation plan?
- What process will the team use to review and update the strategic evaluation plan?
- Who will be involved in strategic evaluation plan review and updates?
- How will the team document revisions to the strategic evaluation plan?

3. Prioritized Evaluations

This section lists the evaluations that were prioritized for implementation. It provides information on each of the prioritized evaluations along with a comprehensive evaluation timeline and details of evaluation capacity building activities. As an evaluation team implements this plan, they will likely need to revise it so that it reflects lessons learned while conducting the evaluations and the program's evolving information needs and evaluation resources.

Prioritized Evaluations

- Provide a rank-ordered list of prioritized evaluations

Table SEP.3 Rank-ordered List of Prioritized Evaluations by Programmatic Area

Program Area 1 (e.g., partnerships)	Program Area 2 (e.g., surveillance)	Program Area 3 (e.g., interventions)
Prioritized evaluation 1	Prioritized evaluation 1	Prioritized evaluation 1
Prioritized evaluation 2	Prioritized evaluation 2	Prioritized evaluation 2

Overarching Timeline

- Provide a timeline for conducting evaluations over the relevant timeframe (e.g., five-year grant cycle). Include program milestones for which evaluation findings should be available (e.g., legislative session, funding proposal, meeting with health plan association). Also note any capacity building activities or resources that will be required to implement the evaluations. (See Chapter 4 for additional information about implementation.) Finally, consider where stakeholder participation or data from one evaluation can be used in another evaluation and account for potential participant burden.

Table SEP.4 Sample Timeline with Sequencing of Proposed Evaluation Activities

	Year 1	Year 2	Year 3	Year 4	Year 5
Program Milestones					
Evaluations					
Capacity Building Activities					

4. Summarize Each Prioritized Evaluation

*This section provides the rationale for including each prioritized evaluation in the strategic evaluation plan. Include enough context and detail so that the selection can be adequately evaluated with each plan revision, but not so much detail that the Strategic Evaluation Planning Team becomes too invested in an evaluation that may no longer be appropriate in several years. This prioritization process will also be useful as the Evaluation Planning Team (individuals who craft the detailed plans for each specific evaluation, the topic of **Chapter 4**) begins its work to refine or revise the evaluation questions.*

In addition to the narrative, you may choose to complete **Table SEP.5** to present an abbreviated version of the information. This table may be useful in looking across all the proposed evaluations.

- What is the purpose of the evaluation and what evaluation questions would it address?
- Why is it a priority?
- Who are the potential audiences for the evaluation?
- How would stakeholders use the information produced by the evaluation?
- What evaluation design would be appropriate for answering these questions?
- What data collection method(s) and data sources would be appropriate?
- What contextual factors should be considered in the evaluation's design and implementation?
- When would the evaluation be conducted?
- What would the evaluation cost, roughly?

Table SEP.5 Evaluation Profile (create one for each prioritized evaluation)

Activity Name	Identify the prioritized activity for evaluation.
Program Component	List the programmatic area to which the activity belongs, e.g., partnerships, surveillance, interventions, other.
Evaluation Justification	Note relevant factors the Strategic Evaluation Planning Team considered in prioritizing this activity for evaluation.
Evaluation Purpose and Use	Identify the evaluation's purpose and potential uses of its findings, including decisions the findings should inform.
Possible Evaluation Questions	List the potential evaluation questions to be addressed
Timing of Evaluation	List the proposed or anticipated start and end dates and any related milestones
Suggested Evaluation Design	Describe potential evaluation design(s) for answering the evaluation questions.
Potential Data Sources	List data sources that could be used, noting any barriers to obtaining them or concerns about their quality.
Potential Data Collection Methods	Describe potential data collection methods, frequency of collection, and the staff who would be responsible for collecting the data.
Cultural or Contextual Factors	Describe how contextual factors and culture will influence the design and implementation strategies.
Potential Audiences	Describe potential audiences for the evaluation findings.
Possible Uses of Information	Describe how the anticipated information could be used.
Estimated Evaluation Cost	Provide a rough estimate of evaluation costs overall or annually, including funds from all sources; specify what portion, if any, comes from partner contributions.

Capacity Building Activities to Support Evaluation

- What additional evaluation capacity will be required to successfully complete the proposed evaluations—including commissioning the evaluations, planning and implementing the evaluations, and making use of the findings? (See the evaluation timeline in **Table SEP.4**.)
- How will you obtain or build that capacity? For example, what sorts of training, conferences, technical assistance, group facilitation, or involvement in evaluation might be needed?
- Who are the audiences for this capacity-building support (i.e., leadership, management, partners, staff members, human resources)?
- When will the capacity-building activities occur?

5. Communication Plan

This section provides guidance about how information about the strategic evaluation planning process and results will be shared.

Communicating

- What information about the strategic evaluation planning process will the team share? For what purposes?
- How often will the team share information about planning and implementation?
- With whom will the team share the information?

- What formats or methods (e.g., in-person meetings, emails, newsletters) will the team use to share information?
- Who is responsible for information sharing?
- How will the team summarize and share the results of the overall process?

Table SEP.6 Communication Plan Summary Matrix

Information and Purpose	Audience(s)	Possible Formats	Possible Messengers	Timing	Person Responsible

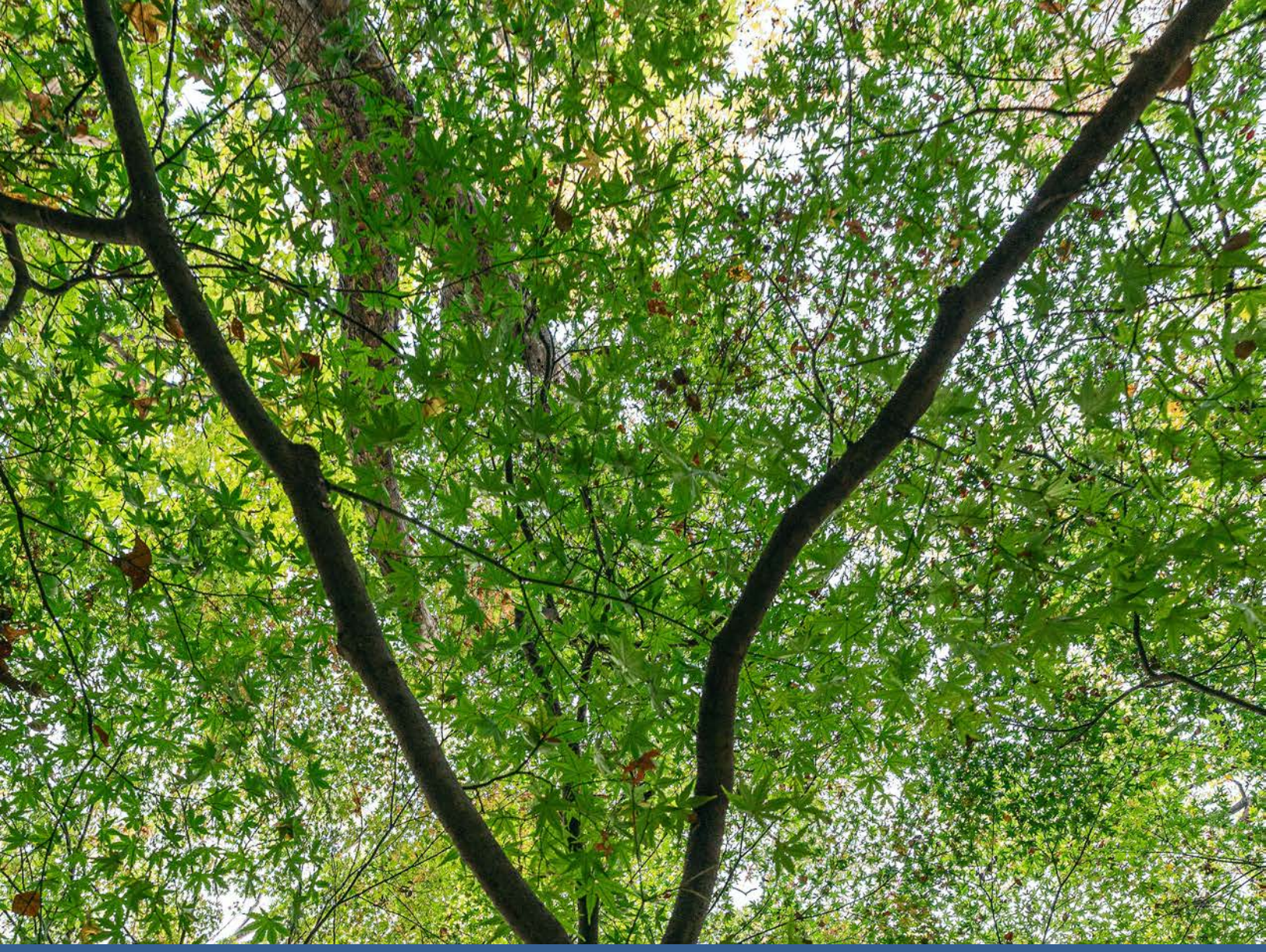
6. Wrapping Up

This section provides guidance on closing out the evaluation activities at the end of the period covered by the strategic evaluation plan.

- At the end of the timeframe covered in the strategic evaluation plan, how will you acknowledge the contributions of Strategic Evaluation Planning Team members and others who contributed to the successful implementation of the plan?
- How will you document evaluation lessons learned in the course of implementing the strategic evaluation plan?

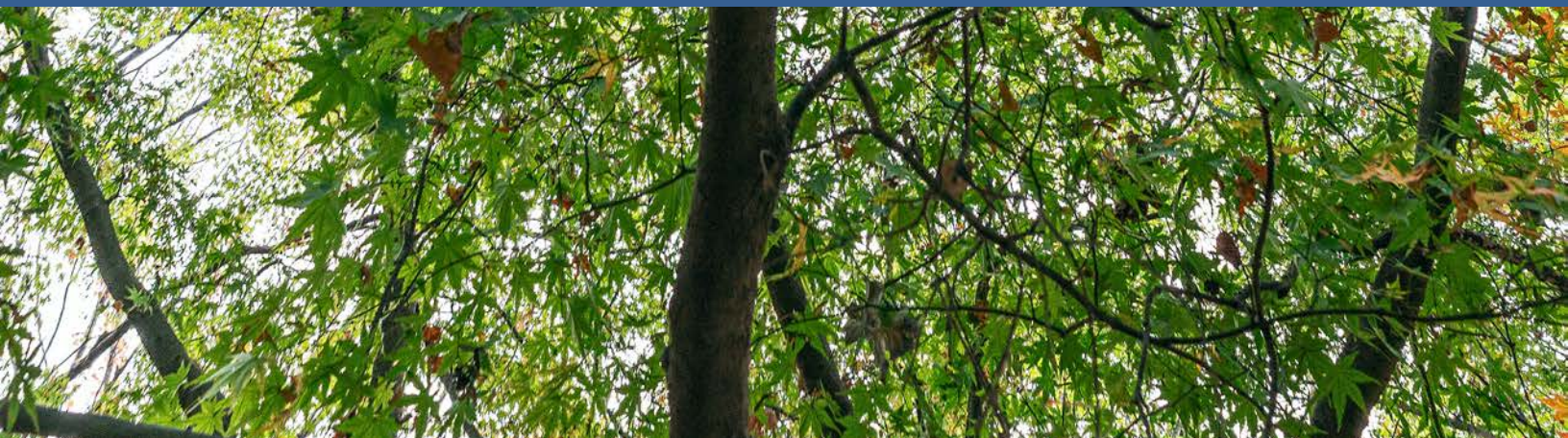
This strategic evaluation plan outline can also serve as tool to document revisions to the strategic evaluation plan. Inserting the following brief checklist after each section may help with this process.

___ Implemented as planned
___ Changes made (describe changes as well as the rationale for changes)



CHAPTER FOUR

Digging into the Details: The Individual Evaluation Plan



CHAPTER FOUR: Digging into the Details: The Individual Evaluation Plan

In this chapter we provide step-by-step instructions for developing what we call individual evaluation plans. Individual evaluation plans capture the specific details that will ultimately be used by the person or team that implements the evaluation. Such plans are relatively common in the practice of evaluation and are often simply referred to as an evaluation plan or evaluation protocol. Individual evaluation plans will be crafted after a strategic evaluation plan is in place (if one has been created) for each of the priority evaluations proposed. Like the process of developing a strategic evaluation plan presented in **Chapter 3**, we follow the steps of the CDC Framework. However, this time, a one-to-one correspondence exists between the steps and activities. At the end of the chapter we provide a detailed outline of an individual evaluation plan.

By the end of Chapter Four, the reader will be able to

- Explain the general process for developing individual evaluation plans.
- Describe the main components of an individual evaluation plan.
- Lead or effectively participate in an individual evaluation planning process.

Introduction

The strategic evaluation plan developed at the end of **Chapter 3** provides preliminary insights about the composition of each evaluation to be conducted in the upcoming years. These details can be used to develop a plan for carrying out each specific evaluation. As seen in **Figure 4.1**, the process for developing this plan closely follows the CDC Framework and produces a written document that captures key elements that will be needed during the implementation phase.

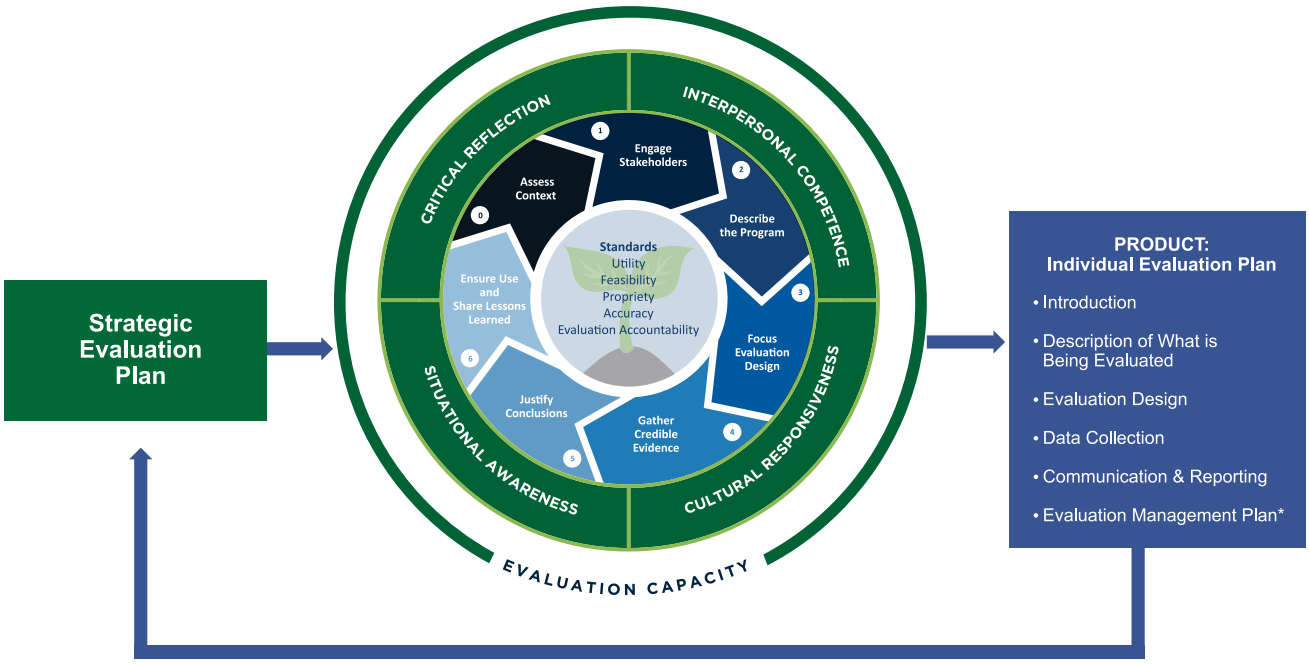


Figure 4.1 Individual Evaluation Planning Process and Product

Step 0. Assess Context

As indicated in **Chapter 2**, the first task in planning an evaluation is to assess the context. Each evaluation context is unique. It is important to gain an understanding of what it means to be in this place at this time. It is also important to understand the history that brought us to the current moment so an evaluation can be appropriately attuned (King & Stevahn, 2013). As with any form of inquiry, this involves collecting and analyzing information from various sources, both formal (e.g., demographic data, historical records) and informal (e.g., personal observations, general conversations).

Evaluators can begin by familiarizing themselves with the nuts and bolts of the program such as its purpose, origin, location(s), intended beneficiaries, basic operations, and influential players. Following this, evaluators should expand their assessment to the broader environment with a particular focus on the historical, economic, health, and social dimensions of the communities in which the program exists. Understanding how power and privilege are distributed across communities and whose perspectives tend to be heard more or less often are critical pieces of information when planning and conducting an evaluation.

For the most part, contextual information can be accessed from various sources such as newspapers, websites, program communications and reports, conversations, and historical documentation. Evaluators should also consider visiting one or more program sites to gain a feel for the program and respective communities through first-hand observations and personal interactions. Site visits also provide a great opportunity for evaluators to develop relationships and build trust and rapport with those involved with or interested in the program or community (King & Stevahn, 2013).

The "How" of Practice
Interpersonal Competence

Evaluators should not assume that all requests for information and conversation will be welcomed, particularly within communities who have been studied extensively or have had a negative experience with evaluation (or research for that matter). Thus, it is imperative to exercise sound judgement and interpersonal skills when interacting with people. Among other qualities, evaluators should be

- honest and upfront about their intentions,
- respectful in how they ask questions,
- aware of potentially sensitive topics and when and how to broach them,
- present during conversations,
- aware of social cues, and
- grateful and considerate of people's time.

These early interactions could result in big gains for the entire evaluation if they are conducted in a manner that facilitates positive perceptions of the evaluation and toward the evaluator.

How can evaluation theory help?

In **Chapter 2**, we highlighted strategies proposed by Greene et al. (2011) to assess context. King and Stevahn (2013) also encourage evaluators to get to know the contextual factors that will affect the evaluation and seek opportunities to leverage them as part of planning. Similar to Greene et al. (2011), King and Stevahn (2013) propose a variety of techniques that evaluators can use to learn about the evaluation context such as conversations with stakeholders, reviewing program or organization websites and other communications, conducting a windshield survey (i.e., driving around the setting to learn about the location), and site visits (p. 202). King and Stevahn (2013) have also compiled a list of contextual factors evaluators should consider when assessing the evaluation. Among the factors, listed in the blue shaded box, King and Stevahn (2013) encourage evaluators to remain sensitive to other relevant factors that are unique to the specific context.

Step 1. Engage Stakeholders

When a variety of stakeholders are involved in evaluation planning from the outset, it is possible to (a) plan and conduct evaluations that more closely fit the stakeholders' collective needs; (b) foster a commitment to use the evaluation results; and (c) promote accountability, transparency, and equity. Many of the causes of misunderstandings about evaluation and misunderstandings about barriers to the productive use of evaluation findings can be avoided or minimized when program stakeholders are included in discussions at various points throughout the evaluation. Including evaluation stakeholders in planning conversations—listening to and understanding their hopes for what will come from the evaluation and what comprises good public health programming—can result in their meaningful inclusion. It can also decrease anxiety they may experience regarding the possible results.

There are three major categories of evaluation stakeholders to consider (Russ-Eft & Preskill, 2009, p. 165–168):

- ***Primary stakeholders.*** Individuals who have the ability and authority to use evaluation findings to alter the course of the evaluand. Examples of primary stakeholders include staff members, managers, and funders of public health programs.
- ***Secondary stakeholders.*** Individuals who are affected in some manner by the evaluand and, therefore, are likely to be affected by any changes made because of the evaluation findings. One example includes participants in a public health program.
- ***Tertiary stakeholders.*** Individuals who are not directly affected by the changes that result from acting on the evaluation findings, but who might have a general interest in the results. Examples include legislators and similar public health programs in other jurisdictions.

Factors to consider when assessing context

- Mission/vision/values
- History/Site/location
- Size/scale
- Structure
- Norms/routines
- Logic models
- Budget/funding
- Content/field
- Wider environment
- Other relevant factors

King & Stevahn, 2013, p.198

Benefits of working with stakeholders

- Develops support among program leadership and other stakeholders for the evaluation
- Facilitates appropriate timing of evaluation in relation to information needs
- Leads to the development of relevant evaluation questions, which in turn supports use or action
- Promotes findings that are credible and understood by stakeholders
- Develops evaluation capacity, including **evaluative thinking**

One way to consult with a subset of stakeholders is to form an **Evaluation Planning Team** for each of the evaluations prioritized in the Strategic Evaluation Plan. You may also hear the evaluation planning team referred to as an evaluation advisory group. In deciding whom to ask to participate, include diverse perspectives on the team. The varied interests and perspectives of different stakeholders will be valuable in conducting a rigorous and useful evaluation. It is certainly possible to identify individuals to invite based upon their representation of the stakeholder categories previously discussed. However, this is also a good opportunity to establish a team that represents many voices.

Throughout each step of the evaluation planning process consult with and consider stakeholder needs. **Box 4.1** lists several items that can be helpful to discuss with the Evaluation Planning Team as well as other stakeholders. In fact, a first task for the Evaluation Planning Team could be brainstorming additional stakeholders to include, as well as how to include them in the evaluation process. Questions that may be helpful to discuss as a team include those asked when assessing the context and establishing the potential membership of the Evaluation Planning Team. Other potentially helpful questions include

- What individuals and groups have an interest in the outcomes of this evaluation?
- How will stakeholders who reflect the diversity of those who may be affected by the evaluation's findings be engaged? (For suggestions see *Practical Strategies for Culturally Competent Evaluation)*
- What aspect of the evaluation are the stakeholders most interested in? For example, are they interested in the evaluation from a cost angle, effectiveness of the program, possible improvements, or something else?
- What role would we like these stakeholders to play in developing or implementing this evaluation? Examples include serving on the Evaluation Planning Team, asking them to review early versions of the Evaluation Plan or products from evaluation implementation as an external reviewer, collecting data, interpreting findings, or using results.

Planning an Evaluation?
Ask stakeholders about…

- Previous experiences with evaluation
- Hopes for the evaluation
- Program priorities
- Information needs
- Evaluation questions to explore
- When information is needed
- What they view as credible evidence
- How they will use evaluation findings
- Other perspectives to consider
- Privacy and confidentiality considerations
- Cultural sensitivity
- What they will do if findings suggest an immediate need for program modifications
- How negative findings may affect the program or community
- Stakeholder preferences and availability regarding engagement

The results of these discussions can be summarized in a table, similar to **Table 4.1**, and included in the individual evaluation plan. The Evaluation Planning Team may find it helpful to include additional columns in this table such as how and when to engage the stakeholder or a column that includes a note for any special considerations that need to be made in engaging a specific individual or group (e.g., times of year that are naturally busy for the stakeholder such as the start-up of a school year for teachers).

Table 4.1. Stakeholder Assessment and Engagement Plan

Stakeholder Name	Stakeholder Category	Interest or Perspective	Role in the Evaluation	Other Considerations
May be an individual or group	*Primary, secondary, tertiary*	*Program participant, staff members, etc.*	*Evaluation Planning Team, external reviewer, etc.*	*Cultural, logistical, historical, or other factors that need to be considered to facilitate meaningful engagement*

How can evaluation theory help?

Engaging stakeholders is a specific instance when drawing upon the **evaluation theories** covered in **Chapter 2** can be helpful. Evaluation theorists prescribe a range of possible approaches for engaging stakeholders in evaluation planning and implementation (Christie & Alkin, 2013). Typically, this variation corresponds with two types of decisions: (1) how many and what types of evaluation stakeholders to engage and (2) in what steps of the CDC Framework should stakeholders be engaged. The range in these approaches is vast. With respect to the number and types of stakeholders, some theories recommend engaging as many individuals within an organization, program, or community (whichever applies best to the evaluand) as possible. Others recommend limiting the engagement to managers or others who have direct decision-making authority. With respect to what steps to engage the stakeholders in, some theorists suggest including stakeholders in extensive conversations to describe the evaluand, prioritize evaluation questions, identify credible methods, and finally to discuss the evaluation findings. Others suggest limited to no inclusion of stakeholders. Still others recommend inclusion of

stakeholders in every step, including the development of data collection instruments, gathering of data, and data analysis and interpretation.

Deciding which approach to take is based upon a variety of factors, including an evaluator's beliefs about the role of evaluation and the evaluator, the organizational culture, and logistical constraints on the evaluation (e.g., resources in terms of time and funding, feasibility of engaging stakeholders, availability of stakeholders). Any evaluation requires a careful balance of activities to maximize, to the extent possible, quality with respect to the standards of utility, feasibility, propriety, accuracy, and evaluation accountability. In making determinations about which stakeholders to engage in planning the evaluation and, ultimately, throughout the implementation, it will be important for the Evaluation Planning Team to discuss how, if at all, the decisions made will impact achievement of the evaluation standards.

Step 2. Describe What is Being Evaluated

Developing a clear description of the evaluand is critical to creating a useful evaluation and in strengthening the evaluand itself. Step 2 can be invaluable for identifying any gaps in logic about how the evaluand is intended to operate, revealing divergent views between stakeholders about intended results, and in later steps, considering topics to focus the evaluation on and reasoning through the most appropriate timing of measurement.

In **Chapter 3**, we suggested including a logic model for the overarching program in the Strategic Evaluation Plan. When developing an individual evaluation plan, it is important to zoom in on the aspect of the larger program that is specific to the evaluand. A logic model, or another visual depiction, that specifically describes the evaluand should be included in the individual evaluation plan. A text-based description of the logic model is also often helpful to include. This description can explain how what is being evaluated contributes to accomplishing the intended outcomes. It may also describe important features of what is being evaluated, such as the context in which it operates, the characteristics of the population it is intended to reach, and its stage of development. Consider the following questions in formulating the narrative:

- **Need:** What need is the program designed to meet?
- **Context:** What is the program's context? What contextual or cultural factors may affect its implementation or effectiveness?
- **Population Addressed:** For whom are the activities intended?
- **Stage of Development:** How long has the program been in place? Is it in the planning or implementation stage?
- **Resources or Inputs:** What resources are available to support the program (e.g., personnel, money, space, time, partnerships, technology)?
- **Activities:** What specific activities are conducted (or planned) to achieve the program's outcomes?
- **Outputs:** What do the activities produce (e.g., materials, services delivered)?
- **Outcomes:** What are the program's intended outcomes? What do the program designers ultimately want to change as a result of the activities (long-term outcomes)? What occurs between the activities and the point at which the ultimate outcomes are realized (short-term and intermediate outcomes)?

Box 4.2. Love them or hate them: Logic models

It seems that people either love or hate logic models. Despite the strong feelings on either side, logic models have proven to be a useful tool over the years for describing evaluands. As the discipline of evaluation has evolved, critiques have continued to arise about logic models, with one of the most common being their overly linear and often simplistic nature.

Though we are keen on logic models as a tool, we recognize that they are not always the tool for the job. We also recognize that the typical and often used box and arrow model subscribes to a Western frame that is not appropriate in many contexts. Following are some other types of visual depictions that you may find helpful to explore in the event you feel a different tool is necessary

- Theory of Change
- Action & Change Model (Chen, 2015)
- Concept Maps (Trochim, 2020)
- Interactive Logic Models
- Rich Pictures

A visual depiction and accompanying narrative can be valuable for the primary stakeholders of the evaluation as well as stakeholders of other programs who might be interested in implementing similar activities to the ones evaluated. With a clear description of the activity and context in which it resides, other programs will be better able to determine how likely it is that the evaluation results relate to what they would see if they chose to implement this same activity. In addition, it is often the case that developing a logic model highlights areas of a program that are missing or underdeveloped. Perhaps the logic model describes a training that the program conducts and stakeholders realize that all of the component pieces are in place to *do* the training but no activity is in place to actually market or make individuals aware of the training. Stakeholders can start taking immediate action on this observation by building out the activity—a clear instance of process use.

How can evaluation theory help?

The development of a logic model, or similar visual description, is widely accepted as an important step in the evaluation process and therefore is discussed within several evaluation theories. One theory, Theory Driven Evaluation, can be very helpful in considering what inputs to use in developing these descriptions. For instance, Chen (2015) and Donaldson (2007) note that consulting stakeholders for their perspective on the intended activities, outcomes, connections between activities and outcomes, and relevant program or organizational documents is important. Chen (2015) and Donaldson (2007) also stress the added value of integrating insights and evidence from existing social science theories and from research or evaluation studies on similar evaluands into these models. For instance, social science theories or evidence from existing research may suggest that one type of outcome is more likely to precede another in a causal sequence or indicate that the intended dose of a specific activity is unlikely to give rise to an anticipated outcome. This would suggest a needed change in the proposed programming or outcome pathway. Program Driven Evaluation theorists add that the integration of social science (or other) theory and research findings can also help in later stages of planning an evaluation, because they provide insights on what to measure, how to measure, and when it is most likely that intended outcomes will arise post-intervention (Donaldson and Lipsey, 2006).

Evaluation theory, specifically Theory Driven Evaluation as described in Chen (2015), also provides suggestions for visual depictions that may be helpful beyond the traditional logic model. Specifically, Chen introduces us to the Action Model and Change Model, as well as the connections between these. The Action Model depicts the details of the program plans. As described by Chen (2015),

> In the action model are found the bases for answering questions such as the following: What are the crucial elements of the intervention? What kind of organization is needed to deliver the services? Who is best qualified to deliver them? How will implementers be trained? What is the target group? How will the target group be reached? (p. 68)

The Change Model describes several things: the specific intervention implemented, the changes this intervention is supposed to make with respect to the factors that contribute to the public health problem, and the ultimate intended outcomes of the intervention. The action and change models offer alternatives to the more traditional logic model and can be particularly helpful in circumstances where a program is implementing an evidence-based intervention that needs to be evaluated for factors such as **fidelity** (i.e., adherence to a planned delivery model) (Mowbrey et al., 2003).

Step 3. Focus the Evaluation Design

For each evaluation they prioritized, the Strategic Evaluation Planning Team generated some initial ideas about the evaluation questions to answer, as well as the possible evaluation designs that could be used to respond to the questions. In developing the individual evaluation plan, revisit these draft ideas with the Evaluation Planning Team and other evaluation stakeholders to make final decisions about the specific questions that will be answered through the evaluation.

Establishing the Purpose of the Evaluation

The first step when focusing the evaluation design is to clarify the purpose of the evaluation. This will help to create a shared vision for the general scope of the evaluation and its intended uses. An evaluation purpose statement is a clear and succinct statement, often two to three sentences in length, that describes why the evaluation is being performed, the general emphasis of the evaluation, and its intended uses (Russ-Eft & Preskill, 2009).

Sample Purpose Statement

The overarching purpose of this evaluation is to provide formative insights about the program operations during the first two years of implementation. It is anticipated that program management and staff members will use interim findings to adjust the program operations during these first two years and use the final results to make decisions about any programmatic redesign needed in subsequent years of programming.

Evaluation purpose statements can explain if the intention is to deliver insights for making developmental (i.e., continuous improvements that align with ongoing innovation in programs), formative (i.e., program adaptations or modifications), summative (i.e., program continuance, cancellation, or expansion) decisions, or a mix. They also may explain if the expectation is for the evaluation questions to focus on processes, outcomes, or costs of the program. Ultimately, the purpose statement "…serves as the basis for the evaluation questions, design, and methods" (DHHS, 2005, p. 40).

Refining Evaluation Questions

The evaluation questions developed as part of the Strategic Evaluation Plan are truly drafts. The purpose of drafting these questions was to have enough detail about the potential evaluation to prioritize limited resources across the proposed evaluations and to develop a general timeline for their implementation. Given this, it is important to revisit these questions with several evaluation stakeholders. Are the questions proposed those that will be most informative for their work given the timing of the evaluation? Is there anything that might be more important, or more pressing to answer given the timeframe? Are the questions generally on target, but require some sub-questions or additional nuance to be most informative? Conversations addressing these items may lead to different questions and perhaps a longer list of evaluation questions than originally listed in the Strategic Evaluation Plan.

Once you have the revised list of evaluation questions, it is important to engage in two additional steps. First, examine the questions for any overlaps and for their phrasing. Perhaps there are two or more proposed evaluation questions that are asking for similar information; if the information request is similar enough there may be an opportunity to combine the questions. After reviewing the questions for redundancies, consider the exact wording. Draft questions are often yes or no questions, but often (and we might argue, always) evaluation stakeholders would find a yes or no response less helpful than needed. Consider revising the wording so the question is more open ended, and asks for the type of insight the stakeholders are seeking: "To what extent…," "In what ways…," "How…," or "Why…"

Evaluation questions are *not* the same as survey or interview questions; evaluation questions provide the general boundaries within which the evaluation is conducted, much like hypotheses provide the framework for research studies. Ultimately, it is important to have evaluation questions that are narrow enough in scope to be answerable while also being broad enough to produce useful insights. Using the Good Evaluation Questions Checklist can increase the likelihood that the evaluation questions selected will produce actionable information.

Aiming for three to five questions will generally result in an evaluation with a reasonable scope. After refining the list of evaluation questions, you may find that you have many more than this. If so, you may find it helpful to engage in a prioritization process, much like that described in **Chapter 3**, with the Evaluation Planning Team (and other stakeholders as appropriate) to select the final evaluation questions. Once the Evaluation Planning Team has identified the specific questions the evaluation should address, the next step is to decide how to answer those questions. This decision-making process has two main steps: (1) deciding on the evaluation's overarching design, which is covered in Step 3 of the CDC Framework (*Focusing the Evaluation*) and (2) deciding how to collect the data, which is covered under Step 4 (*Gather Credible Evidence*).

Selecting an Evaluation Design

Like with the evaluation questions, the Strategic Evaluation Planning Team made some initial assumptions about each priority evaluation. The assumptions allowed them to identify one or more potential study designs. Following the Evaluation Planning Team's discussion about the evaluation questions, it may be the case that the revised or refined set of evaluation questions is different enough that the evaluation design proposed in the Strategic Evaluation Plan needs to be reconsidered.

As described in **Appendix E** there are many evaluation study design options to select from, whether experimental, quasi-experimental, non-experimental, mixed-method, or cost-related. With all these possible designs (not to mention their potential variants), you might be wondering how to go about choosing the most appropriate one. Unfortunately, no cookbook is available to help us decide which design options to use for a given evaluation. However, two general principles can be very helpful in trying to make this decision:

1. Select the study design that is best suited to answering the evaluation questions and that is viewed as credible by the stakeholders
2. Consider the alignment between the design and the likely achievement of each evaluation standard (i.e., utility, feasibility, propriety, accuracy, evaluation accountability)

Although it is often tempting to select a design based solely on its familiarity or feasibility, first consider the evaluation questions the Evaluation Planning Team has identified. The questions often suggest a specific design option. For example, suppose the Evaluation Planning Team discovers that the program is implementing a public health intervention very similar to one offered by program partners. Given the similarity between the intended outcomes of these interventions, the evaluation stakeholders would like to know if they would be better off focusing their efforts on improving and expanding one of the interventions, and if so, which one. A cost-benefit or cost-effectiveness analysis may answer such questions well.

The evaluation questions will not always suggest one specific design. When more than one design is possible, referring to the standards for program evaluation—utility, feasibility, propriety, accuracy, and evaluation accountability—may help to sort through the available options. It may be the case that different designs are needed to respond to different evaluation questions presented by the Evaluation Planning Team or that the collective use of more than one design will enhance the quality of the answers (Takashori & Tedlie, 2009). Ultimately, the decision about which design(s) to use requires carefully balancing multiple ideas, perspectives, and criteria. We encourage employing creativity and flexibility when selecting the design that is most appropriate to the information needs the Evaluation Planning Team identifies.

How can evaluation theory help?

Evaluation theory provides several helpful ideas for identifying and selecting from possible evaluation questions. In the early 1990s, Shadish, Cook, and Leviton (1991) published a book on evaluation theory and grouped a set of seven influential evaluation theorists within three phases of evaluation across the field's history. The last phase comprised the work of two theorists, Cronbach and Chen, who Shadish et al. considered

"contingency" theorists. The defining feature of their theories was providing advice to evaluators about how to navigate the many practice options they have. Cronbach and Chen made several suggestions about what to consider in making important decisions (e.g., which evaluation questions to prioritize).

Cronbach (1982), for instance, described a rather extensive process for creating an exhaustive list of evaluation questions (divergent phase) which was followed by a process for refining the list (convergent phase). As part of the convergent phase, he suggested factors for consideration such as the maturity of the program (How long has it been in place? What is its relative stage of development and what questions are most important to answer at that time?), questions that are likely to have greater influence with specific audiences (leverage), which have not been answered in the recent past, and the balance between the cost of answering the question and the information that will likely result (Shadish et al., 1991).

Other evaluation theorists, such as Greene et al. (2011), call our attention to the value-laden nature of evaluation questions. In Values-Engaged Educative Evaluation (Greene et al., 2011), equity questions are always included in the evaluation. Greene provides evaluators with a better understanding of what evaluation questions may look like when the focus and intent of the evaluation relates to social justice and equity.

With respect to evaluation designs, some theorists challenge our thinking regarding the common use of social science research methodology in evaluation. Michael Scriven, regarded as the father of modern-day evaluation, recognizes the use of social science research methods for evaluation but also offers alternatives in thinking through design. For instance, his popular *Key Evaluation Checklist* (2015) guides evaluators through many details to consider when designing an evaluation. In addition, his **Modus Operandi** approach (1976) calls our attention to the promise of using investigatory methods in responding to evaluation questions.

Step 4. Gather Credible Evidence

Once the Evaluation Planning Team has identified evaluation questions and decided on the most appropriate evaluation design, the next task is to make decisions about the data needed to answer those questions. In emphasizing the need for credible evidence, the CDC Framework reminds evaluators to cast a wide net, considering the types of evidence various stakeholders will find convincing or relevant. Given the variety of people who may be invested in the results of an evaluation, it is likely that there will be a range of perspectives on what counts as credible. For example, stakeholders with scientific backgrounds will likely expect the data to meet the standards of their disciplines. Program advocates will expect data to be sufficiently reflective of community perceptions and values. People of differing cultural and educational backgrounds will bring a multiplicity of assumptions, expectations, and levels of knowledge about the methods and strategies for determining evidence. It is the evaluator's role to work with these stakeholders to come to an agreement about what constitutes credible evidence and how it should be obtained.

Identifying Criteria of Merit and Indicators

The first step in understanding how credible evidence can be gathered is to identify criteria of merit. Once evaluation questions are clarified, the Evaluation Planning Team should spend some time discussing the dimensions of performance (i.e., criteria of merit) that align with the evaluation questions. For instance, imagine you have chosen to evaluate a relatively new intervention designed to educate healthcare practitioners about appropriate disease management practices, and the evaluation question pertains to the success of the intervention. When you engage in a conversation with the evaluation stakeholders about what might constitute success, they identify *attendance* and *completion* as being important. These characteristics constitute what evaluators call criteria of merit – "…the aspects of what is being evaluated that define whether it is good or bad and whether it is valuable or not valuable" (Davidson, 2005, p. 239).

By clearly articulating the criteria of merit, the team is defining what they mean by the ambiguous words that sometimes appear in evaluation questions. Once this is established, measurement becomes an easier process and the team can move on to establishing one or more **indicators** for each criterion (also often called "performance

measures") (DHHS, 2005) with the evaluation stakeholders. Expanding on the example above, when considering the criteria of merit the Evaluation Planning Team knows it will be important to understand the extent to which the intended participants are attending and completing the training. In thinking this through further, they define the intended participants: nurses, physicians assistants, and physicians. The team suggests that the following indicators will provide stakeholders with the most actionable information:

1. Attendance rate
2. Attendance rate by type of practitioner (nurses, physician assistants, physicians)
3. Proportion of attendees who complete the training
4. Proportion of attendees who complete the training by type of practitioner

You can see from this list of indicators that it will be important to have a question on the attendance sheet that asks attendees what type of healthcare practitioner they are. If the Evaluation Planning Team had not discussed the indicators that will be used to determine the success of this intervention, it is possible this important piece of information would have been left off the attendance log.

Once the Evaluation Planning Team has identified criteria of merit and indicators, it may be helpful to review the draft data collection strategies outlined in the strategic evaluation plan. Are there new data sources that may be helpful to incorporate? Are there existing data sources that can be used to calculate the proposed indicators? Do the methods meet the evaluation stakeholders' needs for credible information? Does the draft data collection timeline need to be adjusted? The Evaluation Planning Team should be creative at this stage and not be afraid to question assumptions that underlie each criterion of merit, proposed indicator, and possible data source; doing so will strengthen the final outcomes of the evaluation.

The "How" of Practice
Critical Reflection

There are values underlying criteria of success, just as there are for all methodological decisions. How you define what does (and does not) qualify as success is informed by your worldview and lived experience. Brookfield (2017) and others argue that worldviews are susceptible to dominant ways of thinking that promote existing power relationships and inequities. As such, how we interpret and perceive is often not oriented toward advancing change but reinforcing what currently exists.

So what can we do? Engaging diverse perspectives in developing criteria can be helpful, but the evaluation team will likely have the most influence on what becomes the final set of criteria. As such, critical reflection is key. Here are a couple questions that can help guide you:

- What important things happen, or will happen, when this program is successful? What sources have instilled in me that these things are important? How credible are these sources in this context?
- Who is most likely to be impacted positively and negatively by these criteria? To what extent, will this outcome reinforce existing inequities? How does that relate to my intentions as a professional and as a person?

Data Collection Methods

As the Evaluation Planning Team works with the evaluation stakeholders to build out the final approach to data collection, it is critical to keep in mind why the specific data need to be collected and how the evaluation stakeholders intend to use the findings that will stem from the analysis. Being explicit about the use of data before it is collected helps conserve resources and reduces respondent burden by ensuring only the data that is absolutely needed is collected.

Wondering what the realm of possibilities are? **Table 4.2** shows a range of data collection methods and lists some advantages and disadvantages of each. Because the perfect data collection method is rare, the Evaluation Planning Team will have to carefully weigh the options. In some situations, it will be possible to compensate for the disadvantages of one data collection method by selecting multiple methods. Furthermore, each method has its own inherent biases; mixing methods can alleviate such concerns. See **Appendix E** for a detailed discussion of mixed methods and the various rationales for using this approach.

Table 4.2. Data Collection Methods

Method	What is it?	Some Advantages	Some Disadvantages
Surveys & Questionnaires	"A questionnaire is a set of questions for gathering information from individuals. You can administer questionnaires by mail, telephone, using face-to-face interviews, as handouts, or electronically (i.e., by e-mail or through web-based questionnaires)" (CDC, 2018a, p.1).	• Reasonably inexpensive to administer. • Multiple options for distribution: mail, phone, in person, email, internet. • Respondent privacy can be protected—allows for gathering sensitive data. • Accommodates different types of questions: open ended, closed ended. • Can feasibly administer to many people across large geographic area.	• Time consuming to develop, pilot, and conduct follow-ups. • Low response rates and non-response. • Sampling frame sometimes difficult to identify. • Might not get careful feedback. • Impersonal. • Level of detail provided is limited and may be insufficient for informing subsequent efforts. • Overused.
Method	**What is it?**	**Some Advantages**	**Some Disadvantages**
Interviews	"An interview is a method of asking quantitative or qualitative questions orally of key participants. Quantitative questions are closed ended and have specific answers to choose among that can be categorized and numerically analyzed. Qualitative questions are open-ended; that is, the respondent provides a response in his or her own words. Interviews conducted for program evaluation are typically qualitative but may also include some quantitative questions" (CDC, 2018b, p.1).	• In-depth information obtained. • Option to clarify questions is available. • Accessible to low-literacy populations. • Able to develop rapport and potential trust with participants.	• Often resource intensive to conduct and analyze. • Potential for interviewer bias. • Some respondents find intrusive. • Trained interviewer needed. • Scheduling.
Focus Groups	"A focus group is a group interview of approximately six to twelve people who share similar characteristics or common interests. A facilitator guides the group based on a predetermined set of topics. The facilitator creates an environment that encourages participants to share their perceptions and points of view. Focus groups are a qualitative data collection method, meaning that the data is descriptive and cannot be measured numerically" (CDC, 2018c, p.1).	• Data obtained can have good depth and breadth. • Building upon ideas from other participants can enhance richness of comments provided by participants. • Quick way to obtain common impressions across multiple individuals.	• Analysis is resource intensive. • Potential for facilitator bias and group influence on one another. • Possible for a few individuals to capitalize on time. • Requires skilled facilitator. • Group composition should be selected carefully to ensure comfort in responding (e.g., managers and staff members). • Scheduling.

Method	What is it?	Some Advantages	Some Disadvantages
Observations	"Observation is a way of gathering data by watching behavior, events, or noting physical characteristics in their natural setting. Observations can be overt (everyone knows they are being observed) or covert (no one knows they are being observed and the observer is concealed). The benefit of covert observation is that people are more likely to behave naturally if they do not know they are being observed. However, you will typically need to conduct overt observations because of ethical problems related to concealing your observation" (CDC, 2018d, p.1).	• Direct observation rather than self-report. • Recall bias is not present. • Obtain real-time information. • Obtain information about the context.	• Resource intensive regarding time for observing as well as establishing inter-rater reliability. • Potential for observer bias. • Individuals may behave differently due to presence of observer. • Requires trained observers
Document Reviews	"Document review is a way of collecting data by reviewing existing documents. The documents may be internal to a program or organization (such as records of what components of an asthma management program were implemented in schools) or may be external (such as records of emergency department visits by students served by an asthma management program). Documents may be hard copy or electronic and may include reports, program logs, performance ratings, funding proposals, meeting minutes, newsletters, and marketing materials" (CDC, 2018e, p.1).	• Unobtrusive. • Fairly inexpensive. • Helpful for understanding history. • Often readily accessible.	• Quality of information may be unclear, difficult to assess, or incomplete. • Potentially time consuming. • Reasons for originally collecting data may not align with current needs.

*Information contained in this table comes from CDC's *Evaluation Briefs* series (CDC 2018a-e).

If the Evaluation Planning Team decides that new data (i.e., primary data) need to be gathered to respond to the evaluation questions, it will be important to include a pilot test of the data collection instruments and associated procedures. This is a critical step in upholding the accuracy standard and is important whether the plan includes data collection through surveys, observations, interviews, focus groups, or through abstracting data from archival sources (e.g., medical records). During the pilot test, consider items such as the clarity of instructions, appropriateness and feasibility of the questions, sequence and flow of questions, and feasibility of the data collection procedures. Techniques such as **cognitive interviewing** (Willis, 2005) can help to evaluate question validity and response error by exploring participants' comprehension and interpretation of questions.

In addition to pilot testing instruments, the Evaluation Planning Team should consider how data collectors will be trained. If multiple individuals will be responsible for gathering the same type of data for the evaluation, it is critical that they receive training to uphold the evaluation standards of utility, feasibility, accuracy, propriety, and evaluation accountability. Training can be formal or informal depending on planned activities and the experience level of the data collectors, but all training should aim to ensure (1) that standards and procedures will be applied consistently and (2) that data collectors and their supervisors understand how the data will be used in the evaluation, how planned activities will be carried out, their respective roles and responsibilities, and how to handle events that may arise. **Appendix F** includes details about what to consider in training data collectors as well as the many variants possible in this process.

The "How" of Practice
Cultural Responsiveness

Remember, a culturally responsive evaluation embraces the cultural context of the program. As you plan, consider the relevancy and sensitivity of your proposed strategies to stakeholders and communities. Some questions to think about may be

- Are you allowing for authentic input?
- Do your methods accommodate multiple worldviews?
- What barriers are inherent to your proposed designs and data collection strategies?
- What can you do to better align your data collection strategies with the values and interests of stakeholders?

Keep in mind, many communities feel over-burdened and under-rewarded by research and evaluation. It is everyone's responsibility to change the direction of this relationship. Being culturally responsive is an important step in the right direction.

Sampling Considerations

One important decision point with respect to gathering credible evidence is what data collection methods will be used (assuming new data collection is required). Another equally important consideration is what data will be collected. It is not always possible to collect data about everything that is of interest (e.g., every patient in a clinic; every medical record). In this case, the Evaluation Planning Team will need to consider how to draw a sample. Although sampling is frequently mentioned in the context of collecting data through surveys or questionnaires, it is equally applicable to other forms of data collection, such as record reviews, observations, interviews, and even focus groups. There are two main types of sampling: (1) probability sampling and (2) purposive sampling.

Probability Sampling. Probability sampling relies on randomly selecting units from a larger listing of units that you would like the sample to represent. The general idea of probability sampling is to randomly select units from this list (or within specific, non-overlapping subgroups that exist within this list) so that each unit has an equal and non-zero probability of being selected (Crano et al., 2015). There are multiple ways to obtain a probability sample:

- Simple random sampling is the most basic form of probability sampling. Each unit included in the sampling frame has an equal likelihood of being selected.
- Systematic random sampling is often used instead of simple random sampling. It consists of five steps: (1) creating an unordered list from which the sample will be drawn (i.e., sampling frame), (2) deciding how many units to select from this list (i.e., sample size), (3) calculating the "Nth number" by dividing the

number of the entire sampling frame by the desired sample size, (4) selecting one random number, and (5) selecting every "Nth" numbered item beginning with the item on the list corresponding with the random number (Crano et al., 2015, p. 225).

- Stratified random sampling is often used when evaluators want information about specific subgroups within a population. The basic idea of stratified random sampling is to select the sample in a way that will allow the evaluator to say something about the overall sample and about specific sub-groups within the sample. In this method, the sampling frame is broken down into non-overlapping categories known as strata, and a random sample is taken from each (Crano et al., 2015; Trochim, 2020).

The purpose underlying the use of probability sampling techniques is to gather data from a sample that is representative of the larger population from which the sample came. As such, the statistics calculated from this sample may be a bit off from the values you would have obtained if you were able to collect data from the entire population. The amount you are off because of sampling (known as standard error) gets smaller as you get a larger sample. The smaller standard error occurs since you are getting closer to the number of units that comprise the full population (Trochim, 2020). When you are using quantitative data, you need to account for the standard error when calculating and reporting descriptive statistics (i.e., frequencies, percentages) and inferential statistics (e.g., t-tests; F-test from ANOVA; odds ratios from logistic regression). The precision of estimates of population parameters (e.g., mean) and the power to detect differences that actually exist within or between the populations through the use of inferential statistics are, in part, dependent upon the size of your sample.[8] It is also possible to use probability sampling for qualitative data, however you will not need to calculate a standard error to use when reporting the analyses.

Purposive Sampling. Not all sample selection is based on the principles of probability. In purposive sampling the units are not sampled with the intention that each unit has an equal probability of selection but rather with some sort of intended purpose in mind. Michael Patton (2002) offers a helpful comment regarding this type of sampling,

> The logic and power of purposeful sampling lie in selecting information–rich cases for studying in depth. Information-rich cases are those from which one can learn a great deal about issues of central importance to the purpose of the inquiry, thus the term purposeful sampling. (p. 230)

There are many different types of purposeful sampling strategies that can be used. Patton (2002), for example, describes 15 types: extreme/deviant case, intensity, maximum variation, homogeneous, typical case, critical case, snowball/chain, criterion, theory-based, confirming and disconfirming cases, stratified purposeful, opportunistic/emergent, purposeful random sampling, politically important cases, and convenience (pp. 243–244). The purposes of each strategy vary greatly. For instance, in *extreme sampling*, the purpose is to learn something about very special cases, or outliers, whereas in *typical case sampling* the purpose is to learn about the experience of the normal or average case. The most important thing about purposive sampling is to be very clear about the purpose or intention of a sample; once you have that, the most appropriate sampling strategy often becomes clear.

Concerns regarding sample size are very different when using purposeful sampling and probability sampling. Since purposive sampling does not strive to obtain a representative sample of a larger population, the typical concerns that we experience when talking about obtaining precise estimates of a population value (e.g., mean) from a sample simply do not apply. Patton (2002) explains the considerations behind sample size when it

[8] The standard deviation of the characteristic of interest within the population is also a factor that affects the standard error. The larger the standard deviation, the larger the standard error.

comes to purposeful sampling: "Sample size depends on what you want to know, the purpose of the inquiry, what's at stake, what will be useful, what will have credibility, and what can be done with available time and resources" (p. 244). When calculating sample size for a purposive sample, we encourage you to consider the (1) appropriate balance between the breadth and depth of the type of data you will obtain using a given sample size, and the (2) extent to which the resulting information will be useful for taking actions based on the evaluation findings.

How can evaluation theory help?

Michael Scriven calls our attention to the central importance of valuing in evaluation. He was the first to explain the *logic of evaluation*, which consists of identifying criteria, setting standards for these criteria, collecting data and comparing the results for each criterion to its standard, and synthesizing across the criteria to make a determination about the performance of the evaluand (2015). This logic is articulated within his *Key Evaluation Checklist* and takes a starring role in *Evaluation methodology basics: The nuts and bolts of sound evaluation* by Jane Davidson (2005).

As explained by Christie and Alkin (2013), the approach used by Scriven tends to develop criteria of merit using an objective lens. In Scriven's approach the evaluator logically deduces the criteria for which the evaluand must perform well to be considered good. As he explains, "Bad is bad and good is good and it is the job of evaluators to decide which is which" (Scriven, 1986, p. 19). Other theorists, such as Jennifer Greene (2011), have developed evaluation approaches (Values-Engaged Educative Evaluation) that also stress the importance of establishing criteria (the topic of Step 5—Justify Conclusions). However, in Greene's approach there is a specific emphasis on understanding what various stakeholders value about the evaluand and ensuring these values are articulated in the criteria. The importance of making values explicit as criteria in an evaluation and the importance of including various perspectives is articulated well in the following quote from Greene et al. (2011),

> In evaluation, the criteria used for judging program quality are fundamental to evaluation practice, yet they are all too often assumed (commonly, in stated program goals and objectives) or remain implicit in the evaluation process. These criteria are in an important sense the heart of the evaluative enterprise; they distinguish evaluation from other forms of applied social inquiry. Criteria of quality directly engage cherished beliefs and values and are thus contested or at least legitimately open to multiple perspectives (Mark, Henry, & Julnes, 2000). For example, what constitutes a "good" mathematics education program is different for a mathematics teacher, a student interested in becoming an engineer, a student with artistic talents and sensibilities, a parent struggling just to keep her child in school, and a school board member deeply worried about low test scores in the district. (p. 47)

Step 5. Justify Conclusions

It is important for the Evaluation Planning Team to think ahead to how the data will be analyzed, what results would constitute good or poor performance on the indicators and any additional analyses, how indicator determinations can be made in the future, and who will be involved in interpreting results. Keep in mind, perspectives on appropriate analyses and judgments on performance are normative and value based. As such, the team should take time to critically reflect on their planned analyses to identify the underlying biases and take measures to address them. Additionally, early planning can help ensure that the right data are available to fully answer the evaluation questions. Early planning can also ensure adequate time to interpret the data analyses and to draw conclusions that are meaningful and useful.

Data Analysis

Part of the planning process is considering how the team that implements the evaluation might analyze the data collected. The Evaluation Planning Team should review drafts of any instruments developed to ensure that the data required to calculate the proposed indicators will be available. For instance, if the focus of an evaluation is on

health equity, the team will need to ensure that data are available that allow analysis by sub-categories. Depending upon the level of specificity and complexity of the evaluation, the team may decide to articulate how to calculate these indicators by developing a one- or two-page document that outlines items such as data sources, relevant definitions of terms, and specific calculations required. This process is often followed with respect to articulating indicators for the purpose of public health surveillance; templates that the Evaluation Planning Team may find helpful to adapt are available on the Council of State and Territorial Epidemiologists (CSTE) website (e.g., hospitalizations attributable to drugs with potential for abuse and dependence, human cases of West Nile Virus disease, asthma emergency department visits).

Analyses should be well articulated in advance of data collection. The Evaluation Planning Team should discuss what analytic method(s) the implementation team will need to use to respond to the evaluation questions. For instance, will they need to calculate descriptive statistics as well as inferential statistics? If the latter, which are most likely to be useful given the type of quantatitive data being collected (assuming all underlying assumptions hold)? Does the team want to encourage examination of outliers as potential information rich sources rather than as anomalies? What types of qualitative analysis may be most helpful to perform (e.g., content analysis, thematic analysis)? Where feasible and helpful, the Evaluation Planning Team could provide examples of tables or templates that specify the output for each type of planned analysis. If *a priori* codes are likely to be used as part of planned qualitative data analyses, the Evaluation Planning Team could even begin building out this list with associated definitions.

Establishing Standards

Another part of this step is establishing standards of performance to which the indicators identified in Step 4 can be compared or, at a minimum, articulating the process that will be undertaken at a later date to discuss what constitutes good or poor performance. Standards are often referred to as "performance benchmarks." They may include comparisons over time or comparisons to an alternative approach (e.g., no action or a different intervention). The purpose of establishing standards is to provide an indication of what constitutes success for each aspect of the evaluand examined. It is important to develop such standards with the evaluation stakeholders; doing so can facilitate the use of evaluation findings by (a) creating a shared vision of success, (b) making sure the interpretation of the evaluation findings is clear and transparent, and (c) adding credibility to the results.

A common way of explaining standards is to set a quantitative number that an indicator needs to exceed for the activity it represents to be considered successful. Building on the example provide in Step 4, a quantitative benchmark for the indicator "proportion of attendees who complete training" might be "More than 60% of attendees complete the training." At the end of the evaluation, if more than 60% of attendees finish the training, the stakeholders would regard this as good performance; anything less would suggest improvements are necessary.

Quantitative indicators, such as the one just presented, are important to consider. However, there are several nuances that the Evaluation Planning Team should keep in mind while crafting the plan. First, standards do not have to be quantitative in nature. In fact, sometimes they cannot be. It depends on what is appropriate from a data collection standpoint. For example, standards might need indicate the types of themes coming up in interviews as indicative of good or poor performance. Second, quantitative standards do not have to be a specific number, they may instead capture a range of what is considered good, acceptable, or poor (or another qualitative label that is most appropriate to the situation). Third, whether they are articulated as formal indicators or not, what constitutes success with respect to the results of inferential tests should also be considered by the Evaluation Planning Team and other stakeholders. For instance, if a t-test is statistically significant, what magnitude in the difference between means will be considered meaningful? How large does this difference have to be to constitute success? Finally, the Evaluation Planning Team may feel that not enough existing knowledge or experience is available to set a standard for a given indicator. Our recommendation in this case is not to force a standard to fit. Setting standards that are irrelevant can be unhelpful, and in some cases (depending upon the stakes of the evaluation) even harmful. In this case, the Evaluation Planning Team could instead describe the process that should be undertaken with

stakeholders during the implementation phase to continue understanding what constitutes success and how they will collectively assign value to the evaluation findings.

The Evaluation Planning Team may also find it helpful to view the indicators as a suite. This may be particularly true if several indicators are used for a single criterion of merit. In this case, the team may consider how to decide if performance on the criterion is good, fair, or poor, and to do so they may need to roll-up across several indicators. As part of these discussions the Evaluation Planning Team should set aside time to consider the relative importance of indicators: Are the indicators of equal importance, or should some be weighed more than others? Are there one or more indicators where a certain level of performance must be achieved otherwise the evaluand for the criterion of interest to be considered successful? (Scriven, 1981; Davidson, 2005)

Synthesis and Interpretation

The Evaluation Planning Team must include sufficient time in the evaluation plan for the implementation team to work with program stakeholders to synthesize and interpret findings. There is little utility in articulating and implementing a high-quality data collection and analysis strategy if insufficient time is available to make meaning of the findings. As the team develops the timeline for the evaluation, they should check and double check that an appropriate amount of time remains for this task. Discussing the following questions and adjusting the evaluation plan may strengthen this aspect of the evaluation process:

- What proportion of the timeline is dedicated to synthesis and interpretation (this should be in addition to report writing or producing other products discussed in Step 6)? How does this compare to the amount of time allocated to other tasks?
- Has sufficient time been allocated to examining the results of data analysis with stakeholders to formulate interpretations and to consider the parameters within which recommendations will need to be developed to ensure their feasibility? Might you need to meet with stakeholders more than once for this task? Are there likely to be any competing items on stakeholders' schedules?

The Evaluation Planning Team should consider what adjustments can be made to the proposed evaluation timeline to allow for adequate time in interpreting the findings. Perhaps another step in the evaluation process can be shortened through facilitating more efficient processes or perhaps there is flexibility in the endpoint for the evaluation enabling an extension of the timeline.

How can evaluation theory help?

The evaluation approach espoused by Scriven (2015) and Davidson (2005) help us to understand the general process included in the logic of evaluation. The approach also describes the variants to the process. These variants include such items as

- Any minimum performance levels that if not attained would automatically indicate poor performance, the "*absolute* minimum acceptable performance standards on particular value dimensions" (Scriven, 2015, p. 19).
- Whether performance on a specific dimension needs to be weighed, "…the weights, i.e., the relative or absolute (depending on the context) *importance* of the dimensions of merit (or worth or significance). They are usually best limited to 1–3 or 1–4, or, better, not definitionally tied to numbers at all" (Scriven, 2015, p. 22).
- Examining the data by subgroups, especially with respect to those who are marginalized or for whom the outcome or process may disproportionately effect.

Greene et al. (2011) and House and Howe (2000) call our attention to the role of conversations among stakeholders in making implicit values explicit, particularly in Step 4 when defining indicators and in Step 5 when considering what constitutes success, and why, from multiple vantage points. Facilitating meaningful conversations about values is integral to evaluation practice, and both sets of authors help us to better understand this process through their approaches of Values-Engaged Educative Evaluation and Deliberative Democratic Evaluation in which there is ongoing, regular engagement with a wide array of stakeholders who hold diverse perspectives. Such conversations allow the stakeholders' values to surface, and provides space for them to reflect on, and in some cases challenge, underlying value constructions.

Step 6. Ensure Use of Evaluation Findings and Share Lessons Learned

Research examining evaluation practice indicates that facilitating evaluation use begins with evaluation planning (Johnson et al., 2009). An important part of the evaluation process involves engaging evaluation stakeholders throughout the evaluation planning process. This is a topic that we have weaved throughout the previous five steps and in the strategic evaluation planning process. For instance, engaging stakeholders in defining or refining the evaluation questions helps to ensure that the findings from the evaluation meet the stakeholders' information needs; discussing what evidence stakeholders find credible in terms of evaluation designs, types of data, and indicators that align with what they value lays a solid foundation for delivering insights that will be meaningful; and discussing the timeline of data collection, analysis, and synthesis and interpretation affords an opportunity to consider *when* the evaluation insights are needed for important upcoming decisions.

Another important aspect of facilitating use relates to sharing lessons learned, or, as we refer to it here, communicating and reporting. Such communication should not be an afterthought. Rather, the Evaluation Planning Team should spend a significant amount of time during the planning phase considering whom the Evaluation Implementation Team will need to communicate with as well as the purpose, format, frequency, and timing of each message (Russ-Eft & Preskill, 2009). The results of these conversations can be summarized and described in a table such as **Table 4.3**. This table is similar to one used in the Strategic Evaluation Plan (**Chapter 3**), however it includes communication that occurs during the actual implementation of each evaluation.

The "How" of Practice
Cultural Responsiveness

The CDC's *Practical Strategies for Culturally Competent Evaluation* (2014, p. 20) offers guiding questions to facilitate use and share lessons learned:

- Are communication mechanisms culturally appropriate?
- Does the reporting method meet stakeholder needs (both the message and the messenger)?
- Are the data presented in context, with efforts made to clarify issues and prevent misuse?
- Has the community benefited as anticipated? How?
- How has cultural responsiveness increased both the truthfulness and utility of the results?
- Do the action plans draw on community strengths and capacity? Are the action plans consistent with the purpose of the evaluation?

Additional information on these strategies can be found at https://www.cdc.gov/asthma/program_eval/other_resources.htm

Table 4.3. Sample Communications Strategy for Individual Evaluation Plan (partially completed) *

Audience 1 – Program Implementers			
Purpose	**Possible Formats**	**Timing**	**Notes**
Keep informed about implementation milestones	Email	Quarterly	Request email distribution list from program manager
Request assistance in soliciting participation from program attendees to complete upcoming online survey	Staff newsletter	January edition	
Help with interpreting findings and possible implications	Virtual sense-making sessions	3 months prior to end of evaluation	
Document and share findings	Final written report Tailored fact sheet *-Day-to-day implications*	1 month prior to end of evaluation	Busy time of year for this audience; ensure multiple touchpoints to get the message out, especially regarding fact sheet
Foster use	Action planning sessions	1-2 weeks following receipt of written report and fact sheet	Plan session with assumption that attendees have not had sufficient time to review the report or fact sheet
Audience 2 – Community representatives			
Help with interpreting findings and possible implications	Virtual and in-person sensemaking sessions	3 months prior to end of evaluation	Solicit suggestions in advance for convenient locations to host in-person sessions

*Table adapted from Russ-Eft & Preskill, 2009, p. 407-411

In developing the communications strategy depicted in **Table 4.3**, it is important to recognize what fits into the realm of possible topics. It is natural to put our attention on how findings from an evaluation will be communicated, and this is important. However, it is also important to make sure that stakeholders are informed about the evaluation's progress (Patton, 2013). To that end, the Evaluation Planning Team should also have in their purview the types of communications that will be necessary *during* the evaluation implementation to keep stakeholders engaged and to share interim findings. With this in mind, we provide details about several other aspects of the communications process that Evaluation Planning Team members should discuss in crafting a communications strategy that facilitates use.

Identifying Audiences

You may recall that in Step 1—Engage Stakeholders, we discussed three different types of stakeholder groups: primary, secondary, and tertiary. In Step 6, the Evaluation Planning Team should revisit these groups; each can be considered a potential audience for evaluation communications.

- ***Primary stakeholders*** are the individuals who have the ability and authority to use evaluation findings to alter the course of the evaluand. This group frequently includes program staff members, managers, or the evaluation sponsors. These will often be the primary audiences for the evaluation, and when resources are tight may be the only group of stakeholders that the Evaluation Planning Team feels can be considered in the communications strategy.

- ***Secondary stakeholders*** are individuals who are likely to be affected by any changes made because of the evaluation findings, such as participants of a public health program. Such stakeholders may be the audience for communications informing them about upcoming data collection efforts in which their involvement may be solicited or in communications sharing findings from the evaluation, perhaps even including what programmatic changes are on the horizon as a result.

- ***Tertiary stakeholders*** are individuals who are not directly affected by the changes that result from acting on the evaluation findings but may have a general interest in the results, such as similar public health programs in other jurisdictions. Though communications with this audience may seem to be a luxury, relatively simple communications strategies could be considered to share relevant findings with this group such as presentations at upcoming conferences in which staff members normally participate. Sharing findings with tertiary stakeholders can aid in collective learning about a field.

When thinking about these different audiences, remember that they are likely to need different types of information and may prefer specific formats for messages. For this reason, the Evaluation Planning Team should carefully consider the messages and formats for each audience and describe these choices in the communication strategy section of the evaluation plan.

Timing of Communication

As noted earlier, the Evaluation Planning Team will need to consider communication that will occur *during* and toward the *end* of the evaluation when formulating the communications and reporting strategy. Consider the following when planning communications during evaluation implementation: (1) Times during the evaluation when you need help from stakeholders to implement the evaluation successfully, (2) Who needs to be kept in the loop of the implementation process so that important buy-in is not lost when the evaluation findings arrive, (3) Opportunities for sharing interim results.

It is not uncommon to need assistance from stakeholders during the evaluation implementation process with contacting potential respondents. For instance, you may need help from program staff members or managers to secure a list of potential interviewees or survey participants, or you may find it helpful to have them send out an advance email to notify individuals about the importance of participating in an upcoming data collection effort. The Evaluation Planning Team can include within the communications strategy the dissemination of an email to staff members or managers that informs them of the status of the evaluation implementation and notifies them about an upcoming need to collaborate on data collection efforts. Such information is critical to effective implementation of the evaluation. Including stakeholders in the evaluation plan can help ensure these important details are not forgotten.

The Evaluation Planning Team should consider who needs to be kept in the loop regarding the status of evaluation implementation efforts. Communicating with primary stakeholders throughout the implementation process, about items such as how activities are rolling out, early successes in implementation, and upcoming events, ensures that they remain engaged. Delivering an evaluation plan to primary stakeholders and waiting to communicate with them until the evaluation findings are available can seriously reduce feelings of buy-in. In addition, waiting can cause individuals to forget that an evaluation is underway that may be helpful to future decision making. All these factors combined reduce the likelihood that findings will be used.

Finally, for evaluation findings to be useful, they must be communicated in a timely manner and in alignment with relevant decision-making processes. Sharing interim findings at strategic points keeps stakeholders engaged in the evaluation. Interim findings may bring to light an implementation problem, such as the need to recruit a certain population that is currently being overlooked, allowing time to modify the recruitment approach. Interim findings may prove to be useful in decision making as well. It is not uncommon for a set of data collection

efforts (e.g., surveys or interviews) implemented early in the evaluation to provide the sole insights for answering one or two evaluation questions posed in an evaluation plan. Sharing these findings prior to the end of the evaluation has the potential to lend useful insights for decision making.

The "How" of Practice
Interpersonal Competence

Communicating interim findings can be tricky, particularly when they are likely to be perceived as negative by some stakeholders. Interpersonal skills, such as empathy, active listening, and motivation, can help evaluators to not only navigate this tricky terrain but also reframe negative findings as opportunities for learning and improvement.

The decision to share interim findings needs to be weighed against the associated risks of doing so. In many cases, interim findings do not provide a complete picture and may be more or less representative of certain groups or aspects of the evaluand. As such, evaluators must be transparent about the inherent limitations and biases. Stakeholders' reactions to interim findings can also bias their perspectives on subsequent findings or the evaluation, particularly if the findings appear consistent or in conflict with their interests and expectations. Evaluators should anticipate stakeholder reactions and take action to avoid any pitfalls.

The key is to think strategically and lay out plans for effectively communicating with the various audiences at appropriate intervals. Surprises at the end of the evaluation are never a good thing. You might find that the optimal time to communicate key or interim findings is during routine functions, such as at quarterly staff meetings or an annual retreat for decision makers. Remember that the more engaged you keep your audiences, the more ownership they have of the process and, consequently, the more likely it is that they will use the findings.

Format of Communications

The most common way to share evaluation findings is through a formal, written evaluation report. Though these reports are often required by funders, a strong recognition exists in the field of evaluation that even though written reports may be a necessary evil, they may actually *be* evil when the ultimate desire is to facilitate the use of the information contained within them. As such, much has been written about other options for sharing findings with various audiences. In considering which communication formats to use, the Evaluation Planning Team should consider (1) which is most likely to result in achieving the intended purpose, (2) what format is best suited for the audience in question (e.g., How do they prefer to receive communications? Are there specific norms within their profession, organization, or community around communications?), and (3) the level of resources (time and funds) to support communications. **Figure 4.2** depicts several possible options for communication formats depending upon the level of interaction that may be needed or is most appropriate for the audience.

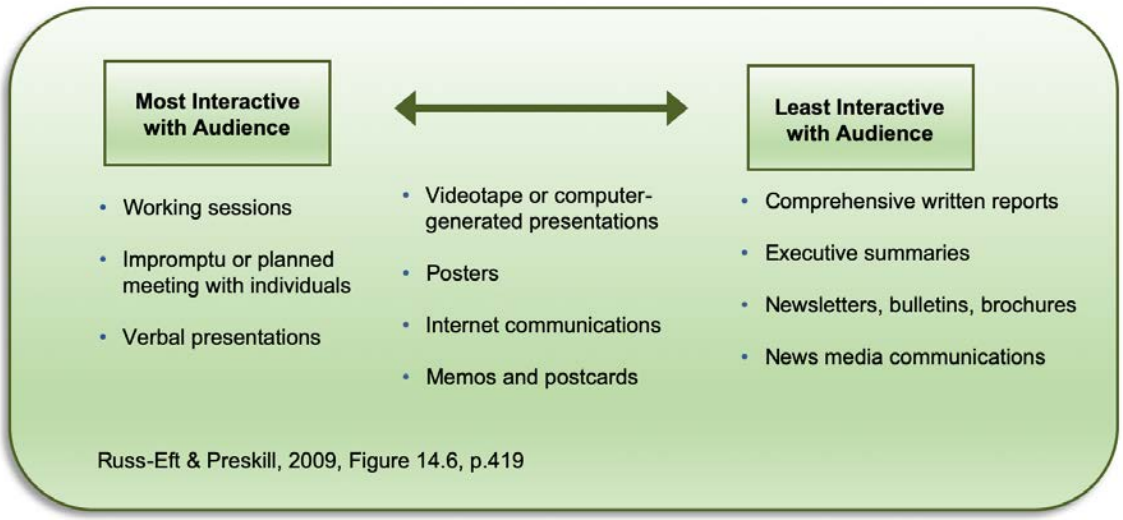


Figure 4.2. Communication Formats

Each communication format has its strengths and limitations. Our suggestion is for the Evaluation Planning Team to consider each one fully with respect to the intended audience and to engage with the audience to learn more about their communication preferences and norms. Following is a description of several communication formats to inform Evaluation Planning Team conversations.

- ***Evaluation reports.*** These reports range in size, formatting, and structure. Often, they are requested from funders so that sufficient documentation is available later for how the evaluation was performed. These can serve as effective legacy documents for evaluations, but due to their comprehensiveness and static nature they often serve a limited purpose. It is often important to couple evaluation reports with other formats of communications. Evaluation reports can potentially be improved by using one rule of thumb shared by the Canadian Health Services Research Foundation (BetterEvaluation, 2014): the 1:3:25 formula. The calls for one page of main messages, a three-page executive summary, and no more than 25 pages presenting the findings.
- ***Executive summaries.*** These short, concise summaries of evaluation reports are usually presented at the beginning of evaluation reports so that individuals who have limited time to review the full report get the important highlights quickly. Though they are typically part of the evaluation report, they can also be disseminated separately from the full report, offering another format for communicating with key audiences.
- ***Newsletters.*** Newsletters provide a format through which evaluation findings can be shared with a broad group of individuals. Often evaluation communications do not need to fill the entire newsletter; consider including important messages about evaluation implementation (e.g., upcoming activities that the audience may need to be aware of) or findings in a section of an existing newsletter that is a common mode of communication for the intended audience.
- ***Fact sheets or briefs.*** These documents are typically brief, one to two pages in length, and can easily be tailored to a specific audience. Such communications will often summarize what the key take home messages are for this audience, why the findings apply to them, and what actions they might take based upon the findings to make a positive difference.
- ***Infographics.*** Infographics can provide a short, at-a-glance, depiction of the key evaluation findings and recommendations. When designed well, audiences can readily see the highlights from the evaluation. As their name implies, infographics provide visual images that communicate findings quickly and simply, allowing the reader to skim insights and find which ones are most interesting and relevant to them (Hutchinson, 2017).
- ***Social media***. Social media is a typical part of daily life, and the Evaluation Planning Team should consider how, or if, to leverage such communication formats to share insights about the evaluation implementation or findings. Facebook, YouTube, and Twitter are platforms that can increase the timely dissemination and potential impact of evaluation; leverage audience networks to facilitate information sharing; expand reach to include broader, more diverse audiences; and facilitate interactive communication, connection, and public engagement. Such platforms enable the easy dissemination of short messages that key audiences can easily consume and the ability to share insights through alternative modes of communication (e.g., videos).
- ***Working sessions.*** The power of conversation should never be underestimated in evaluation. Working sessions can take many forms, from virtual (e.g., using online conferencing platforms) to in-person round-table discussions. Such forums may be especially helpful for developing action plans (see **Chapter 5**) from evaluation findings and recommendations.

The Evaluation Planning Team should be creative in thinking through possible communication formats. So many options are available, from short write-ups in trade journals to poems and skits. The team should also

consider a range of products; Hutchinson (2017) recommends layering a variety of communication formats. A layering approach gives stakeholders the option to go as deep into the findings as they need to or desire. For instance, stakeholders who want to fully understand the evaluation process and findings may seek out the evaluation report, while others may be satisfied with a newsletter or infographic. It is the team's job to consider which message formats are the most appropriate fit for the audience and purpose.

How can evaluation theory help?

Several evaluation theories provide prescriptions about how to facilitate the use of evaluation findings by stakeholders. These include Michael Patton, Bradley Cousins, Hallie Preskill, and Jean King. The majority of these theorists provide insights about communication and reporting. For instance, in his 17-step checklist for conducting Utilization-Focused Evaluation, Patton (2013) calls our attention to the importance of fostering communication with stakeholders throughout the evaluation planning and implementation process. He also encourages evaluators to be flexible with communications and reporting. Though we may have a strategy in place, Patton suggests that we "be alert to unanticipated pathways of influence that emerge as use and dissemination methods unfold" (p. 18). Preskill has also written extensively on communications and reporting (Russ-Eft & Preskill, 2009; Torres et al., 2005) and many of her suggestions are reflected in the current chapter.

Managing the Evaluation

In putting the final touches on the individual evaluation plan, it is important to think about implementation logistics. The outline for the individual evaluation plan at the end of this chapter includes suggestions for elements to include with respect to managing the evaluation. At this stage, consider who will comprise the Evaluation Implementation Team as well as who will hold the responsibility for each major task in the evaluation.

Circling Back: Revisiting the Strategic Evaluation Plan

Once one or more individual evaluation plans have been created, it may be necessary to revisit the Strategic Evaluation Plan with the Strategic Evaluation Planning Team. It is most likely that the Strategic Evaluation Plan will be drafted and then one or two individual evaluation plans, rather than all that have been proposed, will be created and carried out. Lessons learned from implementing the individual evaluation plans can be used to adjust the Strategic Evaluation Plan. Ultimately, the lessons learned may suggest that more or fewer evaluations can be conducted over time or that a different sequence of evaluations is necessary. It is important to remain flexible with the Strategic Evaluation Plan and to formally include a mechanism for annually updating it.

Chapter Summary

In this chapter we walked you through the process of engaging with an Evaluation Planning Team to develop individual evaluation plans. At the end of this process, the Evaluation Planning Team will have produced a solid plan of action covering everything from the purpose of the evaluation to how the findings will be disseminated. The next step is to put this plan into action, and to be ready, or as ready as one can be, to adjust these plans as needed throughout the implementation phase. This is the focus of **Chapter 5**.

Review Questions and Skill-building Exercise

Review Questions

1. In what ways does the membership of the Evaluation Planning Team differ from that of the Strategic Evaluation Planning Team? What are some considerations to reflect upon when comprising a list of individuals to invite to the Evaluation Planning Team?
2. In **Chapter 2** we discussed the importance of interpersonal competence in evaluation. Review Steps 1–6 in this chapter and consider how interpersonal competence contributes to successfully carrying out each of these steps.
3. What are some formats for communicating evaluation results beyond a written, formal evaluation report? Can you think of other options beyond those we covered in this chapter? If so, what comes to mind?

Skill-building Exercise

Identifying and refining evaluation questions is an important aspect of *focusing the design* of an evaluation. Review the evaluation questions in the table that follows and use the Good Evaluation Questions Checklist from CDC and the tips we provided in Step 3 of this chapter to improve the question wording. Remember, you want to have evaluation questions that are focused enough that they are answerable, but broad enough to allow room for inquiry that produces insightful findings. Also remember that in practice, you will work with your stakeholders through this process.

Example evaluation question	Refined evaluation question
Did we uphold the fidelity of the intervention?	
Are partners satisfied with the coalition?	
What is the effect of our intervention?	
Did the intervention cause the outcomes of interest?	

Individual Evaluation Plan Outline

1. Introduction and Stakeholder Engagement

This section provides information about the purpose of the evaluation and identifies stakeholders who are, or need to be, involved in the evaluation.

Evaluation Purpose

- What is the purpose of this evaluation?
- How do you anticipate the findings of this evaluation will be used? By whom?
- How does this fit with the overall Strategic Evaluation Plan for the program?

Stakeholders

- What individuals and groups have an interest in the outcomes of this evaluation? Examples include program participants, staff members, and critics.
- What aspect of the evaluation are they most interested in? For example, are they interested in the evaluation from a cost angle, effectiveness of the program, possible improvements, etc.?
- What role did they play, or what role will they play, in developing or implementing this evaluation plan?

Examples include serving on the planning team, acting as an external reviewer, collecting data, interpreting findings, or using results.

Table IEP.1 Stakeholder Assessment and Engagement Plan

Stakeholder Name	Stakeholder Category	Interest or Perspective	Role in the Evaluation
{May be an individual or group}	{Primary, secondary, tertiary}	{Program participant, staff member, etc.}	{Evaluation Planning Team, external reviewer, etc.}

Culturally Responsive Evaluation

- How will the implementation team engage stakeholders who reflect the diversity of those who may be affected by the evaluation's findings? For suggestions, see CDC's Practical Strategies for Culturally Competent Evaluation at https://www.cdc.gov/asthma/program_eval/other_resources.htm
- How will team ensure that you successfully elicit and incorporate the various perspectives?
- How will team account for the influence of context and culture in the evaluation design, implementation, and reporting?

2. Description Of {Insert Name of Evaluand}

This section provides detailed information about what you are evaluating (for example, the program's strategies, processes, policies). For ease of reference, we use the term program below to refer to what you are evaluating, though you may be evaluating something other than a program.

In this section, describe the need for the program, its context, intended audience, and stage of development. You will also provide information about its inputs, activities, outputs, and outcomes and will develop a logic mode or similar visual depiction. In the narrative portion, include information that might not be obvious when using the shorthand of the logic model.

Need

- What need is the program designed to meet?

Context

- What is the program's context? That is, what contextual or cultural factors may affect its implementation or effectiveness?

Population Addressed

- Who is included in the population for whom activities are intended?

Stage of Development

- How long has the program been in place?
- Is it in the planning or implementation stage?

Resources/Inputs

- What resources are available to support the program (e.g., personnel, money, space, time, partnerships, technology)?

Activities
- What specific activities are conducted (or planned) to achieve the program's outcomes?

Outputs
- What do the activities produce (e.g., materials, services delivered)?

Outcomes
- What are the program's intended outcomes? (Intended outcomes may be short-term, intermediate, or long-term and are changes that occur in something outside of your program—those within your sphere of influence.)
- What does the program ultimately propose to change with its activities (long-term outcomes)?
- What occurs between the activities and the point at which these ultimate outcomes (short-term and intermediate outcomes) occur?

Organizing information about the program in a table can be a useful first step in creating a logic model. You may choose to use only a table; however, developing a diagram that includes boxes and arrows will provide a better sense of the important pathways the program intends to set into motion through the activities performed. You may find such detail more helpful in communicating with program stakeholders about the program, designing the evaluation, and understanding opportunities for using the evaluative insights.

Table IEP.2 Program Description Template

Resources/ Inputs	Activities		Outputs	Outcomes	
	Initial	Subsequent		Short-Term/ Intermediate	Long-Term

Logic Model
- Provide a logic model for the program.

3. Evaluation Design

This section describes the evaluation design. Provide information about stakeholder information needs, the evaluation questions, and the evaluation design that the implementation team will use to answer those questions.

Stakeholder Needs
- Whom will use the evaluation findings?
- What do they need to learn from the evaluation?
- What do intended users view as credible information? How will they likely use the findings?
- What evaluation capacity will need to be built to engage these stakeholders throughout the evaluation?

Evaluation Questions
- What three to five major questions do are to be answered through this evaluation?
- Do the questions align with the Good Evaluation Questions Checklist?

Evaluation Design

- What is the design for this evaluation? (e.g., experimental, pre-post with comparison group, time-series, case study, post-test only)
- What is the rationale for using this design?

4. Gather Credible Evidence

This section describes how the implementation team will gather data for the evaluation. Provide information on methods that the team will use to compile data and how those methods relate to the evaluation questions.

Data Collection Methods

- Will the implementation team collect new data to answer the evaluation questions or will secondary data be used?
- What methods will the team use to collect or acquire the data?
- Will there be a sample? If so, how will it be selected?
- How will the team identify or create data collection instruments?
- How will the team test instruments for readability, reliability, validity, and cultural responsiveness?
- How will the team determine the quality and utility of existing data?
- From whom or from what will the implementation team collect data (source of data)?

Table IEP.3 Evaluation Questions and Associated Data Collection Methods

Evaluation Question	Data Collection Method	Source of Data
1.	1.	
	2.	
2.	1.	
	2.	

5. Data Analysis and Interpretation

In this section, provide information on the indicators and standards that the implementation team can use to judge the success of the program (or policy, etc.); how the team will analyze the evaluation data; and how they will interpret and justify conclusions.

Indicators and Standards

- What are some measurable or observable elements that can serve as markers of the program's performance? What are the criteria of merit and associated indicators?
- What constitutes success on the indicators? That is, to what standards will the implementation team and stakeholders compare the evaluation findings? Alternatively, what process will the team engage in to understand and interpret performance on this indicator?

Table IEP.4 Questions, Criteria, Indicators, and Standards of Success

Evaluation Question	Criteria and Associated Indicator(s)	Standards (What Constitutes Success?)
1.	1.	
	2.	
2.	1.	
	2.	

Analysis

- What method(s) will the team use to analyze the data (e.g., descriptive statistics, inferential statistics, qualitative analysis (i.e., content analysis, thematic analysis, etc.))?
- Provide example table shells, templates, or qualitative codebook that specifies the output for each type of analysis planned.

Interpretation

- Who will be involved in drawing, interpreting, and justifying conclusions? Does this group include program participants or others affected by the program?
- What are the plans, including evaluation capacity building activities, to involve them in this process?

6. Use and Communicate Findings

This section provides information about how information from the individual evaluation planning process and results will be used and shared.

Use

- How will the evaluation findings be used? By whom?
- How does the timeline for reporting findings and potential recommendations align with key events for which information from the evaluation will be needed (e.g., grant application, partner meeting)?
- Who is responsible for creating and monitoring an action plan to guide the implementation of evaluation recommendations? What follow up is needed?
- What lessons learned, including those about evaluation and evaluation capacity building, should be shared? How will they be documented?

Communication

- Which evaluation stakeholders will the implementation team communicate with and for what purpose (e.g., update on status of evaluation, invite to meetings, share interim or final findings)?
- What methods (e.g., in-person meetings, emails, written reports, newsletter article, presentations) will the team use to communicate with evaluation stakeholders?
- Who is best suited to deliver the information (e.g., evaluator, program manager, coalition leader)? Why are these methods appropriate for the specific evaluation stakeholder audience?

7. Evaluation Management

This section provides information about how the evaluation will be managed and implemented and includes a timeline for evaluation activities. You may find that some of the tables suggested here fit better in other sections

of your plan. Regardless of how you structure your plan, carefully think about each of these implementation steps, noting the individual(s) responsible and deadlines for each task.

Evaluation Implementation Team

- Who will manage and implement this evaluation?
- What evaluation skills or approaches are needed to successfully conduct this evaluation?
- At what point(s) will the team pause to examine the extent to which they are upholding the evaluation standards?
- What strategies will the team use to foster a critically reflective perspective throughout the evaluation?
- Has the team identified an external reviewer to provide feedback on the evaluation plan?

Table IEP.5 Roles and Responsibilities of the Evaluation Implementation Team Members

Individual	Title or Role	Responsibilities

Data Collection Management

- What data will be collected? {From **Table IEP.3**}
- What activities are needed to carry out the data collection successfully? When should each of these activities be completed?
- Who is responsible for conducting each activity?
- Who will oversee data collection to ensure appropriate implementation?

Table IEP.6 Data Collection Plan

Evaluation Question	Data Collection Method	Activities Needed	Person(s) Responsible	Due Date
1.	1.	1.		
		2.		
	2.	1.		
		2.		
2.	1.	1.		
		2.		
	2.	1.		
		2.		

Data Analysis and Interpretation

- How will the team ensure the security of the data?
- What data will be analyzed, how, and when?
- Who is responsible for conducting the analyses?
- How will the team engage stakeholders in confirming analysis results and interpreting them?

Table IEP.7 Data Analysis Plan

Analysis to be Performed	Data to be Analyzed	Person(s) Responsible	Due Date

Communication and Report Management

- Who are the audiences for reporting the progress made on the evaluation or evaluation findings?
- What is the purpose of communications with this audience?
- What is the most appropriate type of communication method to use with this audience? Who is the most suitable messenger?
- When will the communication take place?

Table IEP.8 Communication and Reporting Plan *

Audience 1: {insert name of audience}					
Applicable? (√)	**Purpose of Communication**	**Possible Formats**	**Possible Messenger**	**Timing/ Dates**	**Notes**
	Include in decision making about any changes to evaluation design or activities.				
	Inform about specific upcoming evaluation activities.				
	Keep informed about progress of the evaluation.				
	Present initial or interim findings.				
	Present final findings.				
	Document the evaluation and its findings.				
	Document implementation of actions taken because of the evaluation.				

*Adapted from Russ-Eft & Preskill, 2009, p. 407–411.

Timeline

- How much time is needed to gain the requried clearances and approvals (e.g., Institutional Review Board)?
- When will planning and administrative tasks occur? When will training for data collectors occur?
- When will the team pilot test data collection instruments?
- When will formal data collection, analysis, and interpretation tasks occur? When will information dissemination tasks occur?
- Upon mapping all of the above on a single timeline, are there any foreseeable bottlenecks or sequencing issues?

Evaluation Budget

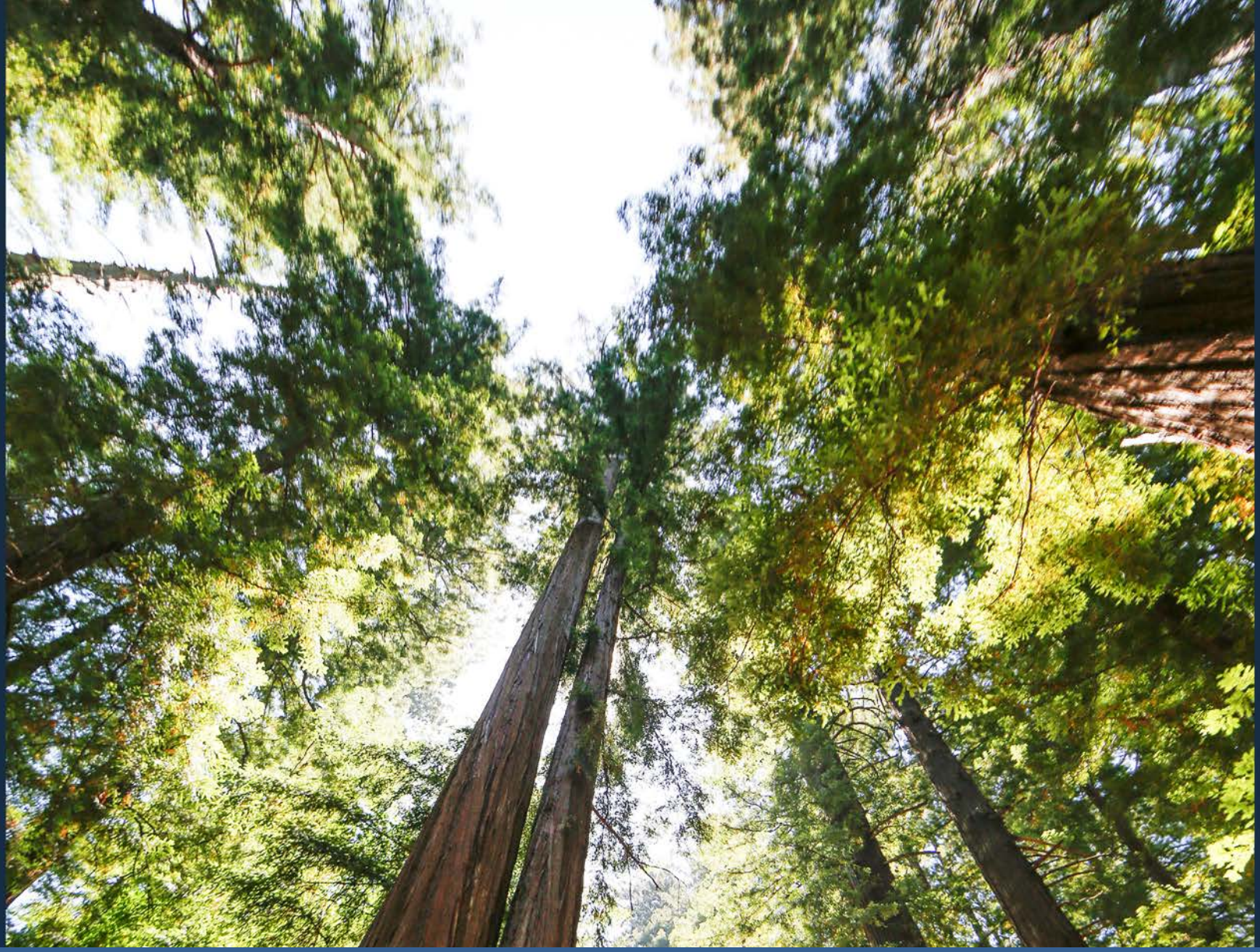
- What is the anticipated cost for this evaluation?
- Where will the funding come from to support the evaluation?
- Are any in-kind, volunteer, or partner resources being contributed?

8. Wrapping Up

- At the end of the evaluation, how will you acknowledge the contributions of Evaluation Planning Team members and others who contributed to the successful implementation of the plan?
- How will the team reflect on and document evaluation lessons learned in the course of implementing the evaluation?
- How and where will the team archive relevant documents, instruments, and data?

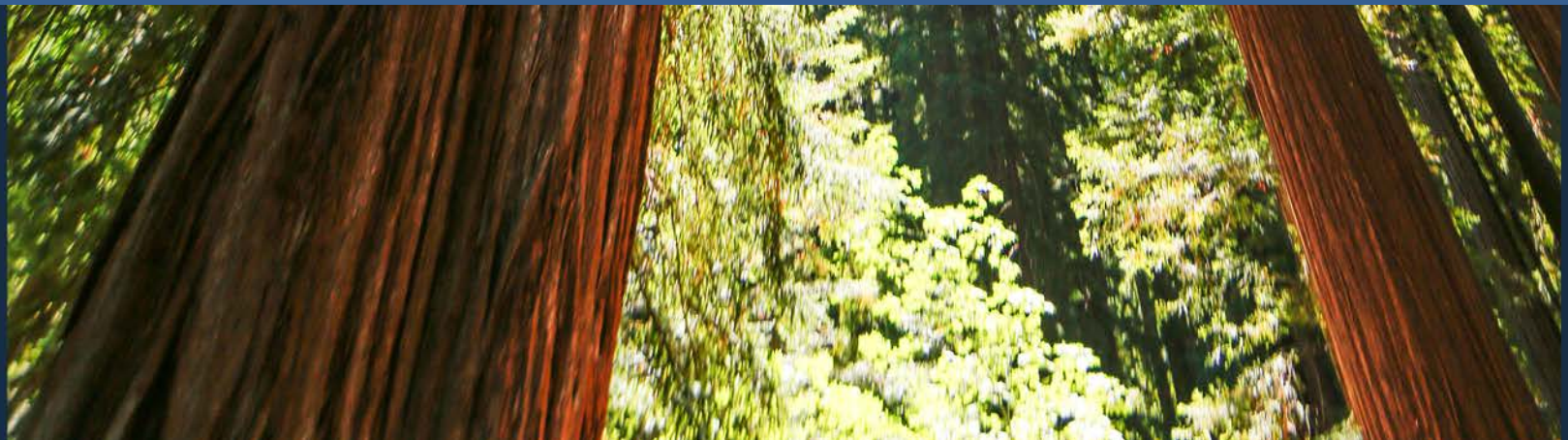
This evaluation plan template can also serve as a tool to document evaluation implementation (as required by the evaluation accountability standard) and can also provide information to internal or external people conducting meta-evaluations. Inserting the following after each section may help with this process.

___ Evaluation was implemented as planned
___ Changes made to the plan(describe changes as well as the rationale for changes)



CHAPTER FIVE

Where the Rubber Meets the Road: Implementing the Evaluation



CHAPTER FIVE: Where the Rubber Meets the Road: Implementing the Evaluation

Introduction

In **Chapters 3** and **4** we emphasized the importance of evaluation planning and introduced tools for planning at both the macro level (strategic evaluation planning) and the micro level (individual evaluation planning). In this chapter, we provide information to help you transition from the planning process to implementation. We also share tips and tools for maintaining the flexibility that is often necessary when the rubber of evaluation planning meets the road of evaluation implementation.

The chapter begins by providing insights that will help you identify and secure involvement from the team, or individual, who will implement the evaluation. Following these insights is a section on project management in evaluation, which includes some helpful tools to facilitate a smooth and successful implementation. Next, we will address the challenging topic of evaluation budget development and provide guidance on developing more accurate and comprehensive budgets. We then provide insights into the many types of challenges that evaluation teams can encounter when implementing an evaluation, along with ways to prepare for and mitigate such issues. Lastly, we describe how to close the loop on the evaluation process to maximize not only learning, but the use of evaluation findings through action planning

By the end of Chapter Five, the reader will be able to

- Establish an Evaluation Implementation Team with the collective knowledge and abilities to implement an evaluation.
- Use several tools for managing evaluation implementation.
- Identify several challenges that can emerge during evaluation implementation, and one or more actions that can be taken to address each challenge.
- Develop an action plan to make changes based upon evaluation findings.

Selecting an Evaluation Implementation Team

In crafting a management plan for the evaluation (see **Chapter 4**), the Evaluation Planning Team will have discussed potential members for the Evaluation Implementation Team. The individuals with direct responsibility for implementing the evaluation need to be selected carefully, as they have a critical role to play in obtaining access to existing data, overseeing new data collection, analyzing the data, synthesizing the findings, preparing the evaluation report or other communication materials, working with stakeholders to create a plan of action once the evaluation is complete, documenting any deviations from the evaluation plan, and gathering lessons learned to improve future evaluations. In some cases, these individuals will be the same people who served on the Evaluation Planning Team; however, they may represent only a subset of the Evaluation Planning Team or, be a different set of people entirely. In fact, it may be the case that you were serving as the evaluator for the planning functions, but someone else steps in to lead the implementation.

Project Management in Evaluation

Running an evaluation is much like running any other project. The things you worry about may be a little different for an evaluation than for other kinds of projects, but the good management practices that help elsewhere also work well with evaluation. Good management includes thinking ahead about what is most important, determining which activities precede other activities, who will accomplish what tasks, what agreements and clearances are needed, when important products are due, and how far the budget will stretch. It is also important to anticipate factors that could present problems down the road, something we discuss in greater detail later in the chapter. When potential challenges are identified, safeguards should be put in place to prevent them from disrupting the implementation process, with contingency plans in mind in case the evaluation does not go as planned.

In the remainder of this section on project management, we provide a few simple management tools that are helpful with evaluation implementation. These tools are meant to be illustrative; they are by no means the only tools that may be useful for managing an evaluation. It may be that you are aware of similar tools that work well for project management, or that one of the tools that we present here can be modified to suit your specific needs.

Evaluation Overview Statement

The Evaluation Planning Team should have included an evaluation purpose statement within the individual evaluation plan. An evaluation overview statement is similar to the purpose statement, but can be slightly longer, providing more detail about the evaluation. The evaluation planning and implementation teams can widely circulate this statement to a variety of audiences that have an interest (see the "Example Evaluation Overview Statement" box). Having a common statement readily available increases the likelihood that messages disseminated by partners are accurate and consistent. The format of the evaluation overview statement may vary, but basic elements include

- Name of the evaluation
- Name of the element or aspect of the program being evaluated
- Time period of the evaluation
- Overall goal of the evaluation
- Identified uses for the evaluation results
- Evaluation design and major data collection activities
- Intended audience(s) for the evaluation
- Special considerations for data collection or analysis

Example Evaluation Overview Statement

Evaluation of Asthma Self-Management Training

Evaluation Sponsor:
Evaluation Funding: XXX

The purpose of this 18-month evaluation is to determine whether participants' asthma self-management knowledge and skills increase as a result of asthma self-management training. Adults who obtain emergency department asthma care at a large urban hospital in [city] will be referred to the training program. Trained asthma educators will deliver the training in a small-group classroom setting. Trainings will be periodically monitored to ensure a standard curriculum and consistency across asthma educators. We anticipate training 810 adults in this program over a six-month period from March 20xx to August 20xx.

Data will be collected through self-administered questionnaires, which will be collected prior to and after the training. The questionnaire will ask about participant demographics, asthma self- management knowledge, asthma self-management skills, and intentions for changing behavior. Questionnaires will only have an ID number to protect patient privacy, and the hospital will not have access to survey results of individual patients. The analysis will focus on changes in knowledge, skills, and behavioral intention from pre- to post-training. We will also look at subgroups by race/ethnicity, age, and gender to understand whether the training was more beneficial for certain groups than others. The results will be used to determine whether to continue this training in the future, and if so, who should enroll.

Roles and Responsibilities Table

A "roles and responsibilities table" is a useful tool to ensure that all aspects of the evaluation are assigned to a specific individual or individuals, to reduce confusion about roles, and to gain agreement from everyone involved in the evaluation about who will do what. These tables can be created in standard word processing or spreadsheet software or with the assistance of project management software such as Slack, MS Project, or Smartsheet. The information typically included in a "roles and responsibility table" is presented in **Table 5.1**.

As shown in the template, column one should list everyone involved in implementing the evaluation, including the dissemination of findings. The second column lists the individual's role in the evaluation, this may or may not be the same as their job title or other program role. The "Responsibilities" column should be a brief but comprehensive bulleted list of what each team member will do throughout the evaluation period; this may include both evaluation and programmatic tasks (that relate to the evaluation). Responsibilities related to coordination or oversight along with direct involvement in evaluation tasks should be included.

Depending on the complexity of the evaluation, it may be useful to also include a "Tasks" column that explicitly lists the major evaluation activities an individual will be involved in (see **Table 5.2** for an example of specific tasks). The roles and responsibilities table should be updated as staff members, consultants, or partners join or leave the project. This helps ensure that no activity falls through the cracks as personnel change over time. The roles and responsibilities table can be linked to the evaluation timeline. Doing so will ensure that Evaluation Implementation Team members assigned to activities are available at the appropriate times.

Table 5.1. Roles and Responsibilities Table

Name	Role in Evaluation	Responsibilities	Tasks
S. Carter	Evaluation Leader	• Develop data collection instruments. • Obtain IRB clearance from hospital IRB. • Facilitate team reflections. • Assess and build needed evaluation capacity. • Train data collectors. • Monitor trainings on quarterly basis. • Oversee data analysis and interpretation. • Write up evaluation interim and final results. • Update stakeholders on evaluation progress. • Conduct briefings on findings with stakeholders. • Work with stakeholder group on action plan for use of results.	1-8
D. Gonzales	Data Analyst	• Purchase analytic software or upgrades. • Compile statistics on attendance and response rates throughout data collection. • Enter data from training forms. • Check data quality. • Conduct main analyses. • Analyze subgroup data. • Write up results.	5,6,8
E. Phillips K. Franz M. Smith	Educator & Data Collector	• Attend data collector training. • Conduct self-management trainings. • Conduct pre- and post-data collection.	3,4
R. McEwan N. Sanders R. Chan J. Francis	Evaluation Planning Team	• Review and suggest changes to instruments. • Provide feedback on interim and final reports. • Participate in action planning for program improvement.	1,6,7

Table 5.2. Task Table

Task Number	Task Title
Task 1	Data Collection Questionnaire Development
Task 2	Obtaining Necessary Clearances
Task 3	Data Collector Training
Task 4	Data Collection and Monitoring
Task 5	Data Management and Analysis
Task 6	Communicating Evaluation Progress and Findings or Reporting Results
Task 7	Action Plans for Program Improvement
Task 8	Evaluation Capacity Building

Evaluation Project Timeline

A timeline is a critical management tool. The implementation team can document when evaluation activities should occur and reference the timeline regularly to track whether activities are going as planned, or if they are behind schedule. Timelines can also serve as a useful communication tool for stakeholders who are interested in or may benefit from a high-level overview of the evaluation activities. When developing the timeline, it is wise to consider known resource constraints, namely financial constraints, staffing limitations, the timeline for other concurrent evaluations, or program activities in which the implementation team may be involved. Will the right staff members be available when they are needed for the evaluation? There are several different types of timelines that may be useful for documenting the proposed trajectory of evaluation activities.

The "How" of Practice
Situational Awareness

When scoping a timeline, evaluators should be mindful that the evaluation context will undoubtedly change throughout the evaluation implementation process. This ever-changing context will impact the timing and duration of events. It is sage advice to build some flexibility into the timeline to accommodate the unexpected.

Basic Yearly Progress Timeline. As seen in **Table 5.3**, this timeline lists major evaluation activities in a tabular format. It includes when key activities are expected to occur, the data source or audience for planned data collection or communication activities (this represents the how for each planned activity), the team members involved in the activity, and progress toward accomplishing the planned activities (e.g., completed, delayed). Though these timelines primarily focus on evaluation activities it can be helpful to weave in critical moments in the program implementation so that evaluation activities are appropriately planned around these moments. You can also add notes about actions that may be needed to support moving the activity forward. Additional rows can be added if an evaluation spans multiple years.

Table 5.3 Timeline: Basic Yearly Progress

Timeline Year 1 (20XX-20XX)				
Month (When)	**Evaluation Activity (What)**	**Data Source(s)/ Audience (How)**	**Person(s) Responsible (Who)**	**Status/Notes**
Ongoing	Monthly progress reporting	Program leadership	Evaluation Lead	
	Monthly email update of evaluation progress	All stakeholders	Evaluation Lead	
	Document and share reflections about evaluation process	Evaluation Implementation Team	Evaluation Implementation Team	
January	Pilot testing of survey instrument	Cognitive interviews and pilot with subset of respondent population	Evaluation Lead	
February	Obtain approval from IRB	IRB forms	Evaluation Lead	
	Train data collectors on protocol	Training materials	Evaluation Lead	
	Purchase data analysis software or appropriate upgrades		Data Analyst	
March– August	[Program Milestone: Intervention is implemented]	[With referred adult participants]	[Program Educators]	
	Conduct data collection/ monitoring	Survey of participants	Program Educators Evaluation Lead	
	Conduct data management	Survey of participants	Data Analyst	

September	Data cleaning	Data quality assurance	Data Analyst	
	Mock evaluation findings session	Primary intended users	Evaluation Lead	Strengthens capacity to use evaluation
October	Conduct analysis and interpretation		Data Analyst Evaluation Lead	
November	Develop evaluation reports and briefings		Evaluation Lead Data Analyst	
	Provide feedback on report		Evaluation Planning Team	
December	Action planning	Reference findings in evaluation report	Evaluation Lead Evaluation Planning Team	
	Reflection session and documentation of lessons learned about evaluation process	Evaluation Planning Team	Evaluation Implementation Team	

Milestone Tables. Table 5.4 depicts an example milestone table. Key products or events, and the dates by which they should be completed, are listed in this type of timeline. In preparing this table, think about the entire evaluation process, from planning through data collection and analysis to the dissemination of findings and subsequent action planning. Milestones may include fixed dates (e.g., a scheduled partnership meeting or training where you plan to collect data), or more dependent dates (e.g., two weeks after approval of new funding). Keeping the table up to date allows you to track progress in meeting milestones as well as keep track of any schedule changes or deviations.

Table 5.4 Evaluation Milestone Table

Date	Description	Status
10th day of every month	Monthly progress report	
Last day of every month	Email update of evaluation progress to stakeholders	
10th day, bi-monthly	Discuss and document team reflections on evaluation process	
2/5/20xx	Submit completed IRB material (including questionnaire)	
2/15/20xx	Obtain IRB clearance	
2/28/20xx	Data collector training	
3/1/20xx – 8/31/20xx	Conduct asthma self-management training	Once dates are scheduled add to table
	Collect data	
Monthly from 3/1/20xx – 8/31/20xx	Check quality of data from one training session per month	Once dates are scheduled add to table
September 15, 20xx	Host mock evaluation findings session with primary intended users to build capacity for using findings	
10/1/20xx-11/1/20xx	Analyze data, interpret findings, write draft report	
November 10, 20xx	Conduct briefing with Evaluation Planning Team	
November 30, 20xx	Submit final evaluation report to stakeholders	
December 15, 20xx	Complete action plan for use of evaluation results	Add in dates for implementation of action plan

Gantt Charts. A Gantt chart is a valuable way to display the overall project timeline and activities. There are many ways to construct a Gantt chart, but the basic structure calls for lists of activities and the duration of each. More complex Gantt charts can convey dependencies between activities (e.g., an activity that cannot start until after another is completed) or relative estimates of labor hours or other resources across activities. A key consideration in constructing a Gantt chart is the level of resolution needed. During periods in which many activities will be ongoing simultaneously, it may be helpful to set up the chart by days or weeks. For a longer-term view, months or quarters may be sufficient. Templates for constructing Gantt charts are available online, and they can also be created in commercially available software products (e.g., MS Project, MS Excel).

Table 5.5 Gantt Chart

YEAR 1 Timeline (20XX)														
Activity	Start	Stop	J	F	M	A	M	J	J	A	S	O	N	D
1. Pilot test survey	1/1/20xx	1/31/20xx												
2. Obtaining clearances	2/5/20xx	2/28/20xx												
3. Data collector training														
4. Data collection and monitoring	3/1/20xx	8/31/20xx												
5. Data management and analysis	9/1/20xx	10/31/20xx												
6. Communicating findings/ Reporting results	11/1/20xx	11/30/20xx												
7. Action planning	12/1/20xx	12/31/20xx												
8. Evaluation capacity building	1/1/20xx	12/31/20xx												

Shared Calendar. A final suggestion is to create a shared calendar for the evaluation. This calendar can be used by all members of the evaluation planning and implementation teams to focus on key dates for the evaluation. Electronic calendars are useful for keeping all team members up to date. Online calendars are particularly useful because they can easily be created and shared with team members and accessed from any location (e.g., Google calendars, Outlook calendars, calendars within SharePoint).

An evaluation timeline should be a living document. The sequence and timing of many activities are dependent on prior actions (e.g., you cannot analyze your results before your data are collected). This means you may need to adjust along the way to keep your evaluation moving forward.

Periodic Evaluation Reports

Keeping a systematic record of your evaluation activities and sharing this information with key stakeholders on a periodic basis can help ensure stakeholders have access to the information they need to play an active role in your evaluation. Periodic evaluation reports also help you maintain a history of your evaluation while it is in progress rather than trying to reconstruct events once the evaluation is complete. These reports can vary in format and audience, depending on project requirements and needs. Two types of reports that you may want to consider include evaluation progress reports and evaluation status reports.

Evaluation Progress Report. An evaluation progress report is a record of progress and accomplishments during a specific period. (See the "Evaluation Progress Report" box.) These reports can be prepared monthly or quarterly throughout the evaluation period and represent a valuable record for developing more detailed annual reports to funders (e.g., continuation applications) or other stakeholders. They can also be a good way to get new staff members or partners up to speed on progress. These reports can either use a narrative format or rely on bullet points but are generally no more than one to three pages in length. Evaluation progress reports usually include multiple types of information:

- Header, including name of evaluation and person reporting
- Time period for report
- Accomplishments during time period
- Problems and proposed or enacted solutions during time period
- Personnel changes
- Progress in meeting planned schedule or deviations from schedule
- Planned activities for next reporting period
- Financial reporting for labor and other expenditures incurred during the time period and percent of budget expended

You may also want to include additional items during certain periods of the evaluation, such as

- Response rates for data collection activities during the data collection period
- Planned or actual requests for information received and any response
- Planned or actual communication activities related to evaluation findings
- Evaluation successes or lessons learned about the evaluation

Evaluation Progress Report

Time Period: [month, year]
Evaluation Title: Evaluation of Asthma Self-Management
Prepared by: Evaluation Lead

Progress and Accomplishments for [Current Reporting Period]
- Program conducted eight trainings with 150 trainees
- Evaluation lead monitored one training
- Data analyst began entering data from completed questionnaires

Problems and Solutions
- Trainees for this month are 70% female. To have greater male participation, ensure that males are being appropriately referred to program; consider asking men who refuse the referral about their barriers to participation; strengthen recruitment materials for men; consider other times or dates for future trainings that may be more acceptable for men or male-only sessions.

Personnel Changes
- None; consider adding male trainer

Schedule Progress
- Trainings were well attended this month; evening trainings were most popular. Informal feedback suggests more weekend trainings are needed.

Planned Activities for [Next Reporting Period]
- Conduct additional trainings
- Monitor one (randomly selected) training
- Continue data entry of new questionnaires

Financial Report
- Hours incurred in month: XX
- Cumulative hours: XX
- Costs incurred in month: XX
- Cumulative costs: XX

Evaluation Status Report. An evaluation status report (**Table 5.6**) is similar to an evaluation progress report, but its primary focus is on tracking where project activities are in relation to plans. Making this regular comparison can help in readily identifying accomplishments and deviations from the plan so the Evaluation Implementation Team can proactively address them.

In our example, to determine your status during data collection, you would need to keep track of the number of people who were referred to the program, the number who were actually trained, and the number of participants who completed the questionnaire each month. This will tell you how well you are doing toward your goal of collecting information from 810 trainees. Evaluation status reports that contain breakdowns by specific subcategories can also identify ways to strengthen the intervention and evaluation as it progresses. Extending the example, the team may find that a need exists to enhance recruitment procedures among certain subgroups, retrain data collectors, or change the logistics of training to ensure that trainees do not leave before filling out the post-survey. You may want to combine elements of the evaluation progress report and evaluation status report depending on the project needs.

Table 5.6 Sample Evaluation Status Report

Trainee demographics	Expected participants per month (n, %)	Actual participants for reporting month (n, %)
Gender		
Male	67 (50)	45 (30)
Female	68 (50)	105 (70)
Race or Ethnicity		
White	86 (65)	95 (63)
African-American	21 (15)	23 (15)
Hispanic	21 (15)	26 (17)
Other	7 (5)	6 (4)
Age		
18–35	67 (50)	70 (47)
36–50	34 (25)	35 (23)
50+	34 (25)	45 (30)
Total	135	150

Budgeting for Evaluation

Conducting an evaluation requires careful allocation of resources, including time and money. It is easy to underestimate implementation costs, and in such cases, there is rarely an opportunity to acquire additional funding. This section summarizes different approaches to budget development and the costs typically associated with evaluation work. This information will help you to develop budgets, or to review budgets of others, with an eye toward comprehensiveness and appropriateness.

Types of Costs to Consider in Budgeting for Evaluation. Generally, the largest cost in conducting an evaluation is personnel time. As you estimate the level of effort required to complete an evaluation, consider each of the CDC Framework steps. Often, we tend to focus on the time it will take to collect data but underestimate the time it takes to complete other tasks such as planning the evaluation (especially when working with larger stakeholder groups), pilot testing data collection instruments, cleaning and preparing the data, analyzing the data, communicating the results, and working with stakeholders on action planning to facilitate use. Monitoring evaluation progress is another aspect of evaluation often missed during budget planning. Remember to allot staff member and contractor time for regular team meetings, staff member reflections, and the preparation of progress or status reports. By carefully thinking through each step, it is more likely the final cost estimate will be realistic. Initially, it may help to think in terms of the tasks that need to be accomplished and the hours they will take to implement. Then, translate the hours into dollars, while assessing the level of expertise required for each specific task.

In addition to staff member time, there are several additional costs you may incur. We list some categories of costs frequently encountered below.

- **Consultants or contractors:** Consultants or contractors used to extend staff capacity or to provide special skills or experience.
- **Communications:** Postage, telephone, video conferencing charges.
- **Travel**: Long distance or local travel for evaluation staff members to conduct the evaluation or present the evaluation results. Common expenses within this category include air travel, reimbursement for mileage, per diem for meals, ground transportation (e.g., taxi, subway), car rental, hotel stays.
- **Design, printing duplication, and publishing:** Preparation of documents, websites, data collection materials, reports, and other materials.

- **Materials:** Purchased data collection instruments, library services, or datasets.
- **Supplies:** Office supplies, software (qualitative or quantitative analysis software), or subscriptions (e.g., upgrades to Dropbox, SmartSheet, online survey software with required functionality) that must be purchased or leased for the evaluation.
- **Specialized equipment:** Equipment needed to conduct the evaluation or data collection (e.g., laptop computers, digital recorders).
- **Purchased services:** Services purchased from outside vendors with a fixed per unit price (e.g., transcription or translation). These types of service relationships typically do not require a consultant type of arrangement.
- **Incentives:** Monetary or nonmonetary items provided to participants to encourage participation in the evaluation.
- **Institutional Review Board review:** If Institutional Review Board (IRB) review applies to the evaluation, check with your particular IRB to find out their fee structure. In some instances, there may not be an IRB located in organizations where the evaluation implementation team members work. In these cases, it is sometimes necessary to seek review through an external, private IRB; this may add expense.
- **Training:** Training associated with building staff capacity (e.g., analysis training, data collection training, software training) or providing specific training for the evaluation (e.g., expenses associated with providing a data collection training workshop).
- **Dissemination (e.g., conferences):** Costs associated with meeting or conference registration or for local facilities if you plan to convene a stakeholder session.
- **Other:** Any other costs necessary for conducting the evaluation.
- **Overhead costs and fees:** Any overhead fees or costs associated with staff member time or other resource usage.

Not all of these costs apply to every evaluation. In addition, funders of grants or contracts will often have policies regarding allowable expenses. Be sure to consult these policies in advance of budget planning if possible. If you are unable to, you may need to revise the budget line items once the project is underway or during any negotiations with the funder prior to being awarded the grant or contract.

If you hire an external evaluator for part or all of an evaluation, be sure to request an itemized work plan and budget that details labor hours or costs and other expenses using similar categories. Having these documents will avoid later misunderstandings about what was and was not included in the consultant or contractor's scope of work and budget.

Want to learn more?

Find additional information to guide budget planning in *A Checklist for Developing and Evaluating Evaluation Budgets* or the *W.K. Kellogg Foundation Evaluation Handbook*

Also, you may obtain in-kind contributions (i.e., services or support that is free of charge to the project) to help with an evaluation, whether in the form of staff member time (e.g., facilitation support) or material support (e.g., meeting space, incentives, telephone, copying). In-kind contributions should be carefully recorded at each stage in the evaluation. This will help you document the actual costs of the evaluation and will serve to illustrate the support and buy-in for the evaluation. It will also ensure that people or organizations are not later overlooked when the implementation team wants to acknowledge contributions.

Lastly, systematically keep track of your time and expenditures as you go along. By recording labor, expenditures, and in-kind contributions on every evaluation that you conduct, your ability to accurately estimate

an evaluation budget will improve with time as you feed these data back into future estimation processes. The record you keep will also help you answer to your funders, managers, and other stakeholders about how program resources were used.

Budgeting Methods. Developing accurate budgets for evaluation takes both planning and expertise. How do you come up with an appropriate estimate or judge one that has been provided to you from a contractor, especially if you have little direct experience? Though other methods exist, the two approaches to budget estimation you may want to consider are the historical and roundtable methods. For both approaches, think through the justification for the estimate, the assumptions made, and the known requirements for the current evaluation.

Historical Budgeting Method. One means for estimating expenses for the current evaluation is to base them on actual expenditures from prior evaluations. Think carefully about the assumptions and requirements of the prior evaluations compared with the current requirements. Where do you need to make adjustments? Are you now conducting data collection over four months rather than six months? Do you have more or fewer staff members than in the past? Do you have more or fewer respondents? As a rule of thumb, you may want to develop a "per unit" price from prior efforts (e.g., if five focus groups were conducted in a prior effort, how much did it cost to recruit, conduct, analyze, and report on each group?). The more evaluations you conduct or fund over time, the more historical budget data you will have to work with and the more accurate your estimates will become.

Roundtable Budgeting Method. If you do not have historical data available as a guide for estimating the costs of your new evaluation or if prior evaluations were too different from current efforts, you can use expert opinions to help develop your budget. Bring together three to four experienced staff members or partners with knowledge of the level of effort required. For example, you may want to bring in a staff member who has experience working with your priority population to estimate how much time will be needed for recruitment. Or you may want to engage an epidemiologist to think about the analysis requirements. As you work with these experts, carefully document and describe the elements of the evaluation that will affect the costs (e.g., How many units? How long will each last? Who will be involved? What experience level is needed? How many and what types of supplies, equipment, and materials will be required? Are there any fixed costs? What are the variable costs?). Work as a group to come up with your best estimates of personnel time and additional resources needed for each component of the evaluation.

The "How" of Practice
Situational Awareness

Regardless of method, or care in budgeting, unforeseen expenses or cost overruns will often arise during an evaluation. Managing a shortfall can be challenging, especially when budgets are tight to begin with. To avoid this challenge, evaluators often allow for contingencies in the budget to help them adapt to changes in expenses. This measure is a good example of being situationally aware.

It may be advantageous to combine the historical and roundtable budgeting methods to arrive at a more accurate estimate. A roundtable group can usually provide a better estimate when it is based on historical data. Historical estimates can benefit from the input of several experienced staff members to assess where and how to make adjustments based on the requirements of the current evaluation.

Meeting Evaluation Implementation Challenge

Good planning can help with anticipating and minimizing potential evaluation challenges. Yet, no matter how good the plan, challenges can and will occur. By promptly identifying and actively addressing evaluation challenges, evaluations are more likely to meet the evaluation standards of utility, feasibility, propriety, accuracy, and evaluation accountability. In **Table 5.7** we present five challenges that we have witnessed frequently in evaluation practice and provide some practical suggestions for ways these could be addressed during the planning or implementation phase. Many more challenges than the ones listed in **Table 5.7** can emerge in practice. Several are summarized along with possible ways to address them in **Appendix G**.

Table 5.7 Meeting Challenges in Evaluation

Evaluation Challenge	Possible Action	Phase		Relevant Standard(s)
		Planning	Implementation	
Changes in program priorities	When priorities shift, frankly discuss whether the evaluation should continue as planned or whether modifications need to be made.		X	Utility Feasibility
	If evaluation continues to completion, discuss implications for program priorities with stakeholders after completion.		X	
Difficulty communicating with evaluation team members or stakeholders	Develop a communication plan about whom you will need to communicate with at various stages of the evaluation and the best modes of communication for each audience type.	X		Utility Feasibility Propriety Accuracy Evaluation Accountability
	Consult regularly with the communication plan to make sure you are on track.		X	
	Develop ways to obtain regular feedback from various audience types to ensure adequate communication.		X	
Difficulty engaging some stakeholders due to anxiety toward the evaluation	Meet with stakeholders to learn about and work through their past experiences with evaluation. This may include making modifications to the current evaluation to avoid situations where stakeholders feel they are reliving the past.	X		Utility Accuracy Evaluation Accountability
	Determine the program “psychologic” (i.e., how the success or failure of the program being evaluated will affect stakeholders personally; Donaldson, Gooler, & Scriven, 2002, p. 265).	X		
	Clarify the purpose of the evaluation, evaluator and stakeholder roles, and the value of stakeholder participation (Donaldson, Gooler, & Scriven, 2002, p. 265).	X		
	Be a role model: be open to and accepting of stakeholders’ evaluation and criticism (Donaldson, Gooler, & Scriven, 2002, p. 265).	X	X	

Evaluation off track in terms of timeline, budget, or scope	Examine commitments across proposed evaluations during strategic evaluation planning.	X		Feasibility
	Consider involving an evaluation colleague in reviewing and providing feedback on the practicality of the evaluation workplan outlined in your individual evaluation plan.	X		
	Monitor timeline, budget, and scope.		X	
	Hold regular meetings with the Evaluation Implementation Team to discuss progress and emerging issues.		X	
	Respond quickly to emerging issues; include Evaluation Implementation Team members in developing solutions.		X	
	Revise timeline, budget, and scope as feasible, considering any fixed deadlines.		X	
	Document effectiveness of procedures used to address emerging issues.		X	
Ineffective data collection instruments or data collection strategies	Consider using or modifying existing instruments that have already been tested.	X		Feasibility
	Where new instruments are needed, include stakeholders and individuals experienced in evaluation in the design of effective, culturally sensitive instruments.	X		
	Consider meeting with members of respondent population to inform instrument development.	X		
	Pilot test instruments before launch of full data collection and revise instruments as needed.	X	X	
	Use multiple methods, where possible, to triangulate findings and in case any one method or instrument does not work well.	X	X	
	Train all data collection personnel (even those with extensive experience) on written data collection protocol for this evaluation.		X	
	Regularly monitor data collection activities to ensure that the processes are proceeding smoothly and to detect any emerging problems.		X	
	If problems surface, particularly early in data collection, consider modifying the instrument or data collection strategy.		X	

Findings not used	Hold discussions with program stakeholders upfront about information needs and intended use of evaluation findings.	X		Utility
	Incorporate stakeholder information needs into the evaluation plan.	X		
	Develop plans for dissemination of findings (including interim findings).	X		
	Ensure that findings are communicated to decision makers in useful formats at strategic times; share interim findings if appropriate.		X	
	Hold post-evaluation discussion with program stakeholders about evaluation findings and implications for the program.		X	
	Document proposed strategies to address evaluation findings in an action plan with clear roles, responsibilities, timeline, and budget.		X	

In addition to being aware of specific challenges and planning for them in advance (i.e., those identified in **Table 5.7** and **Appendix G**), it can be helpful to consider time points or specific evaluation activities that may present challenges. Here we examine five aspects of evaluation and the challenges that tend to surface within each: evaluation context, evaluation logistics, data collection, data analysis, and dissemination of findings. Since these challenges tend to surface during implementation, we address them within this chapter, however, similar to the challenges presented in **Table 5.7**, the suggested actions to mitigate some of the challenges takes place during the implementation *and* planning phases of the evaluation.

Evaluation Context

No program—and no evaluation—occurs in a vacuum. Public health programs, and evaluations of these programs, exist within an organizational hierarchy and are embedded within a community that can influence their conduct and their ultimate success. Public health interventions occur in multiple settings (e.g., homes, schools, workplaces, hospital emergency departments, and clinics). Therefore, evaluations of interventions may require access to these places to collect critical data. To gain this access, identify and cultivate champions for the evaluation within relevant organizations and the community at large. Champions can lend credibility to the evaluation and foster trust among others, including those who may be feeling anxious about the evaluation. These champions can also encourage key program stakeholders to consider and eventually act upon evaluation findings. This type of political will in support of evaluation is extremely valuable and should be thoughtfully and actively fostered.

Evaluation Logistics

An evaluation needs to be managed like any other project. Those working on the evaluation need to know who is doing what, when, how, and why. They also need clear guidelines about how many hours and other resources can and should be spent on individual work assignments. Evaluation progress should be carefully monitored through a variety of means, and contingency plans should be developed if evaluation components, such as the timeline, budget, or scope, lose their trajectory. Good project management processes and tools such as the ones presented earlier in this chapter will support those managing the evaluation in reaching a successful conclusion.

Data Collection

There are many aspects of data collection activities to consider while planning for and implementing an evaluation. This is true whether you are collecting new data through surveys, interviews, or focus groups; systematically reviewing archival data such as medical records; or compiling and analyzing data from existing sources. Any of these types of data collection activities require that you have a clear plan, or protocol, for how the work will proceed and as many safeguards as necessary to ensure the consistency, accuracy, and reliability of the findings. Furthermore, you want to ensure that any data collection upholds the evaluation standards of propriety and accuracy to protect stakeholders while also gathering valid information.

Some important safeguards include documenting procedures to use, pilot testing procedures and instruments, training individuals involved in data collection or compilation, and carefully cleaning the data in preparation for analysis. In addition, you will want to ensure procedures are in place to monitor the quality and consistency of incoming data. Protecting the rights of any participants involved in the evaluation is another consideration that must be planned for upfront and managed carefully during implementation.

Data Analysis

Nothing is more frustrating than approaching the end of an evaluation, only to discover that the data collected cannot be analyzed or do not meet the needs of program staff members and stakeholders. With so many precious human and monetary resources invested in an evaluation, planning ahead for data analysis and use—and documenting these in the individual evaluation plan—is critical. To the extent that such plans are developed in consultation with program leadership and stakeholders, the likelihood that evaluation findings will meet their information needs increases.

Dissemination of Evaluation Findings

We conduct evaluations to improve our programs and provide accountability to funders, program participants, and other decision makers. Yet evaluation findings that are not believable or come too late to meet a specific information need are unlikely to be able to inform programmatic decision making. Fortunately, while planning for and implementing an evaluation, there are things that can be done to help ensure use.

Closing the Loop: Turning Insights into Action

To gain the maximum benefit from evaluation, it is imperative that the results of your efforts be used, whether to support program improvements or to guide other decision making. This aligns with the evaluation standards of utility and feasibility to ensure that evaluation findings are useful and practical. We know from experience that evaluation results are more likely to be used if the evaluation implementation team, stakeholders, and program staff members take the time to develop an action plan. Doing so provides an active mechanism for translating the findings into action and documenting clearly who will be accountable for these actions. Additionally, documenting lessons learned about the evaluation process itself helps to build evaluation capacity and supports both current and future programmatic improvements. Last, it is important to link evaluation findings to the Strategic Evaluation Plan. Not only does this support use, but it provides insight on necessary modifications to planning and potential evaluation candidates.

Developing an Action Plan

An action plan is an organized list of tasks that, based on the evaluation findings, should lead to program improvement, by addressing weaknesses and augmenting strengths, and sustainability. It differs from a to-do list in that all the tasks focus specifically on achievement of the program improvement objectives. It can and should serve as a program's roadmap to ensure that evaluation findings are used to improve the program.

If you identify more than one program area that is ready for modification based on a single evaluation, we recommend creating an action plan for each improvement objective. Your evaluation may also identify program components that should be eliminated or components that are working well and should be sustained. Action plans are appropriate to guide these follow-up activities as well.

Since some stakeholders will be charged with implementing changes based on the evaluation findings, it is of critical importance that you and stakeholders who were involved in designing the evaluation work together to create the action plan and agree upon who is responsible for implementing any appropriate changes. Their involvement can help ensure that planned activities are both desirable and feasible, and they are more likely to participate in implementing changes if they have been involved in identifying actions to be taken. It is very important that program personnel and other stakeholders take part in thinking through and creating the action plan and, even more importantly, that they take responsibility for carrying out activities therein.

We provide an example of an action plan in **Table 5.8**. The format directly connects program improvement objectives to evaluation findings by including a summary of relevant evaluation findings, evidence upon which findings are based, and proposed changes to respond to findings. Most of the plan focuses on specific action(s) you and your stakeholders will take to achieve the stated objectives. The plan identifies the person responsible for each activity, resources they need to accomplish it, and a timeline for completion. The action plan template also includes an area to list the information you will use to monitor the implementation of your action plan. Finally, the plan specifies data you will use to determine whether the improvement(s) you want to make actually occur.

Regularly reviewing the results of your action plan with your stakeholders will help you better use evaluation findings. If you have evidence that your program has improved, this marks an occasion for joint celebration. If more work needs to be done, your stakeholders can help focus your energies and support necessary changes.

Table 5.8 Evaluation Results Action Plan

Program Component (e.g., partnerships, surveillance, specific intervention): **Evaluation Purpose:** **Programmatic Change Sought:**								
Evaluation Result	Describe the key evaluation result that necessitates action.							
Supporting Evidence	Describe the evidence that supports action.							
Plan of Action to Achieve Change							**Monitor Change**	
Change Needed	**Activities to Implement Change**	**Person Responsible**	**Resources Required**	**Due By**	**Indicators that Change is Implemented**	**Data Sources**	**Indicators to Monitor Success of Change**	**Data Sources**
Describe key change(s) you want to achieve based on this finding.	*List activities that need to be carried out to make the change happen in the program*	*List the person(s) who will ensure each activity occurs*	*List resources required for the activity*	*Assign a due date by which the activity will be completed. (The final date should be when the change will be in full effect.)*	*Describe how you will know that the change is implemented as planned*	*Describe what data you will need to know the change was implemented.*	*Describe how you will know whether the change to the program is working*	*Describe the data you will need to measure success.*

Documenting Lessons Learned from Implementing Evaluations and Linking Back to the Strategic Evaluation Plan

History repeats itself because we were not listening the first time. That is as true for evaluation as it is anywhere else. You can begin building historical knowledge about the evaluation process itself, to pass on to future generations, by documenting lessons learned from one evaluation for use in future evaluations. Consider adopting the habit of closing Evaluation Implementation Team meetings by discussing the following questions:

- What have we learned?
- What can we do better next time?
- What steps can we take now to prepare?

Document these discussions to help future evaluations go more smoothly. In this way, you encourage team members to reflect on their evaluation practice, which will lead to professional growth and learning. This practice also upholds the standard of evaluation accountability.

Linking evaluation findings back to the Strategic Evaluation Plan is a critical final strategy in ensuring evaluation use and promoting ongoing program improvement. It is not uncommon for an evaluation report to raise more questions than it answers. When questions are raised, they may result in a new priority for evaluation. For instance, findings from an evaluation may suggest that the aspect of the program evaluated was functioning well, but that another aspect that was touched on tangentially is not functioning well and should be investigated more closely. Or, findings may demonstrate that one aspect of the program is not working well, yet not really explain why that is so or how the problem could be remedied. The why and how of what isn't working may then become grist for the mill of a new evaluation. Additionally, the Evaluation Implementation Team may encounter issues with the logistics of an evaluation that suggest using an alternative approach in an upcoming evaluation. Questions raised in prior evaluations and lessons learned about conducting evaluation itself should be considered by the Strategic Evaluation Planning Team when they update the Strategic Evaluation Plan, ideally on an annual basis.

The "How" of Practice
Critical Reflection

In evaluation, it is common to hear of evaluators or evaluation teams reflecting at the end of the evaluation. While these reflections may generate useful insights on evaluation processes and structure, they are unlikely to lead to any meaningful change in how the team approaches the next evaluation or how team members understand themselves as evaluators and individuals. To truly learn from our experiences, we must reflect critically on our roles, work, and relationships, both individually and as a team. You might ask: how did the evaluation take into account the perspectives of stakeholders? Whose perspective was ignored, and why? Who will be impacted negatively by the actions we plan to take? Will these actions promote equity or reinforce the status quo?

Some benefits of documenting lessons learned about evaluation itself

- Helps to avoid repeating past mistakes.
- Builds evaluation capacity.
- Transfers knowledge to those who come after you.
- Creates an archive of good evaluation practices over time.

Chapter Summary

In this chapter we provided information to ease the transition from planning into implementation. Even the best laid plans can encounter challenges when the rubber meets the road of implementation. Therefore, you must think of tools that can help you more effectively manage evaluations (e.g., timelines, roles and responsibilities tables, comprehensive yet flexible budgets). In addition, you must prepare to identify and respond to common challenges encountered during implementation. When you reach the end of an evaluation, it is important to close the loop. We offered suggestions for putting evaluation findings into action and increasing the likelihood that there will be follow through on these actions by putting accountability measures in a formal action plan. We also recommended revisiting the Strategic Evaluation Plan after each evaluation is finished to identify whether adjustments are needed.

Review Questions and Skill Building Exercise

Review Questions

1. Which of the evaluation timeline tools presented in this chapter do you feel would be most helpful to you if you were conducting an evaluation? What do you view as some of the pros and cons of the different formats?
2. Of the evaluation challenges presented in this chapter and **Appendix G**, were there any that surprised you (i.e., that you didn't anticipate would be challenges in an evaluation)? Select one of these surprise challenges and ask how this challenge could be addressed or alleviated through planning or implementation?

Skill-Building Exercise

You/your team was asked by a national non-profit to design and implement an evaluation study of a telephone helpline. The helpline was designed to connect individuals who recently were admitted to the hospital for asthma to resources in the community that can help them manage this chronic health condition (e.g., asthma self-management educational sessions, medical referrals). The hospital is located in a suburban county in Kansas. Prior to discharge, a nurse at the hospital informs the patient about this helpline and makes the first call with them. After discharge, the patient can reconnect with the Helpline to seek services or assistance, as needed. Small incentives are offered to encourage patients to reconnect with the Helpline after discharge.

The evaluation is to be conducted over the next 24 months. You will be awarded $100,000 for your efforts over the two years. The funder is interested in learning about the current strengths and limitations of how the helpline operates as well as about the effectiveness of the helpline. There is no current evaluation plan in place, so your team will be responsible for developing the plan and implementing the evaluation during these two years. Your initial thinking is that a mixed-methods approach will be needed.

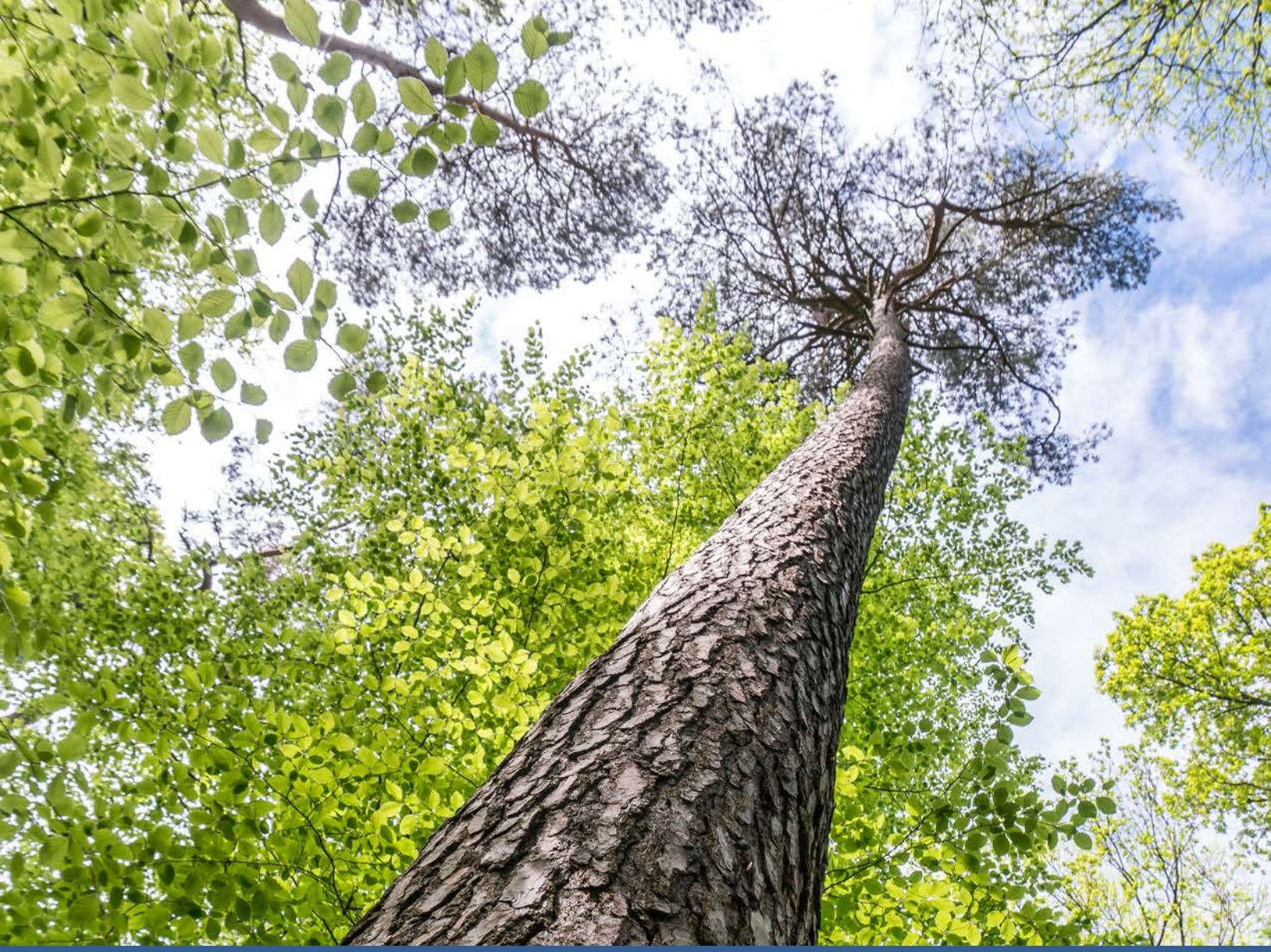
Your task

1. In this chapter we shared numerous project management tools for evaluation. Pick two tools and, using the scenario above, develop draft documents for each. You may use the templates included in this chapter or ones from other sources. When you're done, share your draft documents with a friend for their feedback.
2. Developing accurate evaluation budgets is critical to effective implementation. Based on the description above, create a budget for the evaluation. It's okay if you're unsure of some expenses or the amounts. The point is to start thinking through these items and what sources you may need to reference to obtain accurate numbers. You may use the template below or one from another source.

Evaluation Budget Template

Resource	Year 1 ($)	Year 2 ($)
Evaluation staff salary and benefits		
Consultants or contractors		
Communications		
Travel		
Design, printing, duplication, and publishing		
Materials		
Supplies		
Specialized equipment		
Purchased services		
Incentives		
IRB Review		
Training		
Dissemination		
Other		
Overhead costs and fees		
Total		

*Adapted from Worthen, B. R, Sanders, J. R., & Fitzpatrick, J. L. (1997). *Program evaluation: Alternative approaches and practical guidelines*. (2nd ed.). Longman Inc.



SECTION III

Special Topics in Public Health Evaluation

Section III: Special Topics in Public Health Evaluation

In Section II you were introduced to the evaluation planning and implementation process. Leveraging the CDC Framework, we described the steps to take in developing an overarching Strategic Evaluation Plan and provided you with an example of the process and product in **Appendices C** and **D**. We also described the process of developing an individual evaluation plan for each priority evaluation highlighted in the strategic evaluation plan. Recognizing that even the best designed plans can encounter challenges when implemented, we shared thoughts on how to maintain flexibility during the implementation process. We also gave examples of common implementation challenges along with actions that might help to address the challenges.

In this section we turn our attention to two important topics for evaluation in public health settings: (1) evaluating public health infrastructure, namely partnerships and surveillance (Chapter 6) and (2) evaluating public health interventions (Chapter 7). Instead of repeating the information provided in Section II, we focus on the unique aspects of evaluating these special topics. For instance, what stakeholders might be likely candidates to engage in a partnership evaluation and how does this differ from those you may engage in an evaluation focused on public health surveillance or interventions? We also highlight specific tools or resources that may be particularly useful in these types of evaluations. We will share several data collection instruments that relate to partnership evaluation and introduce you to the concepts of evaluability assessment and program theory in the context of evaluating interventions.

CHAPTER SIX

Public Health Program Infrastructure

CHAPTER SIX: Public Health Program Infrastructure

Public health programming consists of many different types of activities. When we think about evaluating a public health program, it is natural to gravitate toward interventions—whether the provision of a health service or a health education intervention, development of broader systems or policy changes, or implementation of a multi-component intervention leveraging several of these techniques. Program stakeholders, including funders, often want to know whether a difference was made in the health or welfare of the populations who were supposed to benefit from the intervention(s). This is certainly an important area of investigation and is therefore the subject of **Chapter 7**.

There are other common elements of public health programming that support the effective development and implementation of interventions. In this chapter, we discuss two of these common infrastructural supports: public health partnerships and public health surveillance. We view these two important public health activities as part of the infrastructure of programming because they provide the framing through which public health interventions are built. These two activities also provide supports to implement, and sometimes evaluate, the interventions. For instance, results of public health surveillance often highlight pressing needs within specific sub-populations in a nation, state, locality, territory, or tribal nation. These insights often fuel the development of public health interventions to address the identified need(s). Public health partnerships not only help to validate and explain the needs identified through surveillance data, but also often engage in coordinating and carrying out interventions. If partnerships and surveillance are not functioning well, it is unlikely that the interventions they support will be effective.

By the end of Chapter Six, the reader will be able to

- List two to three stakeholders of partnership and surveillance evaluations.
- Describe the different uses/roles of overarching and nested or "zoom in" logic models.
- Develop evaluation questions that are specific to the topic of partnership evaluation and surveillance evaluation.

Highlighting the Nuances of Infrastructure Evaluation

The steps involved in planning for and implementing evaluations of public health partnerships and surveillance are the same as those described in **Chapters 4** and **5**. However, each evaluation must be tailored to the topic, and ultimately the evaluand. As a result, we can think through the types of individuals who might be well suited to assist with planning an evaluation focused on partnerships or surveillance, what logic models or similar visual depictions might include for these public health functions, the types of questions teams might find most applicable, and any particular methods (i.e., data collection, analysis) that may be uniquely well suited to these topics

Partnership Evaluation

Public health has a rich tradition of using partnerships to pursue shared goals (Price et al., 2020; Rowitz, 2001). The purpose of partnerships is to mobilize members' commitment, talents, and assets to effect change (Butterfoss, 2006). Most public health programs involve partners in developing and implementing plans, but there can be significant variation in a partnership's purpose, membership composition, size, structure, and stage of development.

> Irrespective of the name or form, a core function of a partnership is to facilitate collaboration among distinct entities for the purpose of working toward outcomes of mutual interest and pooling abilities, expertise, and resources. Strengthening, supporting, and mobilizing communities and partnerships is one of ten Essential Public Health Services (PHNCI, 2020).

Whether they are called partnerships, coalitions, collaboratives, task forces, or some other name, the published literature points to several factors that contribute to their effectiveness. No commonly agreed-upon definition of effectiveness exists, but both the success of partnerships in engaging and sustaining the involvement of members (process), and the outcomes they achieve have been studied. We define effective partnerships as "those that bring together important program stakeholders, and then organize and engage them so as to achieve the mission, goals, and objectives of both the public health program and its partners."

Whom to engage in planning and implementation (Step 1). As with other aspects of public health programming, multiple stakeholder perspectives can contribute to a rich and comprehensive description of a partnership, while also facilitating a well-balanced and useful evaluation. Involving stakeholders with a variety of perspectives in planning and implementing an evaluation will enrich the experience, increase partner buy-in, and facilitate the use of findings. But who is most likely to be interested in an evaluation of a partnership? Leveraging the categories of evaluation stakeholders presented previously in **Chapter 4**, we offer the following as possible **partnership evaluation** stakeholders:

- ***Primary stakeholders.*** These are stakeholders who have direct involvement in the partnership. They may include staff members of the public health program that is helping to lead or coordinate the partnership, co-leads of the partnership, leaders of working groups that comprise the partnership, and funders. Remember, these are individuals who have the ability and the authority to use the evaluation findings about the partnership to modify how the partnership operates.
- ***Secondary stakeholders.*** Consider who is most likely to be affected by changes that occur to the partnership as a result of taking action on the evaluation findings. These may include members of organizations who are represented on the partnership, the partnership working groups (e.g., data subcommittee, health disparities subcommittee), or individuals affected by interventions conducted by partners.
- ***Tertiary stakeholders.*** These are stakeholders who may have an interest in how the partnership evaluation is performed or the results. This may include other health-related coalitions in the same jurisdiction, coalitions for other public health programs, and regional or local public health coalitions that were not the focus of the specific evaluation

It will be important to consider who within these three categories of evaluation stakeholders may be most helpful to engage on the Evaluation Planning Team, and subsequently, on the team implementing the evaluation. When selecting stakeholders to work with on planning and implementation, one should consider how to foster cultural responsiveness and promote equity. It may not be immediately obvious how to do this through a partnership evaluation. We suggest engaging stakeholders who reflect the diversity within the partnership and those experiencing health inequities. Researching and understanding the contextual factors that influence the partnership or stakeholders, as well as recognizing power and privilege among stakeholders, including the evaluation team, can also be ways to promote equity. Failure to include multiple perspectives can result in a

skewed or incomplete evaluation, and thus, a skewed or incomplete picture of the partnership itself.

Describing the partnership (Step 2). In previous chapters, we have shared examples of logic models. However, you might be wondering what a logic model, or other visual depictions of a public health partnership, might look like. Because public health program partnerships vary, especially in their structures, no two jurisdictions' logic models will look alike, and because partnerships evolve over time, the logic models depicting them will change as well. **Figure 6.1** provides a logic model for a fictitious partnership that is focused on improving asthma management practices in a state.

The "How" of Practice
Interpersonal Competence

In this and other chapters, we have emphasized the value of engaging various perspectives in the evaluation. It is worth noting that diversity is often accompanied by conflict, particularly when partners share differing world views, concerns, or proximities to the program. If not managed correctly, conflict can limit partner input and participation, consume time and other resources, and even derail the evaluation process. To avoid these and other consequences, evaluators must become adept at managing conflict in a way that satisfies all partners. This may involve facilitating difficult conversations, helping partners to articulate their views respectfully and understand other points of view, and leading shared problem-solving and consensus-making.

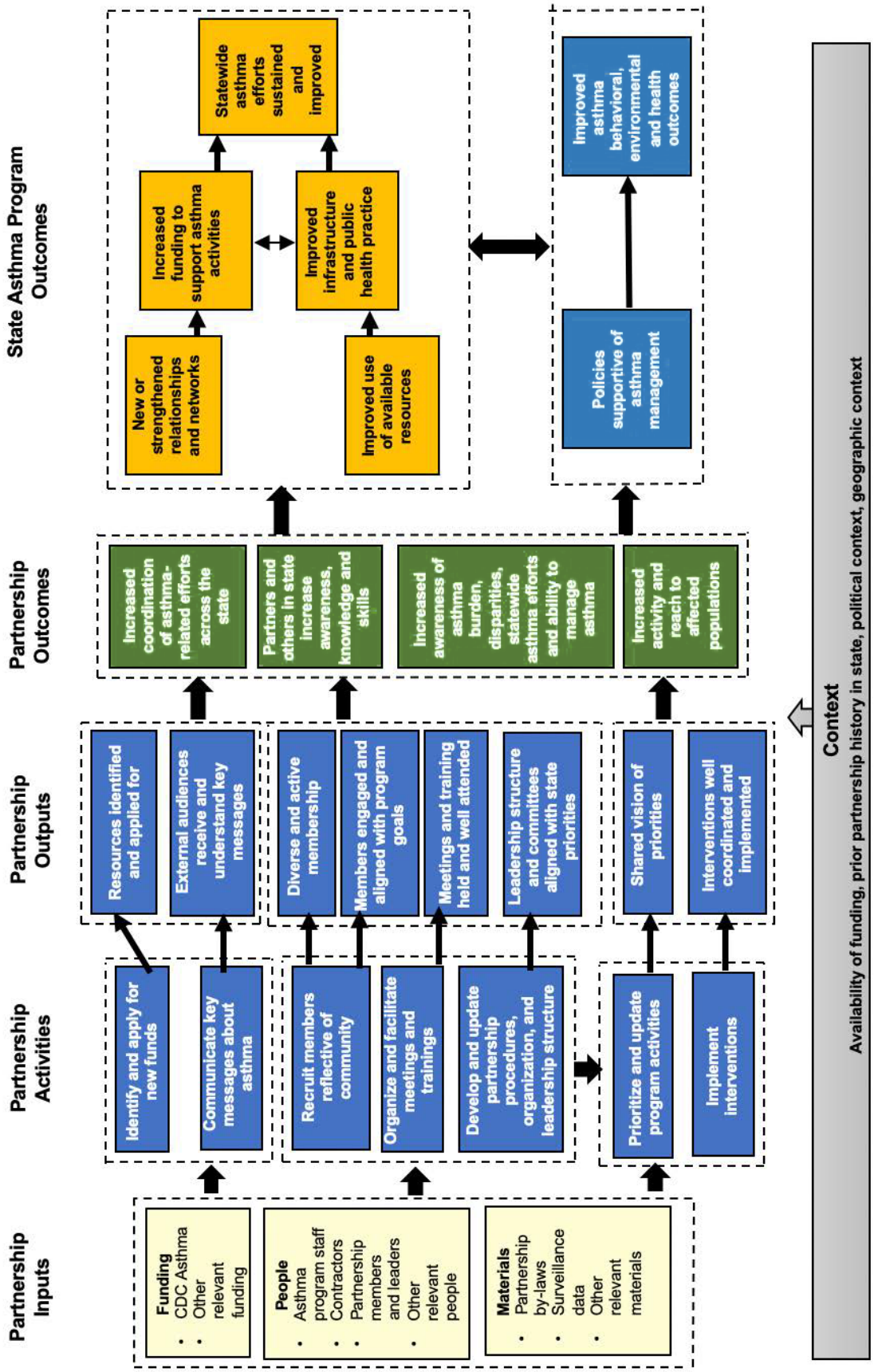


Figure 6.1 Partnership Logic Model for Fictitious State Asthma Program

In reviewing **Figure 6.1**, we can see the main activities that comprise the overarching asthma program partnership, how these connect to direct outcomes associated with the partnership (e.g., increased coordination of asthma-related efforts across the state), and how these partnership-specific outcomes contribute, overall, to the broader outcomes of the state asthma program. Thus, we see how the partnership, an infrastructural element of the overarching program, helps to foster the ultimate intended outcomes of improved asthma health outcomes and its precursors. This is a helpful diagram for explaining the partnership as a whole and arriving at a common vision of the partnership with various stakeholders.

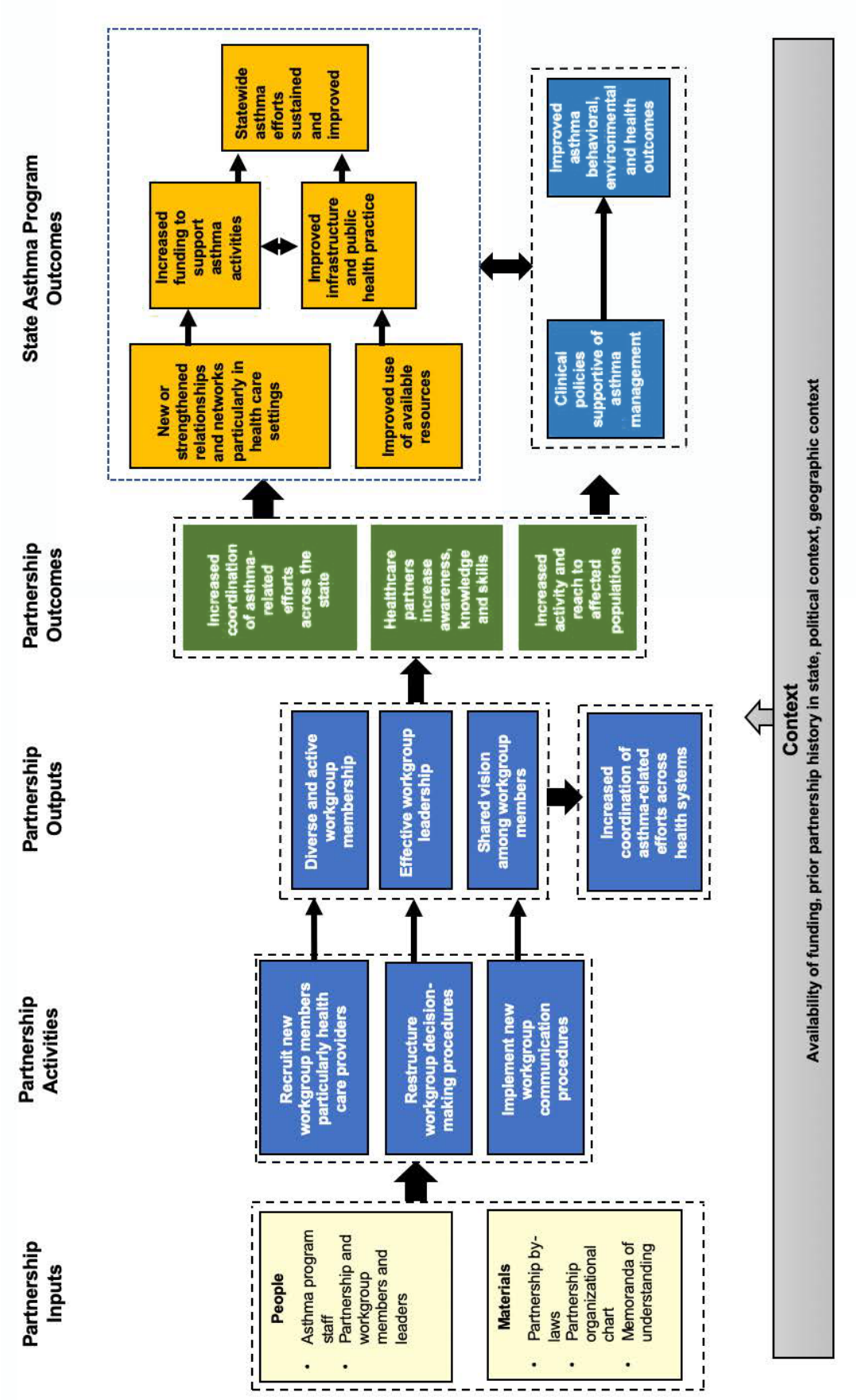
It is likely that an evaluation will not focus on the entire partnership at one time. Rather, the focus for a single evaluation will likely be on one or a couple of aspects of the partnership depending on the information needs (utility), resources available (feasibility), and quality of the resulting information given other constraints (accuracy). For instance, we might focus on the reorganization process of a specific working group that is part of the overall partnership such as one focused on engaging with the healthcare system and create a logic model that zooms in on this specific component of the model.

A Glimpse into Practice
Describing the Specific Partnership Components of Interest

Consider the case of a public health program that focuses on improving asthma management practices in their state. This program has an active partnership that helps coordinate efforts throughout the state. A key feature of this partnership is their Healthcare Systems Workgroup which focuses on (1) building healthcare practitioner's knowledge and skills in accurately diagnosing and treating asthma according to existing guidelines and (2) coordinating care throughout the healthcare system for people who have asthma. Over the past year, staff members of the asthma program and others within the partnership observed that the workgroup was making less progress than anticipated. In an effort to improve the workgroup's performance, the partnership opted to reorganize this workgroup.

The Strategic Evaluation Planning Team decided to prioritize the reorganization efforts for evaluation because of the central importance of this workgroup in achieving the state asthma program's intended outcomes. Furthermore, the Strategic Evaluation Planning Team determined that it would be possible to deliver high-quality results (accuracy) from this evaluation to the partnership in time for them to make future decisions (utility) about the partnership structure and function given the available resources (feasibility). The evaluator invited a small set of stakeholders to participate in the Evaluation Planning Team— two workgroup members who are actively planning the reorganization, another workgroup member who is not involved with it, and a member of a separate workgroup that the Healthcare Systems Workgroup affects. Other stakeholders are invited to review the evaluation plan: a workgroup member who supports the reorganization, one who is critical of it, a leader from the Data and Surveillance Workgroup, and a member of the Strategic Evaluation Planning Team.

In reviewing the overarching logic model for the state asthma partnership (**Figure 6.1**), the Evaluation Planning Team noticed that there are two activities with specific relevance to this evaluation – (1) Develop and update partnership procedures, organization, and leadership structure and (2) Recruit members reflective of community. Since these activities are the primary focus of the reorganization effort the Evaluation Planning Team decided to create a new logic model that "zooms in" on these specific items (**"Healthcare Systems Workgroup Reorganization"**).



Healthcare System Workgroup Reorganization Logic Model

Identifying the focus of the evaluation (Step 3). All too often evaluations of partnerships focus on examining the satisfaction of members through administering a simple survey at partnership meetings. There are, however, many possible evaluation questions that can be asked of public health partnerships. If you examine the boxes and arrows in **Figure 6.1**, how many questions arise? It is likely you identified questions related to the efficiency of the partnership, the composition of its membership, the appropriateness of the activities performed, whether the partnership efforts led to any or some of the outcomes proposed, and many more. In 2006-2007, CDC's Asthma and Community Health Branch (ACHB) formed a workgroup with representatives from 10 state asthma programs, staff members from ACHB, and Battelle Centers for Public Health Research and Evaluation to develop guidance on evaluating asthma program partnerships. A starting point for this work was the development of a conceptual map that described the common features of state asthma program partnerships (**Figure 6.2**) and, most importantly for this step in the evaluation process, articulated several overarching evaluation questions that are likely to be helpful to other public health programs crafting questions for partnership evaluations – Who is involved? How do they interact? What do they do? What are the results? As demonstrated in the section that follows it is possible to stimulate many specific evaluation questions about partnerships under each of these overarching evaluation questions.

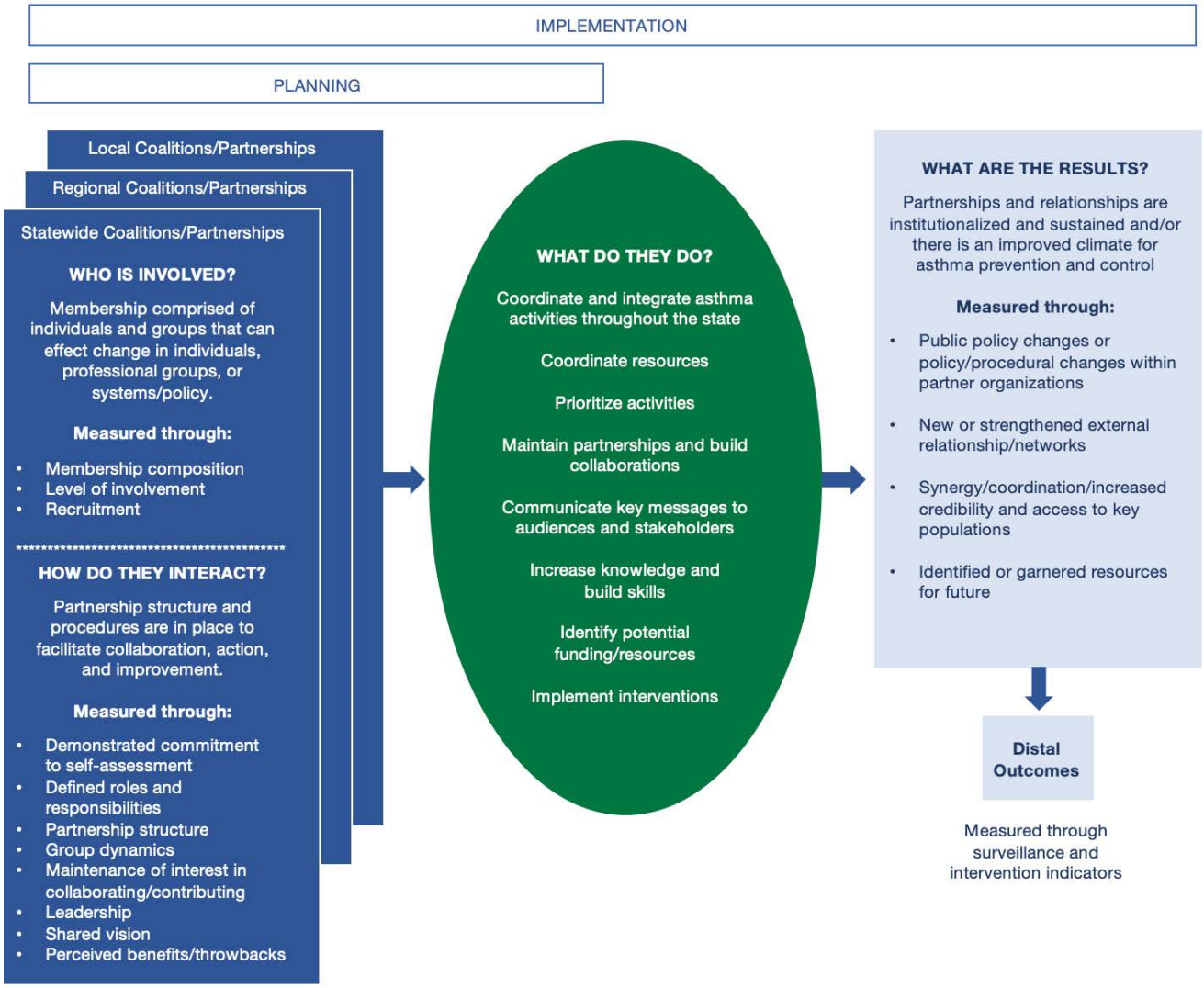


Figure 6.2 Partnership Concept Map for CDC's National Asthma Control Program

Who is involved? The left side of the partnership concept map illustrates the variation in the structure of partnerships; in the case of the asthma programs, they were organized at the state, regional, or local level. Research indicates that effective partnerships include people who understand the problem (e.g., asthma) and are able to stimulate local responses and solutions. Several concepts are relevant to partnership involvement such as the membership composition, the level of involvement of members, and membership recruitment. Evaluation questions related to each of these concepts are presented in **Table 6.1.**

How do they interact? The left side of the partnership concept map also shows how partners interact with one another. Research indicates that partnerships with formalized procedures, structures, and roles and responsibilities are more likely to engage members and pool resources. Partnership structures that are action oriented (e.g., composed of workgroups or committees) tend to be effective in mobilizing resources and implementing strategies. Additionally, research highlights the importance of leadership, communication, shared vision, positive group dynamics, and the ability to resolve conflicts.

Partners are more likely to remain interested in evaluation when they view the benefits of engagement as outweighing the costs (Butterfoss, Goodman, & Wandersman, 1996). Partners typically describe benefits such as skill acquisition, exposure to new ideas and groups, a strengthened ability to meet individual and collective goals, empowerment, capacity building, new relationships, and the opportunity to contribute to a shared vision. A commitment to **self-assessment** is also considered important for a partnership. **Table 6.1** lists evaluation questions that align with these dimensions.

Table 6.1 Sample Evaluation Questions – Who is Involved? How do they Interact?

Who is involved?	
Membership composition	• Who are the members of the public health program partnership? • To what extent does the expertise of these partners align with current and upcoming program plans? • How well represented are the people most affected by the public health program?
Level of involvement	• How engaged are partners? To what extent does engagement vary by type of member? • To what extent do they assume leadership roles? • What types of actions are they most likely to take and how do these actions align with our needs?
Membership recruitment	• How does our membership compare with partnerships within similar public health program? …To the composition of the broader community we represent? • What additional partners should we add to support our efforts? • How timely are gaps identified and addressed in our partnership?
How do they interact?	
Demonstrated commitment to self-assessment	• How frequently does the coalition or partnership conduct a self-assessment? • How is information from these self-assessments used? • How might the use of the results be improved?
Defined roles and responsibilities	• What is the role of staff members in the partnership? • To what extent does the role of staff members align with the culture of this partnership? • Are there additional or different roles that the members feel are necessary and within the constraints of available resources?
Structure	• What roles do committees and subcommittees play? • To what extent do these roles support attainment of the goals of our programs? • How might these committee roles change to better align with the program priorities?

Group dynamics	• What is the level of trust among the partners in this group? In what ways, if at all, does level of trust vary by type of member? • To what extent do members feel they can openly share their comments and ideas? • What level of influence do people affected by the public health problem hold?
Maintaining interest in collaborating or contributing	• To what extent has the partnership been able to maintain the membership's interest? • What techniques have been most successful in maintaining member interest?
Leadership	• What is the leadership structure? • What is the leader's role? • In what ways, if any, is leadership shared amongst members? • In what ways might the leadership structure be improved?
Shared vision, mission, and planning	• To what extent does the partnership have a clearly articulated vision? • To what extent is this vision shared among members of the partnership? • In what ways, if any, does this shared vision reflect the needs of the community who is most affected by the health condition?
Perceived benefits and drawbacks	• To what extent have organizations or individuals benefited from group participation? • What benefits did they expect that were not realized?

What do they do? In the center of the partnership concept map we list potential roles that partners may play. Partners take on a wide variety of roles in public health programs, from contributing material resources to actively implementing interventions. Partners may also develop their own knowledge and skills and use these to effect change in the organizations they represent. This is a rich area of inquiry for partnership evaluations and may include questions such as those in **Table 6.2.**

Table 6.2 Sample Evaluation Questions – What do they do?

What do they do?	
Coordinate and integrate activities	• How does the program interface with other related public health activities in local communities? • In what ways can these relationships be improved upon and sustained?
Contribute resources	• What types of resources have partners contributed to accomplishing the goals of the program? • Does the partnership need other types of resources (e.g., money, time, supplies)? • How might these gaps be filled, and by whom?
Prioritize elements of public health program plan	• What role do partners play in identifying priority interventions? • To what extent do partners feel they were appropriately engaged in prioritization activities? In what ways, if any, do these sentiments vary between types of partners?
Implement interventions	• What training or educational interventions are being conducted by partners? • How might these interventions be expanded or sustained to facilitate quicker or fuller accomplishment of the goals of the program?
Identify potential funding and resources	• Of the funding opportunities identified by the partnership over the past year, which ones do members feel are most relevant to accomplishing the program goals? • What characteristics about these relevant funding opportunities do the partners feel have the potential to be most influential or helpful?
Maintain partnerships and build collaboration	• To what extent has the partnership been able to maintain or expand its membership to accomplish priority activities?
Communicate key messages	• What communication techniques does the partnership use to share key messages with its members? • How effective do members perceive these communications to be? In what ways, if any, do these sentiments vary by type of member?

What are the results? According to the literature, when a partnership performs well, a variety of partnership-specific outcomes emerge. The broad engagement of partners mobilized to effect change in multiple community sectors is more likely to lead to sustained environmental change within partners' peer groups and organizations. The strength of networks and relationships built by the partnership may be important for sustaining the partnership itself as well as for helping it achieve long-term programmatic outcomes. Similarly, the ability of the partnership to secure financial resources for its work may predict its sustainability and its ability to influence outcomes. Combining the perspectives, knowledge, and skills of diverse partners can enable the partnership to think in new ways, plan more comprehensive programs, and strengthen relationships with the broader community. This synergy is believed to be an important indicator that a partnership will be effective in reaching its ultimate goals. **Table 6.3** lists evaluation questions related to the results of partnerships.

Table 6.3 Sample Evaluation Questions – What are the results?

What are the results?	
Public or organizational policy change	• In what ways have partners contributed to discussions about public policy that promotes better health outcomes that are the focus of this program? • What is needed to create an atmosphere in the jurisdiction that is conducive to facilitating this type of change?
Synergy	• How effective is the partnership in combining the perspectives, knowledge, and skills of diverse partners in a way that enables members to think in new ways, plan more comprehensive programs, and strengthen relationships with the broader community? • How might this synergy be enhanced?
Identified or garnered resources for the future	• How successful have the partners' efforts been to acquire funds to support the program? • What are some key factors that contributed to this success? • What has hindered this success?
New or strengthened external relationships or networks	• In what ways has the partnership contributed to producing new linkages between the partnership and other coalitions or organizations? • Between entities external to the partnership itself? • How do these new connections contribute to improving program outcomes?

Equity. An additional concept that has direct importance to public health programs and has recently been highlighted in literature as relevant to partnership evaluation is equity. Considering how to promote equity through the partnership itself is important. **Table 6.4** lists evaluation questions that align with the concept of equity in partnerships (Hilgendorf et al., 2020; Price et al., 2020; Stachowiak et al., 2020; Varda & Sprong, 2020).

Table 6.4 Sample Evaluation Questions about Equity

How is equity included in and affected by the partnership?	
Equity	• In what ways does more diversity in a network make it more difficult or easier to manage goals, outcomes, or perceptions? • What is the role that powerful or influential members play in networks? • How do one-on-one meetings with leaders of adversely affected communities contribute to new insights on partnership recruitment and consideration of health equity in the partnership's activities? What are the **health equity** practices of partnership members? How do these health equity practices change over time in the partnership? • To what extent is the partnership addressing the social determinants of health that affect high-risk populations' ability to achieve optimal health?

Gathering and analyzing credible evidence (Step 4). The general strategies associated with gathering credible evidence articulated in **Chapter 4** (i.e., secondary data analysis and primary data collection strategies such as surveys, interviews, focus groups, and observations) all apply to the collection of data when evaluating public health partnerships. Fortunately, several instruments exist for gathering such data about partnerships. These can be readily adopted or adapted to respond to several evaluation questions that correspond with the concepts outlined in **Figure 6.2.**

In **Appendix H**, we provide a table that lists several existing data collection instruments for partnership evaluation, their source, and identify the concepts from **Figure 6.2** that are covered within the instrument. It is important to recognize that the validity of instruments is dependent upon context. Therefore, even if you select an instrument that has been validated it will be important to validate the instrument again within the context of the specific partnership you are working with. If you wish to read more about partnership data collection instruments and their validity, a good source is Granner and Sharpe (2004).

In addition to some of the standard data analyses we have covered thus far, it is important to consider whether there are options for analysis that are particularly well suited to the topic. One potentially useful data analysis method (depending upon the evaluation questions of interest) for evaluating partnerships is Social Network Analysis (SNA). SNA can be useful for examining relationships, understanding how those relationships produce an effect, identifying important members in a network, understanding the capacity of a network to achieve a goal, tracking changes in a network over time, and understanding the connection between a network and outcomes (Honeycutt, 2009; Varda & Sprong, 2020). For more information on how to use SNA in program evaluation, see this brief.

A Glimpse into Practice
Focusing the Partnership Evaluation Questions and Methods

The reorganization of the Healthcare Systems Workgroup has been completed. There are many aims of this reorganization including an increase in the diversity and engagement of the members, enhanced coordination of members' programmatic-related efforts, and improvements in the coordination of public health-related efforts across the jurisdiction. An evaluation of the reorganization was prioritized in the strategic evaluation plan and the Evaluation Planning Team has decided that the specific questions of most use at this point in the reorganization effort relate to the overarching questions of *Who is involved?* Specifically, they aim to answer the following evaluation questions:

- In what ways has the membership of the workgroup expanded as result of our reorganization efforts? Do these align with our needs to engage more individuals from the healthcare sector? Are there any specific groups that we still need and are missing? (membership composition)
- Among those who have newly joined, what is their level of engagement over the first 6 months? How does this compare with existing member's involvement during the same 6-month window? (level of involvement)

The Evaluation Planning Team refers to **Appendix H** and notices that there are several existing instruments that may include questions pertaining to the constructs of interest. The instruments include the Coalition Self-Assessment Survey II, which includes questions about membership composition and level of involvement. The team moves forward with reviewing this and other instruments to see if the questions match their needs, and whether the type of data collection approach (i.e., survey, interview) aligns with the culture of the partnership.

Justifying conclusions (Step 5) and Ensuring use and sharing lessons learned (Step 6). The final steps of the evaluation process will be quite similar for partnership evaluations as they are for other types of evaluations. Following the general procedures we shared in **Chapter 4** will help ensure that the Evaluation Implementation Team generates an accurate synthesis of the findings and disseminates this synthesis to key audiences in a manner that is most useful to them. As a reminder, when developing products to communicate about the evaluation findings be sure to loop back to the stakeholders identified under Step 1 to ensure coverage of potential audiences.

A Glimpse into Practice
Justifying Conclusions & Acting on the Partnership Evaluation Findings

Approximately three months after the reorganization of the Healthcare Systems Workgroup began, the effort seemed to be making a difference. Anecdotally, members of the partnership and the workgroup shared insights with each other about the rather remarkable increase in attendance since the reorganization and the rich conversations taking place. The partners who were aware of the evaluation were eager to see if the evaluation findings aligned with these observations, but were a little disappointed that they would have to wait until after the six-month mark to have enough data in to learn more about whether the level of involvement among new members was as high as they perceived.

Fortunately, the partnership was able to acquire insights from the evaluation about another important topic relevant to the reorganization at this time—the new membership composition. Several key health professionals had been missing from the workgroup prior to the reorganization, and efforts to improve the workgroup composition were successful in securing participation from the local nursing association, a representative from two of the most frequently used insurance companies in the state's most populous cities, and a well-known emergency physician who also currently served in several leadership roles for the state medical association. These changes were viewed as successes by the broader partnership because the individuals who agreed to participate had the authority and connections within the state's healthcare community to influence change – something that had been missing from the previous workgroup composition.

Accompanying these successes in the new membership composition were some areas for improvement. For instance, the partners viewed the primary school system as an essential player in successfully coordinating asthma care in the state. Participation from the State Department of Education or individuals influential within the educational sector was not secured during the reorganization. It was unclear from the data collected to date why representatives from the educational sector were so difficult to engage in the partnership. Partners and the Evaluation Implementation Team realized that the Evaluation Planning Team had not foreseen this information need: the evaluation questions did not ask why individuals who were invited declined.

The group collectively agreed that this evaluation question needed to be added into the mix, and that it was feasible to do so given the remaining budget – a few brief semi-structured telephone interviews with those who declined may point to ways to incentivize engagement from this, and other, groups that were proving challenging to engage. The Evaluation Implementation Team crafted this new evaluation question—Why do potential partnership members, when invited, decline the invitation?—and started to draft the semi-structured interview protocol right away and logged the change to the original evaluation plan along with the rationale so this could be flagged with the Strategic Evaluation Planning Team during its annual meeting in the Fall.

Surveillance Evaluation

According to CDC (2014), public health surveillance "is the ongoing, systematic collection, analysis, and interpretation of health data, essential to planning, implementation and evaluation of public health practice, closely integrated with the timely dissemination of these data to those responsible for prevention and control" (p. 8). Surveillance data help public health program staff members and community partners with designing, refining, and targeting interventions, that is, in using data to guide strategic action. Surveillance data also clarify trends and patterns in the incidence and prevalence of health outcomes and identify associated risk factors across populations, places, and time. Results from surveillance analysis are used to raise awareness among key stakeholders about the impact of health conditions. Often, public health programs create surveillance products like reports, fact sheets, briefs, newsletters, and online maps and tables to share information with stakeholders to highlight the patterns and trends identified through public health surveillance. CDC's Morbidity and Mortality Weekly Report (MMWR) surveillance summaries provide examples of one type of surveillance output.

Public health surveillance differs vastly in its data sources, structure, and timing depending upon the type of health condition under surveillance. For instance, there are several health conditions within the U.S categorized by public health authorities as nationally notifiable conditions. Most of these conditions are the result of infectious diseases (e.g., cholera, dengue, and measles), however, some are not (e.g., cancer and carbon monoxide poisoning). Healthcare professionals working within medical facilities, at a local level, are required to report instances of these health conditions to their local or state health departments. The data then flows from these health departments to CDC. Furthermore, several of these health conditions have a definitive laboratory test, so confirmation of the disease is possible.

Several other health conditions are also under public health surveillance. However, these other conditions require the collection and analysis of many data sources from multiple organizations, outside of the state or local health department, to paint as complete a picture as possible about the occurrence of the health condition and associated risk or protective factors. Examples of such health conditions include chronic diseases such as asthma. Collaboration with outside data owners can make this type of surveillance process complex.

In 2001, CDC published *Updated Guidelines for Evaluating Public Health Surveillance Systems* (Guidelines), which described a model evaluation of a public health surveillance system. The Guidelines suggested that evaluations of surveillance systems should examine surveillance system usefulness and nine surveillance system attributes. The attributes are simplicity, flexibility, data quality, acceptability, sensitivity, predictive value positive, representativeness, timeliness, and stability. The 2001 Guidelines are a useful resource for planning **surveillance evaluations**. However, these guidelines focus on the data and the system used to collect and manage it. For many public health programs, surveillance typically refers to a more expansive set of activities, such as appropriately sharing information and using it to guide program decisions. We leverage this more expansive description of public health surveillance in this section, while also recognizing the utility of the Guidelines for evaluation within this context.

Identify who to engage in planning and implementation (Step 1). The first step in evaluating public health surveillance activities is to engage stakeholders. But you might wonder who the stakeholders would be for an evaluation focused on surveillance efforts. Similar to partnership evaluation, it can be helpful to identify stakeholders within the three overarching categories described in **Chapter 4**.

- ***Primary stakeholders.*** Who is responsible for the public health surveillance efforts in your jurisdiction? Who holds the authority to make changes should they be necessary? These individuals are likely to include the epidemiologist(s) who oversees the surveillance system; leaders within the health department; and, if the sources of data contained within the surveillance system are distributed among several organizations, these various data owners (e.g., Behavioral Risk Factor Surveillance System (BRFSS) coordinator for the state, hospital association). Should the evaluation's findings indicate that changes are needed in the data, the analyses performed, or the products then these individuals may need to be involved in some manner.

- ***Secondary stakeholders***. It may be challenging to consider who would be affected if the surveillance system changed because of findings from the evaluation. If the focus of the evaluation is on the use of the surveillance products, consider who may use these: possibly program managers, public health partners, public or organizational policy makers. If the focus of the evaluation is on the quality of the data included in the surveillance system, consider who makes use of the data. If the surveillance data are located at the state level it is possible that local-level jurisdictions make use of the data for their public health planning, making them a potential secondary stakeholder group. Researchers may request access to the data or obtain it through a special portal online, thus making them another possible secondary stakeholder group.
- ***Tertiary stakeholders.*** Who might be interested in the results of the evaluation because of the potential to leverage the lessons learned to make changes in their own processes or procedures? Possible tertiary stakeholders lie within the answer to this question. Depending upon the focus of the surveillance evaluation it is possible that public health programs that use similar data sources or have a surveillance system that is structured similarly to the one under review may have an interest. Specifically, epidemiologists who work in the same or other organizations on the same type of surveillance system certainly may learn from your efforts. Additionally, do not be afraid to think broadly. It may be that professional associations such as the Council of State and Territorial Epidemiologists could leverage the findings from surveillance evaluations to inform several efforts including the development of new, or adjustment of existing, recommendations for public health surveillance including case definitions.

A Glimpse into Practice
Surveillance Evaluation Stakeholders

The Strategic Evaluation Planning Team for a state-based asthma program decided, after much discussion and deliberation, that it would be important to better understand the extent to which various intended audiences around the state were using the annual surveillance report in ways that were likely to impact the burden of asthma. Now the task is to identify members of the Evaluation Planning Team who will help to further refine the details of the evaluation drafted by the Strategic Evaluation Planning Team.

The evaluator who will be leading the evaluation effort met with the program coordinator to consider who to engage. The epidemiologist of the program was their first consideration since this individual knows the surveillance data better than anyone and was the first to raise questions about whether and how the report was being used. The team also considered inviting individuals who comprise the audience for the report – there are a few members of the statewide partnership who previously indicated that they have used the information contained in the report to help write grant and a few who noted that they don't make use of the report. For an outsider's viewpoint, they decided to request participation from a potential tertiary stakeholder—an epidemiologist who is in charge of a surveillance system for heart disease—since this public health program also disseminates an annual burden report based upon similar public health surveillance data.

Describe the surveillance effort (Step 2). What might a logic model for a surveillance system look like? What are the intended outcomes? How is this change facilitated? The answers to these questions may not be readily apparent. We provide an example logic model for the surveillance efforts pertaining to asthma, in **Figure 6.3**. This logic model is the result of the efforts of a CDC-State Surveillance Evaluation Workgroup convened by CDC's Asthma and Community Health Branch between 2006 and 2008. Public health programs that include a robust surveillance function and a system that consists of several different types of data from multiple sources may find this model useful to adapt for their purposes. Some aspects of the model may also be helpful for programs that house systems focused on infectious diseases or nationally notifiable conditions.

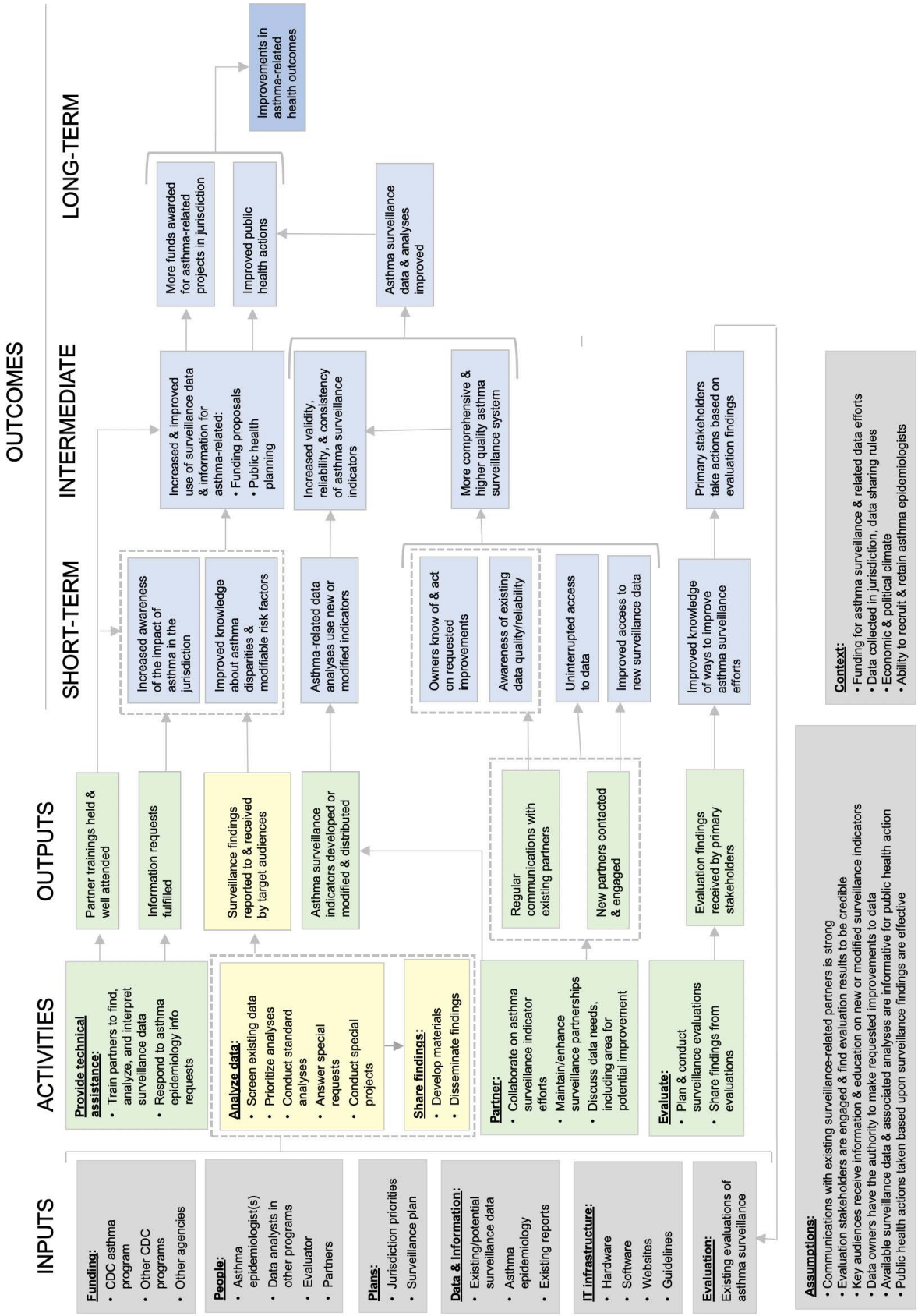


Figure 6.3 Example Logic Model for Public Health Surveillance, Asthma

An overarching surveillance logic model, like the one depicted in **Figure 6.3** can play many roles throughout the evaluation process. It can help the Evaluation Planning Team identify an evaluation focus as they design the evaluation in Step 3. It can also be used as a communication tool to inform surveillance stakeholders about the scope and processes of a program's surveillance efforts. However, similar to what we saw for partnership evaluation, for any one specific evaluation it may be helpful to develop a nested logic model that zooms in on the specific inputs, activities, outputs, or outcomes that are most closely associated with the focus of the evaluation. **Figure 6.4** provides an example of a nested surveillance logic model that portrays details about sharing surveillance findings through a report.

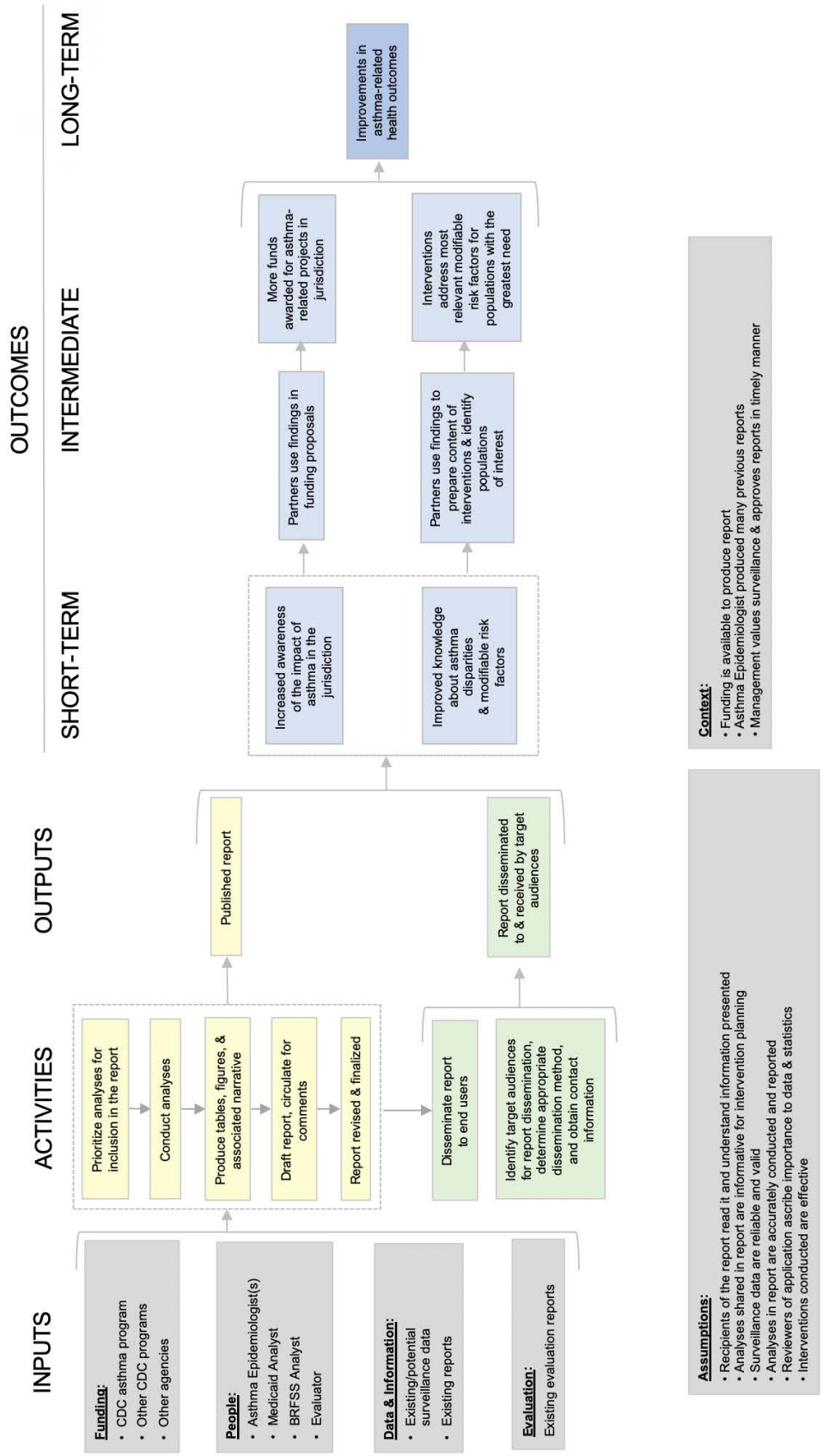


Figure 6.4 Example Zoomed-In Logic Model for Asthma Surveillance Report Dissemination and Use

Identifying the focus of the evaluation (Step 3). It is tempting to think of public health surveillance as just data. However, a quick review of the details included in **Figure 6.3** demonstrates that there is much more to surveillance—surveillance comprises an entire effort on behalf of public health programs that is multifaceted and offers several avenues of exploration in an evaluation. We provide **Table 6.1** to stimulate thinking about the realm of possible evaluation questions for an evaluation focused on public health surveillance.

Table 6.5 Sample Surveillance Evaluation Questions

Logic Model Component	Evaluation Question Examples
Inputs	• To what extent is the existing staffing structure sufficient to carry out the planned activities? Are there opportunities to supplement existing staff members in a manner that is not too costly? • How could the timeliness, completeness, accuracy, and consistency of our existing surveillance data be improved? • In what ways might the existing IT infrastructure be improved for better data collection and management?
Activities	• How could we better educate our stakeholders on interpreting data or findings through our technical assistance efforts? • What data or analyses are missing or lacking? How could the analysis of surveillance data be improved (e.g., faster turnaround, additional indicators)? • How could the surveillance data and analyses be enhanced to support identification of factors that may influence disparities in access to high-quality care? Disparities in health outcomes? • Are the surveillance data presented in a manner that is easily understood? How could we improve the methods or formats used to disseminate data or analytic findings? How could we improve the reach of our dissemination efforts? • How might we more effectively engage our partners and recipients in using our surveillance data? In what ways can we improve upon our existing partner relations and communications?
Outputs	• Were partner trainings held, and if so, which partners attended and why? What additional trainings are needed? In what ways did or didn't the trainings meet the attendees' needs? • Which intended audiences do our surveillance materials reach? Are these materials further disseminated or shared by our partners? If so, with whom are they shared? • What evaluation findings of our surveillance efforts were shared this year? With whom? How? How could we change these communications to better meet our partners' preferences?
Outcomes	• For what purposes are our surveillance data used? How, if at all, are they used in planning and guiding strategic action? • To what extent has the use of surveillance data by our key stakeholders improved because of our partner training? • In what ways has the use of our surveillance data resulted in increased funding for asthma-related projects? • In what ways have our standard surveillance indicators improved since the publication of our last report? How might we continue to increase their validity and reliability?

Gather and analyze credible evidence (Step 4). Unlike in partnership evaluation, where there are a variety of partnership-specific data collection instruments available and particular analytic approaches (such as social network analysis) may be appropriate, the data collection and analysis tools used for surveillance evaluation will typically consist of those already covered in **Chapter 4**. We encourage you to select the data collection and analysis methods that are most closely aligned with the evaluation questions. Sometimes examples can be useful in thinking through our options when it comes to evaluation, including the types of data to collect and analyses

to conduct. Therefore, we offer the following as examples that you may wish to consult when formulating surveillance evaluations.

- For an example of the use of all 10 attributes outlined in the CDC Guidelines in evaluating a chronic disease surveillance system, see Reeves, M. J., Lyon-Callo, S., Brown, M. D., Rosenman, K., Wasilevich, E., & Williams, S. G. (2006). Using billing data to describe patterns in asthma-related emergency department visits in children. *Pediatrics*, 117(4), S106-S117. https://pediatrics.aappublications.org/content/pediatrics/117/Supplement_2/S106.full.pdf
- For a detailed example of a surveillance evaluation examining the attribute of data quality, see Brunner, W. M., Ross, S. K., & Johnson, J. E. S. (2009). Review of the asthma mortality rate for Minnesota residents aged 55 years or older, 2004-2005: When death certificates deserve a second look. *Preventing Chronic Disease*, 6(3), A92. https://www.cdc.gov/pcd/issues/2009/jul/08_0154.htm

A Glimpse into Practice
Surveillance Evaluation Questions and Methods

The Evaluation Planning Team identified several questions to examine as part of the surveillance evaluation focused on the use of the annual report: (1) Who accesses the report?" (2) "Which, if any, intended audiences are able to access the report and other surveillance data on our website?" (3) "To what extent, and in what ways, are the report and other surveillance data products easy to understand?" and (4) "How do stakeholders use the report and other data?" After some additional discussion they opt to fill in another information gap, and include a fifth question: (5) "What specific information do the intended audiences of the report need and what opportunities exist to make the information presented easier to understand?

In arriving at this final set of questions, the team considered a critical element: how the evaluation findings can be used to improve the annual report in meaningful ways. They discussed a variety of topics: how much flexibility the program has, or doesn't, to make changes to future reports, how to balance modifications that might be beneficial to some audiences and a drawback for others, and whether funds are available to share the report in a different format if this is a recommendation that surfaces from the evaluation findings. Considering how the findings can be used is important. What is feasible in the given context? What is on the table in terms of changes and off the table?

Alongside these important conversations about the utility of answering specific evaluation questions the team also had to consider what data collection methods to use in responding to the questions. Ultimately, they decided to leverage existing web analytics maintained in the state health department's information technology office to get a sense of who was accessing the report. They also used an online survey of intended audience members to learn more about all of the evaluation questions. Having few monetary resources and limited time to support this evaluation forced the team to think critically about what information would be nice to know versus information they needed to know. The evaluation methods had to be feasible and capable of producing insights that were accurate enough to inform decision making.

The team recognized that more detailed insights would be needed than could be produced through a survey; having read Chapter 4 of this text they realized that surveys were best for obtaining high-level information from many people. They decided that selecting a small, purposeful sample from all intended user groups to participate in a 20-minute telephone interview was possible given the available resources and the likeliness of providing invaluable insights.

The Evaluation Planning Team also determined that the organizational capacity was sufficient to uphold the protection of the survey and interview data provided by participants, namely the ability to secure such data on password protected computers and using encryption software when necessary to share files among Evaluation Implementation Team members. After one more reflection on the extent to which they had balanced the standards of utility, feasibility, propriety, accuracy, and evaluation accountability, the Evaluation Planning Team documented the evaluation questions and proposed data collection methods in the individual evaluation plan.

Justifying conclusions (Step 5) and ensuring use and sharing lessons learned (Step 6). As was the case with partnership evaluation, the final steps of the evaluation process will be quite similar for surveillance evaluations as they are for other types of evaluations. Following the general procedures we shared in **Chapter 4** will help to ensure that the Evaluation Implementation Team generates an accurate synthesis of the findings and disseminates this synthesis to key audiences in a manner that is most useful to them. Again, it is important to remember when developing products to communicate about the evaluation findings to loop back to the stakeholders identified under Step 1 to ensure coverage of potential audiences.

A Glimpse into Practice
Justifying Conclusions & Acting on the Surveillance Evaluation Findings

The evaluation results were finally in after six months of data collection activities and analytic efforts. To facilitate interpretation and use of the findings, the Evaluation Implementation Team invited the asthma epidemiologist, the health education specialist for the asthma program, and several members of the state asthma partnership to join them in a 2-hour working session. Fortunately, the Evaluation Planning Team had thought about communications and reporting and in the evaluation plan, they documented several ideas for who to invite to such as session– saving some time and energy.

The working session was well attended, likely in part due to the participants having been regularly updated about the evaluation progress by the Evaluation Implementation Team. Invited participants knew the time was coming to interpret the findings and were excited to be digging in. The Evaluation Planning Team had identified some ideas for how to interpret "success" using the annual surveillance report. The indicators and qualitative standards were reviewed at the beginning of the working session. However, it was also made clear that the attendees' perspectives on what constituted successful use could and should be raised during the day's discussions.

The findings from the evaluation generally suggested that the report and other surveillance products were very easy to access. Most individuals who accessed the products strongly agreed that they were very user-friendly and easy to understand. Survey respondents and interviewees reported that they used the information contained in the surveillance products primarily for writing grants to acquire additional funding to support asthma-specific programming across the state. The working session attendees all agreed that these insights highlighted that the time, energy, and thoughtfulness of the state asthma program's efforts to analyze and disseminate relevant surveillance findings to key audiences had paid off.
However, one finding emerged from the evaluation that pointed to an area that needed further investigation and action. Working session attendees from two local health departments along with representatives of community-based organizations in two underserved areas of the state noted a consistent pattern in survey responses (both quantitative and qualitative). These participants noted inadequacy in the underlying surveillance data to disaggregate findings based on (1) social determinants of health that have an impact on individual's ability to acquire adequate medical care and access resources helpful to asthma self-management (i.e., income, transportation availability/ affordability, insurance coverage), (2) race or ethnicity—the team noted that several respondents called to attention to the fact that they anecdotally knew sub-populations were hard-hit by asthma in their communities (in particular community members of Hmong and Puerto Rican origin), but these populations were not represented in the analyses, and (3) geography—respondents highlighted that there were no findings available at the zip-code or community level. Interviewees from localities throughout the state noted that the lack of information about these factors meant that although the accessibility and understandability of the surveillance products were amazing, they were not useful for local-level planning.

The Evaluation Implementation Team thanked the working session participants for their time and critical insights. As a follow-up item, the team reached out to members of the Strategic Evaluation Planning Team to schedule a meeting to discuss the findings from the evaluation. Their hope was that a more in-depth evaluation examining the utility of state surveillance data for local-level decision making could be added to the priority list of evaluations in the revised strategic evaluation plan. The team hoped the new evaluation would help enhance existing data sources

Chapter Summary

In this chapter we provided insights that can be used to tailor evaluations specific to two important components of public health programs—partnerships and surveillance. We articulated the importance of evaluating each of these program elements, given that they comprise the infrastructure that helps to identify the needs to be addressed through public health interventions and support the conduct, and in some cases the evaluation, of these interventions. To provide ideas for future evaluations on these topics, we provided examples of logic models and an array of potential evaluation questions to stimulate thoughts about what these might look like in your context.

Review Questions and Skill Building Exercise

Review Questions

1. As we have learned in previous chapters, there can be a wide range of stakeholders for any evaluation. For both public health partnerships and surveillance, can you list up to three potential stakeholders for each evaluation? What might the stakeholder interests in these evaluations be?
2. Articulate at least one process evaluation question that could be of interest in a partnership evaluation? How about for a surveillance evaluation?
3. Articulate at least one outcome evaluation question that could be of interest in partnership evaluation? How about for a surveillance evaluation?

Skill-building Exercise

You have been asked to lead an Evaluation Planning Team for an evaluation identified as a priority in the Strategic Evaluation Plan (SEP) for your county health department. The evaluand is the county health department's coalition that is charged with putting into place measures to adapt to the effects of climate change. The county has received a grant from the state to help create a 10-year county-wide plan and to start some early implementation efforts. The grant requires that the coalition conduct at least one evaluation during the five-year grant cycle. Given that the success of the initial stages of this effort require a high-functioning coalition, the focus of the evaluation will be on how the coalition interacts (see **Figure 6.2**).

Your task is to engage members of the Evaluation Planning Team in focusing the evaluation. Specifically you need to (1) identify the purpose of the evaluation and articulate this in a written purpose statement (see **Chapter 4** for additional details), (2) prioritize no more than five evaluation questions for examination in the evaluation, and (3) design one or more data collection methods to provide answers to each of these evaluation questions. Remember, the data collection methods proposed must be feasible to implement and produce accurate insights. The evaluation can take no longer than six months to implement and must be conducted with $10,000 USD or less.

CHAPTER SEVEN

Evaluating Public Health Interventions

CHAPTER SEVEN: Evaluating Public Health Interventions

Interventions are the lifeblood of public health programming. They are "action[s] or programme[s] that aim to bring about identifiable outcomes…" and are "…applied to many, most, or all members in a community, with the aim of delivering a net benefit to the community or population as well as benefits to individuals" (Rychetnik et al., 2004, p. 540). As we will discuss in this chapter, the landscape of public health interventions is vast. Interventions are designed to reach many different types of audiences at varying levels of geography, they are conducted in a wide variety of settings, and they use one or more strategies to facilitate positive change in the public's health.

Given their central importance in public health programming, and their potential to lead to large, lasting changes in health outcomes it is essential to evaluate interventions. The use of findings from these evaluations can promote greater efficiencies in the use of public health resources, lead to critical changes in program delivery that increases the reach and impact, and foster evidence-informed decision making about whether to scale the intervention or reduce or eliminate funding for it. In this chapter, we build upon the basics presented in Section II by sharing additional thoughts about how to approach evaluation when the evaluand is an intervention.

By the end of Chapter Seven, the reader will be able to

- Describe why evaluability assessments are important in intervention evaluation.
- Explain what to look for when assessing the context of an intervention.
- List up to two types of visual depictions of program theory.
- Write process and outcome evaluation questions.

The Wide Variety of Public Health Interventions

Public health interventions are conducted in a wide variety of settings (e.g., communities, worksites, schools, and medical clinics). Interventions span geographic levels, with implementation taking place at the level of cities, zip codes, and counties to one or more global regions. In addition, interventions may use multiple strategies to foster intended change. For example, an intervention may use public policies (e.g., smoke-free policies), organizational policies (e.g., worksite wellness policies), health communication or social marketing campaigns, healthcare services (e.g., vaccinations), and health education sessions (e.g., asthma self-management trainings in schools). Public health programs may implement one intervention strategy or use a multi-intervention approach. In a multi-intervention approach, a suite of interventions is implemented to effect positive change.

The intended outcomes of these interventions are equally as vast as the intended audiences and intervention strategies. Examples of commonly anticipated outcomes at the individual-level include changes in motivation, attitudes, knowledge, skills, and behaviors. Similar outcomes may also be desired at a population level (e.g., improvements in disease incidence, prevalence, morbidity, and mortality). Contextual conditions that impact population health status are also often the intended focus for public health interventions. Contextual conditions include environmental exposures (e.g., indoor and outdoor air quality), social determinants, and access to high-quality, affordable health care. Additional outcomes of interest often incorporate cost savings for healthcare systems and sustained capacity to support public health programming. In the sections that follow, we leverage the general process described in **Chapter 4** to describe how you can plan and implement a high-quality evaluation for any intervention you may come across.

Pre-evaluation

It is important to understand that it is possible to design and conduct an evaluation of an intervention (or any other evaluand) at any point in time, but it may not be a wise decision to do so. Understanding the context and conditions in which the intervention and evaluation will be implemented can help evaluators and stakeholders decide if the time is right for evaluation to proceed. For instance, the contextual conditions are not right for conducting an outcome evaluation if the goals of the intervention are unclear or it is not possible to articulate a shared understanding and realistic description of how intervention activities will lead to the intended outcomes. Other conditions that warrant a pause prior to committing to evaluation include requesting results of an outcome evaluation at a time when the intervention is not mature enough to have produced the intended outcomes, commissioning an evaluation when there are serious implementation challenges taking place (e.g., insufficient funds, staffing, participation, or need to support the intervention), sponsoring an evaluation that has a larger scope than the resources available (e.g., funding, staff member availability, data), and embarking upon an evaluation when anticipated end users are uninterested or do not see clear value.

Evaluability assessments (EA) help to detect these types of issues and more (Wholey, 2015). They are conducted prior to engaging in a full evaluation to determine the utility and feasibility of conducting a full evaluation. Since conducting an evaluation requires significant resources, EAs can assure that the investment will be made wisely. If the intervention is ready for evaluation, the results of the EA can be used by the evaluator and stakeholders to inform evaluation design and promote clarity of evaluation use. If the intervention is not ready for evaluation, the evaluator can work with the intervention designers, implementers, and other stakeholders to leverage the EA-generated insights to make the adjustments needed. For instance, perhaps the evaluator finds that the clarity of the intervention description is lacking. The goals of the intervention are vague, and it is challenging for stakeholders to articulate a common vision for how the intervention activities will produce desired changes. Lack of clarity can make it difficult, if not impossible, to design and successfully implement an evaluation of an intervention that aims to improve processes or examine the extent to which an intervention is making a difference. With a lack of clarity, how would we know what to measure? Or what success looks like? Recognizing this early in the evaluation process is critical. With the information gleaned from the EA, the evaluator and stakeholders can work together to develop a common vision, process, and roadmap for the intervention—making it more likely that the evaluation will ultimately be successful. See additional information on evaluability assessments including steps to take in carrying them out.

Assess the Context (Step 0)

Public health practitioners design and implement interventions in many different contexts. Deeply embedded in these contexts are rich histories that lead to entrenched power structures, values, beliefs, and norms that can affect how the intervention is designed, implemented, and evaluated. Evaluation teams should not be naïve to these contexts. Ignoring context can lead to evaluation results that are potentially damaging, inaccurate, misleading, or that, at a minimum, provide recommendations or suggestions for action that are unrealistic or not as helpful as they otherwise could have been.

Evaluators who are unfamiliar with the context(s) where the intervention will be implemented, might try to acquire a basic understanding of the context by visiting the location(s) in person prior to or during the very early stages of planning the evaluation. Evaluators should take time to explore and observe different neighborhoods, historical landmarks, and informally chat with people. What do the people they encounter like or dislike about their community? Where do people gather? What issues are they facing? What is the story of this place (historically)? Evaluators can also look for and examine publicly posted signs. What is highlighted?

It may not be feasible to visit the location(s) where the intervention will be conducted. Perhaps it is a multi-site intervention that spans multiple jurisdictions, states, or even nations. Perhaps there is insufficient funding or time available to make a visit. In these cases, take time to consider other possibilities. For example, you could engage in informal phone conversations with key community members, visit websites that describe the history of the location, review postings on social media sites hosted by the community, or review recent local news online

(e.g., read local newspapers, view recent news broadcasts, view community meetings online or review minutes from city council, or similar, meetings). Evaluators may also find it valuable to analyze existing, publicly available data that provide insights about the demographics of the area. Such data may have already been analyzed by reputable organizations and exist in online dashboards or reports.

It may be the case that the Evaluation Planning or Implementation Team members are already knowledgeable about the contexts within which the intervention takes place. This is certainly possible when the team members reside within a local or state health department, community hospital, or other organization which regularly gathers data about local contexts including existing health disparities and related social determinants. In these cases, it is important to reflect upon the comprehensiveness of the picture one has about the context. Does the information primarily stem from analyses of existing data sources? If so, these analyses may present a fairly complete picture of the general distribution of characteristics of an area but will not lend insights about people's lived experiences, which can still be explored using several of the previously mentioned techniques. Does the individual live in the context where the intervention will take place? If so, the lived experience of one individual is not the same as others within a community. There may be a need to fill in gaps by visiting establishments not regularly visited, talking with other community members who engage in different types activities or work or volunteer in different sectors or parts of the community than the individual is familiar with.

The "How" of Practice
Situational Awareness

According to AEA's Evaluator Competencies (2018), context involves numerous factors:

- site/location/environment
- participants/stakeholders
- organization/structure
- culture/diversity, history/traditions,
- values/beliefs
- politics/economics
- power/privilege (p.3)

This may seem like a lot of factors to consider, but the benefits of conducting a comprehensive assessment far outweigh the time and effort needed to do one. Importantly, knowing the context enhances situational awareness, which enables evaluators to understand program dynamics and foresee potential issues, thereby saving time and other resources at later stages of the evaluation.

These types of activities can give the evaluator a better sense of the places in which the intervention will be conducted including factors that may or may not affect its implementation and effectiveness in producing the desired outcomes. As part of a critical-reflection exercise, the evaluator might consider the ways in which their lived experience would be an asset or present challenges to performing this evaluation (AEA, 2011), and with respect to the later, consider how (or if) they might overcome the challenges. Coupled with critical-reflection about an evaluator's own culture, history and experiences, values, and beliefs the evaluator can more effectively determine whether they are well suited to perform this evaluation, if they might need additional team members to design and implement the evaluation successfully, or if they should make a recommendation for another evaluator.

Who to Engage in Designing and Implementing an Intervention Evaluation (Step 1)

There is no set formula for identifying who to engage in an intervention evaluation. Given the breadth of interventions in public health programming, evaluation stakeholders will vary greatly depending upon the intervention that is the focus for the evaluation. Like the approach used in partnership and surveillance evaluation, evaluators can identify stakeholders of an intervention evaluation by considering the following categories:

- ***Primary stakeholders.*** When the evaluation results are finalized, who is positioned to make use of them to modify the intervention? Who has the authority to continue, cancel, or scale the intervention? The answers to these questions point to the primary stakeholders of the evaluation and often include the funders of the intervention, the leadership within the organization delivering the intervention, as well as the designers and implementers of the intervention.
- ***Secondary stakeholders.*** In the event that evaluation findings are used by the primary stakeholder to modify, cancel, or scale the intervention, who stands to benefit? Who may be harmed by such decisions? Most often the answer to these questions will be the individuals whom the intervention is designed to serve. However, depending upon the circumstances, other groups, such as the staff members delivering the intervention, may also fall into this category.
- ***Tertiary stakeholders.*** This stakeholder category, in the case of interventions, is quite broad. As we defined in **Chapter 4**, tertiary stakeholders are individuals who may have a general interest in the results of the evaluation. Other jurisdictions, similar to the one in which the intervention is implemented, may be tertiary stakeholders of the evaluation. Other tertiary stakeholders may be evaluators and researchers who perform meta-analyses in order to synthesize existing evidence about the effectiveness of interventions (e.g., individuals involved in the Guide to Community Preventive Services).

The people who are served by an intervention are critically important evaluation stakeholders. We have included them above as *secondary stakeholders*, however, the nature of the intervention may influence which category they fall into. Irrespective of the type of stakeholder group they are classified under, it is the Evaluation Planning and Implementation Team's job to ensure stakeholders' voices are heard within the evaluation. Additionally, though they may not fit neatly into one of the stakeholder categories, it is also critical to include individuals who reflect the diversity within the community or communities where the intervention is implemented. Try to employ strategies to assure that all perspectives are respected in the design, conduct, and use of the evaluation. Doing so is respectful of the community within which the evaluation act is being conducted, can promote the cultural responsiveness of the evaluation, and increases the likelihood of fostering all evaluation standards that reside at the center of the CDC Framework. Remember to discuss with all stakeholders their specific needs for information about the intervention, when or if they can use the information, and any contextual constraints on the evaluation implementation timeline. For example, evaluations in school settings need to account for the school calendar.

The "How" of Practice
Cultural Responsiveness

Engaging stakeholders in all aspects of an evaluation is a lofty feat and may be unattainable in some contexts. CDC's *Practical Strategies for Culturally Competent Evaluations* suggests creating a diverse advisory team to support planning and implementation and to advise on when to engage the larger group of stakeholders. Ensuring that the advisory group reflects the diversity of the community increases the likelihood that the perspectives of all cultural groups in the community are represented (CDC, 2014, p.11).

Describing the Intervention (Step 2)

> "A theory of change explains how to produce desired outcomes. It is explanatory. A logic model just has to be sequential (inputs before activities, activities before outcomes), logical, and reasonable. In contrast, theory of change must explain why the activities produce the outcomes" (Patton, 2014, p. 6).

Describing the intervention involves making implicit ideas explicit about why the proposed intervention activities will result in, or contribute to, the *intended* outcomes. As we mentioned in **Chapter 4**, describing the intervention typically involves developing a visual depiction of the *intended* causal pathways between inputs/resources, activities, and outcomes. This visual is often accompanied by a narrative description that provides further explanation. These visual depictions describe the **program theory** underlying the intervention. Program theory is a "plausible and sensible model of how a program [intervention] is supposed to work" (Bickman, 1987 as cited by Donaldson & Lipsey, 2006, p. 64).There are many different ways that program theory can be portrayed for an intervention, two common methods are logic models and theories of change.

There are several sources of information that may be helpful when identifying what elements should be included in the program theory diagram. Programmatic documents that describe the intervention, program designers' and stakeholders' experiential knowledge, published and grey literature describing evaluations of the same or similar interventions, evidence-based models of the same or similar interventions, and social science theories (e.g., Health Belief Model, Theory of Planned Behavior, Self-Determination Theory) are some potential sources of insight (Donaldson, 2007). Ultimately, these sources can be used to better understand how the proposed intervention activities (when implemented well) are supposed to result in the intended outcomes.

Figures 7.1 and **7.2** depict a hypothetical intervention that aims to decrease the occurrence of adverse asthma events (e.g., emergency department visits, hospitalizations, deaths) by improving the indoor air quality (IAQ) in individuals' homes. Both figures include the same elements; however, the elements are arranged differently. **Figure 7.1** depicts the intervention in a tabular format. The tabular format provides a line-listing of the processes and outcomes anticipated. This tabular figure may be helpful in situations where a program has been asked to monitor or account for the activities performed, outputs produced, and outcomes realized to date. Such requests typically include counting what has occurred to keep a record for accountability purposes. We generally do not recommend such models for use in evaluations as they do not provide a sufficient level of detail about the interconnections or intended pathways between inputs, activities, and outcomes.

Figure 7.2 presents a fairly detailed logic model that uses individual boxes and arrows to demonstrate the presumed causal pathways between the inputs that support the intervention, the intervention activities and their outputs, and a series of outcomes that ultimately contribute to the desired long-term change of having fewer adverse asthma events in the jurisdiction. In some respects this model resides in a grey zone between a logic model and a theory of change because the reader can readily see the causal sequence that is set in motion when the appropriate level of resources are available and the activities are conducted well. This model is often more helpful for framing process and outcome evaluation questions rather than fostering the development of monitoring questions only about "Did the trainings for code officers on the new IAQ housing codes take place as scheduled?" or "How many trainings were held?" Using a model similar to that reflected in **Figure 7.2**, stakeholders may be more inclined to ask questions about "In what ways, if any, did training for code officers on the new IAQ housing codes result in improved knowledge about how to enforce these new codes among those who attended?" or "In what ways, if any, did the adoption of a model housing code in the jurisdiction foster improved enforcement of the housing code?"

Process			Outcomes		
Inputs ➡	Activities ➡	Outputs ➡	Short ➡	Mid ➡	Long
Funding Staff Partners	• Conduct training on IAQ for: Apartment owners, maintenance providers, tenants • Conduct smoking cessation program for tenants • Hold collaborative meetings with city officials to enhance housing code • Training for code officers on new IAQ housing codes	• IAQ & cessation training materials • IAQ & smoking cessation trainings conducted & attended • Meetings held • Model IAQ housing codes • IAQ housing code training materials • IAQ housing code trainings conducted & attended	• Increased awareness of indoor asthma triggers • Improved understanding of methods to reduce exposure to triggers • Improved knowledge of harms of smoking & ETS exposure • Model IAQ housing codes adopted by the city housing authority • Increased awareness of new IAQ housing codes • Knowledge of how to enforce new IAQ housing codes	• Better maintenance of apartment complex • Tenants engage in activities to improve IAQ • Improved enforcement of housing codes • Reduced exposure to asthma triggers	• Fewer adverse asthma events

Figure 7.1 Basic Tabular Logic Model for a Hypothetical Indoor Air Quality (IAQ) Improvement Intervention

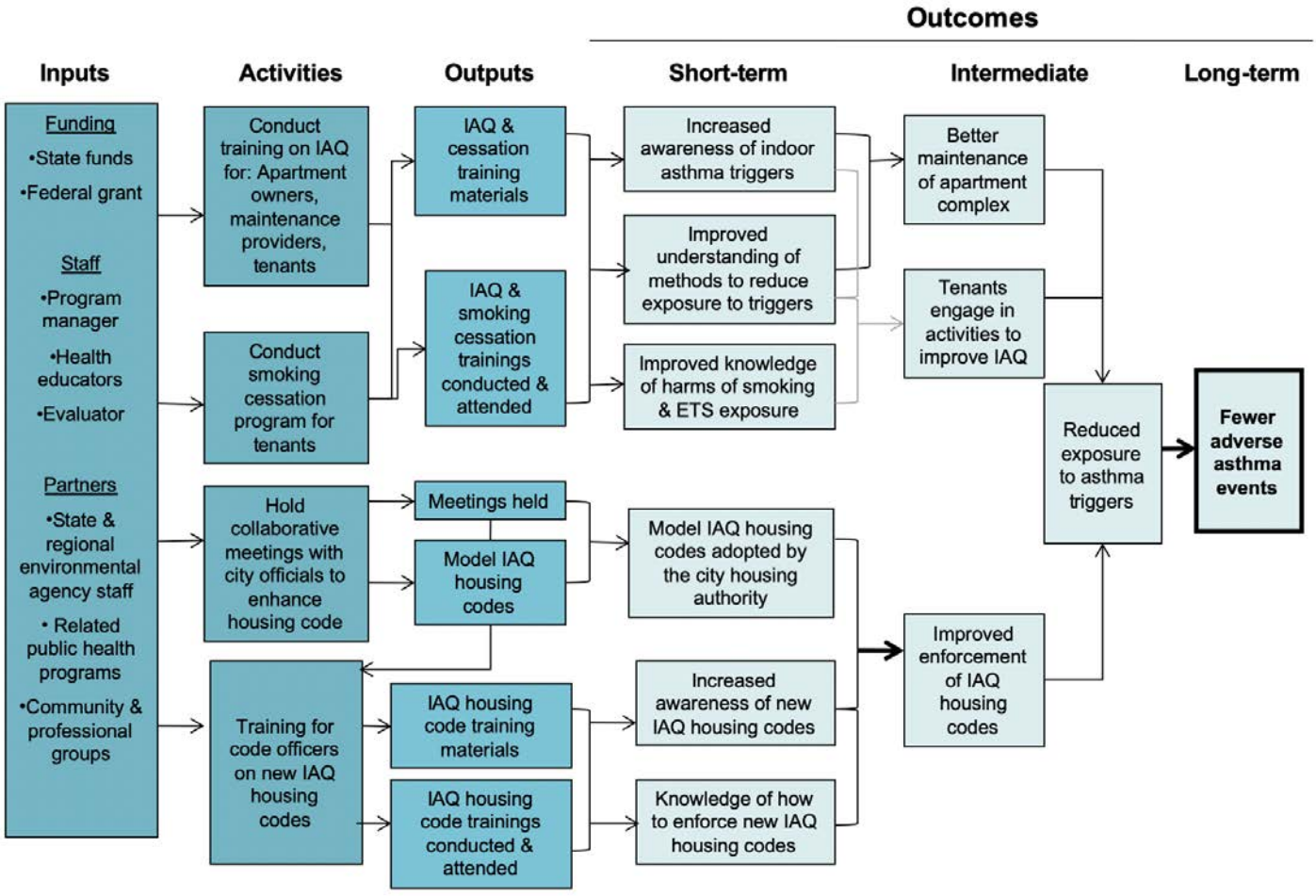


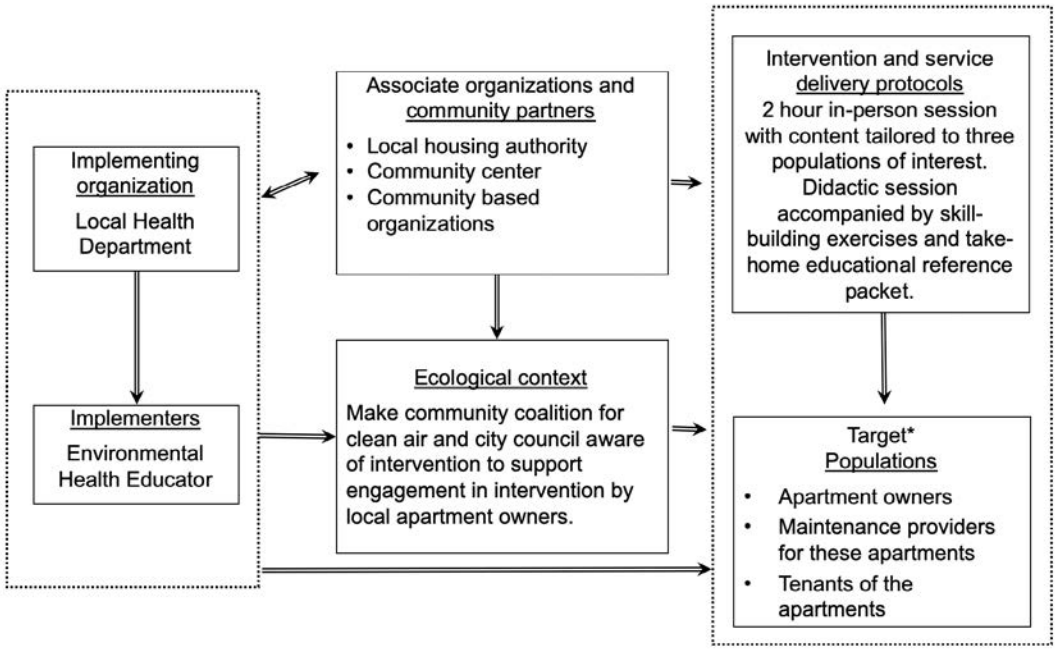
Figure 7.2 Detailed Logic Model for a Hypothetical Indoor Air Quality (IAQ) Improvement Intervention

Similar to our discussion of logic models in **Chapter 6**, you may find it helpful to develop a logic model, or other depiction, that zooms in on a specific component of the intervention. For intervention evaluation, we describe this type of zoomed in model in a few different ways. First, the intervention depicted in **Figure 7.2** is a *multi-component* intervention. A suite of interventions is described, each of which contribute uniquely to reducing individuals' exposure to indoor air asthma triggers; ultimately leading to fewer adverse asthma events. If the evaluation is focusing on one of the intervention strategies, such as the training for code officers on the new IAQ housing codes, it would be beneficial to either highlight or pull out that piece of the logic model on its own.

Second, it is often the case that the evaluation will focus on the process or outcome aspects of the intervention strategy. In these instances, evaluators and their stakeholder may find it helpful to zoom in on either the right- or left-hand side of the logic model and use other types of visual depictions to delve into additional details about each. We describe some of these possible approaches next.

Zooming-in on the Process

As mentioned in **Chapter 4**, the Action Model presented by Chen (2015) as part of Theory Driven Evaluation can be helpful in zooming in on program plans. The Action Model captures different aspects of an intervention that are not typically described in a logic model. Instead of focusing solely on the activities to be performed, an action model describes several elements of the plan for implementing the intervention. Specifically, Chen (2015) explains that action models could include information about the implementing organizations, the individuals implementing the intervention, any organizations that the implementing organization needs to coordinate with to successfully carry out the intervention, nuances about the context (ecological context) within which the intervention is conducted, details about intervention protocols or intended operational plans, and the populations who are to receive the intervention. **Figure 7.3** presents one possible action model for the IAQ training with various audiences (i.e., apartment owners, maintenance providers, tenants) included in the hypothetical intervention represented in **Figures 7.1** and **7.2**. We provide a description of the general plan for the intervention in the "Action Model Description" box.



Derived from template for basic Action/Change Model; Chen (2015), Figure 3.5

**We use the term "target population" here to align with Chen 2015. However, it is important to note that we generally do not recommend use of this term, and instead gravitate towards alternatives such as "intended population" or "intended audience".*

Figure 7.3 Example Action Model for Hypothetical Indoor Air Quality (IAQ) Training Intervention

Action Model Description

A ***health educator focused on environmental health*** and employed by the ***local health department (LHD)*** will deliver the Indoor Air Quality (IAQ) intervention. The local health department would like individuals who need to be aware of the potential impact that IAQ can have on individual's health (in particular asthma) and what can be done to improve IAQ in multi-unit dwellings to attend the training– a***partment owners, individuals who provide maintenance in these apartment buildings, and tenants of these apartments.*** Given this is their first time delivering these trainings, the LHD is starting small and will focus on the five largest apartment buildings in a small city within their jurisdiction.

To successfully carry out the intervention, the LHD will need to coordinate with other organizations in the community to make contact with or market the upcoming trainings to potential attendees. They envision working with the local housing authority to identify apartment owners and encourage their participation. They also believe that finding ways to market the training through the c***ommunity/recreation center and two active community-based organizations*** will solicit greater participation from residents of the apartment buildings in the event apartment owners are not supportive of the intervention. To further bolster support for the intervention and create an ecological context in which apartment owners feel more incentive or pressure to participate, the LHD will attempt to make influential groups within the community aware of the upcoming intervention –a ***community coalition*** that has served an educational and advocacy role in the community regarding clean air for the past 20 years and the ***city council.***

The training consists of one two-hour session. The session will have one hour of didactic training followed by one hour of skill-building exercises with questions and answers. The training content will be tailored for the audience and includes the same content regarding asthma as a health condition, trends in asthma morbidity and mortality, asthma risk factors, asthma self-management, and the role of indoor air quality in asthma control. Specially tailored modules for apartment owners and maintenance staff include policies and procedures for monitoring and improving common IAQ issues. Tenants receive tailored content regarding steps they can take to improve IAQ in their homes, resources for learning more about asthma self-management, and advocacy training related to tenant rights. All attendees receive a hard-copy or e-packet of materials based upon their preference.

In our experience, we have found that the Action Model stimulates different kinds of evaluation questions among team members designing the evaluation and stakeholders than logic models. For instance, the Action Model pushes individuals to carefully think about the many components of an intervention, that is, what it really takes to carry the intervention out. In fact, Chen (2015) describes in detail how the Action Model can be used as a program planning tool to check and test assumptions and therein facilitate better program design. Given its emphasis on the nuances of implementation, stakeholders may be more likely to pose evaluation questions that relate to the intervention operations such as What training and skills are required of the health educator to deliver this intervention effectively? How receptive were the city council members and the community coalition for clean air to encouraging apartment owner participation? Did this encouragement reap rewards or were there other, more significant factors, that stood in the way of apartment owner participation? Because logic models do not focus on detailed *operational* issues of intervention delivery (such as who the trainers are, what protocols are in place for delivering the intervention) but rather on the general activities used to implement the intervention (market the training, design and implement the training), such nuances can often be overlooked.

Zooming in on the Change in Outcomes

Missing from Figures **7.1** and **7.2** is any mention of potential contextual factors that may impede or enhance the causal pathways depicted. We know that there are many factors that can affect the implementation of an intervention and several conditions under which one or more presumed outcomes may be more or less likely to occur, these are known as **moderators** or **effect modifiers**. Theories of change afford the opportunity to depict these aspects of an intervention more clearly. **Figure 7.4** provides another depiction of the IAQ training, this time on the component of the training that is specifically tailored to tenants of the apartment buildings. In this graphic, the primary focus is on the interconnections between the outcomes. Notice there is no unpacking of the activities that comprise the intervention(s) in this figure rather just a simple mention that the intervention took place. As such, we would typically consider this to be a **program impact theory** (Donaldson, 2007), which is closely related to Chen's (2015) **change model**.

The model provide much more detail than what is depicted in **Figure 7.2** about *why* changes come about, as well as what might facilitate or impede progress in the causal chain depicted (i.e., moderators). The pathway draws upon the Health Belief Model (Glanz, Rimer, & Viswanath, 2014) and the empirical literature that supports this model. Here we see that participants who attend the tenant training are expected to experience several changes in their perceptions—about the possible threat that IAQ poses to their health and welfare, the benefits that can come from addressing IAQ in the apartment dwelling, the ability to overcome any barriers that stand in the way of improving IAQ. Furthermore, through the training they become more knowledgeable about what it takes to improve IAQ and perceive they can take action to make improvements. Factors that can affect these early stage changes are depicted in the gray boxes between the intervention implementation initial changes in perceptions (i.e., age) (Glanz et al., 2014).

Collectively, these changes in perceptions lead to actions that either directly improve IAQ or influence others to take actions to improve IAQ. If apartment owners and maintenance providers are not affected, the effectiveness of the intervention, in facilitating the intended distal outcomes, is diminished or even extinguished. Thus, the intervention designers included trainings for other audiences that have a vital role to play in facilitating positive change.

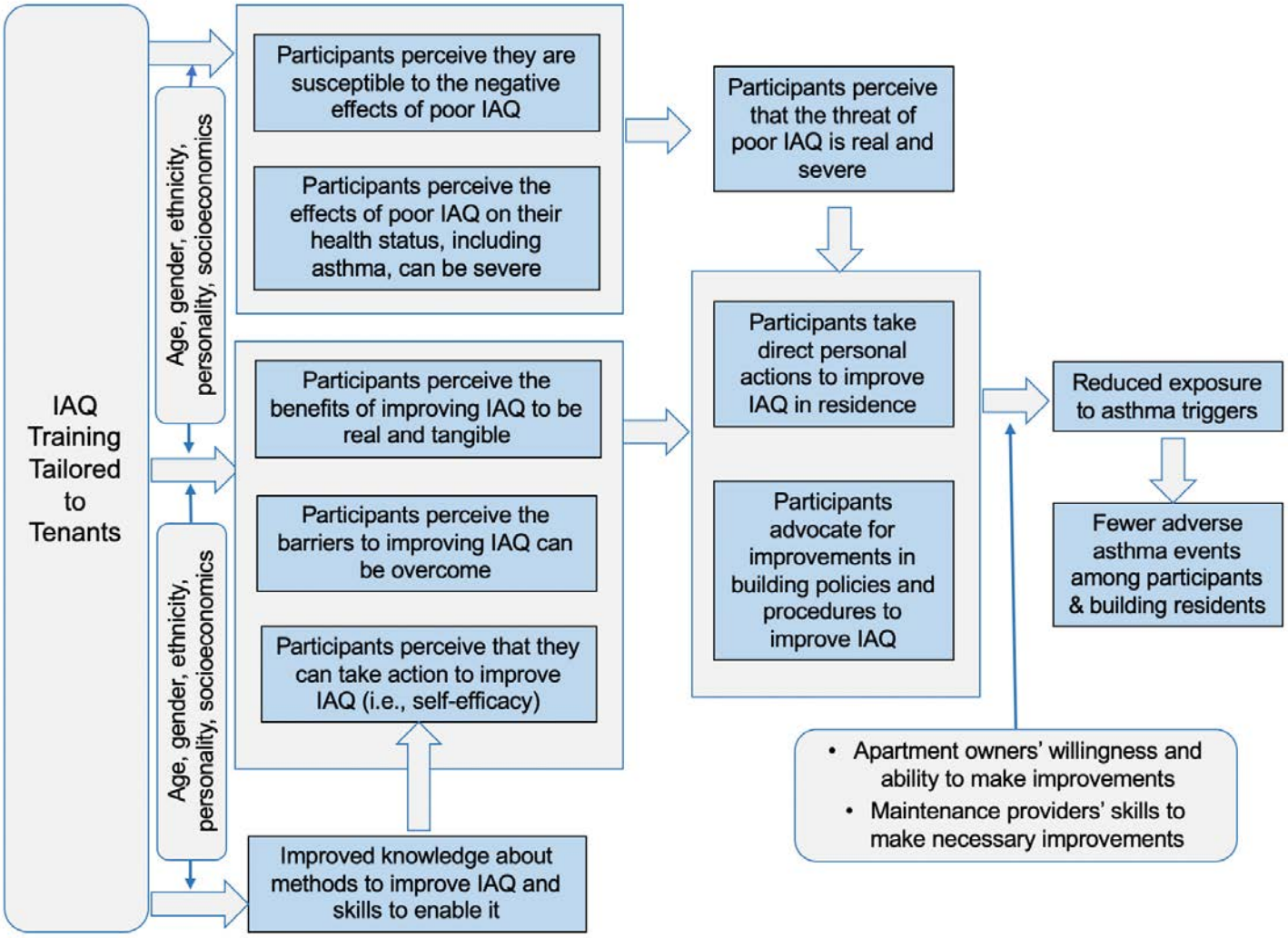


Figure 7.4 Program Impact Theory for Indoor Air Quality (IAQ) Training Tailored to Tenants

Focusing the Evaluation (Step 3)

As we discussed in **Chapters 3** and **4**, developing good evaluation questions is critical to any evaluation. Many aspects of an intervention that can be evaluated however, given time and resource constraints, it is important to focus the evaluation to ensure that it is useful and feasible. Selecting the right evaluation questions for public health interventions can be tricky –there is often the temptation to include more questions than can reasonably be answered with available resources (i.e., time, labor hours, funding). Though many of the principles used to prioritize evaluation questions are the same between evaluations focused on public health infrastructure (**Chapter 6**) and public health interventions, a few criteria are especially important to spend time discussing when the evaluand is an intervention: (1) Length of time the intervention has been in place (e.g., intervention's stage of developments) and (2) What is already know about the implementation or outcomes associated with the intervention.

Discussing these two items can help the Evaluation Planning Team narrow in on the focus for the evaluation. Perhaps the focus is on the intervention *process*, perhaps the focus is on the *outcomes*, or maybe a need exists at that moment in time to understand a little of both or some of the connections between processes and outcomes. Next, we describe process and outcome evaluation questions in more detail, explain how the decision of which to pursue connect back to the stage of an intervention's development and the existing evidence base, and offer up examples of what these types of questions may look like in practice.

Process Evaluation Questions

As described in **Chapter 3**, process evaluation questions focus on the items depicted on the left-hand side of a logic model. Evaluation questions about inputs and activities such as the efficiency with which resources were used for intervention activities, the dose of the intervention delivered and received, the fidelity to the original intervention plan, and the reach of the intervention are characteristic of a process evaluation (Steckler & Linnan, 2002). However, aspects of the intervention process can also be examined as part of an outcome evaluation to better explain how outcomes were, or were not, reached. **Table 7.1** lists some generic process evaluation questions.

Depending on an intervention's stage of development and the results from pre-evaluation activities (e.g., evaluability assessments), the Evaluation Planning Team may decide only to answer process evaluation questions. For instance, if the anticipated outcomes of the intervention have not yet had a chance to materialize, then an outcome evaluation does not make sense and, in fact, could be harmful. Conducting an outcome-focused evaluation before outcomes would likely appear as a result of the intervention could produce null findings and lead to the conclusion that an intervention is ineffective, when in fact it was just too soon to tell. However, it may make sense to set up data collection activities that position the program for conducting an outcome evaluation of the intervention in the future (e.g., thereby making it possible to calculate baselines or conduct a longitudinal analysis of key outcomes).

Some benefits of documenting lessons learned about evaluation itself

Evidence-based interventions are implemented in many different contexts under a variety of circumstances. Reviews that deem an intervention as evidence-based do not always consider the ability of the intervention to be implemented well across a wide array of settings. Furthermore, these reviews sometimes, though not always, weigh the evidence produced by randomized trials higher than other study designs that are likely to be more affected by the context in which they are conducted.

It is very important to evaluate evidence-based interventions *within the contexts where they are being implemented*. We cannot assume that fidelity to the implementation plan will be feasible across all settings or that modifications to implementation will result in the same level of effectiveness seen in the existing evidence-base. We also cannot safely assume that even if fidelity to implementation is upheld, that the intervention will remain effective in settings not covered in existing studies.

Process evaluations can be helpful at any stage of an intervention's lifecycle. For an intervention that has only been operating for several weeks or months, process evaluation can assess items such as alignment between program implementation and context as well as the fidelity to the intended model. Evaluations of both items may

show ways to improve and enhance the intervention, thereby promoting the likelihood of positive outcomes. For interventions that are further along, process evaluation can provide a check in to ensure that implementation fidelity has not strayed off course or that the intervention is still addressing relevant community needs.

Table 7.1 Example Process Evaluation Questions

Inputs
What are the starting conditions and contexts?
To what extent are the resources available for intervention implementation adequate?
What needs are being addressed by the intervention?
In what ways, if any, does the intervention fit with the overall public health strategy?
Activities and Outputs
What key activities were implemented?
Was the intervention implemented as planned?
To what extent is the intervention reaching the appropriate priority population?
How well was the intervention administered?
To what extent is fidelity to the intended implementation model being upheld?
How can intervention administration be improved?
Is the intervention acceptable to the intended participants? Is it culturally appropriate? Is it feasible?
Is the intervention being used by the intended participants?
What are the major barriers facilitators to implementing the successfully?
Context
What support has been mobilized for action?
What progress has been made in capacity building to support increased implementation?
How sustainable is the intervention over time?
How well coordinated is the intervention with other interventions in the community or system?
Multiple aspects of process
How efficient is the implementation of the intervention? (Inputs-Activities)
In what ways, if any, can the efficiency of implementation be improved? (Inputs-Activities)

Outcome Evaluation Questions

Questions that focus on the boxes and arrows located on the right-hand side of logic models—short-, medium-, or long-term changes—are typically characteristic of outcome evaluations. Changes may occur at the level of an individual, group or team, organization, or even a system. Common outcomes include changes in shorter-term outcomes such as individual-level awareness, attitudes, knowledge, or skills; mid-term changes such as the behavioral manifestation of gains in knowledge or skills which can include one or more individuals changing their own health-related behavior or taking steps to modify existing policies and procedures; and longer-term changes including reductions in environmental or societal risk or protective factors (e.g., improved air quality, better access to high-quality medical care) or health outcomes. Mid- and long-term outcomes often result from a convergence of outcomes that emerge before them in time sequence. **Table 7.2** includes examples of some generic outcome evaluation questions.

Table 7.2 Example Outcome Evaluation Questions

Short-Term Outcomes
To what extent, if any, has awareness about the public health issue increased?
To what extent, if any, has knowledge about how to address the health issue improved?
To what extent, if any, have individual's skills to address the health issue improved
How have relationships been formed or changed to promote systems change?
What, if any, positive or negative unintended outcomes emerged from this intervention?
Medium-Term Outcomes
To what extent are individual behaviors related to known risk factors changing in alignment with expectations?
What actions, if any, are decision makers taking to develop or improve existing policies?
What actions, if any, are decision makers taking to develop or improve existing organizational procedures?
What, if any, positive or negative unintended outcomes emerged from this intervention?
Long-Term Outcomes
How effective has this intervention been in improving environmental risk factors associated with the health condition?
To what extent have improvements in the health and quality of life of individuals and families affected by the public health issue been realized?
In what ways, if any, has this intervention contributed to a reduction in health disparities?
What was the cost-benefit or cost-effectiveness of the intervention?
What factors appear to have affected the effectiveness of this intervention?
What, if any, positive or negative unintended outcomes emerged from this intervention?
To what extent, if any, has the intervention contributed to reducing inequities in the public health issue?

Evaluation Designs

To a great extent, the priority evaluation questions developed by the Evaluation Planning Team will drive the selection of an evaluation study design (see **Appendix E** for more information on evaluation designs). For example, if several evaluation questions request information about what happened, such as how many individuals were reached through the intervention or how many communities received the intervention, then it may be the case that **performance monitoring** can be used to answer these questions and that a deeper evaluation study is not warranted. (See the CDC's brief on the differences and similarities between performance monitoring and program evaluation.) If evaluation questions require insights about change over time, a design that includes **repeated measures** (e.g., measurement at baseline, 3-months, 6-months, 1 year) is necessary. If one or more evaluation questions suggest the need for a comparison (e.g., How do the results of the group receiving the intervention compare with results from a group that did not receive the intervention?) it will be important to select a study design that includes a **comparison group** or **control group**. For questions that seek answers about causation (e.g., to what extent did the intervention result in the outcomes observed?), it will be very important to consider designs that will rule out **threats to internal validity** (i.e., account for factors other than the intervention which may have led to the outcomes observed).

It is not uncommon that stakeholders would like to learn more about the cost associated with an intervention. As a result, the Evaluation Planning Team may have prioritized evaluation questions that ask about cost, for example, return on investment, cost-effectiveness, or cost-benefit. If this is the case, it will be important to consider study designs that integrate cost—what we are calling economic evaluations. **Appendix I** provides detailed guidance on different types of economic evaluation, including how to incorporate these elements into evaluation plans, and provides a list of tools, templates, and additional resources. For additional support with economic evaluations, consider collaborating with a statistician, epidemiologist, or health economist within your or a partner's organization.

It may be the case that the Evaluation Planning Team decides to evaluate an entire multicomponent intervention that is intended to affect a system as a whole. Parsing out which components of the intervention contributed most to the outcomes may be difficult in this case. Applying a **systems thinking** lens and understanding **complexity theory** can help to develop evaluations for systems interventions that are not linear or straightforward. Complexity theory recognizes that reality is complex, and outcomes result from multiple causes that are interrelated and interact with each other (Byrne, 2002). The outcomes may be greater or less than the sum of all the components in a complex system because components may enhance or cancel each other out.

Healthcare systems can be characterized as complex systems when there are many components that are highly interconnected (Kannampallil et al., 2011). Such a system is dynamic and has feedback loops that provide information on what is happening within the system. To evaluate complex systems, several methods can be useful depending on your evaluation questions, including Social Network Analysis, Outcome Harvesting, and Outcome Mapping.

Social Network Analysis (SNA) is a method used to understand systemic change, complexity, and connections between individuals or groups within a network (Durland & Fredericks, 2005). SNA can be useful for examining relationships, identifying important members in a network, understanding the capacity of a network to achieve a goal, tracking changes in a network over time, and understanding the connection between a network and outcomes (Honeycutt, 2009). See more information on using SNA in program evaluation.

Outcome Harvesting is an evaluation approach used for dynamic complex systems when the relationship between causes and effects of the intervention are not entirely identifiable or understood (Wilson-Grau, 2015). It can be used when there is an observable change in behavior of an individual or organization. Outcomes are harvested using six iterative steps through a participatory process involving a variety of stakeholders (Wilson-Grau & Britt, 2013).

Outcome Mapping is an approach that can be used for intervention planning, performance monitoring, and evaluation (Earl et al., 2001). Like Outcome Harvesting, this method focuses on behavioral outcomes and assumes contribution to the outcomes, rather than attribution. The approach uses twelve steps through a participatory process with stakeholders. For more information on outcome mapping visit the Outcome Mapping Learning Community.

While developing the evaluation design, consider the data analyses stakeholders will want later. For example, how will they want to sub-categorize information? Will they want to look at separate geographic areas or different demographic variables? The Evaluation Planning Team should consider different scenarios to assure that the design will provide sufficient information.

Gather Credible Evidence (Step 4)

As described in **Chapter 4**, in this step, the Evaluation Planning Team will consider criteria of merit, indicators, and data collection methods and sources. Following the general guidance in **Chapter 4** will lead to successfully implementing this step when evaluating interventions. Though, like other special topics, there are a few issues that are more likely to surface in this step when interventions are the evaluand.

Matching the Anticipated Change to the Correct Unit of Analysis and Data Source

Evaluation stakeholders often want to know whether the intervention changed the intended outcomes (e.g., emergency department visits, quality of life, mortality) for the better. Another common situation in evaluation is that stakeholders would like to leverage existing data as much as possible to answer the evaluation questions. In a public health context, one source of data that is often readily available is the data maintained in public health surveillance systems. It is often tempting to automatically think that data about health outcomes in public health surveillance systems can be used to answer outcome evaluation questions for a given intervention, but this is not necessarily the case.

Data in public health surveillance systems typically exists for a specific population within a rather large geographic area (e.g., national, state, and on fewer occasions county). However, interventions are not often designed to change health outcomes or risk factors at a *population* level. Rather, interventions tend to focus on specific sub-groups within populations (e.g., delivering a health education workshop to individuals who have asthma and were recently hospitalized at the ten largest hospitals in a specific county, or conducting a media campaign about the harms of smoking within specific high-risk zip codes within a county). In these instances, it would not be appropriate to use county, state, or national-level data to evaluate the extent to which the intervention resulted in outcomes. The unit of analysis is incorrect: public health surveillance data at its highest level of resolution likely contains health outcome data for the county population. In the event surveillance data cannot be further disaggregated to the zip code level it would not be informative for an outcome evaluation of the media campaign. For the health education workshop example, data would need to be collected for the individuals who attend the workshop (and if desired or needed, a comparison or control group). People attending the session may be included in the surveillance data but the effect of the intervention would be washed out in the larger data source.

We have surveillance data. Can we use it?

It is often the case that there are limited resources available to conduct evaluations in public health settings. Given this, it can be tempting to use data that are readily available—namely public health surveillance data—to answer the evaluation questions. Typically, public health surveillance data captures population-level data for a relatively large geographic area of interest (e.g., national, state, and sometimes county). Though some interventions (e.g., state-wide policy changes) are intended to effect change in populations, many are not (e.g., health education workshop for people who were recently hospitalized in specific areas of a county). If an intervention is not designed to affect change in populations, it is unlikely that surveillance data is an appropriate data source to answer the question of interest. Using data that are mismatched to the unit of intended change can produce inaccurate and misleading results.

Surveillance data are often well positioned to inform outcome evaluations of interventions that intend to affect large systems or populations. Examples of such interventions include adoption and implementation of a new public policy, modifications to or adoption of public or organizational policies that affect entire healthcare systems, or large public health campaigns such as the roll-out of a vaccination within a state. In these instances, the intended change *is* at the level of a population, and data for the populations of interest can often be acquired through existing public health surveillance.

Using Existing, Validated Data Collection Instruments

In the event the Evaluation Planning Team decides that it is necessary to collect new data to answer the evaluation questions, data collection instruments and procedures may not need to be designed from scratch. Many survey instruments, focus group guides, checklists, and interview instruments are already available. Using instruments that were developed to support other evaluations or studies can help you

- Save effort in designing the evaluation.
- Compare results with those of interventions implemented by others.
- Provide greater assurance of the validity of the data collection efforts.

Given the sheer number of existing resources for data collection, it is important to keep in mind several factors in choosing among them:

- ***Instrument purpose.*** The key consideration in using an existing instrument is whether it will suit the evaluation purposes. Does this instrument include the questions the Evaluation Planning Team is interested in? Does it cover all the topics needed to answer the evaluation questions? The Evaluation Planning Team may need to add questions to cover additional topics.

- ***Making changes.*** When using an existing instrument, especially one that has been validated (tested for validity and reliability among large groups of people), avoid making major changes to question wording, sequence, or answer categories, as you will then lose the benefit of the prior testing. Many survey instruments contain scales (multiple questions related to the same topic) and, if you need to adapt or remove elements, it is better to remove or keep an entire section, rather than remove single questions.
- ***Respondent population.*** Has the instrument been used in a population similar to the respondent population for the current evaluation? Is the instrument appropriate in terms of literacy level, idioms, language, or cultural relevance? If possible, you may want to look for an instrument that has been used with a similar audience rather than adapting an instrument that was designed for a different group.
- ***Instrument length.*** In deciding whether to use an existing instrument, keep in mind how long the respondent population will have to participate in data collection versus how long it takes to complete the existing instrument.
- ***Getting permission.*** It is good to get permission from the instrument developer to use an instrument. This is usually as simple as an email or telephone call. In addition to making sure you are covered to use the instrument, you may also get valuable information not included in public sources, such as information about a new version of the instrument or details about how to analyze results.
- ***IRB and participant protection***. Before administering the instrument, ethical considerations may need to be assessed by an institutional review board (IRB). Initiate data collection after the IRB has determined that minimum risk is involved and specified what protections are appropriate for your priority population.

If the Evaluation Planning Team decides to adopt an existing instrument, it is still critically important to plan for sufficient time to pilot test the instrument. There may be different issues in the context where the current evaluation takes place that were not present in contexts where the instrument was previously used. Additionally, you may find that several instruments you review have been validated through an empirical study. This is certainly positive; however, this validity may not hold up in a new context. Validity is *not* a characteristic of an instrument, it changes as instruments are used in different settings, with different respondent populations, and over time. Do not assume that the instrument is valid in the current evaluation context, instead plan in time to examine the validity of the instrument in the relevant setting.

A Glimpse into Practice
Selecting a Focus & Gathering Credible Evidence
for the IAQ Training Intervention for Tenants

Recognizing that the IAQ training was new, the Evaluation Planning Team decided that the most appropriate focus for the evaluation was on the intervention process and, possibly, early intended outcomes. They were most curious about whether the strategy proposed to get the word out about the tenant training would be sufficient. To what extent would the marketing efforts lead to the intended audience of tenants attending the training? For those who did attend, how did they find out about the training and why did they choose to attend? For those who live in the apartment buildings and may have attended but didn't, why didn't they attend? Assuming sufficient attendance was reached (a target of roughly 25% of all residences in the five building being present at the training was set by the Evaluation Planning Team), the team aimed to answer additional priority evaluation questions: to what extent did the participants' perceptions change regarding IAQ as depicted in the impact model (Figure 7.4)?

To answer these evaluation questions the Evaluation Planning Team proposes identifying the number of units within the five apartment buildings and learning from attendance records how many of these units had representation at the training. Brief exit interviews with attendees will be used to learn about how they found out about the training and why they chose to attend, and a brief hard-copy survey on a postcard with pre-paid postage will be disseminated to each apartment not represented at the training to better understand why they did not attend. If more than 20 individuals attend the session, they will be provided with a pre and post instrument to assess changes in perceptions. Though the Evaluation Planning Team would have preferred to include a comparison group when answering the question about perceptions they recognized that the best comparison would be residents in the same apartment buildings that did not attend the training and recognize that it is highly unlikely that this respondent group would be motivated to participate in the evaluation by completing two assessment instruments over the course of a limited time period. Asking this group to fill out the assessment instrument only once would result in a level of internal validity that was not credible to the evaluation stakeholders

Justifying Conclusions (Step 5) & Ensuring Use and Sharing Lessons Learned (Step 6)

The final steps of the evaluation process will be quite similar for intervention evaluations as it was for infrastructure evaluation. Following the general procedures we shared in **Chapter 4** will help to ensure that the Evaluation Implementation Team generates an accurate synthesis of the findings and disseminates this synthesis to key audiences in a manner that is most useful to them. Again, remember when developing products to communicate about the evaluation findings to loop back to the stakeholders identified under Step 1 to ensure coverage of potential audiences.

A Glimpse into Practice
Justifying Conclusions & Acting on the IAQ Training for Tenants - Evaluation Findings

The Evaluation Implementation Team was able to carry out most of the proposed data collection efforts. Unfortunately, less than 20 individuals participated in the tenant training, so the team did not move forward with gathering data about the early intended outcomes. In synthesizing and interpreting the findings about the marketing efforts, the Evaluation Implementation Team appreciated the details the Evaluation Planning Team provided with respect to potential targets and standards by which to judge the resulting data. For instance, 25% of the residences within the five apartment buildings did not attend, leading to the relatively quick interpretation that interest in participation was far less than desired. Additionally, qualitative responses from the exit interviews, about how the attendees learned about the training, would have included comments that clearly indicated attendees were influenced by messages delivered through the community recreation center or community-based organizations. Almost all attendees heard about the training from these sources: from friends who frequent the community center or the attendee was involved with at least one of the organizations. Few responses from the postcard survey were received (10% response rate), but all noted having heard about the training from one of the three marketing sources and mentioned that the time of day was not ideal or no one in their household had asthma.

The Evaluation Implementation Team developed an infographic summarizing these findings to share with primary and secondary stakeholders and invited these individuals along with the training attendees to a 90-minute meeting to discuss next steps. Rich discussions took place at this session (which included five attendees of the training) about how to improve the marketing of the intervention. It was evident that tapping into social networks in these communities did pay dividends in terms of getting the message out, and stakeholders noted that this resonated with their understanding of and experiences in these neighborhoods. A decision was made to run the intervention again within the same city—to include the five apartment buildings originally targeted as well as a few medium and small apartment complexes in areas of the city where asthma prevalence and risk factors are thought to be high. Team members were assigned to engage with additional organizations and networks within the community to disseminate the marketing messages—specifically two small community advocacy organizations, one of which focuses on children's health and another that focuses on chronic health conditions. Additionally, plans for the intervention materials and the next evaluation were adjusted to include communications materials and data collection instruments in both English and Spanish since both are spoken regularly within the community. One evaluation stakeholder offered to do the translation to Spanish, and another offered time to review the translation for accuracy and nuance.

Circling Back to Inform Future Plans

The Evaluation Implementation Team learned many important things about the evaluation process that could be helpful to future evaluation efforts. As a result, they made sure to spend some time at the end of the evaluation cycle reflecting on and documenting these lessons and connected with the members of the Strategic Evaluation Planning Team to share these insights. The team's reflections highlighted the importance of assessing the context to support culturally responsive approaches to evaluation as well as the challenges of soliciting participation in evaluation from individuals who do not attend interventions. In addition, the team shared their plans to continue with the intervention and evaluation with the Strategic Evaluation Planning Team so they could adjust the timeline and schedules in the strategic evaluation plan for the upcoming years.

Chapter Summary

In this chapter, we introduced you to applying the steps of the CDC Framework to evaluating public health interventions. We emphasized some tools that may be particularly helpful in carrying out such evaluations, including evaluability assessment and program theory (including several ways to visualize program theory). We provided some ideas of specific items to consider when selecting evaluation designs and data collection methods. Throughout the chapter we demonstrated the application of the insights provided through an example of a fictional intervention designed to improve indoor air quality in an effort to reduce asthma morbidity events.

Review Questions and Skill Building Exercise

Review Questions

1. What is an evaluability assessment? What benefits can be realized by conducting evaluability assessments?
2. What are five sources evaluators can use to learn about and assess an intervention's context? What could you expect to learn from each?
3. In this chapter, you learned that logic models could be presented in a tabular format or individual boxes and arrows. If your evaluation is interested in information about a program's processes and outcomes, which format are you most likely to use and why?

Skill-building Question

This chapter describes a hypothetical intervention called the Indoor Air Quality (IAQ) Improvement Intervention, which aims to decrease the occurrence of adverse asthma events by improving the indoor air quality in individuals' homes. **Figures 7.1**, **7.2**, and **7.3** provide details regarding the program's components, partners, and intended beneficiaries.

You have been asked to lead an evaluation of the IAQ Improvement Intervention. As a conscientious evaluator, you are committed to ensuring that your work reflects each "How" of Evaluation Practice presented in the Enhanced CDC Framework (**Chapter 2**). For each evaluation step listed in the table, identify at least two actions for each "How" of Evaluation Practice that you plan to integrate into the evaluation. Your selections may include both individual actions and actions that will be taken by the Evaluation Planning Team.

	Situational Awareness	Critical Reflection	Cultural Competence	Interpersonal Competence
Step 0. Assess context				
Step. 1 Engage Stakeholders				
Step 2. Describe the Program				
Step 3. Focus Evaluation Design				
Step. 4 Gather Credible Evidence				
Step.5 Justify Conclusions				
Step.6 Ensure Use and Share Lessons Learned				

APPENDIXES

APPENDIX A. Review Questions Guide

This guide will assist you in developing responses to the chapter review questions. For many of these questions, there is no single right answer. Rather, the question's intent is to deepen your understanding of evaluation, recognizing that evaluation is a context-dependent practice. The guidance provided for each question highlights the key concepts associated with each question and offers suggestions to facilitate your thinking.

Chapter 1 - Introduction to Evaluation

1. What is the professional practice of evaluation, and why is it important? How is it distinct from research and epidemiologic investigations?

The practice of evaluation encompasses the behavior and work required to conduct a professional evaluation (i.e., determine the merit, value, or worth of a program, policy, process, or function). Its importance to society can be described in various ways but, generally, evaluation creates knowledge that can then be used to inform decisions and actions to benefit society in some way. In the absence of this knowledge, we would have no basis to understand if, how, why, and for whom something works, and how we can make it better. This chapter emphasized two important features of evaluation practice: it is a *systematic* process that involves making a *judgment*. While the use of a systematic process of investigation is common to other forms of scientific inquiry (i.e., research and epidemiologic investigations), the act of making a judgement is unique to evaluation.

2. Why is it important to conduct systematic evaluations of public health programs? What risks are present when evaluations are not systematic?

One could approach this question by first thinking about the importance of systematic evaluations generally, and then assessing its relevance to public health programs. Systematic approaches to evaluation, based on scientific methodology, help to ensure that the evidence yielded from an evaluation is accurate and credible. This is partially achieved by controlling for the bias and error that can be introduced when processes are not carried out in a methodical and conscientious manner (think back to the chapter's discussion on people's natural sensemaking abilities and the presence of cognitive bias). Since stakeholders are more likely to use evidence they view as accurate and credible, systematic evaluations are an antecedent to evidence-informed decision making, which is valued in the public health sector.

This chapter discussed the role of evidence in supporting the public's safety from health threats and ensuring public health programs are effective, efficient, equitable, and sustainable. What risks are associated with all or some of these areas if decision makers do *not* have access to accurate and credible evidence? Some that may come to mind are the development of ineffective responses to public health problems, actions that reinforce existing inequities, misallocation of public monies, and missed opportunity to improve program delivery models. In addition, if people are making decisions based on information that is not systematic, then those decisions are more likely to be swayed by personal or group interests and preferences rather than based on evidence.

3. Why is use important to evaluation? How might each kind of evaluation use (i.e., instrumental, conceptual, or enlightenment) improve health programs?

Put simply, use is what brings evaluative evidence to life and connects evaluation to its ultimate goal of social betterment. The use of evaluative evidence supports improved decision making and helps to ensure that programs and interventions meet their intended outcomes. When evaluative evidence is not used, there is a loss in terms of opportunity to improve the public's health and in the use of resources allocated to conduct the evaluation. This chapter introduced two types of evaluation use: instrumental use and conceptual or enlightenment use. You may wish to review that section to familiarize yourself with those terms if you have not done so already.

- Instrumental use is the direct use of evaluation findings. It typically refers to situations where a change is made to a program, policy, or processes, due to knowledge gained from an evaluation.
- Conceptual use occurs when evaluation findings change someone's thinking or understanding.

Consider a public health campaign to pass legislation for a city-wide ordinance to ban outdoor smoking on public elementary school property. The legislation is aimed at reducing the burden of asthma in children by eliminating triggers from second-hand smoke. The local public health department contracted a marketing firm to create billboards and signs that will be posted throughout the city promoting the bill, which will be voted on by the public in six months. The public health department staff conducted a series of focus groups to obtain feedback from community members on the content and look of the marketing materials. In the focus group sessions, participants remarked that the people featured in the ads were all middle-aged and older adults, but that in their experiences young people in their twenties and thirties also smoked. They reasoned that since the bill was aimed at reducing secondhand smoke in children, it would be better to feature younger-looking adults. The public health team used the focus group feedback to change the content of the ads to feature younger adults. This is an example of instrumental use because the findings were used to make a direct change in the public health campaign. The public health director also noted the importance of focusing on adults who have children when intending to generate future marketing materials aimed at reducing the burden of asthma in children. This is an example of conceptual use because the findings from the focus groups influenced the director's thinking for future public health interventions.

4. In what ways is evaluation and the use of evidence for decision making already integrated into public health practice?

This chapter includes a discussion on the role of evaluation in public health, including the conduct of evaluations and use of evaluative evidence. Evaluation is included as one of the Essential Public Health Services and a component of public health practice. The chapter also describes evaluation's connection to evidence-informed and evidence-based decision making in public health contexts as well as the role of evaluation in obtaining a public health degree at an accredited institution or attaining the Certified in Public Health credential. Evaluation is also integrated in the CEPH – MPH foundational competencies for public health practitioners.

Chapter 2 - Approaches to Evaluation Practice

1. This chapter introduced you the Program Evaluation Standards developed by the Joint Committee on Standards for Educational Evaluation (JCSEE) (i.e., utility, feasibility, propriety, accuracy, and evaluation accountability). Select two standards for this question. How might these standards (a) complement and (b) conflict with one another in an evaluation? You do not need to use the same two standards for a and b.

There is no single correct way to answer this question and you will find that in practice-contexts there is often a need to balance the JCSEE standards in different ways. As an example, consider a situation in

which you are a member of the Strategic Evaluation Team, and are discussing potential methods to use in answering a set of draft evaluation questions for a priority evaluation. The evaluation focuses on answering questions about the extent to which an intervention *causes* an improvement in health outcomes. The team is strongly in favor of using an experimental design in which individuals are randomized to treatment and control conditions as the team feels this will generate results that are most valid in assessing causation, thereby upholding the *accuracy* standard. However, while engaging in additional discussions about this evaluation it comes to light that randomizing individuals to treatment and control groups may withhold important treatments from those who need it (severely impacting the *propriety* standard) and that it may be very challenging to ensure that individuals are appropriately assigned to conditions randomly given how busy the health practitioners who will do this assignment are (suggesting a low level of *feasibility* for this design that may inadvertently affect *accuracy*). The team decides this is not the best balance between the JCSEE standards and in fact could be harmful to participants and therefore moves on to discussing other options for an evaluation design.

2. The chapter presented four additional dimensions of evaluation practice: critical reflection, situational awareness, interpersonal competence, and cultural responsiveness. Out of these four dimensions which would you like to improve upon and why? What steps might you take to further develop your abilities in this area over the next year?

The chapter includes a description of each dimension as well as its significance to evaluation. If you're unclear as to which of the dimensions you'd like to improve upon, it may be beneficial to review these descriptions to assess their relevancy to your learning needs and current and future work or study. You may also find it helpful to seek input from a supervisor, mentor, professor, or peer.
Once you have identified two dimensions, you can begin to develop a learning plan to advance your understanding. There are many different learning plan templates available on the internet if you do not already have a plan. Alternatively, you can develop your own plan by addressing some or all the following steps:

- Self-assessment: this could include an analysis of your motivations, goals, current knowledge of the dimensions, time available for learning, and preferred learning style.
- Identification of learning resources and strategies: research the various resources (e.g., scholarly work, training, people) available to support your learning and the means in which you can access those resources; consider how each resource does or does not align with your self-assessment; develop a final list of the resources and strategies (e.g., self-study, course) that you plan to use to support your learning.
- Develop a workplan: this should outline the specific steps that you will take for each dimension as well as a timeline.

3. What is evaluation theory and how does it help evaluators carry out the steps of the CDC Framework?

This chapter describes evaluation theories as prescriptive approaches to evaluation practice (Alkin, 2013), which provide evaluators advice and guidance on how to conduct an evaluation. Evaluators can draw on these theories when making decisions about how to implement and approach each step in the CDC Framework. For instance, evaluation theories provide us with several ideas for how we might *engage stakeholders*. Some suggest engaging with a wide array of stakeholders throughout the entire evaluation (e.g., designing and implementing the evaluation, including data collection and analysis); others suggest engaging with a small number of stakeholders who have the authority to change the program and to engage them to a limited extent throughout each of the subsequent steps of the CDC Framework (e.g., include

in describing the program and identifying key evaluation questions only). Evaluation theories therefore provide us with options to consider in any single evaluation context about how we may engage in evaluation practice.

Chapter 3 - Thinking Strategically: The Strategic Evaluation Plan

1. What are some benefits of strategic evaluation planning? How does a strategic evaluation plan differ from an evaluation plan?

The chapter's introduction lists some benefits of strategic evaluation planning:

- Helping to allocate resources for evaluation
- Facilitating useful evaluation results
- Enhancing alignment of work across an organization

You may also be able to think of additional benefits that are specific to programs and organizations that you familiar with.

A strategic evaluation plan is the product of strategic evaluation planning, which is a process to systematically prioritizing evaluations, that describe a program's or organization's evaluation activities over a defined period. You can think of Strategic Evaluation Plans as a blueprint that provides a high-level description of multiple evaluations as well how those evaluations fit within a comprehensive evaluation strategy. An evaluation plan focuses on a single evaluation, describing specific details about how the specific evaluation will be carried out and communicated.

2. What are some qualities to look for when seeking members of a Strategic Evaluation Planning Team?

As discussed in the chapter, a Strategic Evaluation Planning Team should consist of individuals with diverse knowledge and skills and a shared interest in program evaluation. The Team should also reflect the diversity of the community served by the program. The team is typically lead or co-lead by an evaluator with experience in a broad range of data collection strategies and evaluation designs, knowledge of public health programs similar to the program of interest, and strong communication skills. Other team members should also be knowledgeable about the program, the role of evaluation in program improvement, and the resources available for evaluation.

3. What types of efficiencies might a Strategic Evaluation Planning Team identify when developing a cross-evaluation strategy?

Table 3.6 identifies numerous potential efficiencies that may be realized as part of developing a cross-evaluation strategy. These include situations when it may be possible and appropriate to use a single data collection tool for multiple purposes and using a single data source for multiple evaluation activities.

Chapter 4 - Digging into the Details: The Individual Evaluation Plan

1. In what ways does the membership of the Evaluation Planning Team differ from that of the Strategic Evaluation Planning Team? What are some considerations to reflect upon when comprising a list of individuals to invite to the Evaluation Planning Team?

The role of the Evaluation Planning Team is to provide advice and guidance on a specific evaluation. Collectively, this team should include individuals with diverse perspectives on the evaluation and an understanding of the program's, or evaluand's, context. When considering membership, one can consider drawing representation from the three stakeholder categories proposed by Russ-Eft & Preskill (2009)—primary, secondary, and tertiary stakeholders. Another option is to assess the context of the program to identify stakeholders who are likely to hold a diverse range of perspectives. **Chapter 4** provides some guiding questions to assist in this assessment.

As you'll recall from **Chapter 3**, the Strategic Evaluation Planning Team is responsible for creating a plan that will guide a program's or organization's evaluation work over a defined time period. Members of this team should have knowledge of the program and the resources available for evaluation as well as an understanding of evaluation's role in program improvement. The Team should also reflect the diversity of the community served by the program. Typically, the team is led or co-led by an evaluator with experience conducting evaluations of similar programs as well as utilizing a variety of designs and methods.

2. In **Chapter 2** we discussed the importance of interpersonal competence in evaluation. Review Steps 1–6 in this chapter and take time to reflect on how interpersonal competence is important to successfully carrying out each of these steps?

For this question, it may be helpful to first review the description of interpersonal competence included in **Chapter 2** and the interpersonal domain of the 2018 AEA Evaluator Competencies. You may also benefit from thinking of a specific program context to help guide your reflections. **Table A.1** provides an example for each step of the CDC Framework. Each example includes a reference, in parenthesis, to a specific AEA Interpersonal Domain Competency.

Table A.1 CDC Framework Steps and Interpersonal Competencies

Step	Importance of Interpersonal Competence
Step 1 Engage Stakeholders	Evaluations can cause stress and anxiety for some stakeholders, which can have a negative impact on current and future evaluations. As such, it is important that evaluators are emotionally intelligent and empathetic to the feelings of stakeholders. Being able to understand and talk about emotions, including those that are non-verbalized, can help overcome emotional hurdles and get the evaluation off on the right track (AEA Evaluator Competency 5.1 Fosters positive relationships for professional practice and evaluation use).
Step 2 Describe the Program	Stakeholders will have varying views on how a program operates and why. Each view is important and can help to validate aspects of the program as well as provide unique insights. Listening and valuing different perspectives is an important quality of evaluators (AEA Evaluator Competency 5.2 Listens to understand and engage different perspectives).
Step 3 Focus the Evaluation Design	Prioritizing evaluation questions and deciding on study designs can be challenging, particularly if team members have contrasting views on these topics. Evaluators may need to exercise their negotiation and persuasion skills to facilitate a productive discussion that brings the team into alignment (AEA Evaluator Competency 5.3 Facilitates shared decision making for evaluation).
Step 4 Gather Credible Evidence	Trust and credibility go hand-in-hand. Engaging stakeholders in the selection of indicators, data sources, and methods not only supports the accuracy of findings, but also builds trust during the evaluation process. In turn, stakeholders are more likely to view the evidence yielded from the evaluation as credible (AEA Evaluator Competency 5.4 Builds trust throughout the evaluation).

Step 5 Justify Conclusions	Stakeholders will draw different meanings from evidence based on their values and interests in the evaluand. Evaluators must understand and acknowledge the diverse values and interests while justifying the evaluation findings (AEA Evaluator Competency 5.6 Communicates in meaningful ways that enhance the effectiveness of the evaluation).
Step 6 Ensure Use and Share Lessons Learned	Conflict frequently arises during an evaluation, and it is often up to the evaluator to manage these disputes. For instance, stakeholders may have differing views on how the evaluation findings should be used and by whom. In such circumstances, evaluators must leverage their mediation and negotiation skills to facilitate a resolution in a positive way (AEA Evaluator Competency 5.7 Facilitates constructive and culturally responsive interaction throughout the evaluation).

3. What are some formats for communicating evaluation results beyond a written, formal evaluation report? Can you think of other options beyond those we covered in this chapter? If so, what comes to mind?

Figure 4.2 and the subsequent section references several communication formats, beyond written reports, that can be incorporated into an evaluation. Evaluation Planning Team members are encouraged to be creative when considering their communication options. One strategy to drive creativity is to examine the practices of other fields and disciplines. For instance, could we learn about using story telling from the field of journalism? Or could we look to music and art for new ways to communicate evaluation results?

Chapter 5 - Where the Rubber Meets the Road: Implementing the Evaluation

1. Which of the evaluation timeline tools presented in this chapter do you feel would be most helpful to you if you were conducting an evaluation? What do you view as some of the pros or cons of the different formats?

The chapter describes numerous timeline tools that can be used to manage an evaluation. Determining the right fit for you could depend on several factors such as personal preference, information needs, and the level of detail that needs to be tracked. **Table A.2** below summarizes some pros and cons of each tool discussed in the chapter.

Table A.2 Pros and Cons of Timeline Formats

Timeline	Pros	Cons
Basic Yearly Progress Timeline	• Clearly articulates what activities should be completed each month. • Easy to read.	• Activity-oriented; does not communicate the project status.
Milestone Tables	• Allows you to track the progress of the project. • Easy to read.	• Does not capture the activities related to each milestone. • Could become unwieldy if the project includes numerous milestones.
Gantt Chart	• Typically provides greater detail on the project (e.g., duration of activities, labor hours). • Depicts activities that coincide and dependencies between activities.	• Can be difficult to comprehend. • Could be too much detail for some audiences.
Shared Calendar	• Easy for team members to access. • Could result in less of a learning curve since most people are familiar with electronic calendars.	• May lack detail for some audiences. • Likely needs to be supplemented by another planning tool.

2. Of the evaluation challenges presented in this chapter and **Appendix G**, were there any that surprised you (i.e., that you didn't anticipate would be challenges in an evaluation)? Select one of these surprise challenges. How could this challenge be addressed or alleviated through planning or implementation?

The chapter describes some possible actions that can be taken for each evaluation challenge. You may also be to think of some additional actions based on your experience managing projects.

Chapter 6 – Public Health Program Infrastructure

1. As we have learned in previous chapters, there can be a wide range of stakeholders for any evaluation. Can you list up to three potential stakeholders each for evaluations of public health partnerships and surveillance? What might the stakeholder interests in these evaluations be?

Primary stakeholders for evaluations of public health partnerships may include the leadership of the public health partnership. They may be interested in the findings of the evaluation to make adjustments to the partnership operations or to tailor the goals of the partnership based on the evaluation findings. Partnership member organizations are a potential secondary stakeholder. These may include community organizations, working groups, and local public health departments among others. Partnership members may be affected by changes in the way the partnership operates and thus would be affected by evaluation findings. Lastly, a tertiary stakeholder may include public health partnerships working on similar issues who may be interested in the results of the evaluation to garner lessons learned for their own work.

Primary stakeholders in a surveillance evaluation are those who have the authority to make decisions based on the findings. This may include the epidemiologist who manages the surveillance system. Secondary stakeholders are likely to be affected by the evaluation findings and in this case may include public health program managers, partner organizations, or policymakers who may use surveillance findings and products. Tertiary stakeholders are likely to be interested in gathering lessons learned from the findings for their own programs or policies. These may include public health programs who have similar surveillance systems.

2. Articulate at least one process evaluation question that could be of interest in a partnership evaluation? How about for a surveillance evaluation?

Chapter 6 provides several examples of evaluation questions in **Tables 6.1–6.4**. Process evaluation questions for a partnership evaluation might include

- How does our membership compare with partnerships within similar public health program? What additional partners should we add to support our efforts? How timely are gaps identified and addressed in our partnership?
- To what extent does the partnership have a clearly articulated vision? To what extent is this vision shared among members of the partnership?

Chapter 6 provides several examples of evaluation questions for surveillance in **Table 6.5**. Process evaluation questions for a surveillance might evaluation include

- What data or analyses are missing or lacking? How could the analysis of surveillance data be improved (e.g., faster turnaround, additional indicators)?
- What evaluation findings of our surveillance efforts were shared this year? With whom? How? How

could we change these communications to better meet our partners' preferences?

3. Articulate at least one outcome evaluation question that could be of interest in partnership evaluation? How about for a surveillance evaluation?

Chapter 6 provides several examples of evaluation questions in **Tables 6.1-6.4**. Outcome evaluation questions for a partnership evaluation might include

- How effective is the partnership in combining the perspectives, knowledge, and skills of diverse partners in a way that enables members to think in new ways, plan more comprehensive programs, and strengthen relationships with the broader community?
- In what ways has the partnership contributed to producing new linkages between the partnership and other coalitions or organizations? How do these new connections contribute to improving program outcomes?

Chapter 6 provides several example of evaluation questions in **Table 6.5**. Outcome evaluation questions for a surveillance evaluation might include

- For what purposes are our surveillance data used? How, if at all, are they used in planning and guiding strategic action?
- To what extent has the use of surveillance data by our key stakeholders improved as a result of our partner training?
- In what ways has the use of our surveillance data resulted in increased funding for our or our partners' projects?

Chapter 7 – Interventions

1. What is an evaluability assessment? What benefits can be realized by conducting evaluability assessments?

Evaluability assessments are a form of inquiry that examines a program's and evaluation's context and conditions to determine the utility and feasibility of conducting a full evaluation. These assessments are important for numerous reasons, such as (1) identifying issues that may prevent the successful implementation of a full evaluation, (2) obtaining information to inform the design of a full evaluation, and (3) providing insights on adjustments to the intervention that are needed prior to a full evaluation.

2. What are five sources evaluators can use to learn about and assess an intervention's context? What could you expect to learn from each?

This chapter describes many information sources that can be used when assessing the context of an intervention. Some sources may involve a site visit, while others may be accessed remotely. A sample of sources and potential information are described in **Table A.3**.

Table A.3 Sample Context Sources and Information

Sources	Potential Information
In-person site visit	• Locations for social gatherings • Economic well-being of the community and distribution of wealth • Residents' access to goods, services, and recreational space
Local newspapers	• Current issues or concerns to the community • How and what information is communicated to community members and by whom • Historical information and background
Existing data	• Breakdown on the community's demographics • Data that can be accessed for evaluation purposes

3. In this chapter, you learned that logic models could be presented in a tabular format or individual boxes and arrows. If your evaluation is interested in information about a program's processes and outcomes, which format are you most likely to use and why?

Box and arrow style logic models are typically more useful for framing process and outcome questions because they describe the causal pathways that underlie the program. In other words, they specify the processes through which program components (i.e., inputs, activities, outputs) connect to desired outcomes and the long-term change that the program ultimately intends to achieve. The tabular format is better suited for monitoring and accountability purposes.

APPENDIX B. Commissioning External Evaluators

Throughout this text, we have provided you with information about the task of evaluating public health programs with the expectation that you may start, or continue, leading evaluations in the future. However, even if you are well-trained to carry out evaluations it is entirely possible that you will find yourself in a position where you need to obtain services from an evaluator external to your organization. This could be for many reasons including needing extra person power to complete all of the evaluations that are needed, or finding that there is a need to have someone lead the evaluation who is not as involved in the public health program that is being evaluated. In these cases, it will be very important to understand what to look for in this evaluator and how to contract with them. Identifying and working with the right evaluator can be challenging. How do you know where to begin? Who will be the right fit? Who has the skills and abilities that are required to plan and implement the evaluations given the context?

Working with an Evaluator

First, it is important to understand that whoever commissions the evaluation will be working in partnership with any evaluator they hire or contract. Although hiring or contracting an evaluator may lessen the work involved for the funder and program staff members, they will not be able to turn over all responsibility of an evaluation. An evaluator cannot effectively do their job without involvement from the individual(s) who commissioned them to do the job and broader evaluation stakeholders.

As a result, the entity funding the evaluation should consider how they might be engaged in the evaluation activities and discuss with the evaluator how they envision the program staff members and any other partners contributing to the evaluation. To fully support the evaluation, it is critical that the program leaders make it clear to all staff members that they are expected to engage in evaluation, approximately how much time they should anticipate dedicating to the evaluation tasks, and how they will work with staff members to ensure that this is feasible given their existing workloads. As we described in **Chapter 2** and elaborated in **Chapter 4**, there are many approaches to evaluation. Some require minimal involvement from stakeholders (especially at the implementation stage) and others require extensive involvement.

Selecting an Evaluator

Deciding on a good evaluator for the program will depend on what the commissioner and stakeholders are looking for in terms of the mix of technical skills, familiarity with the program or context, and personal characteristics. Take time to discuss with the evaluator how they practice evaluation and consider the extent to which this aligns with the vision for this specific evaluation and the culture of the program and stakeholders. The evaluation funders may find that they prefer a different mix of control or involvement, especially in the early days when their relationship with the evaluator is just developing. Regardless, be clear about what tasks program staff members and partners will be expected to be involved with and prepared to allocate the time to the tasks.

If you have read **Chapter 3**, you'll recall that when selecting an evaluation, you should consider the following:

- Experience with program evaluation
- Ability to communicate effectively
- Basic knowledge of the topic the public health program focuses on
- Experience with the range of data collection strategies and evaluation designs that will best serve the program or the particular evaluation activity(ies) being planned
- Good references (from trusted sources)

It may also be helpful to review and find means to establish the extent to which the evaluator, or their team, possess the competencies needed for the project. The American Evaluation Associations' list of evaluator competencies (King & Stevahn, 2020) discussed in **Chapter 1** will be helpful in this assessment.

Special Considerations for Working with External Evaluators

Once an evaluator or evaluation team has been selected to implement the evaluation plan, it is essential that the funder draw up a contract to cover the work. This will ensure there is clarity of expectations by both the evaluator and the program. The contract will set out the main terms and conditions and may include the following:

- Who owns the data collected and the material that the evaluation produces
- How data security will be maintained
- How conflicts of interest will be addressed
- A detailed description of deliverables (e.g., presentations of work to stakeholders and others, frequency of communication)
- Timelines for all work and work products
- Budget and a payment schedule (e.g., periodic billing of hours, pay by deliverable)
- Details regarding procedures for budget modifications
- Discussion of sanctions and contract termination

Contract language should clearly describe the deliverables and timeline and should indicate that program staff members have an opportunity to review major deliverables and request modifications if they do not meet expected quality. The terms of the agreement should be tight enough to ensure that the program receives the product they want, but flexible enough to ensure that mid-course changes are possible.

Designate a key member of the program staff to manage the consultant and the evaluation process. This person will have responsibility for these activities:

- Serving as the point person for communications with the evaluator
- Making sure the evaluator has access to the information required
- Troubleshooting problems that arise
- Ensuring that products are delivered, and payments are made

Selecting an evaluator or evaluation team is an important first step in the implementation process. In the following sections we provide guidance about project management and budgeting—two activities that are necessary to translate plans into action.

APPENDIX C. Strategic Evaluation Plan Case Application

*The following description is a case scenario of a state asthma team going through the process of developing a Strategic Evaluation Plan. The case describes the evaluation context, key stakeholders, and the steps of strategic evaluation planning through a series of six vignettes. Each vignette walks through a step in the process of developing a strategic evaluation plan with stakeholders, providing examples of key components of the plan. Following this is an example of a completed strategic evaluation plan (**Appendix D**). As you read each vignette and the example tools provided, think about how you might adapt this process and apply the tools to craft a strategic evaluation plan for an organization with which you are familiar.*

Case Background

The Healthlandia State Asthma Program (HSAP) was awarded funding from CDC's National Asthma Control Program to deliver and evaluate asthma control activities within the state. The program activities relate to infrastructure such as surveillance and strategic partnerships as well as strategy implementation to expand asthma services and optimize systems. The state asthma program consists of a staff of seven, including a director, assistant director, program coordinator, three epidemiologists, and a student intern. The program works with local public health departments, schools, hospitals and clinics, and community organizations to (1) provide health education and asthma prevention and control services to adults and children living with asthma, (2) monitor the burden of asthma within the state, and (3) create policies and environmental conditions to reduce the burden of asthma in local communities, particularly for those disproportionately affected. A portion of the federal funding will support the evaluation of state asthma program activities, including the development and implementation of a Strategic Evaluation Plan prioritizing program evaluation activities over the next five years.

Vignette 1 – Getting to Know You

Sofia is a program coordinator for the State Asthma Program. Just one week ago, Sofia added a new part-time evaluator to her team (Anthony). Sofia is ready to hold her first meeting with Anthony and is anxious to put Anthony to work on the evaluation tasks that have been languishing on her desk. Anthony is eager to get started and learn what he can about the program. Here's a brief synopsis of their conversation:

Sofia: I'm so pleased to have you on board. We're really proud of the program we've developed and have even bigger dreams for the future. One of the first things we need from you is a plan outlining what we should evaluate in the coming five years. Please tell me what I can do to help you.

Anthony: Thanks. I'm looking forward to working with you. I'll rely on your program knowledge and expertise to help me plan an evaluation strategy. In fact, I can't do my job without your input, so I'm relieved you've offered to help.

Sofia: Feel free to chat with me anytime. I see evaluation as a priority, and I'll do what I can to help. How should we start?

Anthony: First, I'd like to get your thoughts on the purpose of this program. What do you think the ultimate goal of this program is? Years from now, how will we know whether or not we were successful?

Sofia: To me, the ultimate goal for this program is to help people who have asthma in our state better control their asthma so they can have a better quality of life. Reducing asthma disparities is critically important, too. Creating a solid infrastructure and strategically implementing the evidence-based strategies CDC identified in areas where asthma burden is highest can help us get there.

Anthony: Those are great goals. I love goals; the only problem is they take so long to achieve. How can we know a little sooner if our program is moving down a path toward success? I wonder if there is anything we can evaluate now to figure out if we're on the right path for the long term. Have you thought about what types of outcomes we might achieve along the way that could tell us if we're headed in the right direction?

Sofia: What a great question! I'm pretty practical so I know that we have to see progress along the way to keep staff morale high and to keep us focused on what makes a difference. One of the documents in this packet I've prepared for you may have some information that can help. CDC included a diagram in the funding announcement we responded to, and it shows how they envision our work at a high level (**Figure C.1**). Before I saw this model, I mostly thought about how different our program is from those in other states. After all, people in our state have different needs and our program has different partners and, unfortunately, fewer resources than some of these other states. This diagram helped me see that we are all working toward similar goals.

Anthony: This is helpful. It's called a logic model. It'll be good to have this as I work with you and the team to ask the right questions and develop a Strategic Evaluation Plan that will be right for this program. This solidifies it for me! This program is clearly committed to evaluation. I'm going to enjoy being a part of its success!

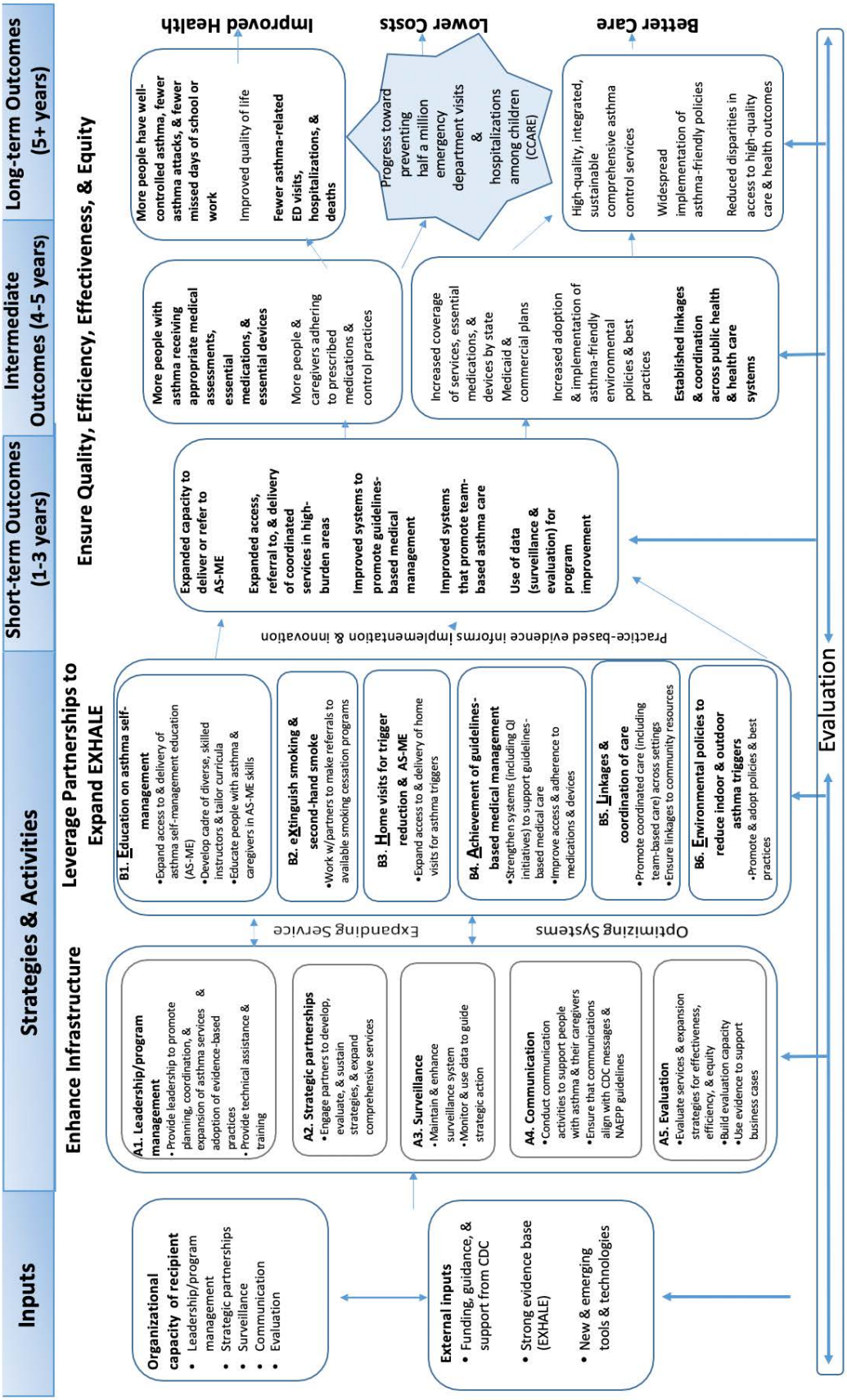


Figure C.1 Asthma Program Logic Model

Did You Notice…?

1. Sofia is clear about what she needs from Anthony in the near future—a Strategic Evaluation Plan for the asthma program covering the next five years.
2. Sofia makes clear that evaluation is a priority for her, and she follows up speech with action. She offers to help Anthony and says he should feel free to contact her at any time. She also shares materials with him that she has received from CDC that may help him, including the asthma program logic model provided in the Notice of Funding Opportunity.
3. Anthony recognizes that he will need to rely heavily on Sofia's knowledge of the asthma program. Not only is he new to the program, but his expertise is in evaluation, not in public health programming. He is open to materials developed by others that will help him understand the program.
4. During this first meeting, Anthony does not use evaluation jargon. He uses terms like ultimate goal instead of long-term outcome; how will we know we did our job well instead of criteria of merit or benchmarks; how can we tell sooner if we're moving down the path to success instead of short-term and intermediate outcomes. In later conversations, once he has a better sense of the program staff's familiarity with evaluation, he can introduce the evaluation jargon while building evaluation capacity.
5. Both Anthony and Sofia understand that while lofty goals help to motivate people, they also need more achievable milestones along the way to keep up their morale, their interest, and their level of engagement. Anthony uses his interpersonal skills to help communicate his commitment to helping the program succeed through evaluation. Evaluation is one way to identify and celebrate small successes along the way to ultimate goals.

Vignette 2 – Where Are We Going?

Anthony and Sofia continue their conversation about the Strategic Evaluation Plan. They review the program logic model and talk about the program's activities and short- and long-range goals. This helps them think about the activities and results they might have questions about that could be answered in an evaluation. We will listen in on their conversation.

Anthony: OK, I can see from this model that the longest-term results of your program are really those long-range goals you mentioned before, aren't they?

Sofia: Yes. We want to make life better for people with asthma, as well as reduce costs, and facilitate better care.

Anthony: OK. That all makes good sense. But I can also see results you expect to occur sooner. For example, you'd expect to see an expanded capacity to deliver or refer people with asthma to asthma self-management education. Knowing that these results are anticipated in the not-too-distant future helps us understand what we might evaluate to tell us if the program is on the right track.

Sofia: Are you saying we could actually start measuring the kinds of things in that first outcome column right now? Do you think we should do a statewide survey about those things? You know we're under a lot of pressure to demonstrate that our program is working. Our funders and partners want to know that.

Anthony: Well, depending on the activities conducted, you might not see much yet at the state level. An intervention in a specific county, for example, is not likely to result in change happening outside that county. But we could look at change within that specific jurisdiction to see if the intervention approaches that are in place are working.

Sofia: That's a good point because the interventions we're doing are different for each county. For example, one county is an industrial hub so we're working on environmental policies that limit the manufacturing exhaust and air pollution that gets out into the surrounding residential communities. We've seen an increase in child asthma in that jurisdiction in the last 10 years. But in a neighboring county that's mostly rural, air quality is much better, so we're focusing more on indoor air pollution. Folks are living in older homes that have mold and pests that can trigger asthma attack, so we're providing asthma self-management education to help people understand the triggers and learn how to manage their asthma.

Anthony: This is such an important context to be aware of as we begin to design evaluations for the interventions the program is conducting. Let's make sure to keep this in mind as we build out the Strategic Evaluation Plan. Getting back to your point about measurement—in terms of deciding on what to measure, within a specific jurisdiction for a given intervention approach we can use both the short-term and intermediate outcomes to help us decide. For example, for those with asthma and their caregivers, you want to see the skills they have acquired through asthma self-management education translate into good asthma management behaviors, because just having a skill doesn't mean you're going to use it.

Sofia: That makes sense. Basically, right now, we shouldn't think too big. Instead, we should use this model to think about what realistic changes we might see based on the actual activities we're conducting.

Did You Notice…?

1. Program logic models are tools that can help an evaluation team determine *what* to measure, *where* to measure, and *when* to measure.
2. Attempting to measure long-term outcomes prematurely can lead to poor or disappointing results.
3. Anthony cautions Sofia about measuring change at the state level if the intervention is more narrowly focused. For example, if you implement a suite of asthma control strategies in a specific county, then you want to measure change in that county where the intervention strategy took place; if possible, you should also look at a similar county where the intervention strategy did not occur.
4. Anthony suggests obtaining partner input to help decide what to evaluate. While a logic model—and an evaluator—can help show what might make sense to evaluate, figuring out what *should* be evaluated must come from the evaluation stakeholders. Only program managers and staff members, in consultation with key evaluation stakeholders, can identify the critical information needs that an evaluation will help address.
5. Often, when we embark on an evaluation, there is a tendency to jump into data collection. Sofia naturally did this by suggesting the use of a statewide survey to measure short-term outcomes. Anthony reinforces the importance of carefully planning evaluations before making any decisions about data collection.

Vignette 3 – Strategy Matters

As Sofia and Anthony continue their conversation, they realize that they need to be strategic when deciding what program activities to evaluate. To make these strategic decisions, they need to form a team of different people who can provide advice and a variety of perspectives on planning evaluation activities. Let's see how they go about deciding who should be involved and what roles they should play.

Sofia: I guess I thought writing the Strategic Evaluation Plan was something you could do for us. I was a little surprised when you mentioned needing to convene a planning team to help with that. Everybody's so busy!

Anthony: I can certainly help you with the Strategic Evaluation Plan, and I'll try to keep people's time commitment to a minimum. But this kind of planning isn't something I can do for you. I know evaluation but I don't know much about your program or your partners yet. Even if I were familiar with the program, it would still be important to include you all in the process, since you all likely hold different, and valuable, perspectives about this program.

Sofia: My own experience with evaluation planning comes from an evaluation we did for a school intervention in a previous funding cycle. We planned ahead about when to collect data, what to collect, and who was doing what. We wrote it all down so everyone was on the same page. Is that what you mean?

Anthony: Not exactly. You're right about wanting to plan each evaluation in advance, but I'm talking about an earlier step that involves how you decide what evaluations to do in the first place. It's thinking strategically about what aspects of your program you want to evaluate over the next five years. I'm guessing that you can't afford to do every evaluation that seems like a good idea. So, you're going to have to pick and choose.

Sofia: You're right about that. But how do I know today what evaluations will be the most important to do three or four years from now?

Anthony: Great question. We don't have a crystal ball. All we can do is develop a Strategic Evaluation Plan based on what we know now, and what we think is important. We'll revisit this Strategic Evaluation Plan at least once a year as we learn from evaluations we've done and as the program grows and changes.

Sofia: Okay, well I'm certainly willing to give this a try. How do we start?
Anthony: As a first step, I'd like to get some documents from you that describe the program goals and activities. I'll look through these and list the activities related to infrastructure such as surveillance and strategic partnerships, as well as asthma strategy implementation that stand out as particularly important to the program. Then, we should invite a small group of stakeholders (half a dozen or so) to help us think through which activities would be best to evaluate over the next five years. They need to be a pretty committed group, as we'll need their input a great deal this year and periodically over the next five years. We want folks who have a broad perspective on the program

rather than stakeholders who are interested in only one activity.

Sofia: OK, I can think of some people who should be involved. You and I will clearly be involved, and I'm sure our epidemiologist will be interested since she's been involved in evaluation in the past and knows our data systems. Maybe someone from the American Lung Association, as they've been a very strong partner from the beginning. Since we want to expand the asthma strategy on linkages and coordination of care, I think it would also be good to have one of the local medical professional organizations involved.

Anthony: Well, that sounds like a good group of folks. It will be important to note how they will contribute to the evaluation planning (**Table C.1**). We should have a name for this group to recognize their contributions. How about the Strategic Evaluation Planning Team?

Sofia: OK, that makes sense. I'll contact stakeholders who might be willing to help us out and set a time for the first meeting.

Anthony: Right. There are a lot of potential things we could start evaluating. I think a good first step would be to sit down with some other partners to think through more details and come up with a clear strategy for what we want to evaluate and when. That way, we'll feel more confident that we're getting the information we need, when we need it.

Table C.1 Strategic Evaluation Planning Team – Contributions, Roles, and Future Involvement

Stakeholder Name	Title and Affiliation	Contribution to Evaluation Planning	Role in Implementing Evaluations
Sofia	Asthma Program Coordinator, Healthlandia State	Oversees development of Strategic Evaluation Plan with evaluator; contributes knowledge of state asthma program	Oversees overall implementation of Strategic Evaluation Plan with evaluator
Anthony	Asthma Program Evaluator, Healthlandia State	Responsible for developing Strategic Evaluation Plan with collaborators; contributes knowledge of evaluation	Responsible for facilitating and carrying out evaluation implementation process; assessing evaluation capacity needs and planning for how to meet them
Leticia	Asthma Program Epidemiologist, Healthlandia State	Provides expertise on state asthma data systems	Oversees data collection and analysis for evaluations
Oscar	Program Director, American Lung Association	Provides expertise on state asthma strategy and community outreach to other stakeholders with relevant knowledge who should be involved; as the team with the longest connection to the program and the most institutional memory, provides insight into the program's context and coalition politics	Is consulted and informed of implementation progress
Adelmira	Program Associate, Healthlandia Medical Association	Provides expertise on expanding the asthma strategy on linkages and coordination of care	Is consulted and informed of implementation progress

Did You Notice…?

1. As much as Sofia might like to turn everything relating to evaluation over to her evaluator, talking to Anthony helps her recognize that she and other program staff members and partners will need to commit time to the strategic evaluation planning process. Sofia, with her program knowledge, and Anthony, with his evaluation knowledge, are both essential to the process.
2. Sofia has specific reasons for each team member she plans to invite. Some are invited because of their past efforts on behalf of the program, others because they represent important new directions.
3. Sofia and Anthony keep the core planning team relatively small so that it will be easier to conduct meetings and make progress on developing the Strategic Evaluation Plan. Others can be called in as needed for their specific expertise.
4. Once the Strategic Evaluation Plan is finished, it should not be considered set in stone. It must be revisited at least annually, and sooner if the program undergoes a major change.

Vignette 4 – Let's Get Picky

*Prior to the second meeting of the team, Anthony prepared a draft list of criteria (**Table C.2**). He also distributed a set of activity profiles (**Table C.3**), which were revised after team discussions during the kick-off meeting, followed by subsequent discussions with stakeholders.*

Anthony: Remember, our task is to choose activities to evaluate across all program components so that we have a broad picture of the program at the end of five years. In our last meeting, we refined the draft logic model. I presented and discussed some of the program's activities in detail, as summarized in the activity profiles in your packets. This time, we'll prioritize the activities as possible candidates for evaluation, according to criteria we develop together. Any questions?

Epidemiologist: Will we have different criteria for the different components? For example, it seems like criteria that fit infrastructure activities might not apply to expanding asthma control strategies.
Anthony: Excellent point. We'll be looking at activities within each of the major program components separately, so there's no reason we need the same criteria for each component. On the first page of your handout is a draft list of criteria I've pulled together (**Table C.2**). Please take a few minutes to look this over. (Group members review draft criteria.)

Anthony: Let's begin with the infrastructure criteria. What's important to consider when deciding which infrastructure activities to evaluate?

Epidemiologist: For an evaluation of our surveillance activities, I'd say "Information Need" is quite important. There are a number of decisions we're trying to make about what data to analyze in the near term versus the long term so I see "Information Need" as a criterion that could help us identify surveillance activities that are high priority for evaluation.
Anthony: That makes sense to me. What about activities relating to the services strategies?

American Lung Association Representative: I'd like to make sure we apply the criterion of Sustainability in our prioritization process for our home visits for trigger reduction and asthma self-management education activities. We really need to think about how to engage payers on this intervention to help ensure that they can exist over the long term.
Anthony: Are there any criteria we should remove or add? Do some apply to all the components?

Medical Association Representative: Sure. "Cost" applies to everything. We could prioritize resource-intensive activities for evaluation. Better yet, we could identify activities that are absolutely essential to our success. I'd vote for dropping "Cost" as a criterion and adding something like "Importance." "Information Need" and "Importance" can easily be applied to all our activities.

Sofia: With my program hat on, I'd like to include the criterion Challenges. If there are activities within our program that have faced difficulties getting launched or sustaining themselves, I'd want to pay some attention there. Evaluation could provide information we need to improve the situation. Also, I would be in favor of dropping other criteria we have discussed in order to add Equity. All of our programmatic activities, whether our strategic partnerships, surveillance, or our interventions need to move us toward greater health equity. Having information from evaluations about activities

that have a strong focus on health equity will not only help us to better understand how to improve that work but also how to leverage any successes to do better in other areas of our programming.

*The group continues until they have selected a final list of criteria and have ranked each activity as high, medium, or low priority against each criterion. Those activities ranked highest across multiple criteria are the evaluation candidates to be considered for inclusion in the Strategic Evaluation Plan (**Table C.4**).*

Table C.2 Potential Criteria for Evaluation Prioritization

Criterion	Information Required for Prioritization
Cost	What financial resources have we invested in this activity?
Labor or time intensive	How much staff member time have we invested in this activity?
Prior evaluation	Have we evaluated this activity before?
Performance	Does information from our performance measurement system indicate a need for more in-depth examination of this activity?
Equity	Does this activity promote health equity?
Maturity	What is the stage of development or implementation for this activity?
Stakeholder interest	How interested are our stakeholders in this activity?
Sustainability	How much does this activity contribute to the sustainability of the program?
Centrality	How connected is this activity to our partners across the jurisdiction?
Plan alignment	How closely aligned is this activity with our plan?
Plausible outcomes	Can this activity reasonably be expected to lead to the intended outcomes?
Disparities	Will this activity reduce health disparities?
Focus	Does this activity affect the population(s) we are interested in reaching?
Reach	How many people in our jurisdiction are (or could be) affected by this activity?
Challenges	Are we (or do we anticipate) struggling with this activity?
Pilot	Do we plan to expand this activity?
Information need	How critical is the evaluation information for making near-term decisions?
Improvements	Would evaluating this activity likely result in recommendations for programmatic improvement?
Use	Is it likely that results or recommendations from this evaluation will be used by the intended audiences?

Table C.3 Program Activity Profile

Program Component	Infrastructure
Title of Activity	Identify and fill gaps in existing surveillance data
Description of Activity	Identify gaps in existing asthma surveillance data, understand missing pieces in reporting, and identify strategies for filling in gaps.
Duration of Activity	On-going
Partner Involvement	State asthma surveillance team is responsible for identifying and filling gaps. Currently engaging with elementary and middle school nurses and local clinic managers to fill in gaps around childhood asthma events.
Cost of Activity	$10,000 annually
Contribution to Intended Program Outcomes	**Short-term:** Existing gaps in surveillance data are known and points of contact for addressing data quality are known. **Intermediate:** Points of contact are aware of system gaps and engage in efforts to improve existing data Availability and use of data that is of higher quality (surveillance) **Long-term:** Better-informed, higher quality asthma control services designed and implemented Reduced disparities in access to high quality care & health outcomes
Known Challenges in Conducting the Activity	High turnover of school nurses in some school districts, leading to breakdown in data entry and reporting.
Performance Measure Data	NA
Prior Evaluation	None to date

Table C1.4 Activities Rank Ordered by Criteria (Partially complete)

Activity	Criteria			
	Information Need	Equity	Importance	Challenges
Surveillance				
Identify and fill gaps in existing data	High	High	High	Medium
Assess data quality	Medium	Medium	High	High
Analyze data	Low	High	Medium	Medium
Disseminate findings	High	Low	Low	Low
Interventions				
Home-based asthma triggers	High	Medium	Medium	Medium
Education on asthma self-management	Low	Medium	Low	Low
School and Clinical care coordination	High	High	High	High

Did You Notice…?

1. As the lead evaluator, Anthony does much of the upfront work to prepare for Strategic Evaluation Planning Team meetings. This helps him to become familiar with the program, while also making sure the meetings run smoothly and don't go over the scheduled time limits.
2. An important role Anthony plays is encouraging discussion and facilitating development of consensus among team members. His strong interpersonal skills are critical here. He also offers his opinion and expertise.
3. The activities Anthony plans for the Strategic Evaluation Planning Team do not require evaluation expertise, but, rather, team members' sound knowledge of the program and its activities. Members of the Strategic Evaluation Planning Team do not need to be trained evaluators. They need to be familiar with the asthma program, willing to learn about evaluation, and ready to commit their time to the strategic evaluation planning process.
4. Anthony gave team members a list of possible criteria to use in choosing which aspects of the asthma program to evaluate. However, he recognizes that only those involved in the program can determine the criteria that are most important to them.
5. The group chooses to select a limited number of criteria in order to make the prioritization process more manageable. In a priority-setting process such as this, deciding which criteria are not important is just as vital as deciding which ones are important to the team.

Vignette 5 – A Balancing Act

*The agenda for this fourth meeting is to discuss the feasibility of conducting the proposed evaluations and potential use of the evaluation findings. Prior to the meeting, Anthony prepared a table that lists the evaluation questions the group developed during Meeting 3 (**Table C.5**), as well as some suggested evaluation designs, data collection methods, and data sources (**Table C.6**) that could be used to answer the evaluation questions posed by the group. Additional columns on the table will be completed by the group to capture when data collection would begin, the date evaluation results are needed, estimates of resources needed, and possible partner contributions. We join the group midway in their discussion.*

Anthony: As we look at all of the evaluation candidates, we see a number that will be resource-intensive or require that we get going right away. It's probably not feasible to do all of these evaluations. Are there some of our candidates where the available data sources may be problematic and the results less accurate or reliable? What about proposed evaluations for which the results may not be all that useful, possibly because they'll come too late or because they don't address the complexity of the activity?

American Lung Association Representative: I think we could simplify the outcome evaluation of the home-based asthma triggers intervention by not having a control group. That would mean less data collection. We'll still have pre-post data.

Epidemiologist: We could do this, and I'm right with you when it comes to reducing the workload. But I'm concerned that eliminating the control group will not provide us with strong enough results to help us answer the causal question we posed.

Sofia: I agree. We'll have to include a control group; otherwise, the findings won't be credible to outsiders who are looking to use or fund this intervention. Where else could we scale back, both in terms of cost and effort required right away?

Medical Association Representative: We are charting some new territory with the School and Clinical Care Coordination intervention, so we definitely could use some information to help fine-tune the intervention itself. I'm not so concerned with doing an outcome evaluation now, as the program itself is too new.

Anthony: That makes sense. A new intervention is likely to go through quite an evolution, which makes outcome data difficult to interpret. At this point, focusing the evaluation on implementation issues will provide the most useful information and cut the costs somewhat.

Table C.5 Example Evaluation Question Development Table for School and Clinical Care Coordination Intervention

Evaluation Candidate	Question Type	Questions	Question Priority (High Med Low)
School and Clinical Care Coordination	Outcome	To what extent are documented improvements in clinical care resulting in health improvements for clinic patients who have been diagnosed with asthma?	Medium
	Process	How engaged are partners in the care coordination activities? To what extent and in what ways are they participating?	Medium
	Process	How efficient is the clinical care coordination, and what opportunities exist to improve efficiency?	Low

Table C.6 Example Evaluation Design and Data Collection Summary for School and Clinical Care Coordination Intervention

Question	Possible Evaluation Design(s)	Potential Data Collection Methods	Possible Data Sources	Data Collection Begins	Final Results Due	Resources Required
How engaged are partners in the care coordination activities? To what extent and in what ways are they participating?	Multi-site case study	Document review; Semi- structured interviews; Online survey	Care coordination meeting minutes (includes attendance); school nurses, clinic office managers; asthma program staff	Year 2	Year 3	Modest
How efficient is the clinical care coordination, and what opportunities exist to improve efficiency?	Multi-site case study	Observations and semi-structured interviews Observations and semi-structured interviews	School nurses, clinic office managers	Year 3	Year 3	Modest

Did You Notice…?

1. Anthony began the evaluation planning process by using normal language to talk about evaluation concepts, but over time he has introduced the evaluation terms that are a kind of shorthand used in the profession.
2. By this fourth meeting, team members are clearly comfortable using the evaluation jargon— terms such as control groups and pre-post data. This shared understanding is an example of process use, which is the development of evaluation knowledge and skills as a result of engaging in evaluation activities. Over time, process use builds the evaluation capacity that helps asthma program staff members and stakeholders become stronger evaluation partners.
3. Team members balance the feasibility of doing an evaluation with the level of evidence desired by intended users of the evaluation findings. The ALA representative suggests removing a control group from an evaluation to help reduce costs. However, Sofia and the epidemiologist are concerned that doing so may compromise the likelihood that intended users will consider the evaluation findings credible enough to take action.
4. In balancing feasibility and utility considerations for the evaluation of the Clinical Care Coordination intervention, the group judged the utility of outcome data to be less important than the process data because the intervention is in the early phases of implementation. An evaluation of this intervention focused on implementation issues may be more meaningful.

Vignette 6 – Work Less, Reap More

Sofia and Anthony continue discussions with their Strategic Evaluation Planning Team to find efficiencies in data collection across evaluations for all components of the state asthma program.

Sofia: It seems we may still be stretched a bit thin conducting all of these evaluations. I'd like to discuss ways to integrate, coordinate, and economize across the entire set (**Table C.7**).

Anthony: Agreed. Looking at our priority evaluations, can we find ways to increase our efficiency?

Epidemiologist: We definitely want to identify how we're doing on filling gaps in our surveillance data. I originally thought that a survey of data users would contribute helpful information. It could, but I worry that it might not give us specific enough information to know how to respond.

Anthony: Focus groups, either in person or by telephone, might be an efficient way to get this information. You can obtain multiple perspectives about what is needed and how best to respond. Also, you may find that the dialogue among participants raises issues and solutions that may not have come to the surface with a survey.

Epidemiologist: Yes, that's a good point. A few telephone focus groups would be fairly inexpensive and would allow us to clarify respondents' comments.

Sofia: You could tack on a few questions about whether the data are used to focus interventions. That would be a way to address some of the other surveillance evaluation questions we had. You know, Melinda on my staff would make an excellent focus group facilitator, especially if she had some focus group training.

Anthony: Let's check on her interest. Maybe we could support her to take a workshop or course on facilitation techniques. What about partnerships? Is there a way to simplify data collection there?

American Lung Association Representative: Yes, I think so. A priority partnership question has to do with how CDC-funded programs leverage additional resources. I think we could make some phone calls to the directors of those programs to find out what they're currently doing to support asthma and what they see as untapped potential.

Medical Association Representative: I confess that I'm not hesitant to request that the School and Clinical Care Coordination intervention monopolize the remaining resources.

Anthony: All of the evaluation questions for that intervention focus on data collected from school nurses and clinic office managers, so that's efficient. I worry about overburdening the school nurses and office managers—we should brainstorm ways to make this as painless as possible for them.

American Lung Association Representative: I'd like us to remain open to the possibility of evaluating the other intervention—the home-based asthma triggers intervention. The initial walk-

thru inspections that are part of the intervention itself serve as baseline data. Some post walk-thru inspections and interviews with the families would be very informative. If we can postpone the decision, there may be some year-end funds we could contribute.

Table C.7. Issues to Consider When Looking Across Proposed Evaluation Strategies

Area	Definition	Issues to Consider
Evaluation Design	What evaluation designs are proposed?	• Will a proposed evaluation design be suitable for answering multiple evaluation questions? • What, if any, unintended consequences may result from implementing the proposed evaluation design?
Data Collection: Respondent Population	From whom is information being collected?	• If several data collection strategies have the same respondent population, can you collect information for more than one purpose using a single data collection tool? • Are data collection activities concentrated too heavily on one respondent population? • Can burden be shared more equitably? • What are respondents' previous experience with evaluation? How will these experiences shape engagement with respondents? • To what extent do these data collection methods align with the values and interests of respondents? Will data collection methods allow for authentic input from respondents?
Data Collection: Timeline	When is information being collected?	• How can evaluation data collection needs be integrated into the program timeline? For example, if baseline data need to be collected, program activities may need to be delayed. • If information on different evaluation activities needs to be collected at the same time, do you have the resources to conduct multiple evaluation activities simultaneously? • What contextual factors need to be taken into account when considering the timing of data collection (e.g., school breaks, holidays, busy periods for respondents)?
Data Collection: Source	From where are data being collected?	• Can the same data source be used for multiple evaluation activities? • Can a single source be modified or enhanced to support your strategies for the future? • How frequently have you used these methods for data collection purposes? To what extent (if any) are personal biases influencing your selection of data collection strategies?
Who	Who will conduct the evaluation activity?	• Do you have the personnel and resources to conduct the prioritized evaluations given the strategy proposed? • Do they have the necessary skills and expertise? If not, how could they obtain these skills? • Can you leverage additional evaluation assistance from partners?
Analysis	How will the data be analyzed?	• Who will do the analysis? • Do they have the necessary skills and expertise? If not, how could they obtain these skills? • Can you leverage additional analytic capability from partners? • How will the results of the analysis be validated?

Use	How will the information from the evaluation likely be used?	• Will the information be provided in time to inform decisions? Who will use the information provided? • In what ways, if at all, will these findings benefit communities? • Are there capacity-building activities that need to be conducted with intended users to increase the likelihood that results will be used? • What is the potential for the misuse of findings and how will this be mitigated?

Did You Notice…?

1. The Strategic Evaluation Planning Team explores how to economize and leverage resources without sacrificing the utility and accuracy of the evaluation findings. For example, Sofia suggests adding a few questions to an already planned data collection activity (the focus groups) to answer a related evaluation question. The epidemiologist thinks through what type of information will be most useful for answering the evaluation questions: realizing that too little information could end up being useless (e.g., survey results that lack specificity). She proposes an alternative, more feasible and more cost-effective approach that has the potential to yield more useful information.
2. In recommending focus groups with data users, Anthony points out that, in addition to being efficient in terms of time and expense, stakeholder focus groups have an advantage over surveys in terms of providing both an exchange of ideas and a critique of proposed options.
3. One way to extend evaluation resources is to build capacity in house. Money that could be spent hiring a professional focus group facilitator to conduct the data user focus groups might better be spent supporting a promising staff member to gain that skill.
4. Paying attention to respondent burden is important. Anthony is conscious that the School and Clinical Care Coordination intervention itself demands considerable extra time from school nurses and clinic office managers beyond their routine responsibilities. Data collection for the evaluation component of the intervention needs to be efficient, possibly even integrated into the intervention itself through participant forms and checklists, for example.
5. Leveraging partner contributions is a good way to extend evaluation resources. For example, the ALA representative identifies an opportunity to evaluate a second intervention. His organization may even be able to contribute if the evaluation timeline can be pushed to year's end.

As we can see from Sofia and Anthony's work with the Healthlandia Asthma Program, developing a Strategic Evaluation Plan requires thoughtful planning and input from key stakeholders. Planning should include discussion of 1)the purpose of the Strategic Evaluation Plan and roles for key stakeholders, 2) the criteria to prioritize evaluation candidates to be included in the Strategic Evaluation Plan, 3)the potential evaluation questions and data collection methods for each evaluation candidate, 4) any activities that may be needed to build evaluation capacity to carry out the Strategic Evaluation Plan, and 5) a communication plan for keeping key actors informed of progress. The process of a developing a Strategic Evaluation Plan may evolve and change over time, which is why it is important to revisit it on a regular basis to make updates. Involving a variety of stakeholders will help ensure you have considered important aspects of planning for evaluations across a period of time and will help you make adjustments as needed.

There is no one right way to develop a Strategic Evaluation Plan. The most important thing is to develop one. We provide an example in ***Appendix D*** *of a Strategic Evaluation Plan for the Heartlandia State Asthma Program that includes the different templates and tools from the Strategic Evaluation Plan Outline presented in* ***Chapter 3****. You may modify these templates and tools for your own Strategic Evaluation Plan or make other modifications to fit your organizational context.*

APPENDIX D. Example Strategic Evaluation Plan

Healthlandia State Asthma Program

Strategic Evaluation Plan for
2020-2025

Prepared by:

Sofia and Anthony

Healthlandia State Asthma Program

December 7, 2020

1.0 Program Background and Purpose of Strategic Evaluation Plan

1.1 Program Background

The Healthlandia State Asthma Program (HSAP) aims to reduce the burden of asthma; coordinate care for people living with asthma; and lower healthcare costs and reduce inequities, morbidity, and mortality due to asthma in the state of Healthlandia. The program aims to reach these goals by expanding the reach, quality, effectiveness, and sustainability of asthma control services through strengthening leadership and infrastructure.

Figure 1 is a logic model of the state asthma program. The overarching program strategies are to enhance infrastructure and leverage partnerships to expand asthma control services among marginalized groups in the state, including immigrant and refugee populations. Enhancing infrastructure includes developing and leveraging leadership and program management to adopt evidence-based practices, using surveillance data to identify and engage high burden areas, and developing and enhancing existing partnerships to expand comprehensive services. Enhancing infrastructure includes identifying asthma health inequities in the state and expanding outreach to priority populations using culturally responsive education on asthma self-management and referrals to culturally sensitive healthcare facilities.

The short-term outcomes of these activities include expanded capacity to deliver culturally responsive asthma self-management education; expanded access, referral, and delivery of coordinated, culturally responsive care in high burden areas; and improved systems to encourage guidelines-based care. These will lead to the intermediate outcomes of more people with asthma receiving appropriate medical assessments, medications, and devices as well as established linkages and coordination across public health and healthcare systems. Ultimately, the long-term goals of the program are well-controlled asthma all individuals in the state who have this health condition, improved quality of life, and reduced asthma disparities within the state.

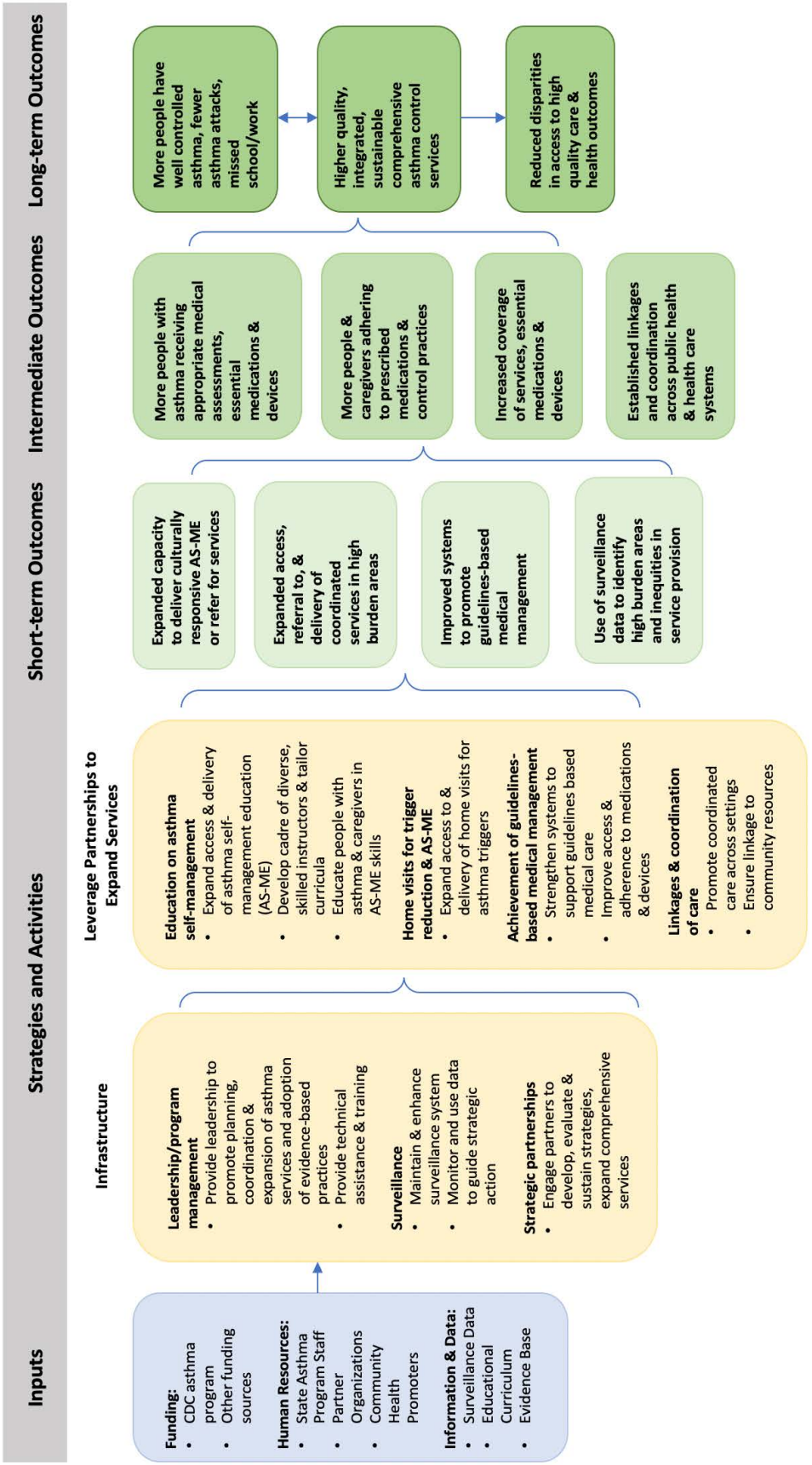


Figure 1. Healthlandia State Asthma Program Logic Model

1.2 Purpose of the Plan

This Strategic Evaluation Plan outlines priority strategies, activities, and outcomes to evaluate over the next five-year period. Evaluation of priority areas will help us gather data to make programmatic improvements and monitor change over time to reach our ultimate goals in a timely and effective manner. By evaluating our efforts, we will understand which components are working as intended, which need to be changed, and how we are achieving our short, intermediate, and long-term outcomes to help people living with asthma and their families.

Findings from our priority evaluations will be used by key stakeholders implementing our asthma program components to identify aspects that are working well, those that need to be modified, and gaps in implementation. Findings will also be used to monitor progress toward outcomes, which will be used to inform our key partners and identify additional resources for continued programming.

2.0 Methods for Developing and Updating the Strategic Evaluation Plan

2.1 Stakeholders

Table 1 lists the Strategic Evaluation Planning Team members and their respective roles and contributions to developing and implementing the strategic evaluation plan. The evaluator interviewed additional staff and coalition members to assess the context of the program and get a sense of the comfort level with evaluation of various stakeholders.

Table 1 Strategic Evaluation Planning Team – Contributions, Roles, and Future Involvement

Stakeholder Name	Title and Affiliation	Contribution to Evaluation Planning	Role in Implementing Evaluations	Considerations to Support Participation
Sofia	Asthma Program Coordinator, Healthlandia State	Oversees development of Strategic Evaluation Plan with evaluator; contributes knowledge of state asthma program	Oversees overall implementation of Strategic Evaluation Plan with evaluator	
Anthony	Asthma Program Evaluator, Healthlandia State	Responsible for developing Strategic Evaluation Plan with collaborators; contributes knowledge of evaluation	Responsible for facilitating and carrying out evaluation implementation process; assessing evaluation capacity needs and planning for how to meet them	
Leticia	Asthma Program Epidemiologist, Healthlandia State	Provides expertise on state asthma data systems	Oversees data collection and analysis for evaluations	
Oscar	Program Director, American Lung Association	Provides expertise on state asthma control strategy and community outreach to other stakeholders with relevant knowledge who should be involved; as the team with the longest connection to the program and the most institutional memory, provides insight into the program's context and coalition politics	Is consulted and informed of implementation progress	Advance notice for requests for information and perspectives from community stakeholders is necessary

Adelmira	Program Associate, Healthlandia Medical Association	Provides expertise on expanding the asthma control strategy on linkages and coordination of care	Is consulted and informed of implementation progress	Quarter 3 of calendar years is especially busy time

2.2 Methods Used to Develop the Strategic Evaluation Plan

The Strategic Evaluation Planning Team used the nominal group technique to identify the criteria used to prioritize programmatic activities for evaluation. Rather than generating possible criteria for each programmatic component (i.e., partnerships, surveillance, and interventions), the team decided to generate one list of potential criteria that could be used across all components. After generating a comprehensive list of 12 criteria, each team member individually ranked each criterion in order of priority. The evaluator tallied each ranking to identify the top four criteria. The team subsequently discussed the ranking and revised the criteria in alignment with their shared values. The final criteria used in the prioritization process are listed in **Table 2**.

The team then applied a similar technique to identify and prioritize the programmatic activities. As seen in **Table 2**, in-depth discussions with staff members and other program stakeholders were the primary means through which we obtained insights on each of the criteria. Based upon the results of these conversations, the team assigned a value of low (1), medium (2), or high (3) to every criterion for *each* activity. The activities with the highest overall score are considered the priority evaluations for the upcoming five years. The results of the ranking process are listed in **Table 3**.

Table 2 Prioritization Criteria

Criteria Used	How Criteria Were Applied	Information Supporting Criteria Determination
Information need	Activities for which the information need is higher in the near term to make decisions were given a higher priority.	Situational analysis based on stakeholder discussions
Equity	Activities with potential to diminish structural supports for inequities were given a higher priority.	Situational analysis based on stakeholder discussions
Importance	Activities viewed as essential to programmatic success (i.e., achieving long-term outcomes) were rated higher than others.	Situational analysis based on stakeholder discussions
Challenges	Activities where staff members have consistently encountered difficulties were rated higher than others.	Discussions with staff members responsible for program implementation

2.3 Proposed Methods for Reviewing and Updating the Strategic Evaluation Plan

We aim to update the Strategic Evaluation Plan on an annual basis beginning in January of each year. When possible, we will convene in person for a one-day session to update the plan. In advance of this session each team member will be responsible for gathering insights from evaluation stakeholders and evaluation implementation teams about the lessons learned from the evaluation *process* as well as any additional questions that arose as a result of the evaluation findings. The team members will review these insights in advance of the one-day meeting and come prepared to reflection on these lessons learned with the intention of updating the plan. We will make updates directly to the document, noting changes made to each section using the brief checklist in the "Wrapping Up" section of this document.

3.0 Priority Evaluations

3.1 Prioritized Evaluations

Table 3 lists the priority evaluations.

Table 3 Rank-ordered List of Priority Evaluation Candidates

Leadership and Program Management	Surveillance	Strategic Partnerships	Interventions
Provision of technical assistance	Identify and fill gaps in existing surveillance data	Acquisition of funds	Home visits for trigger reduction
	Assess surveillance data quality		School & clinical care coordination

3.2 Overarching Timeline

Table 4 presents an overarching timeline for the prioritized evaluation activities related to the surveillance infrastructure activity and **Table 6** refers to evaluation activities for the expanding services component.

Table 4 Timeline with Sequencing of Proposed Evaluation Activities

<table>
<tr><th></th><th>Year 1 (2021)</th><th>Year 2 (2022)</th><th>Year 3 (2023)</th><th>Year 4 (2024)</th><th>Year 5 (2025)</th></tr>
<tr><td>Program Milestones</td><td>School and clinical care coordination procedures developed

Protocol for home visits intervention testing begins</td><td>Acquire new surveillance data

Protocol for home visits intervention testing begins</td><td>Protocol for home visits intervention testing ends
Mid-term partnership reflection meeting (2-day)</td><td>Apply for future funds</td><td>Apply for future funds</td></tr>
<tr><td rowspan="4">Evaluations</td><td></td><td>Provision of technical assistance</td><td></td><td></td><td></td></tr>
<tr><td>Identify and fill gaps in existing surveillance data</td><td></td><td>Assess surveillance data quality</td><td></td><td></td></tr>
<tr><td></td><td colspan="3">Acquisition of funds</td><td></td></tr>
<tr><td></td><td colspan="2">School & clinical care coordination</td><td colspan="2">Home visits for trigger reduction</td></tr>
<tr><td>Capacity Building</td><td>Training with epidemiologists to assist in data collection and analysis</td><td></td><td>Training with staff members on culturally responsive evaluation and data analysis for equity</td><td></td><td></td></tr>
<tr><td colspan="2">Mini trainings at the start of select asthma partnership meetings to orient all participants to evaluation concepts.</td><td colspan="2">Trainings for community members (involved in home visits intervention) to understand purpose of the evaluation and their rights as participants.</td><td></td></tr>
</table>

4.0 Summary of Each Priority Evaluation

Evaluation 1 will focus on the state's efforts to provide technical assistance to local health departments to strengthen their capacities in forming, engaging, and sustaining a local asthma coalition. Information from this evaluation will be used by program staff members to refine the approach to technical assistance. More information on this evaluation is provided in **Table 5**.

Evaluation 2 will examine how well the program is doing at identifying and filling gaps in existing surveillance data. The evaluation findings will help to identify areas for improvement, including suggestions for more effective engagement with data stewards. **Table 6** provides more details on this evaluation.

Evaluation 3 will assess surveillance data quality after some of the gaps identified through Evaluation 2 have been filled. The aim is to use these findings to further strengthen the surveillance infrastructure. More details on this evaluation can be found in **Table 7**.

Evaluation 4 will focus on understanding the extent to which the statewide partnership has, or has not, been successful in acquiring new resources to support the program. Information from this evaluation will be used by the partnership to refine their efforts, with the expectation that this will lead to improvements in acquiring additional funds to sustain the program in years 4-5. Additional details are provided in **Table 8**.

Evaluation 5 will focus on the processes involved in implementing the school and clinical care coordination intervention. This evaluation will be performed relatively early in the implementation phase of this intervention with the expectation that the findings will be used to modify or adjust the implementation efforts. **Table 9** includes additional details.

Evaluation 6 will focus on the outcomes associated with the home visit trigger reduction intervention. A major focus of this evaluation will be examining the relative benefit of the intervention across populations, with a particular lens toward improving health equity. See **Table 10** for additional information.

Table 5 Evaluation 1 Profile

Activity Name	Provision of technical assistance
Program Component	Leadership and program management
Evaluation Justification	The partnership plans to significantly change the content and increase the delivery of technical assistance to local health departments in this funding cycle. Information needs are therefore high to understand how this new plan is working.
Evaluation Purpose and Use	To understanding the strengths and limitations of the new approach to technical assistance, including any facilitators or barriers to its successful implementation. The evaluation will be used by the program staff to refine the approach to technical assistance further.
Possible Evaluation Questions	How well was the new TA approach implemented? What facilitated successful implementation? What factors presented challenges to successful implementation? What aspects of the content was understood by the local health department staff? What aspects of the TA, if any, were confusing or poorly understood by the local health department staff?
Relevant Performance Measures	NA
Timing of Evaluation	February 2022-August 2022
Suggested Evaluation Design	Case study

Potential Data Sources	TA participants
Potential Data Collection Methods	Pre-post survey and follow-up interviews or focus groups
Cultural or Contextual Factors	Local health department staff members are incredibly busy. Data collection efforts need to be brief.
Potential Audiences	Healthlandia Asthma Partnership
Possible Uses of Information	Further adjustments to TA implementation process and content of TA
Estimated Evaluation Cost	$5,000

Table 6 Evaluation 2 Profile

Activity Name	Identify and fill gaps in existing surveillance data
Program Component	Surveillance
Evaluation Justification	High potential to reduce health inequities by facilitating a more robust and comprehensive data system.
Evaluation Purpose and Use	To understand how well the program is doing at identifying and filling gaps in existing surveillance data. The findings will be used to strengthen the existing surveillance system as well as relationships with data stewards to facilitate a more robust and comprehensive surveillance system for asthma. .
Possible Evaluation Questions	What measures have we taken to identify gaps in our asthma surveillance data over the past two years? When have these efforts succeeded and when haven't they? Why? How do data stewards view the asthma program's engagement with them around surveillance? Where are there opportunities for improvement in these relationships?
Relevant Performance Measures	NA
Timing of Evaluation	March 2021 – August 2021
Suggested Evaluation Design	Case study
Potential Data Sources	Surveillance work plans; Asthma epidemiologist; Data stewards from departments that do or could provide data for asthma surveillance
Potential Data Collection Methods	Document review; semi- structured interviews
Cultural or Contextual Factors	Staff members tend to be away during summer months, plan for data collection prior to this time.
Potential Audiences	Asthma program coordinator and asthma epidemiologist
Possible Uses of Information	Improve surveillance system
Estimated Evaluation Cost	$5,000

Table 7 Evaluation 3 Profile

Activity Name	Assess surveillance data quality
Program Component	Surveillance
Evaluation Justification	Additional potential (additional to evaluation #2) to enhance surveillance system, especially with respect to the ability to regularly examine health disparities.
Evaluation Purpose and Use	To obtain insights about how data sets included in the surveillance system can be improved. The findings will be used by the Healthlandia asthma epidemiologist to continue strengthening the quality of existing data used for surveillance.
Possible Evaluation Questions	What is the quality of the surveillance data after some of the gaps identified in Evaluation #2 have been filled? In what ways can the system be improved upon to better examine potential health disparities?
Relevant Performance Measures	% of missing data by data source
Timing of Evaluation	May 2023-December 2023
Suggested Evaluation Design	Case study
Potential Data Sources	Asthma epidemiologist, data stewards from other departments supplying data for asthma surveillance, surveillance data
Potential Data Collection Methods	Secondary data from surveillance system, semi-structured interviews
Cultural or Contextual Factors	Need to ensure access to surveillance data; ensure all interviews are completed prior to holiday season
Potential Audiences	Asthma program coordinator, asthma epidemiologist
Possible Uses of Information	Make adjustments to the data system and reporting protocols to enhance data quality
Estimated Evaluation Cost	$9,500

Table 8 Evaluation 4 Profile

Activity Name	Acquisition of new resources
Program Component	Strategic partnerhsips
Evaluation Justification	Information need is high regarding the extent to which the partnership is able to acquire new resources. Limited to no information is readily available on this topic at the moment, and it is necessary to help sustain the program past the current funding cycle.
Evaluation Purpose and Use	To understand the extent to which the statewide partnership has, or has not, been successful in acquiring new resources to support the program. Information from this evaluation will be used by the partnership to refine their efforts, with the expectation that this will lead to improvements in acquiring additional funds to sustain the program in years 4-5.
Possible Evaluation Questions	What steps has the partnership taken in the past 2 years to acquire additional funding? To what extent have these efforts been successful? When they have been successful, what factors have contribute to the success? What factors have contributed to unsuccessful attempts to secure funding?
Relevant Performance Measures	NA
Timing of Evaluation	April 2024-April 2025
Suggested Evaluation Design	Case study
Potential Data Sources	Written proposals, funder feedback, potential or current funders, partnership members (subset)

Potential Data Collection Methods	Document review, semi-structured interviews
Cultural or Contextual Factors	Funders may be hesitant to share insights. Avoid data collection through semi-structured interviews during buying season.
Potential Audiences	Asthma program coordinator, Healthlandia Asthma Partnership

Table 9 Evaluation 5 Profile

Activity Name	School and clinical care coordination
Program Component	Interventions
Evaluation Justification	Information need is high because this is a new intervention that the program is implementing and has high potential to face significant challenges during the early implementation phase.
Evaluation Purpose and Use	To obtain information about the strengths and challenges faced when implementing the scool and clinical care coordination intervention during its early stages. It is anticipated that the state asthma program coordinator and the school and clinical care working group of the Healthlandia Asthma Partnership will discuss the findings, build an action plan based upon the findings, and implement the plan to improve implementation efforts.
Possible Evaluation Questions	How, and in what ways, is the current implementation effort performing well? What barriers or challenges exist with respect to implementation?
Relevant Performance Measures	Proportion of eligible clinics in state enrolled in care coordination Proportion of eligible schools in state enrolled in care coordination
Timing of Evaluation	March 2022 – August 2023
Suggested Evaluation Design	Case study employing mixed methods
Potential Data Sources	Performance monitoring data, school nurses, clinic office staff members, clinic leadership
Potential Data Collection Methods	Analysis of performance monitoring data Online survey of school nurses and clinic office staff members Semi-structured interviews (or focus groups) with school nurses and clinic office staff members. Semi-structured interviews with clinic leadership.
Cultural or Contextual Factors	School nurses and the clinic office staff are extremely busy and will have limited time to engage in data collection efforts as respondents. Consider offering incentive that is tailored to these respondent populations; for interviews and focus groups consider scheduling during regular break times and when in person offer snack or lunch options.
Potential Audiences	Asthma program coordinator, school and clinical care working group of the Healthlandia Asthma Partnership, schools and clinics currently engaged in care coordination intervention and those potentially interested in joining.
Possible Uses of Information	The asthma program coordinator and school and clinical care working group can use the information provided from the evaluation to take actions to improve implementation. Schools and clinics interested in joining intervention may review the evaluation findings with an eye toward what factors are present in their specific settings that may faciliate or present challenges to implemention and base their decision to join or not in light of this information.
Estimated Evaluation Cost	$25,000

Table 10 Evaluation 6 Profile

Activity Name	Home visit trigger reduction
Program Component	Interventions
Evaluation Justification	This intervention is considered essential to the program's successful achievement of intended outcomes. Furthermore, the intervention was originally designed to increase health equity but there is limited information available to understand if it is doing so.
Evaluation Purpose and Use	To provide insights about the extent to which the intervention is achieving its intended outcomes, especially with respect to health equity and why. It is anticipated that the state asthma program coordinator and the community health workers engaged in the intevention will use the insights from this evaluation to modify implementation strategies in order to achieve greater success in future funding cycles. The Healthlandia State Asthma Partnership may also use the findings from this evaluation to support future proposals to acquire additional resources.
Possible Evaluation Questions	To what extent, if at all, are home visits for trigger reduction leading to reductions in harmful exposures, decreases in acute asthma syptoms, fewer acute visits to the emergency department or urgent care centers, fewer missed work and school days, and improved quality of life? What has contributed to these changes (positive and negative) in intended outcomes? In what ways, if at all, do these differ between groups participating in the intervention? When differences exist, what may be contributing to the difference?
Relevant Performance Measures	NA. All performance measures for outcomes are calculated at the level of the state, this intervention is not implemented statewide
Timing of Evaluation	February 2024 – September 2025
Suggested Evaluation Design	Quasi-experimental design – multiple measures with comparison group. Mixed methods – complement QE design with qualitative component for explanatory purpose..
Potential Data Sources	Participants, community health workers implementing the visits
Potential Data Collection Methods	Paper-based survey for participants Focus group with community health workers
Cultural or Contextual Factors	Need to better understand the literacy level of participants to decide on most appropriate way to administer the survey. Careful consideration regarding timing of community health worker focus groups is needed. Consider options for central location, provision of food if timing aligns with meal-times, and possible child care offering if during non-work hours.
Potential Audiences	State asthma program coordinator Community health workers Healthlandia State Asthma Partnership Other state asthma programs currently engaged in delivering a similar home visit intervention, or considering implementing a similar intervention.
Possible Uses of Information	It is anticipated that the state asthma program coordinator and the community health workers engaged in the intevention will use the insights from this evaluation to modify implementation strategies in order to achieve greater success in future funding cycles. The Healthlandia State Asthma Partnership may also use the findings from this evaluation to support future proposals to acquire additional resources. Other state asthma programs may leverage the lessons learned from this evaluation to modify their intervention approach or to make decisions about implementing a similar intervention within their jurisdiction.
Estimated Evaluation Cost	$40,000

Capacity Building Activities to Support Evaluation

To support the effective implementation of the proposed evaluations, the state asthma program evaluator will design and deliver several activities to build evaluation capacity. Each evaluation capacity building activity is described below, as the details for individual evaluations are developed further additional capacity building needs may arise and need to be added to this list.

- **Building knowledge among Healthlandia State Asthma Partnership members.** The evaluator will deliver mini trainings at the start of select asthma partnership meetings to orient all participants to evaluation concepts. The intention of these trainings is to orient the partnership to evaluation so they can more readily engage in the proposed evaluations over the course of the funding cycle and are better prepared to make use of the evaluative insights that result. The evaluator will conduct these trainings during the first two years of the funding cycle. Should there be a large amount of turnover on the partnership, additional trainings may be necessary in subsequent years.
- **Enhancing propriety through engaging home visit participants.** Individuals who live within the communities where the home visit trigger reduction intervention will be implemented are frequently asked to participate in public health interventions and research. We believe it is important that participants in this evaluation fully understand the purpose of the evaluation, the potential benefits and drawbacks of the evaluation, and their rights as potential participants prior to agreeing to participate. As a result, we will host several community training sessions during the enrollment period for the evaluation.
- **Enhancing evaluation skills among epidemiologists.** To reduce the costs associated with the two evaluations of surveillance, we propose engaging two to three epidemiologists within the Healthlandia State Health Department in implementing the evaluations. Epidemiologists have several relevant skills for evaluation and have expert knowledge of surveillance, therefore they are ideal for inclusion. However, it will be important that all epidemiologists who participate in the evaluation are trained on the common procedures and expectations for the specific surveillance evaluations planned. This training will take place during the first year of the program, prior to the implementation of the first surveillance evaluation (February 2021).
- **Improving cultural responsiveness for evaluation.** In preparing for the implementation of the home visit trigger reduction evaluation, the evaluator will host a training for program staff members (who agree to assist in this evaluation) on culturally responsive evaluation and data analysis for equity. This training will support a higher-quality evaluation aimed at examining health equity. We anticipate that this training will occur in mid-late January 2024

5.0 Communication Plan

We recognize the importance of regular communications throughout the implementation of this Strategic Evaluation Plan. Such communications are important for the effective implementation of this plan as well as for improving the plan each year. As a result, we developed a communication plan that ties directly to the Strategic Evaluation Plan implementation (**Table 11**).

Table 11 Communication Plan Summary Matrix

Information and Purpose	Audience(s)	Possible Formats	Possible Messengers	Timing	Person Responsible
Present final Strategic Evaluation Plan	Healthlandia State Asthma Partnership	PowerPoint presentation	Sofia	End-of year meeting	Sofia & Anthony
Provide general update on status of evaluations as proposed in Strategic Evaluation Plan	Strategic Evaluation Planning Team	Email	Anthony	Quarterly	Anthony
Share information about changes and lessons learned when implementing the Strategic Evaluation Plan	Strategic Evaluation Planning Team	Presentation and discussion	Anthony	Annual meeting	Anthony
Inform about specific upcoming evaluation activities	Asthma program staff members, Healthlandia State Asthma Partnership, Community Health Workers (for home visit evaluation)	Email	Sofia	2-3 weeks prior to evaluation implementation	Sofia & Anthony
Document and share synthesis of findings and lessons learned from evaluations during entire funding cycle	Strategic Evaluation Planning Team	Formal presentation	Sofia	End of funding cycle	Sofia & Anthony
	Asthma program staff members, Healthlandia State Asthma Partnership, Community Health Workers (for home visit evaluation)	Formal presentation and reflection session	Anthony	End of funding cycle	Anthony

6.0 Wrapping Up

Throughout the course of implementing the Strategic Evaluation Plan, Sofia and Anthony will meet on a quarterly basis to document lessons learned from each evaluation conducted. These meetings will include reflections on the *process* involved in carrying out each evaluation as well as the *results* of the evaluations. Anthony will ensure that modifications to the Strategic Evaluation Plan are documented by annotating the Strategic Evaluation Plan using the checkbox below. When changes occur, he will describe the changes and lessons learned during the course of implementation. Anthony will share these modifications with the Strategic Evaluation Planning Team during their annual meeting.

At the end of the funding cycle we will acknowledge the contributions of the Strategic Evaluation Planning Team members through a formal letter authored by the director of our department. In addition, we plan to host a celebration with the Team where we can reflect on our collective efforts and acknowledge and celebrate our successes.

___ Implemented as planned
___ Changes made (describe changes as well as the rationale for changes)

APPENDIX E. Common Evaluation Designs

As you will recall, Step 3 of the CDC Framework (Focus the Evaluation Design) involves identifying the evaluation questions and determining how to collect the information needed to address these questions. This decision-making process has two main components: (1) deciding on the evaluation's overarching design, and (2) deciding how to collect the data. In this appendix, we cover the first component, while **Chapter 4** provides information about the second. Our goal is to present a general introduction to evaluation design, whet your appetite for further information, and direct you to resources that will supply the level of detail required to construct a sound evaluation design.

General Description of Evaluation Designs

Three Categories of Evaluation Designs:

- Randomized/true experiments
- Quasi-experiments
- Non-experiments

Evaluators often join the profession through a back door. Many people engaged in evaluation might describe themselves as an "accidental evaluator" or a "Monday morning evaluator" meaning they have been asked to engage in evaluation work, but their training and professional experiences lie elsewhere (King & Stevahn, 2013).

As a result, the evaluation designs in use today have often come to us through a variety of disciplines, such as psychology, sociology, anthropology, epidemiology, and health services research. While having so many designs to choose from can seem overwhelming, the diversity of designs offers us opportunities to be creative in seeking answers to our evaluation questions.

Given the various origins of evaluation designs, there are many possible ways to categorize them. Here we adopt a categorization scheme proposed by Trochim (2020) for classifying a closely related set of designs—social research method design.[9] Under this framework, evaluation designs fall into one of three categories: (1) randomized or **true experiments**, (2) quasi-experiments, or (3) non-experiments. Descriptions of these categories follow, along with examples of specific design types that fall within each.

Randomized or True Experiments

Experimental designs are characterized by **random assignment** of participants into groups. In evaluation, this type of design is often referred to as a randomized controlled trial. The most basic design of this type consists of random assignment of participants to a group that receives an intervention (**intervention** or **treatment group**) and a second group that does not (control group). However, there are many ways a randomized or true experiment can be designed that move beyond this basic structure. According to Shadish, Cook, and Campbell (2002), "the variations are limited only by the researcher's imagination" (p. 259). For example, these designs can be spiced up in several ways, depending upon the evaluation questions.

- **The number of groups.** Participants can be randomized to more than two groups. For example, a control condition (which receives the standard intervention) could be compared to multiple intervention groups, each with a slightly different "take" on the intervention being evaluated (e.g., two groups might both receive a training intervention, but one would receive a longer training [nine sessions over three weeks] than the other [six sessions over three weeks]).
- **The number of time points when data are collected.** Data can be collected at many different times during a randomized controlled trial. For example, data might be collected at a time after the intervention has occurred (i.e., post-only), before and after the intervention occurs (i.e., pre-post), or at multiple time points before and after the intervention has taken place.

[9] Trochim's (2020) online Research Methods Knowledge Base is a useful resource when planning evaluations. It can be accessed at https://conjointly.com/kb/

- **The number of "factors" that vary (factorial design).** In the basic design, only one factor is intentionally "varied" between the intervention and the control group, namely the administration of the intervention itself. However, you could also vary another "factor" of interest such as the administration of a pre-test. For example, let's assume two groups receive the intervention and two do not. Further, for one intervention and one control group, data will be collected through a pre- and a post-test, rather than only a post-test. This allows you to see whether administering a pre-test is related to a change in the outcome of interest (i.e., Solomon Four Group design).

Shadish and colleagues (2002) discuss, in detail, additional variations on the randomized/true experiment and the pros and cons associated with each in *Experimental and Quasi-experimental Designs for Generalized Causal Inference.*

Quasi-experiments

Quasi-experimental Designs are characterized by the use of one or both of the following: (1) the collection of the same data at multiple points in time or (2) the use of a comparison group. Quasi-experimental designs differ from the experimental design in that they do not include random assignment of participants to conditions. Many designs fall under this heading, including pre-post tests without a comparison group, a nonequivalent comparison group design with a pre-post test or post-test only, interrupted time series, and regression discontinuity.

Similar to the randomized experiment, many variations on the basic quasi-experimental design are possible. For example, the interrupted time-series design includes collection of the same data at many time points for a single group prior to and after the intervention. However, your Evaluation Planning Team may decide it is appropriate to collect these same data on a second group that does not receive the intervention, perhaps a group similar to the first on many factors that have the potential to influence change in the outcome of interest (e.g., age, socioeconomic status, grade level). How you choose to select this comparison group or match on selected factors can also vary; for example, you can decide to perform one-to-one matching using demographic data at the participant level or match on broader factors at the group level (e.g., zip code to zip code, city to city). Shadish and colleagues (2002) discuss other ways an interrupted time-series design can be structured, including measuring additional outcomes; introducing, removing, and reintroducing the intervention to the same group over time; and introducing and removing a treatment to two similar groups at different time points (i.e., switching replications).

For further detail about quasi-experimental design options and their variants, see *Experimental and Quasi-experimental Designs for Generalized Causal Inference* by Shadish and colleagues (2002).

Non-experiments

Similar to designs under the quasi-experimental heading, non-experimental designs do not involve random assignment. Referred to as "observational" or "descriptive," the designs in this category include (1) data collection at a single time point (i.e., one-shot designs) or (2) collection of data over time, although the same indicator is not collected over time as would be the case with an interrupted time-series or pre-post quasi-experiment (i.e., repeated measure). Many evaluation designs are considered "non-experimental" because they do not comfortably fit under the previous two definitions.

Some examples of non-experimental evaluation designs include post-test only, cross-sectional, retrospective pre-tests, case studies (single or multiple), ethnography, and phenomenology. Unlike "randomized/true experiment" or "quasi-experiment" type designs, the designs under the "non-experimental" heading are most frequently considered in evaluations that do not attempt to answer questions of a causal nature. While this is not always the case, it is how they have been viewed historically within the evaluation field. Some evaluation scholars assert that carefully constructed non-experimental designs (e.g., case study) can indeed provide valuable information to answer causal questions (see Campbell, 1978; Yin, 2018).

Whether used to answer causal or non-causal evaluation questions, non-experimental designs are among the most common designs and offer a wide variety of options. Further details on some of the designs in this category are provided in Table E.1.

Table E.1 Summary of Evaluation Design Features

	Randomized/True Experiment	Quasi-Experiment	Non-Experimental
Random assignment	Yes	No	No
Comparison/control group	Yes	Possibly	No
Repeated measures	Possibly	Possibly	No

Additional Options

Some designs do not necessarily fit neatly under any of the three categories mentioned above yet build upon their foundations. We include *economic evaluations* and *mixed-method evaluations* under this heading.

Economic Evaluation

Because many evaluation questions center on the topic of cost, economic evaluation is often a useful approach.[10] The following describe two characteristics of true, full-scale economic evaluations (Drummond et al., 2005):

- There is an alternative examined to the intervention that is the primary focus of the evaluation. In other words, you are examining a choice—is Option A (e.g., intervention of interest) better than Option B (e.g., the status quo)?
- The comparison that is made between the intervention of primary interest and the potential alternatives considers both the costs and the *consequences* of the options. When it comes to evaluating interventions, consequences are typically considered to be the outcomes (short, intermediate, or long-term) believed to be associated with the intervention.

Full-scale economic evaluations include cost-effectiveness analysis, cost-benefit analysis, and cost-utility analysis.

According to Drummond and colleagues (2005), partial economic evaluations solely consider cost, without attending to outcomes. For instance, an evaluator, together with stakeholders, may decide that they wish to only explore the costs associated with one specific program or intervention. However, they may decide that the evaluation will examine the costs for more than one alternative intervention, referred to as a cost analysis. They could also choose to examine both costs and outcomes, but only for a single intervention (i.e., no comparison made to an alternative intervention), this design is known as a cost-outcome description. Further information on economic evaluations can be found in **Appendix I.**

[10] Trochim's (2020) online Research Methods Knowledge Base is a useful resource when planning evaluations. It can be accessed at https://conjointly.com/kb/

Mixed Methods

Earlier we mentioned that evaluation designs have been borrowed from a number of different disciplines. Mixed-method evaluations blend various designs and data collection strategies. In these evaluations, "the investigator collects and analyzes data, integrates the findings, and draws inferences using both qualitative and quantitative approaches or methods in a single study or program of inquiry. A key concept in this definition is integration …" (Tashakkori & Creswell, 2007, p. 4).

Mixed-method approaches to evaluation are still fairly new and so different authors' descriptions of the various approaches vary widely. Some focus on why and how you would use a mixture of qualitative and quantitative data within an evaluation (Greene, 2007), while others present options such as mixing at the level of the *design* itself (Creswell, 2009; Tashakkori & Teddlie, 2010). An example of this mixing would be nesting a case study within a randomized controlled trial to better understand whether a given medical treatment improved particular health outcomes, as well as to understand the treatment experience from the patients' perspectives (Creswell, 2009, p. 215). In the brief overview of mixed-method designs below, we discuss the approaches that are most likely to apply to health program evaluations, those at the data collection level.

One important consideration in selecting a mixed-method evaluation is articulating why a mixture of methods would make sense for a given evaluation. Although it is important to consider the rationale when selecting any evaluation design, it is particularly important when choosing whether to use a mixed-method design. Mixing methods may require more time and effort than other approaches, so you need to think through whether the approach is appropriate to address the evaluation questions. Additionally, evaluators often collect different types of data (e.g., qualitative and quantitative) without stepping back to consider how these data will be used *together*. Understanding the potential purposes behind integrating these different types of data in a specific evaluation may bring to light ways to strengthen an evaluation. Greene (2007) identifies the following considerations: triangulation, complementarity, development, initiation, and expansion. We describe each these, leveraging the explanations provided in her text.

Triangulation. Triangulating data—collecting data about a construct in multiple ways—can increase the validity associated with measuring a specific construct (e.g., attitude toward asthma self-management). Many of the constructs we collect data on for health programs (e.g., attitudes, beliefs) can only be measured indirectly. As a result, each measure has some sort of associated error—in other words, we do not get a perfect picture of the concept we are measuring. How far off the mark we are can be considered a form of error. As a result, collecting data about this construct using multiple methods or from multiple sources can be helpful, such as using both closed-ended questions on a survey and open-ended questions in an interview. Since these data collection methods likely have erred in different ways, combining information from both sources can give us a more complete or accurate measure of the construct. When examining data for the purpose of triangulation, evaluators are often looking for how the findings converge

Complementarity. The purpose of mixed-method evaluations that have a "complementarity" design is to "elaborate, enhance, deepen, and broaden the overall interpretations and inferences from the study" (Greene, 2007, p. 101). As a result, different perspectives are sought on a problem of interest, like in an evaluation examining healthier lunchtime food options in a school cafeteria (Greene, 2007). In this example, the evaluator chooses to collect data by observing the food choices that students make in the cafeteria and then decides to gather additional data from students about these food choices through interviews. The interviews reveal the extent to which peers influence what a student chooses to eat in the cafeteria. Here we see that the topic of interest for both data collection efforts is the choice made in selecting from the available lunch options; however, different aspects of this topic are examined (i.e., the choice and a potential influence).

Development. When development is the purpose for a mixed-method evaluation, one data collection

method informs another. For example, imagine your program conducts an intervention designed to improve physician communication with patients. To answer the evaluation question, "To what extent do physicians in this intervention demonstrate improvements in patient communications," your evaluation team collects survey data from a random sample of patients who visit the participating physicians before and after the intervention. Ten percent of patients who demonstrate the largest and smallest change in quality ratings between the pre- and post-measures are then selected for telephone interviews to learn more about the interactions with their physicians. In this case, the findings from the pre-post surveys provide the sample for follow-up interviews.

Initiation. When mixing methods for the purpose of initiation, we are looking for differences that emerge with respect to a problem of interest. Once differences are uncovered, further exploration is often warranted to understand why these differences exist. For example, an asthma program partner may implement a series of trainings for school nurses to enhance their ability to work with students on their asthma self-management behaviors. The training evaluation includes the collection of data from a subset of attendees through semi-structured interviews and a self-assessment exit survey that is completed by the instructors at the end of the course. Evaluation findings indicate that the attendees' comments about the courses are much more favorable than the instructors.' Furthermore, these differences do not appear to occur specifically within a given training site. Such a finding creates a paradox of sorts—why do such extreme differences exist between the instructors and students? This might encourage the evaluator to dig deeper to better understand the discrepancy.

Expansion. When using a combination of methods for the purpose of expansion, an evaluation team often strives to answer questions about various aspects of the program being evaluated to get an expanded understanding of it. A classic example is the use of quantitative methods to explore the extent to which program outcomes occurred paired with qualitative methods to better understand the process of implementing the program (Greene, 2007, p. 103). In this case, the evaluation team is essentially answering different evaluation questions about a specific program through the use of different methodologies. In the example of a hypothetical healthier eating program for children: "… the evaluator could assess student knowledge gains with a standardized pre-post test of nutrition knowledge, possible changes in lunchroom norms via a modest ethnographic inquiry component, and parental awareness of the program through a random selection of families for phone interviews" (Greene, 2007, p. 104). In this example, we can see that the evaluator is seeking a rich understanding of the program itself and seeks this understanding by collecting data about multiple topics that concern this program (i.e., nutrition knowledge, norms, and parental awareness), rather than focusing on one specific topic (e.g., food choices made by students).

Beyond purpose, other considerations regarding mixed-method designs include the weight given to each method used in a mixed-method evaluation, the timing or sequencing of the various data collection methods, as well as how the evaluator chooses to connect the various methods throughout the course of the evaluation (Greene, 2007).

When to Use Which Design

With all the designs, and their potential variants, just covered, you might be wondering how to go about choosing the most appropriate design(s). Unfortunately, there is no cookbook to help us decide which design options to use for a given evaluation. However, four general principles can be very helpful in trying to make your decisions:

Although it is often tempting to select a design based solely on familiarity or feasibility, it is important to first consider the evaluation questions your Evaluation Planning Team is trying to answer.

1. Always begin this decision-making process with the evaluation questions
2. Consider which designs and methods are best suited for the context

3. Determine which designs and methods will allow for authentic input from participants and produce evidence participants deem credible
4. Refer to the evaluation standards for guidance

The evaluation questions often suggest a specific design option. However, there are instances when multiple designs are plausible. When you find yourself in this situation, referring to the standards for program evaluation—utility, feasibility, propriety, accuracy, and evaluation accountability—may help you sort through your options.

Consider that the Strategic Evaluation Planning Team in your jurisdiction has decided that a high priority is evaluating an intervention designed to increase asthma self-management knowledge and skills among people who have been diagnosed with asthma at a local federally qualified health center (FQHC). A question of interest to the Evaluation Planning Team is, "To what extent did our asthma management training lead to improvements in asthma self-management knowledge among those who completed the training?"

Since this question asks about causation, a number of potential design options are available, including a randomized/true experiment; various quasi-experimental approaches; mixed methods; and, some might even argue, a case study design. Each of these designs has strengths and limitations related to the standards for program evaluation depending upon the context in which the intervention is being conducted. Consider the following examples:

- The primary end users of this evaluation are very familiar with randomized/true experiments and view the results of such evaluations to be highly credible. Where randomization is not possible, these stakeholders may acknowledge that an evaluation design using comparison groups would still be useful for their purposes. Yet they would be uncomfortable making decisions in the absence of some sort of reasonable comparison. This is an example of examining the design with respect to the utility standard.
- Mixed-method and case study designs would include the collection of data through multiple avenues. This would likely make these designs more time intensive than the other options being considered. Stakeholders debate whether the expanded scope of these types of evaluation is worth the greater investment of time and resources. This is an example of examining design with respect to the feasibility standard.
- Some of the evaluation stakeholders may raise concerns about randomly assigning individuals to a control group that receives the standard treatment when there is a convincing argument that the intervention leads to improved self-management knowledge and skills that may translate into improvements in health outcomes. These stakeholders might argue that all patients who receive services from the FQHC should receive the intervention. An option that would address this concern is providing the control group with the intervention at a later time. The Evaluation Planning Team decides that other design options are preferable because they allow for individuals to determine on their own whether they should enroll in the intervention. This is an example of examining design with respect to the propriety standard.
- In our last example, the Evaluation Planning Team would like to know whether an intervention is causing a specific outcome. This requires an assessment of the internal validity associated with each design option. Internal validity refers to the certainty with which we can state that an action (e.g., intervention) results in a change in a specific outcome (e.g., knowledge gain). There are many known threats to internal validity,[11] some of which are better dealt with by using specific evaluation designs (Trochim, 2020). For example, one of the reasons your stakeholder group may find the randomized/true experiment to be more credible

[11] A thorough explanation of numerous threats to internal validity is provided by Trochim (2020) through the online Research Methods Knowledge Base (http://www.socialresearchmethods.net/kb/intval.php).

and useful for their purposes is that, when implemented well, the design itself combats threats to internal validity. Yet discussions may reveal that certain design options also produce levels of internal validity sufficient for the needs of primary stakeholders. This is an example of examining design with respect to the accuracy standard.

As is illustrated by this hypothetical exploration of evaluation design options, numerous considerations go into selecting an evaluation design. The decision requires carefully balancing multiple ideas, perspectives, and criteria. We encourage you to be creative and flexible in selecting the design that is most appropriate to the information and other important needs you identify while planning for the evaluation.

Some Helpful Resources

There are many good resources that discuss the various design options briefly explained in this appendix. Below are some references you may find useful as you continue to plan and implement various evaluation designs.

Experimental and Quasi-Experimental Designs

Shadish, W. R, Cook T. D, & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference.* Houghton Mifflin Co.

Economic Evaluation

Drummond, M. F., Sculpher, M. J., Torrance, G. W., O'Brien, B. J., & Stoddart, G. L. (2005). *Methods for the economic evaluation of health care programmes.* (3rd ed.). Oxford University Press.

Haddix, A. C., Teutsch, S. M., & Corso, P. S. (2005) *Prevention Effectiveness: A guide to decision analysis and economic evaluation* (2nd ed.). Oxford University Press.

Gold, M. R., Siegel, J. E., Russell, L. B., & Weinstein, M. C. (1996). Cost-effectiveness in health and medicine. Oxford University Press.

Mixed Methods

Creswell, J. W., & Plano Clark, V. L. (2018). *Designing and conducting mixed methods research* (3rd ed.). Sage Publications.

Greene, J. C. (2007). *Mixed methods in social inquiry*. Jossey-Bass.

Tashakkori, A., & Teddlie, C. (Eds). (2010). *Handbook of mixed methods in social and behavioral research*. Sage Publications.

Teddlie, C., & Tashakkori, A. (2009). *Foundations of mixed methods research: Integrating quantitative and qualitative approaches in the social and behavioral sciences.* SAGE Publications.

Other Aspects of Design

Creswell, J. W. (2006). *Qualitative inquiry and research design: Choosing among five approaches.* Sage Publications, Inc.

Creswell, J. W. (2009). *Research design: Qualitative, quantitative, and mixed methods approaches.* Sage Publications, Inc.

Patton, M. Q. (2002). *Qualitative research and evaluation methods.* Sage Publications, Inc.

APPENDIX F. Training and Supervising Data Collectors

In order to collect data that meet the standards of utility, feasibility, accuracy, propriety, and evaluation accountability data collectors be trained and supervised. Even if the evaluation plan calls for using existing data, or data that would be collected as part of the intervention or other program activities, it is good to review plans together so that data collectors and supervisors share the same understanding of the purpose of data collection, the data collection procedures, the division of labor, and any special data collection circumstances. It will also promote consistency in data collection procedures across data collectors, thereby increasing the reliability of data. Training should be required whether data collection is being done by a program's own staff members, by partner staff members, or by contractors or consultants.

Identifying Who Needs to Be Trained

Thinking systematically about who should receive training is a critical first step. **Table F.1** provides an example of an asthma education training intervention that not only lists the people who may be directly collecting data for the evaluation, but also those who supervise data collection or whose participation is necessary to gain access to the data, in this case, those who would be referring participants to the intervention. The training needs of each of these groups may not be the same so the evaluator should systematically think through the roles and training needs of each group.

Table F.1. Data Collector Involvement and Training Needs for an Asthma Education Training Intervention

Data Collector/ Stakeholder/ Other	Data Collection Type	Role in Data Collection	Training Needs
Asthma Educators	Pre- and post-intervention survey of asthma education program participants and attendance logs	• Maintain attendance log of all asthma education participants • Administer data collection questionnaire • Collect questionnaire • Keep questionnaires secure until collected by evaluation lead	• Data collection procedures • Attendance log procedures • Data collection logistics • Informed consent • Data handling and confidentiality
Evaluation Lead	Pre- and post-intervention survey of asthma education program participants and attendance logs	• Monitor randomly selected education sessions to assess consistency and quality of delivery • Collect questionnaires and attendance logs from asthma educators	• Data monitoring procedures Data handling and confidentiality
Clinic Staff Members	Pre- and post-intervention survey of asthma education program participants	• Provide referrals to asthma education sessions	• Understand recruitment procedures • Recruitment logistics to reduce burden

Selecting Your Training Method

Training can take many forms from informal to formal and from simple to complex. The choice of methods will depend on the audience(s), the training needs identified, available training resources, and the trainer's personal style. Some training methods to consider include

- **Written instructions:** In some cases, simple instructions on a data collection form may be sufficient.

- **Verbal instructions:** For simple data collection activities, verbal instructions may be sufficient (e.g., place completed forms in the box at the door before you leave); however, we suggest pairing these with written instructions whenever possible.
- **Meetings:** It may be necessary to hold meetings with partners, stakeholders, or decision makers to ensure access to the data needed for the evaluation.
- **Memoranda of Understanding (MOU) or data-sharing agreements:** Depending on institutional needs, it may be necessary to set out formal agreements for how data can be accessed. In such agreements, it is important to work out who will have access to data, under what circumstances, and when it will be available. It is also important to agree on the formats in which data will be made available and to be aware of any restrictions on the use of data. The contents of any agreements should be incorporated into training activities.
- **Train-the-trainer:** In some cases, data may be collected by people who are also conducting an intervention (e.g., teachers conducting training with youth). In this situation the evaluation data collection training could be embedded into the larger training on the intervention itself.
- **Formal data collector training:** For more complex data collection activities specific to the evaluation, and in cases where multiple data collectors are involved, we recommend that you hold formal data collector training. If your situation calls for a more formal data collector training, using a variety of adult learning strategies and techniques will help convey the important concepts. These can include both instructional approaches (e.g., didactic approaches, case examples or narratives, brainstorming) to convey knowledge and hands-on approaches (e.g., modeling, role-playing, small group and peer support, practice sessions, or "on-the-ground" training) to teach skills.

Regardless of the approach, engaging participants in active and interactive learning by asking and answering questions, being enthusiastic, and providing immediate positive and constructive feedback is important (e.g., "I liked how you did X. Next time, I'd like to see you do Y as well."). If the evaluation design involves conducting data collection at different points in time, a training may be needed before each data collection period. If the same data collectors will be used during each time period, the training can serve more as a review of concepts.

Defining Your Training Topics

Although your training will be customized to meet the needs of your evaluation, most training sessions will include the following

- Background material about the data being collected that clarifies the type of data being collected, from whom, and for what purpose
- Instructions for data collection and data management, including roles and responsibilities
- Other topics, as needed, such as staff member safety, team building, and special considerations in working with the intended audience

Background Material

Providing information about the purpose of the evaluation and how the data will be used will make data collectors feel more confident; motivate them to obtain high-quality data; help them make better decisions regarding the data collection; help them troubleshoot, answer respondent questions, and respond to unusual situations; and contribute to a more professional attitude. A broader understanding of the evaluation will help data collectors appreciate how the evaluation standards informed the evaluation design and their role in maintaining those standards during implementation. Background material should include basic information about what kind of data will be collected, from whom, and for what purpose. It should also include information about who is

sponsoring the evaluation and who will use the data to generate evaluation findings. For more formal data collector trainings, consider compiling a data collection handbook that includes the protocols, instruments, instructions, contact numbers, and other supplementary materials that were developed for the evaluation. Data collectors can use this handbook as a reference after the training is completed.

Data Collection Instructions

Data collection instructions should cover every aspect of data collection, from identifying or locating appropriate respondents or records to processing the collected data. These instructions should be detailed written instructions. In addition, all data collectors need to know their own specific roles and responsibilities as well as to whom they report and whom they should call with questions. In some cases, data collectors will be working in teams and may need instruction on how to divide the work efficiently. Supervisors also need to be clear about their roles and responsibilities. **Table F.2** provides additional details on training topics related to data collection and management.

Table F.2. Common Data Collection Training Topics

Topic	Description
Data collection logistics	Training of data collectors should cover the logistics of the data collection: what, when, where, how, and from whom. Be sure to stress the importance of adhering to scheduling requirements that impact the quality of the evaluation, such as the timing of pre- and post-test data collection.
Identifying appropriate respondents or records	For some types of evaluation, it is important to obtain data from only those respondents or records that meet the evaluation requirements. If data collectors understand the importance of adhering to the data collection protocol, they will be less likely to substitute respondents or records inappropriately, thus preserving the quality of the data.
Recruiting participants	Data collectors should be given detailed and explicit information about how to recruit participants or gain access to data. For instance, for survey data collection, high response rates are important. Interviewers or those administering questionnaires should be taught how to encourage a respondent to participate, while at the same time protecting respondents' rights to refuse to participate.
Gaining access to data	Field workers who are abstracting records will need to learn what to say in order to gain admittance and request records. Despite having obtained the necessary organizational agreements or required clearances, data collectors may have to deal with gatekeepers or new staff members who may be unaware of these agreements or who may find it burdensome to retrieve records or share offices.
Introducing the study and obtaining consent or access	Data collectors should know how to provide informed consent to participants and how to gather and maintain the data collected according to ethical considerations and professional evaluation standards. Whenever possible, evaluation materials should include written scripts for how an evaluation should be introduced to participants or stakeholders as well as procedures for obtaining consent to participate in the evaluation.
Collecting unbiased data	Data need to be collected in a consistent and unbiased fashion in order to allow meaningful comparison and interpretation. Ensuring this type of consistency and neutrality in data collection should be a key consideration in training. For complex data collection instruments, it is good practice to develop a "Question-by-Question" (QxQ) manual that provides information about the intent of each question or item (e.g., "when we ask about asthma medications, we mean only prescription medication and not over-the-counter or herbal remedies"). If structured interviews are planned, interviewers should be trained to read the questions as written and in the specified order, use a neutral tone of voice, and avoid interjecting comments or opinions. Focus group moderators need to make sure they do not ask leading questions and that they adequately guide the discussion to keep one person from dominating. For records abstraction, training should focus on which records are to be reviewed and precisely what information from the records is to be obtained.

Recording responses	Accurate recording of data is critical. Data collectors should have opportunities to practice recording and reporting data as part of the training. Encourage data collectors to make notes about any ambiguous responses. This will help data analysts better interpret the data later. You may want to measure the degree to which different data collectors record or code the same data in the same way. For more information on intercoder reliability, see https://conjointly.com/kb/types-of-reliability/
Knowing when to terminate an interview	Sometimes interviewers should terminate or reschedule the interview. For example, if the respondent cannot focus or is experiencing difficulty comprehending or communicating, perhaps due to being emotionally upset, tired, or some other reason, then it is better to terminate or reschedule.
Data handling and security	Data collection procedures and training should address what to do with data once they are collected, how to protect the confidentiality and security of the data, who is allowed access the data, and what to do if any breach in security or confidentiality occurs. Data collectors need to learn these procedures and why data confidentiality and security are important.
Data collection supervision and monitoring	Regardless of who is collecting the data, there needs to be a plan in place for supervision and monitoring to help ensure that data are being collected appropriately and that any issues can be resolved as they arise. Depending on the complexity of the data collection activity, supervisory responsibilities might be limited to training and quality checks, but might also include a range of additional roles such as hiring data collectors, validating samples, supervising data entry, monitoring data collection, and coordinating with data analysts.
Routine methods for gathering feedback from data collectors	Most importantly, ensure that you have a method for routinely gathering feedback from data collectors about any problems they have encountered or about field observations they have that may necessitate reviewing data collection procedures or instruments. Further, devise means to share lessons learned among all data collectors and their supervisors while data collection is in progress. Keeping communication channels open, identifying emerging issues as soon as they arise, sharing critical information among all data collectors, and working together with them to develop effective solutions are among the best ways to safeguard the accuracy, propriety, and utility of any data collected.

Tips for Successful Data Collection Training

In this section, we offer a few tips to keep in mind when developing data collection procedures and the training approach.

- **Always conduct some type of data collection training.** Data collection training (either formal or informal) is needed for all data collection activities in your evaluation. You cannot assume that procedures will be intuitive or obvious to those conducting the data collection. Even with simple data collection procedures, it is better to be explicit to avoid later misunderstandings that can result in data that are not useful.
- **Experienced data collectors need training, too.** Each data collection effort is different, and even experienced data collectors will benefit from the opportunity to think through the specific procedures for this evaluation and having the time to practice.
- **Use trainers of high-quality.** In multi-person data collection teams, Bamberger, Rugh, and Mabry (2006) recommend that, when resources are scarce, you should recruit the best supervisors and trainers possible, even if this means recruiting less experienced data collectors. They point out that poor supervision or training can impede performance of even good data collectors, whereas good supervision and training can improve performance of both poor and good data collectors.
- **Ensure respondent comfort.** Respondents must feel comfortable with data collectors. In some cases, this may mean selecting data collectors of similar racial, ethnic, linguistic, or geographic background to respondents.

- **Build data collection training into your evaluation schedule.** Avoid underestimating the time it may take to be ready for data collection.
- **Think broadly about training needs**. Even if using a secondary data source, think about the procedures needed to access the data and abstract the needed elements. Ensure these procedures are explicit and well-documented.
- **Emphasize to data collectors the importance of reporting problems and observations as they arise.** Data collectors are the members of the evaluation team closest to the evaluation implementation. Their observations can be invaluable.
- **Ensure appropriate documentation.** Even if a formal training is not needed, it is still important to think through all aspects of data collection activities and have procedures in place to deal with anticipated, as well as unanticipated, issues. Being thorough and preparing written instructions help to ensure that the data collection approach is well documented and that others can step in to take over, should it become necessary. The documentation also becomes a historical record of how the evaluation was conducted in case others wish to review the methods or undertake something similar.
- **Monitor the data collection**. Ongoing monitoring will indicate whether data collection is proceeding as planned and will allow the evaluator to intervene or provide additional training or guidance as needed. Situations that may indicate a need for additional training include changes in the protocol, unplanned deviations from the protocol, implementation problems, or complaints about the performance of data collectors.

APPENDIX G. Challenges in Evaluation

As mentioned in **Chapter 5**, proper planning is essential to implementing a successful evaluation. Part of good planning is anticipating challenges that may arise and initiating strategies to avoid or minimize their impact on the evaluation. Additionally, evaluator's most adapt to unforeseen issues that emerge during implementation. **Tables G.1 – G.5** describe common challenges experienced during an evaluation and practical strategies for addressing them. Managing challenges ensures the evaluation is more likely to uphold the evaluation standards, resulting in a better-quality evaluation.

Table G.1 describes common challenges that may occur while gathering information about the evaluation context during the planning and implementation phases as well as possible solutions.

Table G.1 Meeting Challenges in Evaluation Context

Evaluation Challenge	Possible Action	Phase		Relevant Standard(s)
		Planning	Implementation	
Negative community response to evaluation	Learn about past experiences with evaluation and discuss how past efforts could be improved upon.	X		Utility Propriety Evaluation Accountability
	Revisit context assessment to gain insight into resistance.	X		
	Discuss information needs and evaluation plans with stakeholders to ensure the evaluation is not merely extractive but also responds directly to community information needs.	X		
	Include stakeholders from the community, including program participants or intended beneficiaries on the Evaluation Planning Team.	X		
	Meet with stakeholders to better understand the source of their negative response and consider how to address it.	X		
	Listen to stakeholder concerns and consider using evaluation approaches that elevate stakeholder perspectives.	X	X	
	Discuss evaluation findings with stakeholders and explore implications for the program and community.		X	
	Include stakeholders in developing an action plan.		X	
	Use an issues management approach to public relations if findings may reflect negatively on the community.		X	

Limited will to support evaluation	Listen and learn about information needs with stakeholders and incorporate into the evaluation.	X		Utility Feasibility Evaluation Accountability
	Discuss strategies to increase awareness of the importance of evaluation, as well as increase support for the evaluation.	X		
	Conduct early ECB activities to reframe evaluation from a punitive activity to one about improvement.	X	X	
	Discuss evaluation successes, findings, and implications for the program with stakeholders and organizational leadership post-evaluation.		X	
	Consistently send messages about the importance of evaluation.		X	
Lack of support from program leadership	At the start of strategic evaluation planning include frontline program leadership in stakeholder discussions about the evaluation.	X		Utility Feasibility
	Keep leaders informed about the evaluation with progress reports and solicit their input without overburdening them.		X	
	Consider ways to share findings that are most useful to busy leaders and in alignment with organizational culture.		X	
	Should leadership change, inform new leaders about the evaluation and its progress and solicit their input.		X	
	Include leaders in briefings on evaluation results and implications for program improvement.		X	

There are many logistical aspects to planning and implementing an evaluation ranging from clarifying the roles and responsibilities of the evaluation team to ensuring adequate resources for carrying out the evaluation. **Table G.2** describes potential strategies for managing these situations.

Table G.2 Meeting Challenges in Evaluation Logistics

Evaluation Challenge	Possible Action	Phase		Relevant Standard(s)
		Planning	Implementation	
Confusion among team members about roles and responsibilities	Clearly document team member roles and responsibilities in individual evaluation plan or management plan and discuss with the team as part of onboarding.	X	X	Feasibility Accuracy Evaluation Accountability
	Hold regular meetings with the Evaluation Implementation Team to discuss progress and address emerging issues.		X	
Insufficient financial resources to complete the evaluation	Have resource estimates developed by individuals experienced in evaluation.	X		Feasibility
	Consider efficiencies across evaluations during strategic evaluation planning.	X		
	Engage partners in understanding potential additional funding sources.	X	X	
	Consider delays in evaluation schedule to accommodate funding cycle.	X		
	Allow for some wiggle room in the budget in case surprises occur.	X		
	Regularly monitor evaluation budget during implementation and compare against expectations (i.e., the burn rate).		X	
	Consider reduction in scope and other cost-saving measures if budget monitoring indicates a need to economize.		X	
	Document effectiveness of cost-saving measures.		X	
	Keep track of resources spent to generate more realistic estimates in future evaluations.		X	
Inadequate staff resources to complete evaluation	Plan around staff skills and availability, including looking across evaluations during strategic evaluation planning.	X	X	Feasibility
	Consider alternatives if staffing falls short of requirements, such as contracting externally, training existing staff in needed skills, borrowing partner staff members, involving interns from local colleges and universities.	X	X	
	Implement previously developed contingency plans to deal with staff shortages.		X	

Data collection is not only an important part of evaluation but can also require adapting to situations that may arise during the process. **Table G.3** highlights common challenges during the data collection phase and potential remedies for managing these pitfalls.

Table G.3 Meeting Challenges in Data Collection

Evaluation Challenge	Possible Action	Phase		Relevant Standard(s)
		Planning	Implementation	
Lack of access to data	Identify potential data sources and determine the availability and accessibility of any existing data required for the evaluation.	X		Feasibility
	Develop memoranda of understanding and data-sharing agreements for access to required data prior to launch.	X		
	Discuss with Evaluation Implementation Team and stakeholders how to work around failures or divisions in data-sharing agreements.		X	
	If necessary, revise evaluation scope to accommodate lack of access and tap alternative data sources.		X	
Difficulties recruiting participants	Include and listen to stakeholders and individuals experienced in evaluation in planning to maximize respondent participation in the evaluation.	X		Feasibility Propriety Accuracy Evaluation Accountability
	In designing instruments and recruitment materials, consider respondent burden and the costs and benefits to respondents in participating.	X		
	Budget permitting, consider offering incentives to participants.	X		
	Identify and minimize barriers to recruitment or participation in the evaluation (e.g., reduce length of instrument, change data collection strategies to be more appealing or less burdensome).	X		
	Train data collectors in effective recruitment techniques.		X	
	Solicit support from community members in identifying and gaining the cooperation from eligible respondents.	X	X	
	Solicit support from community members and explain to them the importance of the program and the importance of evaluation for improving the program.	X	X	

Difficulties working with contractors	Plan for which evaluation tasks will need to be contracted out and identify funds available for this work.	X		Feasibility Accuracy Evaluation Accountability
	Identify contractors whose evaluation approach and style align with your expectations and preferences.	X		
	Develop detailed agreements that clearly outline contractors' roles, responsibilities, products, timeline, and budget; include requirements and funds for regular meetings and progress reports.	X		
	Train contract staff members to be involved in data collection on the written data collection protocol.		X	
	Monitor contractor timeline, budget, and performance through regular meetings and written progress reports.		X	
	Have a back-up list of contractors to call in the event services cannot be rendered.		X	

Planning for and conducting data analysis can often be fraught with unexpected situations that the evaluation team will need to manage. **Table G.4** presents some common challenges and strategies for resolving them.

Table G.4 Meeting Challenges in Data Analysis

Evaluation Challenge	Possible Action	Phase		Relevant Standard(s)
		Planning	Implementation	
Data collected are not useful	Listen to and discuss with stakeholders their information needs and priorities and incorporate these into the individual evaluation plan.	X		Utility
	Identify what type of data potential end users view as credible evidence (e.g., qualitative, quantitative, mixed).	X		
	Identify a study design that will provide credible evidence for end users (e.g., pre-post, pre-post with control).	X		
	Specify how each data element collected, and data analyses, will help answer each evaluation question.	X		

	Draft table shells to show how data will be displayed.	X		
	Pilot test instruments and revise as needed.	X		
	If feasible, revise data collection strategies or instruments, or clarify instructions, to enhance data quality.		X	
	Conduct preliminary analysis of pilot data to check for usefulness of data and feasibility of analysis plan; revise data collection and analysis plans as necessary (including possible revisions to the sampling plan).	X	X	
Uncertainty about how to analyze the data	Include stakeholders and individuals experienced in evaluation and in data analysis in planning for data analysis and learn from their advice.	X		Feasibility Accuracy Propriety
	Cross check analysis plans with evaluation data collection instruments to ensure data are collected in an appropriate manner for intended analyses.	X		
	Document data analysis approach in individual evaluation plan.	X		
	Ensure availability of individuals (either staff members or contractors) with the requisite skills and experience to implement the analysis plan.	X		
	Consult with analysts on the staff or in partner organizations to learn from them.	X	X	
	If data cannot or will not be analyzed, consider dropping the data elements from data collection instruments.	X	X	
	Do not report data with small cell sizes that might result in inadvertent disclosure of confidential information (e.g., when small numbers of cases further broken down by demographic factors could lead to identification of individuals). In these situations, it may be advisable to note the reason for not reporting on certain analyses and to find other ways of representing the experiences of these groups.		X	
Stakeholders disagree about interpretation of findings	Employ interpersonal skills to engage stakeholders in a dialogue with each other to deliberate findings and arrive at an agreement.		X	Accuracy Propriety Utility

	Hold a data party or sensemaking session with stakeholders to discuss the findings and interpret results together.		X	
Preliminary findings indicate need for program modifications	Discuss with stakeholders how to handle the situation if preliminary findings suggest need for program modification.	X	X	Utility
	Consider preparing for a mock findings session in which possible results scenarios are presented.	X	X	
	Discuss preliminary findings with stakeholders to decide whether the program should be modified immediately or after evaluation concludes.		X	
	If the program is modified, consider with evaluation team and stakeholders any implications this has for the remainder of the evaluation.		X	

Lastly, ensuring use of the evaluation findings is an important step in the planning and implementation of an evaluation (**Table G.5**). These challenges should not be overlooked as it is important to manage them to ensure evaluation use.

Table G.5 Meeting Challenges in Dissemination of Evaluation Findings

Evaluation Challenge	Possible Action	Phase		Relevant Standard(s)
		Planning	Implementation	
Late timing of evaluation in relation to information needs	Discuss with stakeholders when information is needed.	X		Utility
	Monitor evaluation timeline to ensure it stays on track.		X	
	If appropriate, disseminate interim findings prior to completion of evaluation along with caveats that the information is not final.		X	
Findings not viewed as credible	Discuss with program stakeholders and decision makers any design and data collection preferences they may have and incorporate this into the individual evaluation plan.	X		Utility Accuracy
	Document design and data collection strategies in individual evaluation plan and share with intended end users for feedback.	X		
	Post-evaluation, discuss findings alongside methodology used and related rationale.		X	
	Respond to stakeholder questions about methodology used.		X	
Findings from one evaluation have implications for later evaluations in the strategic evaluation plan	During strategic evaluation planning, be aware of potential relationships and interdependencies between the various evaluations proposed.	X		Utility Feasibility
	Plan check-ins with Strategic Evaluation Planning Team and Evaluation Implementation Teams to discuss implications for future evaluations.	X		
	As part of post-evaluation discussions, address whether any of the evaluation findings affect future planned evaluations (e.g., which evaluations to conduct, how to conduct them, or how much of the resources have been expended).		X	
	If necessary, revise the strategic evaluation plan.		X	
Findings not welcomed by some stakeholders	Listen to and discuss with stakeholders what their information needs are and how this can be addressed in the evaluation plans.	X		Utility Feasibility
	Discuss with stakeholders how to handle a situation where findings do not show the program in a positive light or suggest a need for program modification.	X		
	Consider modes of dissemination that are most accessible to stakeholders.	X		
	Communicate with stakeholders throughout the evaluation to avoid surprises at the end.		X	
	Post-evaluation, discuss evaluation findings with stakeholders and explore implications for program and community, emphasizing positive, constructive action that can be taken.		X	

APPENDIX H. Sample Partnership Evaluation Tools

Table H.1 provides a list of data collection instruments for conducting a partnership evaluation and the concepts that each instrument measures. While the tools may need to be adapted to the context of the evaluation, this list provides a good starting place for selecting an appropriate instrument.

Table H.1 Partnership Evaluation Tools and Corresponding Concepts

Tool Name	Source	Population/ Instructions	Terms in Partnership Concept Map
Annual Satisfaction Survey for Community Coalitions Worksheet 1 https://www.tomwolff.com/resources/backer.pdf pp. 28–33	Fawcett, Foster, & Francisco, 1997.	Coalition members and funding partners	• Synergy, coordination, increased credibility, and access to key populations • Group dynamics • Partnership structure • Identified and garnered resources for future • Increase knowledge and build skills • Perceived benefits and drawbacks • New or strengthened external relationships or networks • Communicate key messages to audiences and stakeholders
Assessing Strategic Partnership: The Partnership Assessment Tool https://www.conservationgateway.org/ConservationPlanning/partnering/cpc/Documents/AssessingStrategicPartnership.pdf	Hardy, Hudson, & Waddington, 2003; Office of the Deputy Prime Minister, Strategic Partnering Taskforce.	Partnerships– Developmental tool to assess the effectiveness of a partnership. Checklist approach used with individual partners and discussed to ascertain areas of consensus or conflict in six Partnership Principles areas	• Implement interventions • Synergy, coordination, increased credibility, and access to key populations • Group dynamics • Partnership structure • Perceived benefits and drawbacks • Contribute resources
Climate Diagnostic Tool: The Six R's of Participation, Worksheet 4 https://www.tomwolff.com/resources/backer.pdf p.50–57	Kaye & Resnick, 1994.	Coalition members	• Group dynamics • Partnership structure • Perceived benefits and drawbacks • Maintain partnerships and build collaborations

Tool Name	Source	Population/ Instructions	Terms in Partnership Concept Map
Coalition Effectiveness Inventory Self-Assessment Tool http://coalitionswork.com/wp-content/uploads/coalition_effectiveness_inventory.pdf	Butterfoss, F., 1998a, 1998b; Center for Pediatric Research, South Carolina DHEC.	Partnership members Coalition members complete rating of coalition Can be repeated pre- and post- intervention	• Level of involvement • Implement interventions • Synergy, coordination, increased credibility, and access to key populations • Membership composition • Group dynamics • Partnership structure • Recruitment • Identified or garnered resources for the future • Perceived benefits and drawbacks • New or strengthened external relationships or networks • Maintain partnerships and build collaborations • Contribute resources • Communicate key messages to audiences and stakeholders
Coalition Self-Assessment Survey II http://www.asthma.umich.edu/media/eval_autogen/CSAS.pdf	Allies Against Asthma (2002).	Coalition members Survey administered annually	• Level of involvement • Implement interventions • Synergy, coordination, increased credibility, and access to key populations • Membership composition • Defined roles and responsibilities • Group dynamics • Partnership structure • Recruitment • Leadership • Shared vision • Increase knowledge and build skills • Perceived benefits and drawbacks • Maintain partnerships and build collaborations
Collaboration Checklist https://atrium.lib.uoguelph.ca/xmlui/bitstream/handle/10214/3117/Borden_Assessing_Your_Collaboration_%20A%20Self_Evaluation_Tool_complete.pdf?sequence=1&isAllowed=y	Borden & Perkins, 1999.	Coalition members read a brief description for each of the areas (core concepts) and then rate how well the collaboration is functioning in each area.	• Group dynamics • Leadership

Tool Name	Source	Population/ Instructions	Terms in Partnership Concept Map
Community Group Member Survey: Using the Results https://cdn.shopify.com/s/files/1/0145/8808/4272/files/G3658-09.pdf	Taylor-Powell, 1998; University of Wisconsin Extension	Community group members Survey, also pro-vides examples of how to report on evaluation results to internal and external stakeholders using survey.	• Maintenance of interest in collaborating or contributing • Level of involvement • Implement interventions • Membership composition • Group dynamics • Partnership structure • Perceived benefits and drawbacks
Diagnosing the Health of Your Coalition https://ctb.ku.edu/en/table-of-contents/assessment/promotion-strategies/maintain-a-coalition/tools	Community Toolbox, n.d.	Coalition members (larger group preferable) Survey. Instrument developers suggest reviewing results and making recommendations for changes and conducting an annual review to assess progress.	• Membership composition • Group dynamics • Partnership structure • Shared vision • Perceived benefits and drawbacks • New or strengthened external relationships or networks • Maintain partnerships and build collaboration • Communicate key messages to audiences and stakeholders
Diagnosing Your Coalition: Risk Factors for Participation, Worksheet 2 https://www.tomwolff.com/resources/backer.pdf pp.34-47	Kaye, 1993.	Coalition members	• Demonstrate commitment to self-assessment • Group dynamics • Partnership structure • Perceived benefits and drawbacks • New or strengthened external relationships or networks • Maintain partnerships and build collaborations • Communicate key messages to audiences and stakeholders
Diagnostic Tool for Evaluating Group Functioning https://www.extension.iastate.edu/Publications/PM1844.pdf	Iowa State University Extension, 2000 (based on Taylor-Powell et al., 1998).	Partnership members. Each member is asked to rate what's happening in the group. Then members should have a time out group discussion about what's happening and what to do about it.	• Defined roles and responsibilities • Group dynamics • Recruitment • Leadership • Shared vision • Communicate key messages to audiences and stakeholders

Tool Name	Source	Population/ Instructions	Terms in Partnership Concept Map
Evaluating Community Coalition Characteristics and Functioning: A summary of measurement tools https://academic.oup.com/her/article/19/5/514/571017	Granner & Sharpe, 2004.	Various coalitions. Review article listing a variety of evaluation tools from various articles.	• Maintenance of interest in collaborating • Level of involvement • Implement interventions • Changes to policy, staffing, or funding within partner organizations • Synergy, coordination, increased credibility, and access to key populations • Membership composition • Group dynamics • Partnership structure • Recruitment • Leadership • Identified and garnered resources for future • Increase knowledge and build skills • Perceived benefits and drawbacks • New or strengthened external relationships or networks • Maintain partnerships and build collaborations • Contribute resources • Prioritize elements of the asthma program plans
Instrument for evaluating dimensions of group dynamics within community-based participatory research partnerships	Schulz, Israel & Lantz, 2003.	Partnership members. Compilation from three questionnaires for evaluating group dynamics characteristics and intermediate measures of partnership effectiveness.	• Implement interventions • Synergy, coordination, increased credibility, and access to key populations • Membership composition • Group dynamics • Partnership structure • Leadership • Increase knowledge and build skills • Perceived benefits and drawbacks • New or strengthened external relationships or networks
Inclusivity Checklist, Worksheet 6 https://www.tomwolff.com/resources/backer.pdf p. 63	Rosenthal, 1997.	Coalition members. Coalition members check which of 11 items describe their coalition. Unchecked items indicate areas for improvement.	• Membership composition • Group dynamics

Tool Name	Source	Population/ Instructions	Terms in Partnership Concept Map
Key Informant Interviews www.asthma.umich.edu/media/eval_autogen/key_informant.pdf	Allies Against Asthma, 2003.	Partnership members.	• Synergy, coordination, increased credibility, and access to key populations • Identified or garnered resources for the future • Perceived benefits and drawbacks • Maintain partnerships and build collaboration
Partnership Self-Assessment Tool https://atrium.lib.uoguelph. ca/xmlui/bitstream/handle/10214/3129/Partnership_Self-Assessment_ Tool-Questionnaire_complete. pdf?sequence=1&isAllowed=y	Center for the Advancement of Collaborative Strategies in Health, 2002.	Partnership members of coalitions with the following characteristics: • In existence at least 6 months • Group of people and organizations that continually work together • Have begun to implement plans • Have at least 5 active partners Members fill out a questionnaire. The website provides detailed instructions on how to score, summarize, and report findings.	• Implement interventions • Synergy, coordination, increased credibility, and access to key populations • Group dynamics • Partnership structure • Leadership • Identified or garnered resources for the future • Increase knowledge and build skills • Perceived benefits and drawbacks • New or strengthened external relationships or networks • Contribute resources • Communicate key messages to audiences and stakeholders • Identify potential funding or resources
Sustainability Benchmarks, Worksheet 8 https://www.tomwolff.com/re-sources/backer.pdf pp. 66–72	Center for Collaborative Planning, 2000.	Coalition members.	• Changes policy, staffing, or funding within partner organizations • Synergy, coordination, increased credibility, and access to key populations • Identified or garnered resources for the future • Increase knowledge and build skills • New or strengthened external relationships or networks • Communicate key messages to audiences and stakeholders • Identify potential funding or resources

Tool Name	Source	Population/ Instructions	Terms in Partnership Concept Map
Wilder Collaboration Factors Inventory https://www.wilder.org/wilder-re-search/resources-and-tools	Mattessich & Johnson, 2018.	Partnership members. Members fill out a 44-item questionnaire measuring 22 evidence-based factors for successful collaboration. A paper questionnaire is available, or groups can register to use the tool online to see item averages and open-ended responses.	• Membership composition • Demonstrated commitment to self-assessment • Defined roles and responsibilities • Group dynamics • Leadership • Maintain partnerships and build collaboration • Communicate key messages to audiences and stakeholders • Shared vision • Perceived benefits and drawbacks

Tool Name	Source	Population/ Instructions	Terms in Partnership Concept Map
Collective Impact, Equity, Systems Change, and Population Change Rubrics Appendix A: https://www.orsimpact.com/DirectoryAttachments/10262018_111513_477_CI_Study_Report_10-26-2018.pdf p. A-10 – A-13	Sparks Policy Institute & ORS Impact, 2018.	Coalition member, site, or initiative. Each site is categorized as mature, emerging, or absent for each set of indicators.	• Collective impact rubric: • Demonstrated commitment to self-assessment • Defined roles and responsibilities • Structure • Leadership • Communicate key messages to audiences and stakeholders • Shared vision • Equity rubric: • Membership composition • Membership recruitment • Maintaining interest in collaborating or contributing • Implement interventions • Synergy or coordination • Increased credibility and access to key populations • New or strengthened external relationships or networks • Equity • Systems change rubric: • Defined roles and responsibilities • Structure • Contribute resources • Maintaining interest in collaborating/ contributing • Coordinate and integrate activities • New or strengthened external relationships or networks • Equity • Population change rubric: • Demonstrated commitment to self-assessment • Equity

Tool Name	Source	Population/ Instructions	Terms in Partnership Concept Map
PARTNER (Platform to Analyze, Record, and Track Networks to Enhance Relationships) Tool https://visiblenetworklabs.com/partner-platform/ https://visiblenetworklabs.com/partner-tool-resources/	Visible Network Labs, n.d.	Partnership members. Collect data from respondents using the existing template or your own, score and visualize your network using the tool.	• Membership composition • Level of involvement • Partnership structure • Group dynamics • Leadership • Shared vision • Perceived benefits and drawbacks • Contribute resources • Communicate key messages to audiences and stakeholders • Can import data collected on other constructs to conduct additional analyses
Adverse Childhood Experiences (ACEs) and Resilience Collective Community Capacity (ARC3) Survey http://www.appi-wa.org/wp-content/uploads/2016/07/APPI-White-Paper.pdf	Hargreaves et al., 2016	Coalition members. Measures capacity at the coalition, network, and community-wide levels.	• Membership composition • Demonstrated commitment to self-assessment • Defined roles and responsibilities • Structure • Group dynamics • Leadership • Shared vision • Coordinate and integrate asthma activities • Contribute resources • Maintain partnerships and build collaboration • Communicate key messages to audiences and stakeholders • New or strengthened external relationships/networks • Equity

You can also use the PARTNER survey (https://visiblenetworklabs.com/partner-tool-resources/) for quality improvement purposes to understand if the network is performing well. The PARTNER Quality Improvement Methodology (Varda & Sprong, 2020) prescribes three steps to do this:

- Define the goals of the network.
- Collect data in the PARTNER survey.
- Compare the data against the goals to understand where the network is and where it wants to be and to identify steps to take to address gaps.

For more information on how to use the PARTNER Quality Improvement Methodology and examples, see

- https://www.maxwell.syr.edu/parcc/eparcc/simulations/2008_1_Simulation/
- https://visiblenetworklabs.com/wp-content/uploads/2019/03/Data-Driven-Management-Strategies-in-PHCs.pdf

APPENDIX I. Economic Evaluation

Economic evaluation is the systematic appraisal of costs and benefits of programs and projects to determine the relative economic efficiency of programs (Levin & McEwan, 2001). It is typically used alongside other types of evaluations, opposed to a 'stand-alone' effort. Like any evaluation, it involves asking good questions about the merit, worth, and value of a program (e.g., measurements in costs, consequences, and benefits), and then choosing the best method for answering them.

Economic evaluation assists with decisions that are operational, managerial, or strategic in nature by bringing an additional lens to a discussion of a program's value. It can

- **Provide information that can be very persuasive to policy makers and other decision makers**. An economic evaluation provides information about the stakeholders' perceptions of a program's value or information about alternatives to a program through systematic analysis and the accompanying interpretation of data. Such information is critical to making and defending decisions like whether to keep, expand, or eliminate a program, or how to choose among similar programs.
- **Promote fiscal responsibility and increase fiscal transparency.** When a program collects cost data on program's inputs and outputs during implementation, program managers are able to incorporate fiscal aspects into their decision making and are better equipped to address stakeholder's queries about, for example, how dollars are spent (which groups get what services); the return on investment of the program; and how changing investments in the program will impact the number of people the program might serve.
- **Help set priorities when resources are limited.** Program managers can use cost information in designing or redesigning the program as well as budgeting funds, so that the program realizes the greatest benefits at the lowest cost. When added to a process evaluation, a detailed cost analysis can help determine the costs associated with various delivery steps to aid in assessing the efficiency of program delivery and operations. Added to an outcome evaluation, findings from an economic evaluation delineate choices about a program and between program alternatives based on documented effects.

Similar to other types of evaluation, the CDC Framework can be applied to an economic evaluation. The difference is that the focus for each step progressively leads you and your stakeholders toward accurately assessing program costs and benefits to support making sound judgments about the merit of the program. The Evaluation Standards (utility, feasibility, propriety, accuracy, and evaluation accountability) are equally applicable in economic evaluation. Just as in evaluations that do not include an economic component, the standards are very useful references for when you are faced with decisions or alternatives in planning your evaluation.

When to Use Economic Evaluation

While economic evaluation can provide useful information at any stage of a program's development, it should only be done once an evaluation has demonstrated that the program works in an environment and that it is reaching those who need it. Often, it is an add-on to evaluations that examine efficacy, effectiveness, or availability, expanding the scope of the evaluations and the types of information they can provide.

Before you decide to take on any economic evaluation activities, consider the necessary conditions for successful economic evaluation. First, ensure that sufficient interest in or a need for conducting economic evaluation exists. Who needs this information and why do they need it? What do they want to know? In other words, ask, "What is the purpose of doing an economic evaluation or of adding economic evaluation questions and methods to our existing evaluation?" "How will the results be used and by whom?" These are fundamental questions since adding an economic evaluation typically requires a significant investment from a program.

Once sufficient need for the information is determined, assess the resources available for the economic

evaluation. Can you assemble a team that is both willing and knowledgeable about how to obtain, analyze, and value relevant data? If you don't have the expertise or availability of appropriate resources in-house, is there willingness and a budget to engage outside resources?

Next, consider the data you can access. Although you may not have identified all the data you will need to collect, you should determine how much access to data you can anticipate. For example, if you know you will need work or school absenteeism data, is it available? Is it available to you or any member of your team? Is it in a format that you can analyze? Does the data reflect actual costs or charges? Will you have access to aggregated or individual-level data? Note that some data may come with a cost (e.g., data held by private sources) or may require you to establish a collaboration, partnership, or formal data sharing agreements. In addition, cost data are often considered sensitive information and may require additional activities to protect appropriately.

Finally, you will need to engage in some level of pre-planning with stakeholders to consider the aforementioned aspects. To meaningfully engage stakeholders in discussion about these aspects, you may need to train them and build their capacity to understand the complexities of economic evaluation. This may involve teaching them terminology, design options, data collection and analysis methods, and how to understand the evaluation results. Be sure you have a willing audience for this type of stakeholder education. Then build time for it into your evaluation management plan. Stakeholder input is essential to accurately estimate or determine costs that may be incurred in your program's specific context or setting. A program deemed cost-effective in one community may incur different costs in another.

Developing a Program Description for Economic Evaluation

While clarity in description is important to all evaluations, in the case of economic evaluation, the description, including delineating the scope or boundaries of the program, must be clear to the stakeholders, especially those whose perspective is guiding the evaluation. The description should include information about the program's stage of development—whether it is a pilot, a fairly new program, or a mature program. The description should also include any information about the program's efficacy or effectiveness. These factors will have implications for the types of questions you will ask and the methods you will use to answer them. Refer to Module 5, *Evaluating Services and Health Systems Interventions* for more information.

In economic evaluation, a program description in economic evaluation has two additional program components that have significant impact on program costs and benefits and should be included: the **time frame** and the **analytical horizon** (see **Figure I.1**).

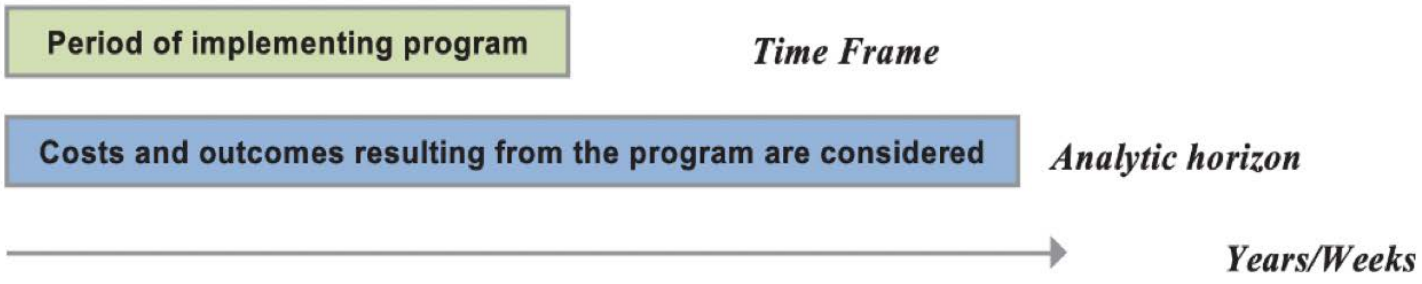


Figure I.1 The Distinction Between Time Frame and Analytic Horizon

The time frame is a specified period during which all program activities are undertaken to produce all planned program outputs. Note that if a program's time frame does not account for seasonal variations, it will misreport the costs that might be expected in a program for which the time frame spans a longer period. For example, if an asthma program is delivered during a period when emergency department visits for asthma exacerbations are typically low. Therefore, an indication of the program's time frame in your program description is crucial not only for thinking about your evaluation design, but also for interpreting, communicating, and most importantly, replicating your results.

The analytic horizon is the period over which all program costs and outcomes of interest are measured and accounted for in the analysis. While the aim of the program is to achieve the long-term outcomes (e.g., reduced suffering and death), it is also important to consider intermediate outcome measures like increased symptom-free days or reduced medication use. These are often more immediately relevant to the program's activities.[12] Additionally, demonstrating a link between your program and the long-term outcomes is challenging, due to long lag periods between the intervention and the health outcome, lack of data, and other confounding influences. An analytic horizon that is long enough to encompass intermediate program outcomes is considered reasonable when evaluating most public health programs. **Figure I.2** shows how a logic model may represent the concepts of time frame and analytic horizon. Remember that programs can incur costs before implementation (e.g., staff training), during implementation, and after implementation (e.g., staffing costs to collect data on program benefits).

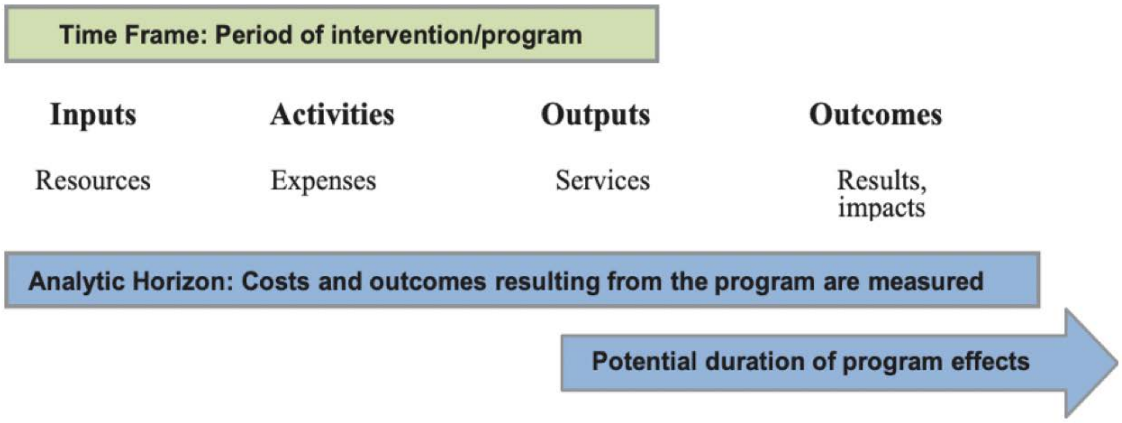


Figure I.2 Time Frame and Analytic Horizon Within a Logic Model Construct

In practice, decisions about the time frame and analytic horizon are largely influenced by available resources and time constraints. You will need to work with your stakeholders to ensure they understand how decisions regarding the time frame and analytic horizon may impact your results.

The Importance of Perspective

An important decision that you and your stakeholders will make when conducting economic evaluation is deciding the viewpoint, or **perspective**, your evaluation will take. The choice of perspective drives decisions about the relevant costs and benefits to include in the evaluation, as well as those that can be excluded. Common perspectives in economic evaluation are program, program participant, payer, and societal.

The **program perspective** represents the point of view of a program and explores specific costs and benefits that can be attributed to the program (Haddix et al., 1996). The program perspective is usually chosen when the results of economic evaluation will be used primarily by the program staff. The **program participant perspective**, also called client perspective, focuses on the costs of participation borne by program participants (e.g., fees, transportation costs, or lost labor); it also typically includes the benefits the client receives from participation (Haddix et al., 1996). The **payer perspective** examines costs and benefits from the viewpoint of the person or entity ultimately responsible for the financial cost of the program. For example, from the viewpoint of an insurance company or a foundation (Haddix et al., 1996). Finally, the **societal perspective** includes costs and benefits of everyone directly and indirectly affected by the program, including taxpayers (Haddix et al., 1996). It is the broadest possible perspective and is typically used when the intent is to share outcomes with policy makers or to resonate with wider audiences.

[12] See http://www.cdc.gov/asthma/nacp.htm

Focusing the Design

The first step in focusing your economic evaluation design will be to develop evaluation questions. A well-defined economic evaluation question specifies the perspective, identifies alternatives, and then makes a succinct statement about the purpose and use of the information that the evaluation will generate. For example: from the perspective of a Medicaid-managed care organization, is a home-based environmental program provided by community health workers more cost-effective than the currently available asthma disease management program provided by healthcare providers? It is typically best to limit the evaluation questions to a single perspective; mixing viewpoints could result in confusion, overlapping, or double counting (Rossi & Freeman, 1993).[13] If varied perspectives are needed, you need to ensure the ramifications of such decisions are clearly understood by your stakeholders. **Table I.1** provides a list of sample questions that are generally encountered at project, program, or policy levels.

Table I.1 Sample Evaluation Questions

Evaluation Question	Possible Corresponding Economic Evaluation Question
To what extent are our initiatives equitably serving people in our priority population?	From the perspective of the program, how are the program funds being expended across the jurisdiction's population?
How well are program initiatives being directed to address disparities among children?	From the perspective of the program, what portion of allocated funds is being expended to address disparities among children?
How can the community health workers (CHW) program be expanded to reach more families?	From the perspective of the home-visit program, how many CHWs can be added to the program, without increasing the current fixed cost?
How can we increase the efficiency of the CHWs in our home visiting program?	From the perspective of the home-visit program, would allowing CHWs to work from home (rather than the office) be cost-effective?
How does the number of symptom-free days among participants vary with changes in the level of home environmental remediation (minor, moderate, and major remediation)?	From the perspective of the program, how cost-effective is it to fund moderate- or major-level home environmental remediation over minor level?
How many symptom-free days did the adult patients experience in the six months after participating in the self-management education (SME) program provided by nurses in the hospital?	From the perspective of a payer, how cost-effective is offering an SME program to adults after an ED visit, as compared to not offering the program?
To what extent did schools realize reductions in student absences after they participated in a health management program?	From the societal viewpoint, what is the net benefit and program cost of having one fewer school day missed compared to nonparticipation in the program?
How many symptom-free days were reported by students after known triggers were removed from school?	From the viewpoint of a payer, what are the net benefits and program cost of having one additional symptom-free day as a result of the trigger-related intervention?
How do symptom-free day outcomes and the net costs compare in a self-management education program when it is provided by school nurses or by pharmacists?	From the societal viewpoint, which is more cost-effective: self-management education in schools or in pharmacies?

[13] For more information about the complexities of using multiple perspectives, see Rossi and Freeman (1993).

Selecting the Right Type of Economic Evaluation

As with all evaluations, the types of evaluations you conduct and the methods you choose depend on the evaluation's focus—its identification of stakeholder information needs balanced with evaluation resources. There are many types of economic evaluation to choose from, each designed to answer specific questions about a program's costs, consequences, or benefits. A **partial economic evaluation** assesses either the costs or the outcomes of a program, but not both; a f**ull economic evaluation** examines both.

Table I.2. summarizes the methods commonly used for partial and full economic evaluations. The most basic type of analysis, cost analysis, looks at a single program and provides insights on the costs incurred and the savings realized by that program. It is the first step in any economic evaluation and goes in the "partial" category. Full-scale economic evaluations include cost-effectiveness analysis, cost-benefit analysis, and cost-utility analysis.

Table I.2 Commonly Used Analytic Methods in Economic Evaluation and Description

Common Economic Analysis Methods	
Cost Description and Cost Analysis	Cost Description and Cost Analysis are the most basic and common types of economic evaluation. Cost Description involves systematic collection, categorization, and analysis of all program costs. Cost Analysis helps with understanding the net program costs – costs that the program incurs (program costs) and costs that it saves or averts (program savings). Cost Analysis is the first step in any full economic evaluation.
Cost Minimization Analysis	Cost-Minimization Analysis measures and compares input costs. It assumes outcomes to be equivalent.
Cost-Effectiveness Analysis	Cost-Effectiveness Analysis relates net program costs to a quantifiable outcome measure of program's effectiveness. Cost-Effectiveness evaluation can determine the net program cost of each unit of effectiveness, or the additional net program cost for each incremental unit of effectiveness.
Benefit-Cost Analysis	Benefit-Cost Analysis (also referred to as cost-benefit) identifies and places dollar values on the costs of programs and weighs those costs against the dollar value of program benefits accrued over a defined period. Return on Investment is a type of Benefit-Cost Analysis.
Cost-Utility Analysis	Cost-Utility Analysis compares net program costs and the net effectiveness of a program, which is measured in terms of health years, number of life years saved, or Quality Adjusted Life Years.

Figure I.3 depicts the types of economic evaluation described in **Table I.2** and may be helpful in narrowing down the type of economic evaluation you would want to consider for an evaluation.

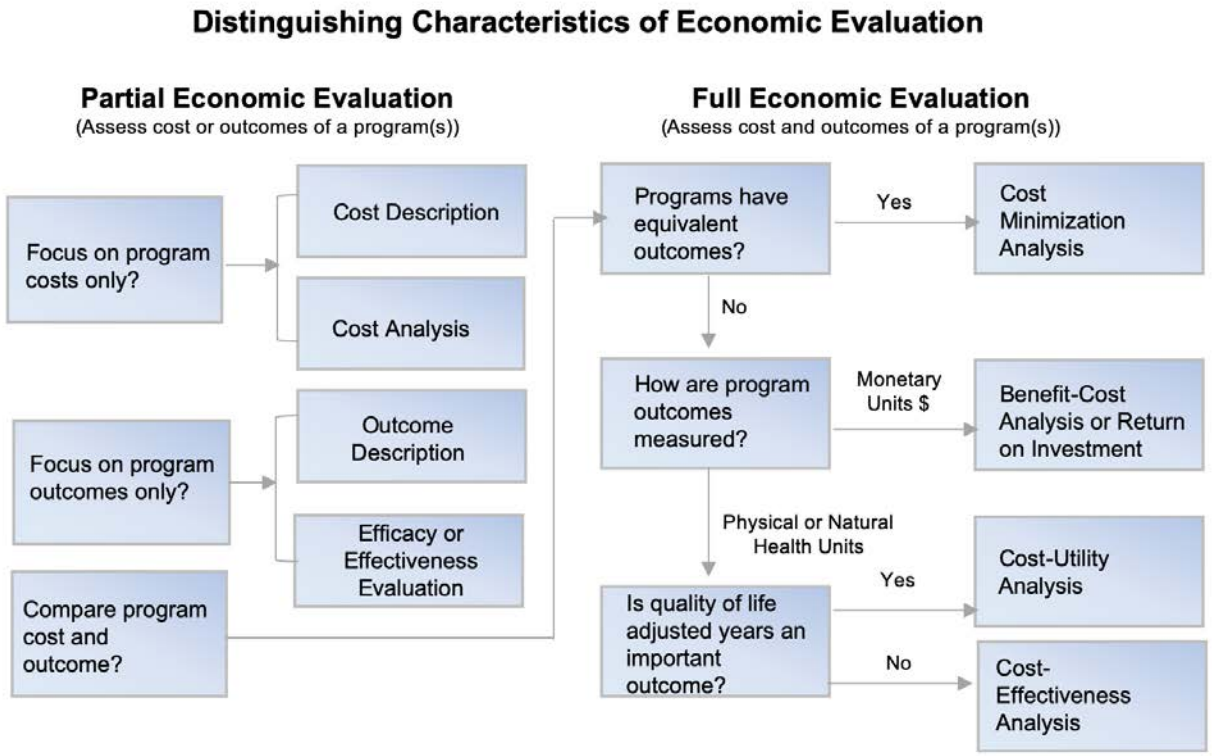


Figure I.3 Distinguishing Characteristics of Economic Evaluation

Return on investment (ROI) is another important way to assess the value of a program. ROI answers the question: "What is the payback for the investment made to implement the program over a defined period?" (Buzachero et al., 2013). A variation of benefit-cost analysis, ROI is often preferred when the focus is on understanding direct program costs and monetized benefits, and when the parameters of the program and the evaluation outcomes are clearly defined (Buzachero et al., 2013). ROI is commonly used for building business cases—as a justification for a program or intervention that is intended to convince a decision maker to approve some kind of action, typically funding. Because of this, you will typically use the payer's perspective for the analysis.

Some argue that judging a public health service by its monetary value is inadequate, and even inappropriate. In the last decade, Social Return on Investment (SROI) has emerged to address these concerns and account for "social value" of a program that explores value in a broader sense, beyond cost, prices, and investments, taking into account social, economic, and environmental factors (Nicholls et al., 2012). SROI takes a systematic and holistic perspective on whether a project is beneficial and profitable. While ROI may be presented as a sole outcome measure, SROI must be considered as an additional piece of information to provide to stakeholders. You will use other information generated through an evaluation to tell the complete story for the program.

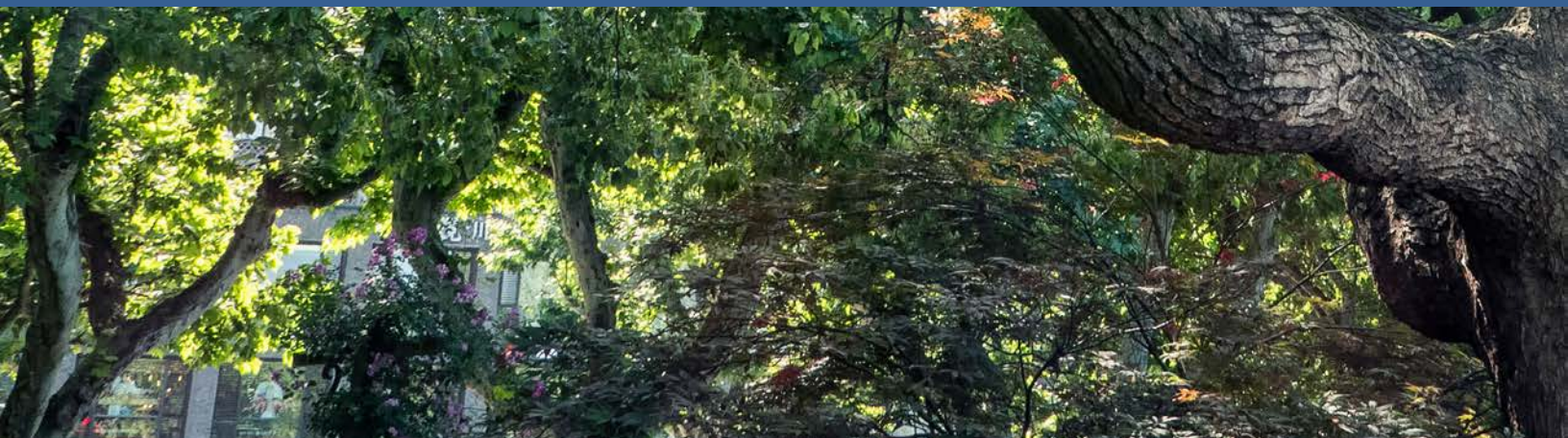
Additional Resources

This appendix is designed to provide a basic understanding of economic evaluation as a complement to other evaluations. It by no means covers all the possible methodologies for economic evaluation and the nuances behind them, nor does it provide all the information needed to conduct one. Those interested in learning more may refer to the *CDC's Learning & Growing through Evaluation: State Asthma Program Evaluation Guide*, which includes a module on Economic Evaluation for State Asthma Programs. In addition to a more extensive description of

economic evaluation, the module includes worksheet and templates to assist with planning and implementation. Other useful resources include the following:

- CDC's Introduction to Economic Evaluation Tutorials: These tutorials teach the basics of economic evaluation, why they should be conducted, forms of economic evaluations, and how to determine which forms to use. http://www.cdc.gov/dhdsp/pubs/docs/CB_January_10_2012.pdf
- Drummond, M. F., Sculpher, M. J., Claxton, K., Stoddart, G. L., & Torrance, G. W. (2015). *Methods for the economic evaluation of health care programmes* (4th ed.). Oxford University Press.
- Five-Part Webcast on Economic Evaluation (Division for Heart Disease and Stroke Prevention): Economic evaluation is a process to identify, measure, value, and compare the costs and outcomes of programs and policies. This webcast series is designed to help you understand the value of economic evaluation and how to incorporate these methods into your programs. The five-part webcast will assist you in choosing the appropriate economic analysis for your heart disease and stroke prevention program or policy. https://www.cdc.gov/dhdsp/evaluation_resources/economic_evaluation/index.htm
- Gold, M. R., Siegel, J. E., Russell, L. B., & Weinstein, M. C. (Eds.). (1996). *Cost-effectiveness in health and medicine*. Oxford University Press.
- Levin, H. M., & McEwan, P. J. (2001). *Cost-effectiveness analysis: Methods and applications* (2nd ed.). Sage Publications.

GLOSSARY



GLOSSARY

Action Model	A set of prescriptive assumptions about a program or policy that indicate the activities or components of an intervention that must be implemented to lead to the desired changes or outcomes in a change model (Chen, 2015). See change model.
Analytic Horizon	The duration of time into the future during which costs and effects that accrue from an intervention are considered (Haddix et al., 1996).
Capacity Assessment	An assessment of the effectiveness and quality of an agency's evaluation and research activities to identify areas in need of improvement and priorities for building capacity (OES, n.d.). The results of the capacity assessment can also inform strategic decision making for allocating resources toward building capacity.
Case Study	"...an empirical method that investigates a contemporary phenomenon (the 'case') in depth and within its real-world context" (Yin, 2018, p. 15). A case study focuses on a particular unit - a person, initiative, or program and typically uses a combination of quantitative and qualitative data. They are often used to gain a deep understanding of how different factors produce observed outcomes.
Change Model	A set of descriptive assumptions about the causal processes that underlie a program or policy and its outcomes (Chen, 2015). This model describes the intended outcomes or changes that result from the set of activities described in an action model. See action model.
Cognitive Interviewing	A way of testing the appropriateness of questions in a questionnaire (Willis, 2005). Specifically, people are asked to complete the questionnaire, thinking aloud and articulating their thoughts about the questions and why they are responding as they are.
Comparison Group	A group not exposed to a program or treatment. Sometimes referred to as a control group, comparison group is a term used more frequently in quasi-experimental design (than in experimental designs; DHHS, 2005; EPA, 2007).
Complexity Theory	A theory that proposes that systems, programs, or an environment are holistic rather than a sum of parts and that elements in the systems interact with each other in a non-linear way that evolves over time (Byrne, 1998). As such, one can guide the system or program toward desired outcomes but must be able to manage the process in different ways depending on how the system evolves.

Control Group	A group whose characteristics are similar to those of a program's participants but who do not receive the program services, products, or activities being evaluated. Participants are randomly assigned to either the experimental group (those receiving program services) or the control group. A control group is used to assess the effect of program activities on participants who are receiving the services, products, or activities being evaluated. The same information is collected for people in the control group and those in the experimental group (EPA, 2007).
Conceptual Use	This refers to the ways that the findings from an evaluation influence or shape the way people think about an evaluand (Rogers, 2005). Also referred to as enlightenment use.
Cost-Effective	Refers to the least cost or greatest effectiveness of a program or policy for a specified level of cost (EPA, 2007). This is measured in nonmonetary outcomes by relating the costs of a program or policy to its performance.
Criteria of Merit	"…the aspects of what is being evaluated that define whether it is good or bad and whether it is valuable or not valuable" (Davidson, 2005c, p. 239).
Critical Reflection	An approach to practice that involves a "sustained and intentional process of identifying and checking the accuracy and validity" of one's assumptions about their knowledge, values, beliefs, interpretations (Brookfield, 2017, p.3).
Cultural Responsiveness	Acknowledges and gives attention to the values, beliefs, and customs of a particular group or community (Hood, Hopson, & Kirkhart, 2015). In an evaluation, cultural responsiveness means attending to the cultural aspects of a program and its stakeholders in a respectful way while also being aware of one's own cultural identity.
Culturally Responsive Evaluation	An evaluation theory that emphasizes the importance of using a culturally responsive approach when conducting evaluations of programs that serve populations in cultural contexts unfamiliar to the evaluator or different than the evaluator's cultural background (Frierson et al., 2010). In addition to being an evaluation theory, culturally responsive evaluation also refers to an evaluator's responsibility to conduct evaluation in a manner that is sensitive to the culture of the program and its participants.
Deliberative Democratic Evaluation	An evaluation theory that engages stakeholders in a dialogue to arrive at unbiased justifiable conclusions by considering all relevant perspectives, values, and interests (House, 2005). This approach recognizes that biases can never be eliminated, but aims to reduce them through inclusion, dialogue, and deliberation.
Economic Evaluation	The systematic appraisal of costs and benefits of projects, normally undertaken to determine the relative economic efficiency of programs (Levin & McEwan, 2001).

Effect Modifier	A third variable (that is neither the treatment or exposure variable nor the outcome variable) that produces different results at each level of the third variable (also called an interaction) (Bovjerg, n.d.). For example, smoking and exposure to asbestos are risk factors for lung cancer. However, the risk of getting lung cancer (the outcome) is higher for people who have been exposed to asbestos (the exposure) and people who smoke (the effect modifier).
Effective	Refers to the extent to which a program or policy produces intended outcomes (Davidson, 2005a).
Efficient	Refers to the extent to which a program or policy produces intended outputs and outcomes without wasting resources, time, or money (Davidson, 2005b).
Enlightenment Use	This refers to the ways that the findings from an evaluation influence or shape the way people think about an evaluand (Rogers, 2005). Also referred to as conceptual use.
Evaluability Assessment	An evaluability assessment (EA) is a systematic pre-evaluation assessment of a program or activity designed to determine the utility and feasibility of conducting a full evaluation. Evaluability assists with determining whether an intervention is at an appropriate stage of development to warrant rigorous outcome evaluation; it also ascertains a program's capacity to carry out such an evaluation (Leviton & Gutman, 2010).
Evaluand	The object of an evaluation, which may include a program, policy, product, or person (Mathison, 2005a).
Evaluation Accountability	One of the program evaluation standards developed by the Joint Committee on Standards for Educational Evaluation. This standard encourages increased transparency in planning and implementation of evaluation as well as how conclusions are drawn through documentation and meta-evaluation (Yarbrough et al., 2011).
Evaluation Candidate	As used in this text, this term refers to a list of potential programs, policies, or evaluands that may be selected for evaluation.
Evaluation Planning Team	As used in this text, this term refers to a group of stakeholders charged with advising the planning of an evaluation. Also referred to as an evaluation advisory group.
Evaluation Policy	An evaluation policy defines the role and use of evaluation within an organization (BetterEvaluation, 2020). The policy document may outline a framework for evaluation within the organization, including the evaluation function, roles and responsibilities, and how evaluations are planned, managed, and used.
Evaluation Question	A question generated by your Stakeholders to ascertain information about a program's implementation, Outputs, or Outcomes, depending on where on the continuum of the logic model the evaluation is focused. The goal of an evaluation effort is to answer one or more evaluation question(s) (Russ-Eft & Preskill, 2009).

Evaluation Theory	An approach or model that provides a body of principles, prescriptions, or guiding frameworks for conducting an evaluation (Alkin, 2013). Evaluation theories serve to provide direction to the practice of evaluation (Mathison, 2005b).
Evaluative Thinking	This is a type of critical thinking used in the context of evaluation that is characterized by a sense of curiosity, inquiry, and valuing of evidence (Buckley et al., 2015). Evaluative thinking entails identifying assumptions, posing thoughtful questions, and trying to understand different perspectives to inform decision making.
Experimental Design	A design that tries to ensure the initial equivalence of one or more control groups to a treatment group by administratively creating the groups through random assignment, thereby ensuring their mathematical equivalence (DHHS, 2003). Examples of experimental or randomized designs are randomized block designs, Latin square designs, fractional designs, and the Solomon four-group. Also referred to as a randomized controlled trial or a true experiment.
Fidelity	The extent to which the delivery of an intervention is implemented as originally planned or developed (Mowbray et al., 2003).
Formative Evaluation	Evaluative activities undertaken to furnish information that will guide program improvement (EPA, 2007).
Full Economic Evaluation	Full economic evaluations are studies in which a comparison of two or more interventions or care alternatives is undertaken and in which both the costs and outcomes of the alternatives are examined (Drummond et al., 2015).
Health Equity	A state where everyone has access to health and where the barriers to health such as poverty, prejudice, and discrimination are eliminated (Braveman et al., 2017).
Indicator	A specific, observable, and measurable characteristic or change that shows the progress a program is making toward achieving a specified outcome (DHHS, 2005).
Instrumental Use	This refers to the direct use of findings from an evaluation to inform decisions or make changes an evaluand (Rogers, 2005).
Time Series Design	A research design that collects data over long time intervals before, during, and after program implementation. This allows for the analysis of change in key factors over time (EPA, 2007).
Intervention Group	The participants in the intervention being studied (EPA, 2007). Also referred to as the intervention group.
Interpersonal Competence	A set of social skills for constructive interactions during the evaluation process, including communication, conflict resolution, and facilitation skills (Garcia & Stevahn, 2019).

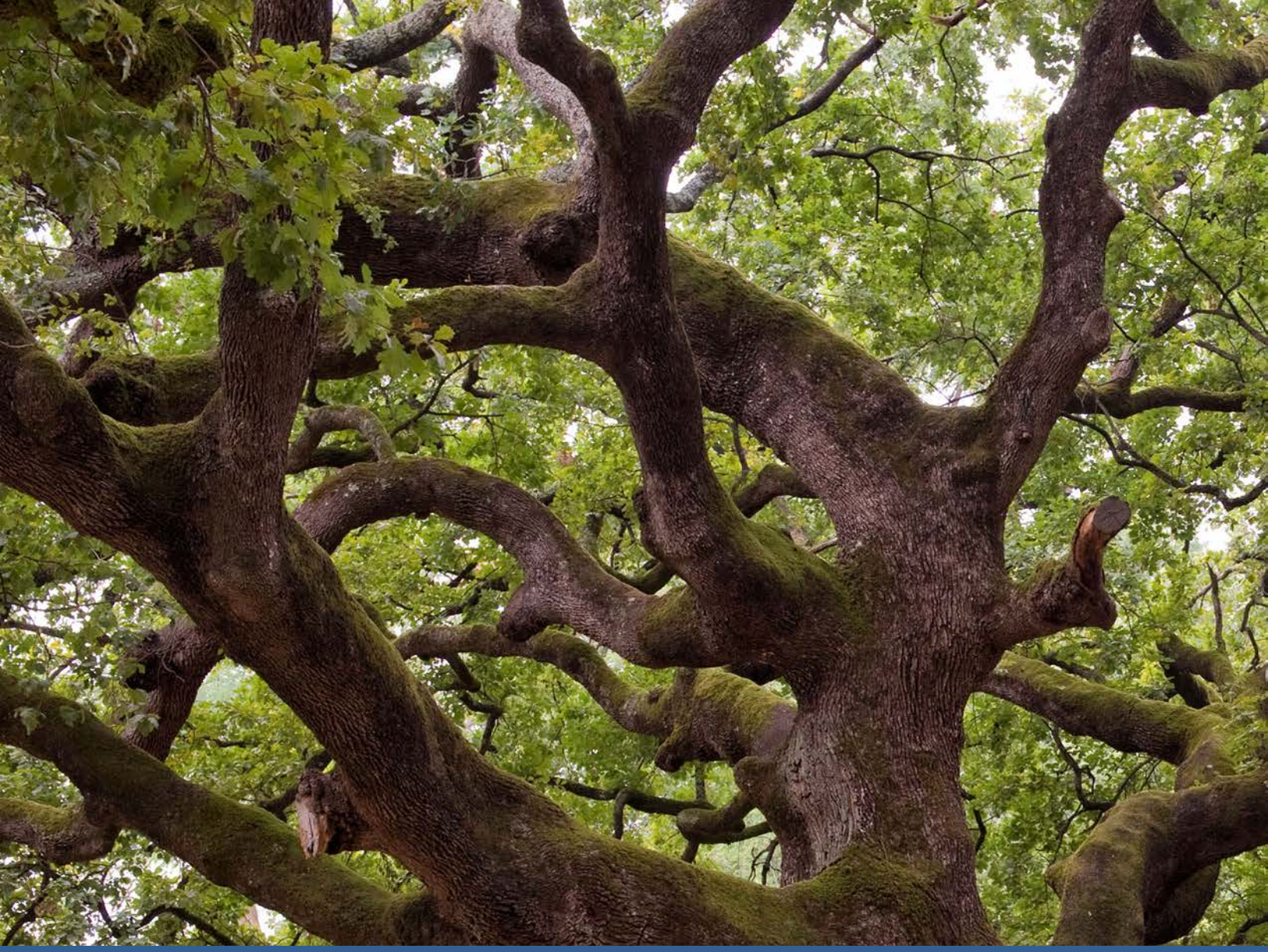
Learning Agenda	A plan for identifying priority questions related to an agency's programs or policies used to engage stakeholders in evidence planning and building (OES, n.d.). The results of a learning agenda can be used to inform decisions about improving an agency's work.
Logic Model	A systematic and visual way to present the perceived relationships among the resources you have to operate the program, the Activities you plan to do, and the changes or results you hope to achieve (DHHS, 2005).
Mediator	A third, intermediary variable between the independent and dependent variables that helps to explain the causal relationship (Crano et al., 2015).
Meta-Evaluation	External and internal review of evaluation processes and outcomes to determine whether procedures were appropriate and conclusions are valid (Stake, 2014).
Mixed-Methods Design	In a mixed-methods design, the evaluator collects and analyzes both qualitative and quantitative data rigorously in response to evaluation questions, integrates the two forms and their results, organizes these procedures into specific evaluation designs that provide the logic and procedures for conducting the study, and frames these procedures within theory and philosophy (Creswell & Plano Clark, 2018).
Moderator	A third variable that can either increase or decrease the cause-effect relationship between the independent and dependent variables (Crano et al., 2015).
Modus Operandi	This term refers to identifying the cause of a particular outcome by searching for clues that uncover the chain of events leading up to that effect (Mathison, 2005c). This is often used in police work when uncovering the cause of a crime. Michael Scriven is credited with introducing the term to evaluation.
Non-Experimental Design	An evaluation design in which participant information is gathered during or after an intervention. There is no Comparison Group, Control Group, or repeated measurements of the treatment group (DHHS, 2005; Salabarría-Peña et al., 2007).
Outcome Evaluation	The systematic collection of information to assess the impact of a program, present conclusions about the merit or worth of a program and make recommendations about future program direction or improvement (DHHS, 2005).
Partial Economic Evaluation	Economic studies which consider costs or consequences but which either do not involve a comparison between alternative interventions or do not relate costs to benefits (Drummond et al., 2015).
Partnership Evaluation	As used in this text, an evaluation that assesses an aspect of public health infrastructure, namely partnerships between organizations or coalitions targeting a particular public health issue. A partnership evaluation may assess the process of recruiting and engaging group members or may assess the outcomes of the partnership and its impact on the public health issue of interest.

Performance Monitoring	The periodic measurement of progress toward short, intermediate, or long-term results often used to provide information to improve performance (Kusek & Rist, 2005).
Personal Factor	A concept in which the presence of an individual or group of people who are interested in evaluation and using the findings results in uptake of evaluation results (Patton, 1997). In contexts where the personal factor is present, evaluations tend to be used for decision making or making improvements to the evaluand.
Perspective	The viewpoint of the Audience vis-à-vis the evaluation (Haddix et al., 1996). • **Program perspective** - the point of view of a program. It explores specific costs and benefits that can be attributed to the program. Usually chosen when the results of economic evaluation will be used primarily by the program staff (Haddix et al., 1996). • **Program participant perspective** - (also known as client perspective) focuses on the costs of participation borne by program participants, such as fees, transportation costs, or lost labor; it also typically includes the benefits the client receives from participation (Haddix et al., 1996). • **Payer perspective** - examines costs and benefits from the viewpoint of the person or entity ultimately responsible for the financial cost of the program, such as an insurance company (Haddix et al., 1996). • **Societal perspective** - includes costs and benefits of everyone directly and indirectly affected by the program, including taxpayers (Haddix et al., 1996).
Post-Only Design	A non-experimental design in which measures (data collection) are taken from the primary population(s) after the activity or intervention (Salabarría-Peña et al., 2007). Since this is a non-experimental design, it does not involve comparison/control groups.
Pre-Post Test Design	This elementary quasi-experimental design involves the measurement of indicators prior to implementation of the treatment, and subsequent re-measurement after implementation (EPA, 2007). Any change in the measure is attributed to the treatment with acknowledgement of validity threats. Also known as a before-after design.
Process Evaluation	The systematic collection of information to document and assess how well the program was implemented (Issel, 2009).
Process Use	This refers to "individual changes in thinking and behavior, and program or organizational changes in procedures and culture, that occur among those involved in evaluation as a result of the learning that occurs during the evaluation process" (Patton, 1997, p. 90).
Program Impact Theory	The conceptual theory for how a program is presumed to solve a problem or problems (Donaldson, 2007).

Program Theory	A set of stakeholder assumptions describing what activities are needed to solve a social problem and why those activities will solve the problem (Chen, 2005). Program theory is the conceptual framework for how the program is intended to work and is useful in evaluation for understanding whether a program is effective or not and why.
Quasi-Experimental Design	Study structures that make comparisons to draw causal inferences but do not use randomization to create the treatment and Comparison Groups. The treatment group is usually given the treatment or program, whereas the comparison group is not; comparison groups may be selected to match the treatment group as closely as possible, selected as non-equivalent comparison groups which must be corrected for statistically, selected based on a specified pre-program cutoff score, or the treatment group may serve as its own comparison group over time to observe changes in an outcome; in this way inferences on the incremental impacts of the program can be made (Campbell & Stanley, 1966; Trochim, 2020).
Random Assignment	The assignment of individuals in the pool of all potential participants to either the experimental (treatment) group or the control group in such a manner that their assignment to a group is determined entirely by chance (GAO, 2012; GAO, 2005).
Randomized Controlled Trial	A design that tries to ensure the initial equivalence of one or more control groups to a treatment group by administratively creating the groups through random assignment, thereby ensuring their mathematical equivalence (DHHS, 2003). Examples of experimental or randomized designs are randomized block designs, Latin square designs, fractional designs, and the Solomon four-group. Also referred to as an experimental design or a true experiment.
Regression Discontinuity Design	A design that assesses the effect of a treatment condition by looking for a discontinuity in regression lines between individuals who score lower and higher than some predetermined cutoff score (Johnson & Christensen, 2008).
Repeated Measures	This quasi-experimental design involves the measurement of outcome indicators over time. This design can include a simple pre and post evaluation design where the indicator in question is only measured once before the intervention and after the intervention is introduced. This design can also be used if you have different versions of the intervention you are testing. In this case, you collect data prior to implementation of the intervention and then after each version of the intervention is introduced. This evaluation design is also useful when a comparison or control group is not available to use (Crano et al., 2015).
Research on Evaluation	The empirical study of evaluation theory and practice (Smith, 1993). This includes descriptive studies of evaluation practice, meta-evaluations, analyses of evaluation theories among others.

Self-Assessment	A personal process that enables an individual to reflect on their knowledge, skills, and abilities to identify strengths and areas for improvement (Ghere et al., 2006).
Situational Awareness	This refers to the ability of an evaluator to understand how contextual factors such as the program history, size, and complexity; the purpose of the evaluation (e.g., formative, summative); evaluator experience, resource constraints; politics; and other factors affect evaluation design and use (Patton, 2005). Being situationally aware enables an evaluator to adapt and respond to these contextual factors by negotiating and implementing an evaluation that fits the intended uses.
Stakeholders	People or organizations that are invested in the program (program stakeholders) or that are interested in the results of the evaluation or what will be done with results of the evaluation (evaluation stakeholders; DHHS, 2005).
Strategic Evaluation Planning Team	As used in this text, this term refers to a group of program Stakeholders charged with directing implementation of the strategic evaluation plan.
Summative Evaluation	A type of outcome evaluation that assesses the results or outcomes of a program. This type of evaluation is concerned with a program's overall effectiveness (EPA, 2007).
Surveillance Evaluation	As used in this text, an evaluation that assesses an aspect of public health surveillance. A surveillance evaluation may assess the appropriateness, effectiveness, or utility of a surveillance system for meeting its objectives.
Systems Thinking	A set of analytic skills used for understanding and predicting system behavior and for developing alterations to introduce to the system to obtain desired results (Arnold & Wade, 2015).
Theory of Change	A visual explanation of the causal processes that link program inputs to expected outputs and outcomes (Weiss, 1998). A theory of change describes how the program is supposed to work to achieve its intended outcomes.
Theory Driven Evaluation	An evaluation approach that is organized by articulated assumptions for how an intervention will effect social change. These articulated assumptions are used to guide the design and execution of evaluation projects by prescribing what factors effect change and the types of change expected. Theory-driven differs from method-driven evaluation in that the latter is guided by the structural goals of a particular method, e.g., qualitative evaluation (Chen, 1990; Donaldson, 2007).
Threats to Internal Validity	The factors that can threaten the validity of the causal relationship established between the intervention and outcomes; threats include: history, maturation, testing, instrumentation, statistical regression, mortality, selection bias, diffusion of treatment information, compensatory treatment equalization, compensatory rivalry, and demoralization of comparison group (Campbell & Stanley, 1966; Trochim, 2020).

Time Frame	The specified period over which the intervention strategies are actually applied or implemented (Haddix et al., 1996).
Transformative Participatory Evaluation	An evaluation theory that aims to stimulate action in the pursuit of social justice (Mertens & Wilson, 2019). This approach seeks to include people who are marginalized to address issues of inequity that inhibit achieving program goals.
Treatment Group	The participants in the intervention being studied (EPA, 2007). Also referred to as the intervention group.
True Experiment	A design that tries to ensure the initial equivalence of one or more control groups to a treatment group by administratively creating the groups through random assignment, thereby ensuring their mathematical equivalence (DHHS, 2003). Examples of experimental or randomized designs are randomized block designs, Latin square designs, fractional designs, and the Solomon four-group. Also referred to as an experimental design or a randomized controlled trial.
Utilization-Focused Evaluation	An evaluation theory or approach in which the use of evaluation findings is key for designing and implementing the evaluation (Patton, 1997). Evaluation is conducted in a manner that aligns with the information needs of the people who are the intended users of the evaluation findings.
Values-Engaged Educative Evaluation	An evaluation theory that pays special attention to the values held by stakeholders in a program, including the values of diversity and equity, and aims to facilitate learning and increased understanding about an evaluand from diverse stakeholder perspectives (Greene et al., 2011).



REFERENCES

REFERENCES

Alkin, M. C. (2013). *Evaluation roots: A wider perspective of theorists' views and influences* (2nd ed.). SAGE Publications.

Alkin, M. C., & Christie, C. A. (2005). Unraveling theorists' evaluation reality. In M. C. Alkin, & C. A. Christie (Eds.), *Theorists' models in action. New Directions for Evaluation, 106,* 111–128. https://doi.org/10.1002/ev.155

Alkin, M. C., & King, J. A. (2016). The historical development of evaluation use. *American Journal of Evaluation, 37*(4), 568–579. https://doi.org/10.1177/1098214016665164

Allies Against Asthma. (2002). *Coalition self-assessment survey II.* http://www.asthma.umich.edu/media/eval_autogen/CSAS.pdf

Allies Against Asthma. (2003). *Key informant interviews.* www.asthma.umich.edu/media/eval_autogen/key_informant.pdf

American Evaluation Association [AEA]. (n.d.). *The 2018 AEA evaluator competencies.* https://www.eval.org/Portals/0/Docs/AEA%20Evaluator%20Competencies.pdf

American Evaluation Association [AEA]. (2011). *Statement on cultural competence in evaluation.* https://www.eval.org/Community/Volunteer/Statement-on-Cultural-Competence-in-Evaluation

American Evaluation Association [AEA]. (2018). *Guiding principles for evaluators.* https://www.eval.org/About/Guiding-Principles

Arnold, R. D., & Wade, J. P. (2015). A definition of systems thinking: A systems approach. *Procedia Computer Science, 44,* 669–678.

Ash, S. L., & Clayton, P. H. (2009). *Learning through critical reflection: A tutorial for service-learning students* (Instructor version.). Ash, Clayton & Moses.

Association of Schools & Programs of Public Health [ASPPH]. (n.d.). *ASPPH core competency model.* Association of Schools & Programs of Public Health. https://www.aspph.org/teach-research/models/mph-competency-model/

Ayoo, S., Wilcox, Y., LaVelle, J. M., Podems, D., & Barrington, G. V. (2020). Grounding the 2018 AEA evaluator competencies in the broader context of professionalization. In J. A. King (Ed.), *The American Evaluation Association's Program Evaluator Competencies. New Directions for Evaluation*, 168, 13–30. https://doi.org/10.1002/ev.20440

Bamberger, M., Rugh, J., & Mabry, L. (2006). *RealWorld evaluation: Working under budget, time, data, and political constraints.* Sage Publications.

BetterEvaluation. (2014). *One-three-twenty-five (1:3:25) principle.* BetterEvaluation. https://www.betterevaluation.org/en/evaluation-options/report_friendly_writing

BetterEvaluation. (2020). *Evaluation policy.* https://www.betterevaluation.org/en/evaluation-options/evaluation_policy

Bickman, L. (1987). *Using program theory in evaluation.* Jossey-Bass.

Borden, L. M. (1999). Assessing your collaboration: A self-evaluation tool. *Journal of Extension, 37*(2), 2TOT1.

Bourgeois, I., Fierro, L. A., Lemire, S., Castleman, A. M., & Cho, M. (2021). Taking stock of evaluation capacity building: Findings from a structured review [Manuscript in preparation].

Bovjerg, M. (n.d.). Effect modification. In *Foundations of epidemiology.* Creative Commons License. https://open.oregonstate.education/epidemiology/chapter/effect-modification/

Braveman, P., Arkin, E., Orleans, T., Proctor, D., & Plough, A. (2017). *What is health equity?: And what difference does a definition make?* Robert Wood Johnson Foundation. https://www.rwjf.org/en/library/research/2017/05/what-is-health-equity-.html

Brookfield, S. D. (2017). *Becoming a critically reflective teacher* (2nd ed.). John Wiley & Sons, Incorporated. https://public.ebookcentral.proquest.com/choice/publicfullrecord.aspx?p=4790372

Brownson, R. C., Fielding, J. E., & Maylahn, C. M. (2009). Evidence-based public health: A fundamental concept for public health practice. *Annual Review of Public Health, 30(1).*

Bruner Foundation. (n.d.). *Commissioning evaluation: Tips for grantmakers and grant seekers.* http://www.evaluativethinking.org/docs/EvaluativeThinking.bulletin.5.pdf

Brunner, W. M., Ross, S. K., & Johnson, J. E. S. (2009). Review of the asthma mortality rate for Minnesota residents aged 55 years or older, 2004-2005: When death certificates deserve a second look. *Preventing Chronic Disease*, 6(3), A92.

Buckley, J., Archibald, T., Hargraves, M., & Trochim, W. M. (2015). Defining and teaching evaluative thinking: Insights from research on critical thinking. *American Journal of Evaluation, 36*(3), 375–388.

Butterfoss, F. (1998a). *Coalition effectiveness inventory.* Center for Pediatric Research, South Carolina DHEC. http://coalitionswork.com/wp-content/uploads/coalition_effectiveness_inventory.pdf

Butterfoss, F. (1998b). *Coalition effectiveness inventory self-assessment tool.* Center for Pediatric Research, South Carolina DHEC. http://coalitionswork.com/wp-content/uploads/coalition_effectiveness_inventory.pdf

Butterfoss, F. D. (2006). Process evaluation for community participation. *Annual Review of Public Health, 27*(1), 323–340. https://doi.org/10.1146/annurev.publhealth.27.021405.102207

Butterfoss, F. D., Goodman, R. M., & Wandersman, A. (n.d.). Community coalitions for prevention and health promotion: Factors predicting satisfaction, participation, and planning. *Health Education Quarterly, 23*(1), 65–79. https://doi.org/10.1177/109019819602300105

Buzachero, V. V., Phillips, J., Phillips, P. P. P., & Phillips, Z. L. (2013). *Measuring ROI in healthcare: Tools and techniques to measure the impact and ROI in healthcare improvement projects and programs.* McGraw-Hill. https://accessmedicine.mhmedical.com/book.aspx?bookID=2316

Byrne, D. (2002). *Complexity theory and the social sciences: An introduction.* Taylor and Francis.

Campbell, D. T. (1978). Qualitative knowing in action research. In M. Brenner, P. Marsh, & M. Brenner (Eds.), *The social context of methods* (pp. 184–209). Croom Helm.

Campbell, D. T., & Stanley, J. C. (1966). *Experimental and quasi-experimental designs for research* (N. L. Gage, Ed.). R. McNally.

Canadian Evaluation Society. (n.d.). *Our history.* Canadian Evaluation Society. https://evaluationcanada.ca/our-history

Center for Collaborative Planning. (2000). *Sustainability benchmarks.* Center for Collaborative Planning. www.tomwolff.com/resources/backer.pdf

Center for the Advancement of Collaborative Strategies in Health. (2002). *Partnership self-assessment tool.* https://atrium.lib.uoguelph.ca/xmlui/bitstream/handle/10214/3129/Partnership_%20Self-Assessment_Tool-Questionnaire_complete.pdf?sequence=1&isAllowed=y

Centers for Disease Control and Prevention [CDC]. (n.d.). *10 essential public health services.* https://www.cdc.gov/publichealthgateway/publichealthservices/essentialhealthservices.html

Centers for Disease Control and Prevention [CDC]. (1999). Framework for program evaluation in public health. *Morbidity and Mortality Weekly Reports, 48*(11), No. RR-11.

Centers for Disease Control and Prevention [CDC]. (2001). *Updated guidelines for evaluating public health surveillance systems: Recommendations from the guidelines working group.* https://stacks.cdc.gov/view/cdc/13376

Centers for Disease Control and Prevention [CDC]. (2018a). *Evaluation brief no. 13: Data collection methods for program evaluation: Focus groups.* U.S. Department of Health and Human Services. https://www.cdc.gov/healthyyouth/evaluation/pdf/brief13.pdf

Centers for Disease Control and Prevention [CDC]. (2018b). *Evaluation brief no. 14: Data collection methods for program evaluation: Questionnaires.* U.S. Department of Health and Human Services. https://www.cdc.gov/healthyyouth/evaluation/pdf/brief14.pdf

Centers for Disease Control and Prevention [CDC]. (2018c). *Evaluation brief no. 16: Data collection methods for program evaluation: Observations.* U.S. Department of Health and Human Services. https://www.cdc.gov/healthyyouth/evaluation/pdf/brief16.pdf

Centers for Disease Control and Prevention [CDC]. (2018d). *Evaluation brief no. 17: Data collection methods for program evaluation: Interviews.* U.S. Department of Health and Human Services. https://www.cdc.gov/healthyyouth/evaluation/pdf/brief17.pdf

Centers for Disease Control and Prevention [CDC]. (2018e). *Evaluation brief no. 18: Data collection methods for program evaluation: Document review.* U.S. Department of Health and Human Services. https://www.cdc.gov/healthyyouth/evaluation/pdf/brief18.pdf

Centers for Disease Control and Prevention [CDC] & Council of State and Territorial Epidemiologists [CSTE]. (2008). *Applied epidemiology competencies.* Centers for Disease Control and Prevention / Council of State and Territorial Epidemiologists. http://www.cste2.org/webpdfs/AppliedEpiCompwcover.pdf

Centers for Disease Control and Prevention [CDC], National Asthma Control Program. (n.d.). *Finding the right people for your program evaluation team: Evaluator and planning team job descriptions.* Centers for Disease Control and Prevention. https://www.cdc.gov/asthma/pdfs/Finding_the_Right_People_for_Your_Program_Evaluation_Team.pdf

Chappelle, E. (2014). *Establishing a baseline as part of your evaluation.* https://www.cdc.gov/dhdsp/pubs/docs/cb_jan2014.pdf

Chen, H. (1990). *Theory-driven evaluations.* Sage Publications.

Chen, H. (2005). Program theory. In S. Mathison (Ed.), *Encyclopedia of evaluation* (pp. 341–342). SAGE Publications, Inc.

Chen, H. (2015). *Practical program evaluation: Theory-driven evaluation and the integrated evaluation perspective* (2nd ed.). SAGE Publications.

Chen, H., & Rossi, P. H. (Eds.). (1992). *Using theory to improve program and policy evaluations.* Greenwood Press. http://www.gbv.de/dms/hbz/toc/ht004498020.pdf

Christie, C. A., & Alkin, M. C. (2013). An Evaluation Theory Tree. In M. C. Alkin (Ed.), E*valuation roots: Tracing theorists' views and influences* (2nd ed., pp. 11–57). SAGE Publications.

Christie, C. A., Quiñones, P., & Fierro, L. (2014). Informing the discussion on evaluator training: A look at evaluators' course taking and professional practice. *American Journal of Evaluation, 35*(2), 274–290. https://doi.org/10.1177/1098214013503697

Ciliska, D., Thomas, H., & Buffet, C. (2008). *An introduction to evidence-informed public health and a compendium of critical appraisal tools for public health practice.* National Collaborating Centre for Methods and Tools. https://www.nccmt.ca/uploads/media/media/0001/01/b331668f85bc6357f262944f0aca38c14c89c5a4.pdf

Community Toolbox. (n.d.). *Diagnosing the health of your coalition.* https://ctb.ku.edu/en/table-of-contents/assessment/promotion-strategies/maintain-a-coalition/tools

Council on Education for Public Health. (2016). *CEPH accreditation criteria for schools of public health 7 public health programs, amended Oct 2016.*

Crano, W. D., Brewer, M. B., & Lac, A. (2015). *Principles and methods of social research* (3rd ed.). Routledge.

Creswell, J. W. (2006). *Qualitative inquiry and research design: Choosing among five approaches.* SAGE Publications, Inc.

Creswell, J. W. (2009). *Research design: Qualitative, quantitative, and mixed methods approaches (*3rd ed.). SAGE Publications.

Creswell, J. W., & Plano Clark, V. L. (2018). *Designing and conducting mixed methods research* (3rd ed.). SAGE Publications.

Cronbach, L. J. (1982). *Designing evaluations of educational and social programs* (1st ed.). Jossey-Bass.

Davidson, J. (2005a). Effectiveness. In S. Mathison (Ed.), *Encyclopedia of evaluation* (pp. 122–122). SAGE Publications, Inc.

Davidson, J. (2005b). Efficiency. In S. Mathison (Ed.), *Encyclopedia of evaluation* (pp. 122–123). SAGE Publications, Inc.

Davidson, J. (2005c). *Evaluation methodology basics: The nuts and bolts of sound evaluation.* Sage Publications.

DeSalvo K., Claire Wang Y., Harris A., Auerbach J., Koo D., & O'Carroll P. (2017). Public health 3.0: A call to action for public health to meet the challenges of the 21st century. *Preventing Chronic Disease, 14*(9). https://doi.org/10.5888/pcd14.170017

DeSalvo, K., O'Carroll, P. W., Koo, D., Auerbach, J. M., & Monroe, J. A. (2016). Public health 3.0: Time for an upgrade. *American Journal of Public Health, 106*(4), 621–622. https://doi.org/10.2105/AJPH.2016.303063

Doll, K. (2020). L*et's get strategic: An exploration of the how-to's, what-not-to-do's, and promising upshots of strategic evaluation initiatives* [Dissertation, Claremont Graduate University]. https://www.proquest.com/openview/ee37fdc703babeee28c781e25e851d45/1.pdf?pq-origsite=gscholar&cbl=18750&diss=y

Donaldson, S. I. (2007). *Program theory-driven evaluation science: Strategies and applications.* Psychology Press.

Donaldson, S. I., Gooler, L. E., & Scriven, M. (2002). Strategies for managing evaluation anxiety: Toward a psychology of program evaluation. *American Journal of Evaluation, 23*(3), 261–273. https://doi.org/10.1016/S1098-2140(02)00209-6

Donaldson, S. I., & Lipsey, M. W. (2006). Roles for theory in contemporary evaluation practice: Developing practical knowledge. In I. F. Shaw, J. C. Greene, & M. M. Mark (Eds.), *The SAGE handbook of evaluation* (pp. 57–75). SAGE Publications Ltd. https://methods.sagepub.com/book/the-sage-handbook-of-evaluation/n2.xml

Drummond, M. F., Sculpher, M. J., Claxton, K., & Torrance, G. W. (2015). *Methods for the economic evaluation of health care programmes* (4th ed.). Oxford University Press.

Drummond, M., Sculpher, M. J., Torrance, G. W., O'Brien, B. J., & Stoddard, G. L. (2005). *Methods for the economic evaluation of health care programmes* (3rd ed.). Oxford University Press.

Durland, M., & Fredericks, K. (2005). An introduction to social network analysis. In. M. M. Durland & K. A. Fredericks (Eds.), *Social Network Analysis in Program Evaluation. New Directions for Evaluation, 107*, 5-13. https://doi.org/10.1002/ev.157

Earl, S., Carden, F., & Smutylo, T. (2001). *Outcome mapping: Building learning and reflection into development programs.* International Development Research Centre. https://central.bac-lac.gc.ca/.item?id=0889369593&op=pdf&app=Library

Eastern Evaluation Research Society. (n.d.). *About EERS.* Eastern Evaluation Research Society. http://eers.org/about-eers/

Eisner, E. W. (2013). The roots of connoisseurship and criticism. In M. C. Alkin (Ed.), *Evaluation roots: A wider perspective of theorists' views and influences* (pp. e1–e7). SAGE Publications. https://studysites.sagepub.com/alkin2e/

Epstein, I. (2009). Promoting harmony where there is commonly conflict: Evidence-informed practice as an integrative strategy. *Social Work in Health Care, 48*(3), 216–231. https://doi.org/10.1080/00981380802589845

Evergreen, S. D. H. (2020). *Effective data visualization: The right chart for the right data* (2nd ed.). SAGE Publications, Inc.

Fawcett, S., Foster, D., & Francisco, V. (1997). Monitoring and evaluation of coalition activities and success. In G. Kaye & T. Wolff (Eds.), *From the ground up: A workbook on coalition building and community development* (pp. 163–185). AHEC/Community Partners.

Fetterman, D. (2017). *Empowerment evaluation.* BetterEvaluation. https://www.betterevaluation.org/en/plan/approach/empowerment_evaluation

Fetterman, D. M. (2013). Empowerment evaluation: Learning to think like an evaluator. In M. C. Alkin (Ed.), *Evaluation roots: A wider perspective of theorists' views and influences* (pp. 304–322). SAGE Publications, Inc.

Fierro, L. A., & Christie, C. A. (2011). Understanding evaluation training in schools and programs of public health. *American Journal of Evaluation, 32*(3), 448–468. https://doi.org/10.1177/1098214010393721

Foster, A., King, L. R., & Bender, K. (2018). Are public health academia, professional certification, and public health practice on the same page? *Journal of Public Health Management and Practice, 24*(Suppl 3), S47–S50. https://doi.org/10.1097/PHH.0000000000000746

Fournier, D. (2005). Evaluation. In S. Mathison (Ed.), *Encyclopedia of evaluation* (pp. 140–140). SAGE Publications, Inc.

Frierson, H. T., Hood, S., Hughes, G. B., & Thomas, V. G. (2010). Chapter 7: A guide to conducting culturally responsive evaluations. In J. Frechtling (Ed.), *The 2010 user-friendly handbook for project evaluation (Revision to NSF Publication No. 02-057)* (pp. 75–96). National Science Foundation.

Garcia, G. L., & Stevahn, L. (2020). Situational awareness and interpersonal competence as evaluator competencies. *American Journal of Evaluation, 41*(1), 107–124. https://doi.org/10.1177/1098214018814941

Ghere, G., King, J. A., Stevahn, L., & Minnema, J. (2006). A professional development unit for reflecting on program evaluator competencies. *American Journal of Evaluation, 27*(1), 108–123. https://doi.org/10.1177/1098214005284974

Glanz, K., Rimer, B. K., & Viswanath, K. (2014). *Health behavior theory, research, and practice.* (5th ed.). Wiley.

Gold, M. R., Siegel, J. E., Russell, L. B., & Weinstein, M. C. (1996). *Cost-effectiveness in health and medicine.* Oxford University Press, Incorporated.

Granner, M. L., & Sharpe, P. A. (2004). Evaluating community coalition characteristics and functioning: A summary of measurement tools. *Health Education Research, 19*(5), 514–532. https://doi.org/10.1093/her/cyg056

Greene, J. C. (2007). *Mixed methods in social inquiry* (1st ed.). Jossey-Bass.

Greene, J. C. (2013). Making the world a better place through evaluation. In M. C. Alkin (Ed.), *Evaluation roots: A wider perspective of theorists' views and influences* (pp. 208–217). SAGE Publications.

Greene, J. C., Boyce, A. S., & Ahn, J. (2011). *A values-engaged, educative approach for evaluating education programs: A guidebook for practice.* University of Illinois, Urbana-Champaign. http://comm.eval.org/communities/community-home/librarydocuments/viewdocument?DocumentKey=f3c734c0-8166-4ba4-9808-a07e05294583

Haddix, A. C., Teutsch, S. M., & Corso, P. S. (2005). *Prevention effectiveness: A guide to decision analysis and economic evaluation* (2nd ed.). Oxford University Press. http://books.google.com/books?id=7BcQAQAAMAAJ

Haddix, A. C., Teutsch, S. M., Shaffer, P. A., & Dunet, D. O. (1996). *Prevention effectiveness: A guide to decision analysis and economic evaluation.* Oxford University Press.

Hall, M. E. (2020). Blest be the tie that binds. In L. C. Neubauer, D. McBride, A. D. Guajardo, W. D. Casillas, & M. E. Hall (Eds.), *Examining Issues Facing Communities of Color Today: The Role of Evaluation to Incite Change. New Directions for Evaluation, 166*, 14-22.

Hardy, B., Hudson, B., & Waddington, E. (2003). *Assessing strategic partnership: The partnership assessment tool.* Office of the Deputy Prime Minister, Strategic Partnering Taskforce. https://www.conservationgateway.org/ConservationPlanning/partnering/cpc/Documents/AssessingStrategicPartnership.pdf

Hargreaves, M. B., Verbitsky-Savitz, N., Coffee-Borden, B., Perreras, L., Pecora, P. J., Roller White, C., Morgan, G. B., Barila, T., Ervin, A., Case, L., Hunter, R., & Adams, K. (2016). *Advancing the measurement of collective community capacity to address adverse childhood experiences and resilience.* Community Science. http://www.appi-wa.org/wp-content/uploads/2016/07/APPI-White-Paper.pdf

Hilgendorf, A. E., Moore, T. R., Wells, A., & Stanley, J. (2020). Positioning health equity within a systems thinking framework to evaluate coalitions and collaborative initiatives. In A. W. Price, K. K. Brown, & S. M. Wolfe (Eds.), *Evaluating Community Coalitions and Collaboratives. New Directions for Evaluation, 165,* 91–102. https://doi.org/10.1002/ev.20403

Hobson, K. A., Coryn, C. L. S., Fierro, L. A., & Sherwood-Laughlin, C. M. (2019). Instruction of evaluation competencies in Council on Education for Public Health (CEPH)-accredited Master of Public Health (MPH) degree programs. *American Journal of Evaluation, 40*(4), 590–606. https://doi.org/10.1177/1098214019845510

Honeycutt, T. (2009). *Making connections: Using social network analysis for program evaluation.* Mathematica Policy Research. https://www.mathematica.org/our-publications-and-findings/publications/making-connections-using-social-network-analysis-for-program-evaluation

Hood, S., Hopson, R. K., & Frierson, H. T. (2015). *Continuing the journey to reposition culture and cultural context in evaluation theory and practice.* Information Age Publishing. http://site.ebrary.com/id/11011898

Hood, S., Hopson, R. K., & Kirkhart, K. E. (2015). Culturally responsive evaluation. In K. E. Newcomer, H. P. Hatry, & J. S. Wholey (Eds.), *Handbook of practical program evaluation* (4th ed., pp. 281–317). John Wiley & Sons, Ltd. https://doi.org/10.1002/9781119171386.ch12

House, E. R. (1993). *Professional evaluation: Social impact and political consequences.* SAGE Publications.

House, E. R. (2005). Deliberative democratic evaluation. In S. Mathison (Ed.), *Encyclopedia of evaluation* (pp. 105–108). SAGE Publications, Inc.

House, E. R. (2013). Work memoir—Ideas and influences. In M. C. Alkin (Ed.), *Evaluation roots: A wider perspective of theorists' views and influences* (pp. 198–207). SAGE Publications.

House, E. R., & Howe, K. R. (1999). *Values in evaluation and social research.* SAGE Publications. http://site.ebrary.com/id/10586606

House, E. R., & Howe, K. R. (2000). *Deliberative democratic evaluation checklist.* Western Michigan University Checklists. https://wmich.edu/sites/default/files/attachments/u58/2014/deliberative-democratic-evalution-checklist.pdf

House, E. R., & Howe, K. R. (2000). Deliberative democratic evaluation. In K. E. Ryan & L. DeStefano (Eds.), *Evaluation as a Democratic Process: Promoting Inclusion, Dialogue, and Deliberation. New Directions for Evaluation, 85,* 3–12. https://doi.org/10.1002/ev.1157

Hutchinson, K. (2017). *A short primer on innovative evaluation reporting* (1st ed.). Kylie S. Hutchinson.

International Organization for Cooperation in Evaluation. (n.d.). *VOPE directory.* https://ioce-vopes.wildapricot.org/directory

International Development Research Centre. (2004). *Selecting and managing an evaluation consultant or team.* International Development Research Centre. web.https://idl-bnc-idrc.dspacedirect.org/bitstream/handle/10625/47282/133628.pdf?sequence=1&isAllowed=y

Iowa State University. (2000). *Purposeful partnerships in the community interest.* Iowa State University, University Extension. https://store.extension.iastate.edu/Product/pm1844-pdf

Issel, L. M. (2009). *Health program planning and evaluation: A practical, systematic approach for community health* (2nd ed.). Jones & Bartlett Learning, LLC. https://books.google.com/books?id=yF0DhTNAxa8C

Johnson, B., & Christensen, L. B. (2008). *Educational research: Quantitative, qualitative, and mixed approaches* (3rd ed.). SAGE Publications, Inc.

Johnson, K., Greenseid, L. O., Toal, S. A., King, J. A., Lawrenz, F., & Volkov, B. (2009). Research on evaluation use: A review of the empirical literature from 1986 to 2005. *American Journal of Evaluation, 30*(3), 377–410.

Kannampallil, T. G., Schauer, G. F., Cohen, T., & Patel, V. L. (2011). Considering complexity in healthcare systems. *Journal of Biomedical Informatics, 44*(6), 943–947. https://doi.org/10.1016/j.jbi.2011.06.006

Kaye, G. (1993). Diagnosing your coalition: *Risk factors for participation.* Community Development Associates. https://www.tomwolff.com/resources/backer.pdf

Kaye, G., & Resnick, I. (1994). *Climate diagnostic tool.* Community Development Consultants. http://www.tom-wolff.com/resources/backer.pdf

Kinarsky, A. R. (2018). The evaluation landscape: U.S. foundation spending on evaluation. In S. B. Nielsen, S. Lemire, & C. A. Christie (Eds.), *The Evaluation Marketplace: Exploring the Evaluation Industry. New Directions for Evaluation, 160*, 81–96. https://doi.org/10.1002/ev.20342

King, J. A., & Stevahn, L. (2013). *Interactive evaluation practice: Mastering the interpersonal dynamics of program evaluation.* SAGE Publications, Inc. https://doi.org/10.4135/9781452269979

King, J. A., & Stevahn, L. (2020). Presenting the 2018 AEA evaluator competencies. In J. A. King (Ed.), *The American Evaluation Association's Program Evaluator Competencies. New Directions for Evaluation, 168*, 49–61. https://doi.org/10.1002/ev.20435

Kumah, E. A., McSherry, R., Bettany-Saltikov, J., Hamilton, S., Hogg, J., Whittaker, V., & Schaik, P. (2019). Evidence-informed practice versus evidence-based practice educational interventions for improving knowledge, attitudes, understanding, and behavior toward the application of evidence into practice: A comprehensive systematic review of undergraduate students. *Campbell Systematic Reviews, 15*(1–2). https://doi.org/10.1002/cl2.1015

Kusek, J., & Rist, R. (2005). Performance-based monitoring. In S. Mathison (Ed.), *Encyclopedia of evaluation* (pp. 301–306). SAGE Publications, Inc.

LaVelle, J. M. (2020). Educating evaluators 1976–2017: An expanded analysis of university-based evaluation education programs. *American Journal of Evaluation, 41*(4), 494–509. https://doi.org/10.1177/1098214019860914

Lemire, S., Fierro, L. A., Kinarsky, A. R., Fujita-Conrads, E., & Christie, C. A. (2018). The U.S. federal evaluation market. In S. B. Nielsen, S. Lemire, & C. A. Christie (Eds.), *The Evaluation Marketplace: Exploring the Evaluation Industry. New Directions for Evaluation, 160*, 63–80. https://doi.org/10.1002/ev.20343

Levin, H. M., & McEwan, P. J. (2001). *Cost-effectiveness analysis: Methods and applications* (2nd ed.). SAGE Publications.

Leviton, L. C., & Gutman, M. A. (2010). Overview and rationale for the systematic screening and assessment method. In L. C. Leviton, L. K. Khan, & N. Dawkins (Eds.), *The Systematic Screening and Assessment Method: Finding Innovations Worth Evaluating. New Directions for Evaluation, 125*, 7–31. https://doi.org/10.1002/ev.318

Lopez, A. A. (2018). *Leadership in evaluation capacity building: Influencing organizational change and learning about evaluation* [Dissertation]. Claremont Graduate University.

Mark, M. M., Henry, G. T., & Julnes, G. (2000). *Evaluation: An integrated framework for understanding, guiding, and improving policies and programs*. Jossy-Bass. https://www.wiley.com/en-us/Evaluation%3A+An+Integrated+Framework+for+Understanding%2C+Guiding%2C+and+Improving+Policies+and+Programs-p-9780787948023

Mathison, S. (2005a). Evaluand. In S. Mathison (Ed.), *Encyclopedia of evaluation* (pp. 139–140). SAGE Publications, Inc.

Mathison, S. (2005b). Evaluation theory. In S. Mathison (Ed.), *Encyclopedia of evaluation* (pp. 143–143). SAGE Publications.

Mathison, S. (2005c). Modus operandi. In S. Mathison (Ed.), *Encyclopedia of evaluation* (pp. 258–258). SAGE Publications, Inc.

Mattessich, P. W., & Johnson, K. (2018). *Collaboration: What makes it work.* Amherst H. Wilder Foundation, Wilder Research Center. https://www.wilder.org/wilder-research/resources-and-tools

McDavid, J. C., Huse, I., & Hawthorn, L. R. L. (2019). *Program evaluation and performance measurement: An introduction to practice* (3rd ed.). SAGE Publications, Inc. https://www.wilder.org/wilder-research/resources-and-tools

Mertens, D. M. (2009). *Transformative research and evaluation.* Guilford Press.

Mertens, D. M., & Wilson, A. T. (2019). *Program evaluation theory and practice: A comprehensive guide* (2nd ed.). Guilford Press.

Mezirow, J. (1990). *Fostering critical reflection in adulthood: A guide to transformative and emancipatory learning* (1st ed.). Jossey-Bass Publishers. https://www.wilder.org/wilder-research/resources-and-tools

Mezirow, J. (2009). An overview of transformative learning. In K. Illeris (Ed.), *Contemporary theories of learning* (pp. 91–105). Routledge.

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. https://doi.org/10.1177/109821400302400303

National Board of Public Health Examiners [NBPHE]. (n.d.). *CPH content outline 2019.* National Board of Public Health Examiners. https://s3.amazonaws.com/nbphe-wp-production/app/uploads/2017/05/ContentOutline-May-21-2019.pdf

National Collaborating Centre for Methods and Tools [NCCMT]. (2009). *A model for evidence-informed decision making in public health.* The National Collaborating Centre for Methods and Tools. https://www.nccmt.ca/uploads/media/media/0001/01/0285efc9fa08b3b0fab06b7c940c2f6767087cfc.pdf

National Collaborating Centre for Methods and Tools [NCCMT]. (n.d.). E*vidence-informed public health.* https://www.nccmt.ca/tools/eiph

National Commission for Health Education Credentialing [NCHEC]. (2020). *Areas of responsibility, competencies and sub-competencies for health education specialist practice analysis II 2020 (HESPA II 2020).* https://assets.speakcdn.com/assets/2251/hespa_competencies_and_sub-competencies_052020.pdf

Neubauer, L. C., McBride, D., Guajardo, A. D., Casillas, W. D., & Hall, M. E. (2020). Examining issues facing communities of color today: The role of evaluation to incite change. In L. C. Neubauer, D. McBride, A. D. Guajardo, W. D. Casillas, & M. E. Hall (Eds.), *Examining Issues Facing Communities of Color Today: The Role of Evaluation to Incite Change. New Directions for Evaluation, 166*, 7–11. https://doi.org/10.1002/ev.20406

Nicholls, J., Lawlor, E., Neitzert, E., & Goodspeed, T. (2012). *A guide to social return on investment.* https://socialvalueuk.org/wp-content/uploads/2016/03/The%20Guide%20to%20Social%20Return%20on%20Investment%202015.pdf

Office of Evaluation Sciences [OES]. (n.d.). *Evidence act toolkits.* U.S. General Services Administration. https://oes.gsa.gov/toolkits/

Office of Management and Budget [OMB]. (2019). *Memorandum for heads of executive departments and agencies—Phase 1 implementation of the Foundations for Evidence-Based Policymaking Act of 2018: Learning agendas, personnel, and planning guidance.* Office of Management and Budget. https://www.whitehouse.gov/wp-content/uploads/2019/07/M-19-23.pdf

Office of Management and Budget [OMB]. (2020). *Memorandum for heads of executive departments and agencies—Phase 4 implementation of the Foundations for Evidence-Based Policymaking Act of 2018: Program evaluation standards and practices.* Office of Management and Budget. https://www.whitehouse.gov/wp-content/uploads/2020/03/M-20-12.pdf

Patton, M. Q. (1997). *Utilization-focused evaluation: The new century text* (3rd ed.). Sage Publications.

Patton, M. Q. (2002). *Qualitative research and evaluation methods* (3rd ed.). Sage Publications.

Patton, M. Q. (2005). Situational responsiveness. In S. Mathison (Ed.), *Encyclopedia of evaluation* (pp. 390–391). SAGE Publications, Inc.

Patton, M. Q. (2008). *Utilization-focused evaluation* (4th ed.). Sage Publications.

Patton, M. Q. (2011). *Developmental evaluation: Applying complexity concepts to enhance innovation and use.* Guilford Press.

Patton, M. Q. (2013). *Utilization-focused evaluation checklist.* Western Michigan University Checklists. https://wmich.edu/sites/default/files/attachments/u350/2018/ufe-patton.pdf

Patton, M. Q. (2014). *Evaluation flash cards: Embedding evaluative thinking in organizational culture.* Otto Bremer Trust. ottobremer.org

Patton, M. Q. (2018). *Principles-focused evaluation: The guide.* The Guilford Press.

Patton, M. Q., Grimes, P. S., Guthrie, K. M., Brennan, N. J., French, B. D., & Blyth, D. A. (1977). In search of impact: An analysis of the utilization of federal health evaluation research. In C. H. Weiss (Ed.), *Using social research in public policy making* (pp. 141–163). D.C. Heath.

Patton, M. Q., & Horton, D. (2009). *Utilization-focused evaluation for agricultural innovation* (ILAC Brief 22). Institutional Learning and Change Initiative. https://hdl.handle.net/10568/70056

Preskill, H., & Boyle, S. (2008). A multidisciplinary model of evaluation capacity building. *American Journal of Evaluation, 29*(4), 443–459.

Price, A. W., Brown, K. K., & Wolfe, S. M. (2020). Editors' notes. In S. M. Wolfe, A. W. Price, & K. K. Brown (Eds.), *Evaluating Community Coalitions and Collaboratives. New Directions for Evaluation, 165*, 9–16. https://doi.org/10.1002/ev.20400

Public Health Accreditation Board [PHAB]. (2014). *PHAB standards and measures, version 1.5 adopted Dec 2013.* Public Health Accreditation Board. https://phaboard.org/wp-content/uploads/PHABSM_WEB_LR1-1.pdf

Public Health National Center for Innovations [PHNCI]. (2020). *The 10 essential public health services.* https://phnci.org/national-frameworks/10-ephs

Reeves, M. J., Lyon-Callo, S., Brown, M. D., Rosenman, K., Wasilevich, E., & Williams, S. G. (2006). Using billing data to describe patterns in asthma-related emergency department visits in children. *Pediatrics, 117*(4), S106–S117.

Rogers, P. (2005). Conceptual use. In S. Mathison (Ed.), *Encyclopedia of evaluation* (pp. 74–74). SAGE Publications, Inc.

Rosenthal, B. (1997). Multicultural issues in coalitions. In G. Kaye & T. Wolff (Eds.), *From the ground up: A workbook on coalition building and community development* (pp. 51–73). AHEC/Community Partners.

Rossi, P. H., & Freeman, H. E. (1993). Evaluation: *A systematic approach* (5th ed.). Sage Publications. http://books.google.com/books?id=NPxGAAAAMAAJ

Rowitz, L. (2001). *Public health leadership: Putting principles into practice.* Aspen Publishers.

Russ-Eft, D. F., & Preskill, H. S. (2009). *Evaluation in organizations: A systematic approach to enhancing learning, performance, and change* (2nd ed.). Basic Books.

Ryan, P. D. (2019, January 14). *H.R.4174 - 115th Congress (2017-2018): Foundations for Evidence-Based Policymaking Act of 2018* (2017/2018) [Webpage]. https://www.congress.gov/bill/115th-congress/house-bill/4174

Rychetnik, L., Hawe, P., Waters, E., Barratt, A., & Frommer, M. (2004). A glossary for evidence based public health. *Journal of Epidemiology and Community Health, 57*(8), 538–545. https://doi.org/10.1136/jech.2003.011585

Salabarría-Peña, Y., Apt, B. S., & Walsh, C. M. (2007). *Practical use of program evaluation among sexually transmitted disease (STD) programs.* Centers for Disease Control and Prevention.

Schulz, A. J., Israel, B. A., & Lantz, P. (2003). Instrument for evaluating dimensions of group dynamics within community-based participatory research partnerships. *Evaluation and Program Planning, 26*(3), 249–262.

Scriven, M. (1976). *Maximizing the power of causal investigations: The modus operandi method* (G. V. Glass, Ed.; Vol. 1). SAGE Publications, Inc.

Scriven, M. (1981). *The logic of evaluation.* Edgepress.

Scriven, M. (1986). New frontiers of evaluation. *Evaluation Practice, 7*(1), 7–44.

Scriven, M. (1991). *Evaluation thesaurus* (4th ed.). Sage Publications.

Scriven, M. (2013). Conceptual revolutions in evaluation: Past, present, and future. In M. C. Alkin (Ed.), *Evaluation roots: A wider perspective of theorists' views and influences* (pp. 167–179). SAGE Publications.

Scriven, M. (2015). *Key evaluation checklist.* Western Michigan University Checklists. http://www.michaelscriven.info/images/MS_KEC_8-15-15.doc

Secretary's Advisory Committee on National Health Promotion and Disease Prevention Objectives for 2020. (2010). *Evidence-based clinical and public health: Generating and applying the evidence.* U.S. Department of Health and Human Services. https://www.healthypeople.gov/sites/default/files/EvidenceBasedClinicalPH2010.pdf

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference.* Wadsworth Cengage Learning.

Shadish, W. R., Cook, T. D., & Leviton, L. C. (1991). *Foundations of program evaluation: Theories of practice.* Sage Publications. http://www.gbv.de/dms/hbz/toc/ht003873485.PDF

Smith, N. L. (1993). Improving evaluation theory through the empirical study of evaluation practice. *Evaluation Practice, 14*(3), 237–242. https://doi.org/10.1016/0886-1633(93)90051-P

Sparks Policy Institute & ORS Impact. (2018). *When collective impact has an impact: A cross-site study of 25 collective impact initiatives.* https://www.orsimpact.com/DirectoryAttachments/10262018_111513_477_CI_Study_Report_10-26-2018.pdf

Stachowiak, S., Lynn, J., & Akey, T. (2020). Finding the impact: Methods for assessing the contribution of collective impact to systems and population change in a multi-site study. In S. M. Wolfe, A. W. Price, & K. K. Brown (Eds.), *Evaluating Community Coalitions and Collaboratives. New Directions for Evaluation, 165*, 29–44. https://doi.org/10.1002/ev.20398

Stake, R. E. (2004). *Standards-based & responsive evaluation.* Sage.

Steckler, A. B., & Linnan, Laura. (2002). P*rocess evaluation for public health interventions and research.* J. Wiley.

Stevahn, L., Berger, D. E., Tucker, S. A., & Rodell, A. (2020). Using the 2018 AEA evaluator competencies for effective program evaluation practice. In J. A. King (Ed.), *The American Evaluation Association's Program Evaluator Competencies. New Directions for Evaluation, 168,* 75–97. https://doi.org/10.1002/ev.20434

Stockdill, S. H., Baizerman, M., & Compton, D. W. (2002). Toward a definition of the ECB process: A conversation with the ECB literature. In M. Baizerman, D. W. Compton, & S. H. Stockdill (Eds.), *The Art, Craft, and Science of Evaluation Capacity Building. New Directions for Evaluation, 93,* 7–26. https://doi.org/10.1002/ev.39

Symonette, H., Miller, R. L., & Barela, E. (2020). Power, privilege, and competence: Using the 2018 AEA Evaluator Competencies to shape socially just evaluation practice. In J. A. King (Ed.), *The American Evaluation Association's Program Evaluator Competencies. New Directions for Evaluation, 168,* 117-132. https://doi.org/10.1002/ev.20435

Tashakkori, A., & Creswell, J. W. (2007). Editorial: The new era of mixed methods. *Journal of Mixed Methods Research, 1*(1), 3-7.

Tashakkori, A., & Teddlie, C. (Eds.). (2010). *Sage handbook of mixed methods in social and behavioral research* (2nd ed.). SAGE Publications. http://srmo.sagepub.com/view/sage-handbook-of-mixed-methods-social-behavioral-research-2e/SAGE.xml

Taylor-Powell, E. (n.d.). *Community group member survey: Using the results* (Publication no. G3658-9). University of Wisconsin Extension. https://cdn.shopify.com/s/files/1/0145/8808/4272/files/G3658-09.pdf

Teddlie, C., & Tashakkori, A. (2009). *Foundations of mixed methods research: Integrating quantitative and qualitative approaches in the social and behavioral sciences.* SAGE.

The White House. (2021). *Memorandum on restoring trust in government through scientific integrity and evidence-based policymaking.* The White House.

Torres, R. T., Preskill, H. S., & Piontek, M. E. (2005). *Evaluation strategies for communicating and reporting: Enhancing learning in organizations (*2nd ed.). Sage Publications.

Trochim, W. (2020). Internal validity. In *Research methods knowledge base.* https://socialresearchmethods.net/kb/internal-validity/

Trochim, W. M. K. (2020). *Research methods knowledge base.* Conjoint.ly. https://conjointly.com/kb/

United States Agency for International Development [USAID]. (2017). *Landscape analysis of learning agendas: USAID/Washington and beyond.* https://usaidlearninglab.org/sites/default/files/resource/files/landscape_analysis_report_04_10_17.pdf

U.S. Department of Health and Human Services [DHHS]. (2014). P*ractical strategies for culturally competent evaluation.* Centers for Disease Control and Prevention. https://www.cdc.gov/dhdsp/docs/Cultural_Competence_Guide.pdf

U.S. Department of Health and Human Services [DHHS], Centers for Disease Control and Prevention. (2003). *A public health action plan to prevent heart disease and stroke.* U.S. Department of Health and Human Services.

U.S. Department of Health and Human Services [DHHS], Centers for Disease Control and Prevention, Office of the Director, Office of Strategy and Innovation. (2005). I*ntroduction to program evaluation for public health programs: A self-study guide.* Centers for Disease Control and Prevention.

U.S. Environmental Protection Agency [EPA]. (2007). *Program evaluation glossary.* https://ofmpub.epa.gov/sor_internet/registry/termreg/searchandretrieve/glossariesandkeywordlists/search.do?details=&glossaryName=Program%20Evaluation%20Glossary

U.S. Government Accountability Office [GAO]. (2005). *Performance measurement and evaluation: Definitions and relationships* (GAO-05-739SP). U.S. Government Accountability Office. https://www.gao.gov/assets/210/202674.pdf

U.S. Government Accountability Office [GAO], Program Evaluation and Methodology Division. (2012). *Designing evaluations.* (GAO-12-208G). Retrieved from https://www.gao.gov/products/gao-12-208g

Varda, D. M. & Sprong, S. (2020). Evaluating networks using PARTNER: A social network framework, survey, and analysis tool. In S. M. Wolfe, A. M. Price, & K. K. Brown (Eds.), *Evaluating Community Coalitions and Collaboratives. New Directions for Evaluation, 165,* 67-89. https://doi.org/10.1002/ev.20397

Visible Network Labs. (n.d.). *PARTNER: The platform to analyze, record, and track networks to enhance relationships.* https://visiblenetworklabs.com/partner-platform/

W. K. Kellogg Foundation. (1998). *Evaluation handbook.* W. K. Kellogg Foundation. https://www.wkkf.org/~/media/62EF77BD5792454B807085B1AD044FE7.ashx

Waapalaneexkweew (Nicole Bowman, Mohican/Lunaape), & Dodge-Francis, C. (2018). Culturally responsive indigenous evaluation and tribal governments: Understanding the relationship. In F. Cram, K. A. Tibbetts, & J. LaFrance (Eds.), *Indigenous Evaluation. New Directions for Evaluation, 159*, 17–31. https://doi.org/10.1002/ev.20329

Weiss, C. H. (1998). *Evaluation: Methods for studying programs and policies* (2nd ed.). Prentice Hall.

Weiss, C. H. (2000). Theory-based evaluation: Theories of change for poverty reduction programs. In O. N. Feinstein & R. Picciotto (Eds.), *Evaluation and poverty reduction: World Bank series on evaluation and development* (Vol. 3, pp. 103–112). Routledge. https://doi.org/10.4324/9781351325325-16

Weiss, C. H. (2000). Which links in which theories shall we evaluate?. In A. Petrosino, Rogers, P. J., Huebner, T. A., & Hacsi, T. A. (Eds.), Program Theory in Evaluation: Challenges and Opportunities. *New Directions for Evaluation, 87*, 35-45. https://doi.org/10.1002/ev.1180

Wholey, J. S. (2013). Using evaluation to improve program performance and results. In M. C. Alkin (Ed.), *Evaluation roots: A wider perspective of theorists' views and influences* (pp. 261–266). SAGE Publications.

Wholey, J. S. (2015). Exploratory evaluation. In K. E. Newcomer, H. P. Hatry, & J. S. Wholey (Eds.), *Handbook of practical program evaluation (*4th ed., pp. 88–107). John Wiley & Sons, Ltd. https://doi.org/10.1002/9781119171386.ch4

Willis, G. B. (2005). *Cognitive interviewing: A tool for improving questionnaire design.* Sage Publications. https://dx.doi.org/10.4135/9781412983655

Wilson-Grau, R. (n.d.). Outcome harvesting. BetterEvaluation. http://betterevaluation.org/plan/approach/outcome_harvesting

Wilson-Grau, R., & Britt, H. (2013). *Outcome harvesting.* Ford Foundation.

Wolff, T. (2000). A practical approach to evaluating coalitions. In T. Backer (Ed.), *Evaluating community collaborations.* Springer Publishing. https://www.tomwolff.com/resources/backer.pdf

Worthen, B. R., Sanders, J. R., & Fitzpatrick, J. L. (1997). *Program evaluation: Alternative approaches and practical guidelines* (2nd ed.). Longman.

Yarbrough, D. B., Shulha, L. M., Hopson, R. K., & Caruthers, F. A. (2010). *The program evaluation standards: A guide for evaluators and evaluation users* (3rd ed.). SAGE.

Yin, R. K. (2018). *Case study research and applications: Design and methods* (6th ed.). SAGE Publications, Inc.