THE MASSACHUSETTS DEPARTMENT OF EDUCATION

Making Knowledge Productive

Skills and Tools for Evaluating Community Service-Learning Programs in Massachusetts

Prepared for:
The Massachusetts Department of Education
Community Service-Learning Program

Prepared by:
Center for Youth and Communities
Heller School for Social Policy and Management
Brandeis University

PLACEHOLDER ONLY

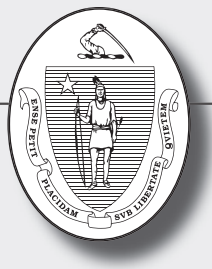
—

THIS PAGE DOES NOT PRINT

Making Knowledge Productive

Skills and Tools for Evaluating Community Service-Learning Programs in Massachusetts

PRINTED ON RECYCLED CONTENT PAPER



Prepared for:
The Massachusetts Department of Education
Community Service-Learning Program

Prepared by:
Center for Youth and Communities
Heller School for Social Policy and Management
Brandeis University

OCTOBER 2005

This publication was prepared by Brandeis University for the Massachusetts Department of Education and is funded by and based upon work supported by the Corporation for National and Community Service under Learn and Serve America Grant No. 03KSAMA001. Opinions or points of view expressed in this document are those of the authors and do not necessarily reflect the official position of the Corporation or the Learn and Serve America Program. The publication is intended to be used in connection with the advancement of service-learning as educational methodology. The Massachusetts Department of Education authorizes the use of this publication for non-commercial educational purposes only. Any publication or distribution of this document must acknowledge that it was produced by the Massachusetts Department of Education, Brandeis University and the Corporation for National and Community Service. It cannot be sold or shown for profit.

The Massachusetts Department of Education Community Service-Learning Program

Kristen McKinnon, Community Service-Learning Specialist
Massachusetts Department of Education
350 Main Street • Malden, MA 02148 • 781-338-6306
kmckinnon@doe.mass.edu

About the Center for Youth and Communities

Since its inception in 1983, the ***Center for Youth and Communities (CYC)*** has established a national reputation as one of the nation's leading research centers, and professional development and policy organizations in ***youth and community development***. CYC is part of the Heller School for Social Policy and Management at Brandeis University. CYC's ultimate goal is to "***make knowledge productive***." We do this by connecting the knowledge gained from ***scholarly research and practical experience*** in ways that help both policy makers and practitioners. This blend of theory and practice provides CYC with a unique perspective and capacity. In our work we view practitioners and policy makers as partners in a practical knowledge development effort in which both the community and the academy bring critical strengths, and in which practical solutions to real-world issues are developed through a collaborative, mutually respectful approach. Center staff have been conducting nationally recognized research on service-learning programs for over a decade.

Center for Youth and Communities
Heller School for Social Policy and Management, Brandeis University
60 Turner Street • Waltham, MA 02453
Phone: 781-736-3770 • Fax: 781-736-3773
cyc@brandeis.edu
http://heller.brandeis.edu/cyc/

PRINCIPAL AUTHORS
Thomas Piñeros-Shields
Lawrence Neil Bailis
Alan Melchior
Faye Cohen

Tool #1

Overview of *MKP*

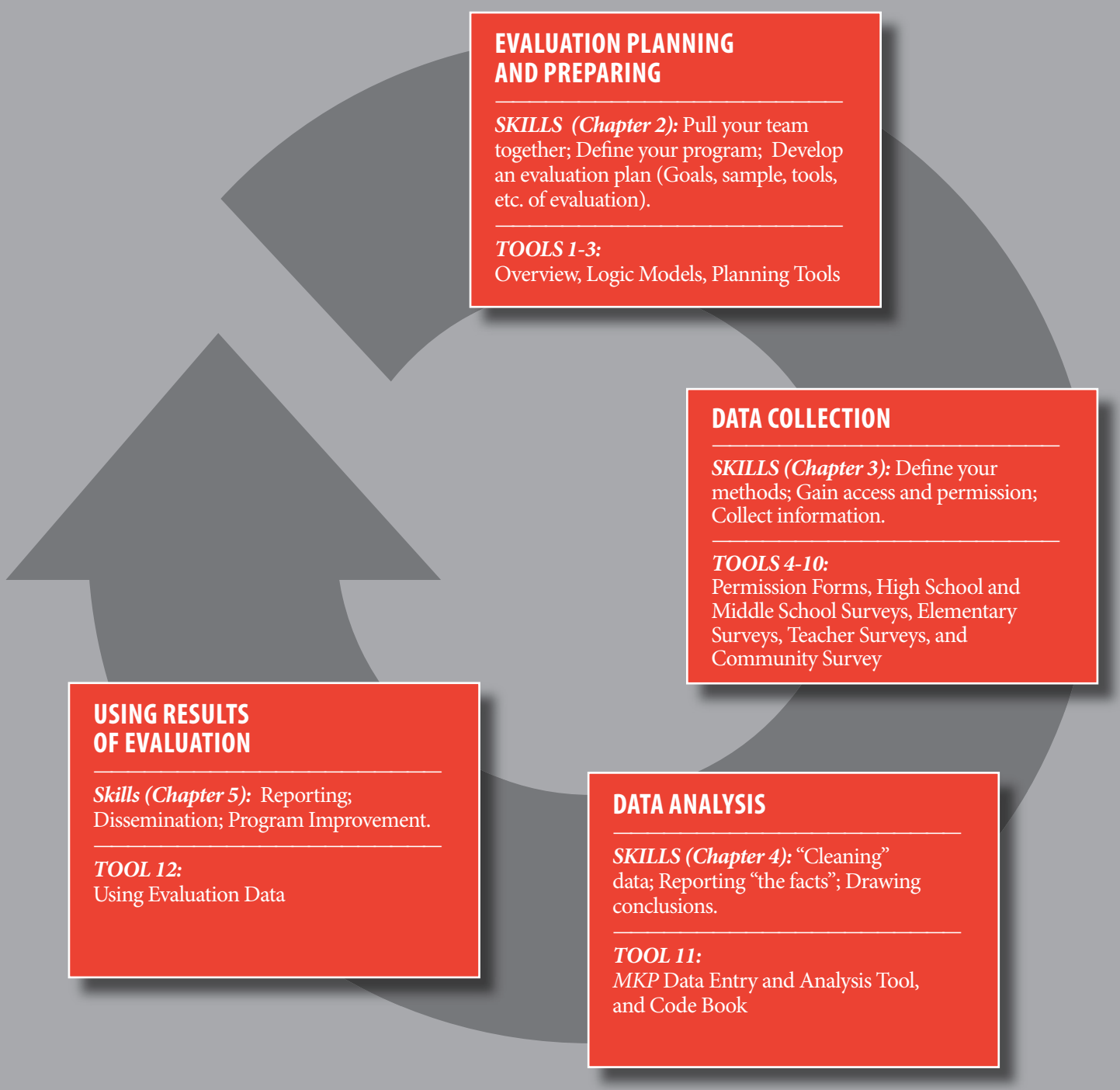


Table of Contents

Introduction

Making Knowledge Productive (MKP) is a guide to service-learning program evaluation for Massachusetts Community Service-Learning (CSL) administrators, coordinators, teachers, and students. It is the product of a collaborative effort between the Massachusetts Department of Education, The Center for Youth and Communities at Brandeis University and a working group of six community service-learning practitioners who met in spring 2004. The *MKP* working group concluded that program evaluation can help build support for programs, respond to funders' needs, inform program improvement, and convince skeptics that service-learning programs do have an impact on students. In short, they saw program evaluations as providing knowledge that can be useful – and they recognized that knowledge can be powerful.

These tools have been developed to meet the needs of teachers, community service-learning coordinators, the Massachusetts Department of Education Community Service-Learning program, and others...

At the same time, the working group recounted many challenges with the program evaluations that they have experienced. They described how program evaluations can be time consuming, expensive, complicated, and pose risks to their programs. Some members of the working group also noted that they did not know how to conduct program evaluations. They had come to understand that program evaluations often required a special set of skills different from the skills required of those who "do" the service-learning work.

We have prepared *Making Knowledge Productive* to meet the needs of and respond to the challenges that community service-learning practitioners face when evaluating their programs. The emphasis of *MKP* is on using evaluation as a way to improve the quality and accountability of service-learning across the state. We hope it will help (a) the state assess CSL throughout the Commonwealth, and (b) local program administrators, coordinators, teachers, and students to assess the CSL activities that they are involved in on a more regular and more effective basis. We hope *MKP* will involve community service-learning coordinators, teachers and administrators of local school districts in the creation and application of useful knowledge gained through evaluation of their service-learning programs.

▶ What is *Making Knowledge Productive?*

Making Knowledge Productive provides an introductory set of skills and tools to help Massachusetts teachers and service-learning coordinators conduct credible and respectable program evaluations as part of their on-going service-learning efforts. Specifically, we have written *MKP* with the following two goals in mind:

1. To provide service-learning practitioners with the skills and tools that they need to assess and improve their own programs.
2. To develop a common base of information across the state about the impact of Community Service-Learning in Massachusetts schools.

MKP is organized into two parts. ***Part One*** is a guide book designed to walk the service-learning practitioner through the general process of preparing, conducting and using a program evaluation. ***Part Two*** provides the tools (e.g., survey instruments) that Massachusetts practitioners can use in implementing a local evaluation. Taken together, Parts One and Two are intended to provide service-learning practitioners with both the skills and tools to develop the knowledge that they need to prove and improve the effectiveness of service-learning programs in Massachusetts.

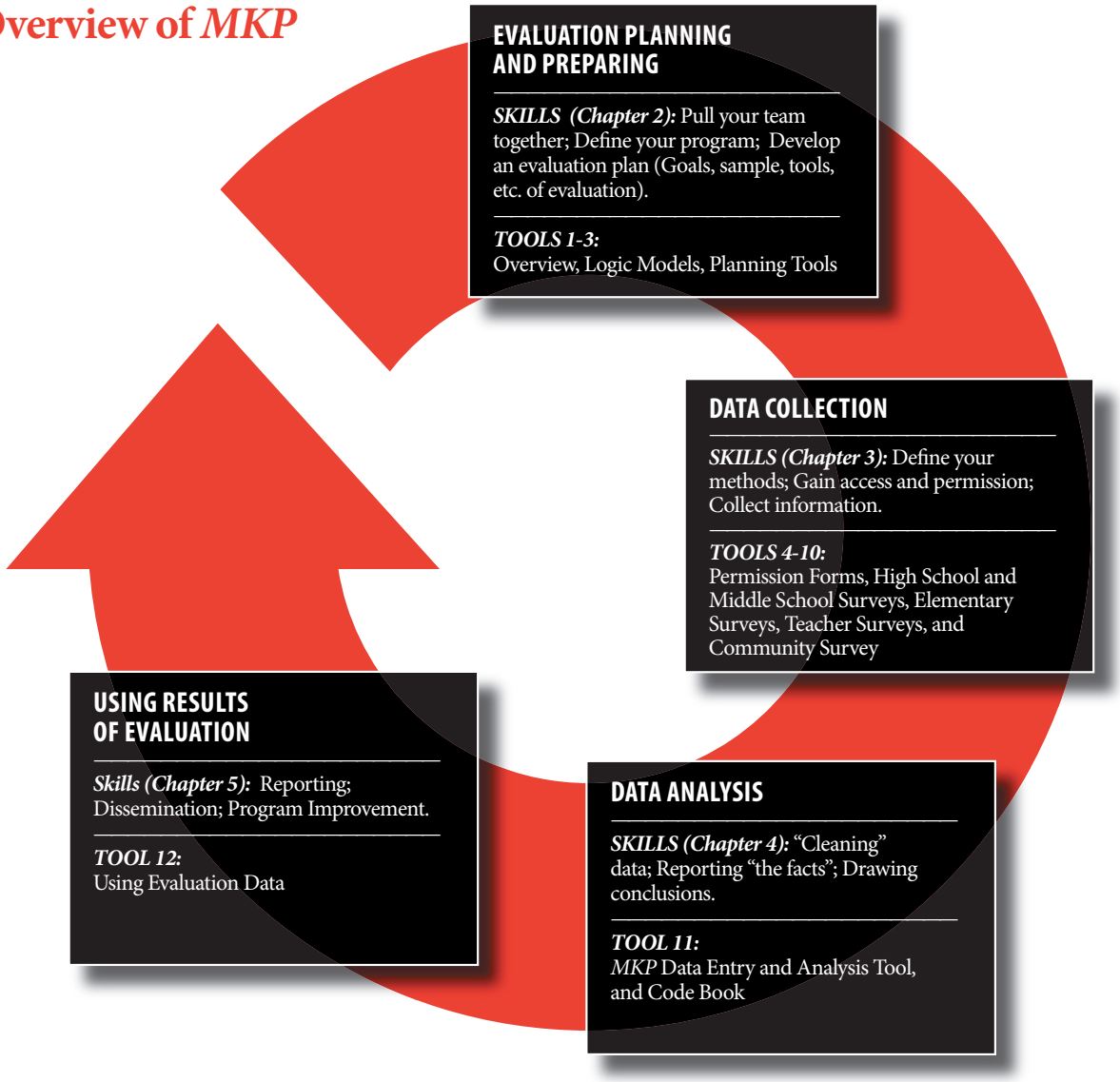
▶ How to Use *Making Knowledge Productive*

We recognize that not every school or program has the resources available to hire a professional program evaluator. And we do not presume that teachers will necessarily have expertise (or the time to develop such expertise) in program evaluation or research and survey design. For this reason, we have included Chapter One, "A Program Evaluation Primer." If you already have a great deal of experience with conducting program evaluations, some of this may seem elementary. Even so, you may want to read through this chapter to refresh your memory around some of the technical terminology.

After you finish Chapter One, some of you may want to familiarize yourselves with the tools in Part Two before reading the rest of Part One. These tools have been developed to meet the needs of teachers, community service-learning coordinators, the Massachusetts Department of Education Community Service-Learning program, and others who are involved with CSL in assessing their activities and using the results to make their programs better.

We recommend that before using these tools, that you finish reading this introduction and the rest of Part One (Chapters Two through Five), which provide an overview of the four stages of the evaluation process: *Preparing for an Evaluation, Collecting Data, Analyzing Data and Using the Results* ***(see below)***.

▶ Overview of *MKP*



► Acknowledgements

MKP represents the combined wisdom of many people. The working group defined the initial concepts for this publication. The *Making Knowledge Productive* Working Group included:

- Lawrence Neil Bailis
  (Center for Youth and Communities, Heller School for Social Policy and Management, Brandeis University)
- Jessica Donner (Massachusetts Department of Education)
- Beth Fleurant (Massachusetts Service Alliance)
- Nancy Holczer (Newton Public Schools)
- Alan Melchior
  (Center for Youth and Communities, Heller School for Social Policy and Management, Brandeis University)
- Tom Piñeros-Shields
  (Center for Youth and Communities, Heller School for Social Policy and Management, Brandeis University)
- Regina Wironen (Leominster Public Schools)

Teachers in three schools shared the survey draft with their students who provided valuable feedback. These districts included Hudson Public Schools (Todd Wallingford), Newton Public Schools (Nancy Holczer) and Leominster Public Schools (Regina Wironen). Students at Leominster Public Schools also completed a (much longer) pilot version of this survey. These teachers and students provided important feedback in defining the purposes, uses and approach of *MKP* and outcome domains (social, civic and academic outcomes) for the survey instruments in this toolkit.

***MKP* represents the combined wisdom of many people...**

The surveys themselves were adapted from a survey that Center for Youth and Communities, Brandeis University developed with and for KIDS Consortium. KIDS Consortium is a nationally recognized non-profit organization that works with teachers, administrators and students to involve students in addressing real challenges faced by their communities. It should also be recognized that the majority of the funding for the Community Service-Learning programs supported by KIDS Consortium and the Massachusetts Department of Education was provided by the Corporation for National and Community Service.

Brandeis based contributing authors of the publication include Lawrence Neil Bailis, Faye Cohen, and Alan Melchior. Tom Piñeros-Shields served as *MKP's* principal author. The title "Making Knowledge Productive" comes from the mission statement of the Center for Youth and Communities, which complements the Heller School's mission of "Knowledge Advancing Social Justice."

CHAPTER ONE

A Program Evaluation Primer

Much has been written about evaluating service-learning programs. Even more has been written about evaluating a wider range of educational and social programs. There are toolkits, handbooks and textbooks that cover a wide array of content areas and structures using various approaches and degrees of sophistication. (See the bibliography provided at the end of Part One.) This chapter of *MKP* is not meant to replace these existing resources – but it does highlight a few of the key lessons and components that are essential to making an evaluation process useful and credible.

"Program evaluation is the systematic collection of information about the activities, characteristics, and outcomes of programs to make judgments about the program, improve program effectiveness, and/or inform decisions about future programming."

— Michael Quinn Patton, Utilization-Focused Evaluation

We recognize that teachers and community service-learning coordinators are usually stretched for time. Most of you would love to better demonstrate success, detect shortcomings and identify changes, but do not have lots of time to invest in planning and conducting an evaluation.

At the same time, we recognize the types of problems that can arise if you "just do it." Evaluating service-learning, like service-learning itself, is a process that requires preparation, action and reflection. Knowing why, how and for whom you are doing an evaluation is as important as administering a survey and collecting information. This chapter helps provide an overview of program evaluation to help you understand what a program evaluation is and whether (and how) you should be doing one.

Finally, it is difficult to go too far into the world of program evaluation without hearing "jargon" and "gobble-de-gook". This is unfortunate. We hope the rest of this chapter will help de-mystify the practice and language used by professional researchers and evaluators and explain evaluation basics in simple, clear and direct language.

► Why do Program Evaluation?

Most people do program evaluation in order to gain knowledge. This knowledge can come from considering both the results – and sometimes the process – of program evaluation. As the title *"Making Knowledge Productive"* implies, we want to help you make your program evaluations useful. The results and the process of a program evaluation can be useful for many purposes, including:

- To track how programs are being implemented in different ways
- To identify areas for program improvement
- To clarify goals of a program and define specific outcomes that you are trying to achieve
- To tell your story to the world
- To help raise money
- To engage stakeholders in the process of program reflection and improvement
- To meet the requirements of a grant
- To celebrate accomplishments.

Whatever your reason or reasons for wanting to conduct a program evaluation, we assume that you want your program evaluation to be useful. Understanding what program evaluation is – and should be – may help.

► What is Program Evaluation?

There are dozens of definitions of program evaluation, including the one that opened this chapter. You may already have a definition of program evaluation. For our purposes, a simple definition of program evaluation may be stated as follows:

Program evaluation is the systematic collection and analysis of data to prove and improve the effectiveness of programs and/or organizations.*

A keen observer will notice that we have even been able to slip some jargon into this simple definition. For example, 'systematic' simply means that the evaluation has a plan to consistently and intentionally collect information by following a set of logical rules that the evaluators create. Systematic data collection and analysis implies that the program evaluation is planned and carried out in an intentional and consistent way across observations.

Data represent the bits of evidence that taken together, help us answer questions or understand something in new ways. Data are all around us. There are data in your classrooms; your students provide data in the form of their questions, behaviors and work every day. You may have heard that data comes in two general types: quantitative and qualitative. ***Quantitative data*** refer to bits of information that have been systematically collected and can be readily translated into numbers. ***Qualitative data*** refers to words (written or spoken), pictures, sounds, visual images or objects that have been systematically collected and can not be readily translated into numbers.

Data represent the bits of evidence that taken together, help us answer questions or understand something in new ways.

The challenge for program evaluation is to collect and transform that data into information that can be productive and useful to you and others. ***Information*** refers to the organized presentation of data. If we think of data as the raw materials, information is the finished product. Program Evaluation is the process for transforming data into useful information. The two parts of this process (from our definition) are **collection** and **analysis**.

You may have already guessed that the systematic collection of data can be done in several different ways. The ways we systematically collect data are called ***methods***. There are many types of research methods – including ***experimental*** (e.g., medical research to test new drugs in which participants are randomly assigned to treatment and control groups), ***quasi-experimental*** (e.g., a study that compares participants in a community based program with a group of non-participants (not randomly assigned) over time, ***naturalistic observations*** (e.g., anthropological studies to observe new cultures), and ***historical research*** (e.g.,- examining documents, records, artifacts and other evidence of past behavior). Program evaluation can include all of these approaches, and others, to answer questions about people, groups of people or policies and programs.

The types of methods used and types of data collected are best determined by the questions that you want to answer in your **analysis**. Analysis refers to using the information from a program evaluation to answer specific questions that you might have. The next section describes different types of program evaluations, which may help you understand the different types of questions for analysis and different methods/data collected.

* *This definition is based on the philosophy of evaluation of the Center for Youth and Communities (CYC), originally published in the W.K. Kellogg Foundation Evaluation Handbook, (1998). Edited by Susan Curnan (CYC Director), Lisa LaCava (CYC), Dianna Langenburg, Mark Lelle, and Michelle Reece.*

▶ Table 1

TYPES OF EVALUATION

	PROCESS/IMPLEMENTATION EVALUATION			OUTCOMES EVALUATION	
	Mapping Context	**Documenting Activity**	**Understanding Process**	**Assessing Outcomes**	**Assessing Impact**
SAMPLE QUESTIONS FOR ANALYSIS	Who is my target group? What are the needs of my target group or community? What personal or collective strengths or assets exist within the target community?	How many participants? What are characteristics of participants? How many hours of service? What types of activities do the participants take part in? What are the costs/ funding for the program?	Was the program implemented as expected (and if not, why not)? What parts of the program worked well, and what parts need to be strengthened? Did the program staff have the resources and training they needed? Did the program's design and activities match the program's goal?	Did the expected changes in participant attitudes and behavior take place? Were there changes in attitudes, instructional strategies, etc. among teachers/leaders during the program? What contributions did the program make to the community?	What difference has the program made (relative to no program)? Would the observed changes have taken place in the absence of the program?
SAMPLE DATA SOURCES/ METHODS	Public data sources (i.e.- census, etc.) Organizational reports. News clippings Traffic counts Existing databases or surveys	Time sheets Attendance records Work products Activity logs Participant reports Enrollment/sign-up forms	Surveys of students, teachers and community partners as appropriate Interviews with stakeholders Focus groups Observations	Portfolios/ projects Student self-report (essays, surveys, etc.) Pre/Post surveys on student attitudes and behavior Grades/school record information Supervisor interview or evaluation Focus groups Observations	Ideally, such evaluations include pre/post assessments of participants compared with a random assignment comparison group (called an "*Experimental Design*" study), or with a comparison group selected in other ways (a "Quasi-experimental design.")

► Determining the Types of Program Evaluation That Are Right For You

Just as there are many typologies to categorize and understand service-learning, so too, there are many typologies to categorize and understand program evaluation. Table One provides one useful approach to your evaluations by illustrating two general types of evaluation: ***process*** and ***outcomes*** evaluations.

Process Evaluations attempt to answer questions about the implementation of a program. They tend to focus on such things as: mapping context, documenting activity and understanding processes. ***Mapping context*** includes evaluation approaches designed to understanding the target group or community needs and assets that your program will address. ***Documenting activities*** provides basic descriptive information on the ***inputs*** and ***outputs*** of the program, as well as describing the actual ***activities*** that take place. ***Inputs*** refer to the resources that a program requires to operate (e.g. - people, space, time), ***Outputs*** refer to the direct products of a service-learning program, such as the number of service hours that students complete. ***Understanding processes*** describes how the program works – the ins and outs of how it is operating. That is, how well is your program working? What are its goals? How were decisions made? Who made them? How were these decisions carried out? While you may use lots of numbers in describing context or documenting activities, understanding the process often relies more on qualitative data, such as information from interviews with staff and participants.

The two columns to the right of the table are examples of ***Outcomes Evaluations***. Outcomes evaluations focus on the successes or shortcomings of your program. An outcomes evaluation asks "How well are you meeting your goals in the light of the program's intended result?" A program's intended result may occur over time in the form of ***outcomes or impacts***. ***Outcomes*** refer to specific changes in the knowledge, behavior, skills, or functioning in the target participant, institutions and/or community of the program. Outcomes may be either short-term (usually within 1-3 years) or long-term (usually within 4-6 years). ***Impacts*** tie back into the mission statement of a program and refer to more general long term changes (usually about 10 years) in the participant, but also in the organizations, community or society that result from achieving outcomes.

Assessing impacts tells us whether the program made a difference.

Outcomes may be measured through a combination of qualitative and quantitative data. Finally, assessing impacts assumes some comparison of the program's outcomes with a measure of what would have happened if the program had not taken place (usually by comparing outcomes for participants and non-participants). Assessing outcomes tells us whether the expected changes took place. Assessing impacts tells us whether the program made a difference: that is, produced changes greater than what might have happened anyway.

Table One outlines some of the questions and methods generally associated with each of the five types. You will notice that the questions to the right of the table generally (though not always) require more sophisticated methods and more resources than those to the left. It should be noted that these types of evaluations are not mutually exclusive. An evaluation may consider both ***process and outcomes*** questions by addressing multiple questions or through multiple methods. In fact, most evaluations include questions and methods reflecting more than one type of evaluation approach. Therefore, *MKP* provides tools that include a mix of documentation, process, outcome, and impact evaluations.

► Collecting Data at the Beginning and End of a Program

"People need to begin early on. Good program evaluation is not an 'add-on.'"
– A MKP working group participant

Perhaps the most fundamental evaluation question is whether there have been any changes or improvements in the knowledge, skills, attitudes or behaviors of students over the course of their community service-learning participation. In order to understand whether this has occurred, we may want to collect data for at least two points in time. We want to collect

data before a student has begun a community service-learning project, class or experience to provide an understanding of the student's skills, knowledge, attitudes or behaviors before the program. This is sometimes called the 'baseline' results.

We also want to collect data after the student has finished the community service-learning program (i.e. "post-program"). In order to collect data at these two times, *MKP* includes two types of student surveys: "Pre-tests" and "Post-tests." By comparing the results of these two surveys, we can see whether there are changes in their self-reported knowledge, skills, attitudes or behaviors during their participation in a program.

► Choosing a Comparison Group

Knowing that a change has occurred during a program is important. But, for many audiences (especially skeptics), simply demonstrating improvements would not be enough to demonstrate the success of your program. That is, some people are likely to point out that over a year, many young people would improve in their academic, social or civic outcomes even without participating in a community service-learning experience. In order to be sure that it is your community service-learning program that is causing the observed change in the students over time, we need a point of comparison.

In order to develop a good comparison group, you may ask "What would the students have done if they had not participated in the community service-learning experience?" Of course, we can never know the answer for sure, but we can select a group of students who are similar to the students in the service-learning class and see how they change during the same period that our students are in the service-learning program. One way to select a comparison group is to ask another teacher (with similar students in the same grade) who is not using community service-learning, to administer the pre- and post-test comparison surveys to their students. If your entire school or grade is using community service-learning projects, then you may want to look for another school in your district, or a nearby district.

If we compare the results of the pre-test and post-test of the students who participated in community service-learning with the change over time of the students who did not participate in the community service-learning program – then we can begin to show whether the program made a difference. This method (see table below) is sometimes called the ***Difference-of-Differences*** approach.

The difference of differences approach takes into account the differences in the pre/post survey and the differences between CSL participants and non-participant comparison students. The key to making effective use of comparison groups is to find comparison students who are as similar as possible to the students in your program – same ages, demographic characteristics, academic backgrounds, etc. The more similar the comparison students are to those in your program, the more credible the comparison. It is important to note that finding a good comparison can be hard – especially if the students in your program have joined because of a special interest or incentive. It is important to think carefully about the comparison you want to be able to make.

	PRE TEST (BEFORE THE PROGRAM)	POST TEST (AFTER THE PROGRAM)	DIFFERENCE
Students in Community Service-Learning Class	Student outcomes measured by Pre-test Survey	Student outcomes measured by Post-test survey	Difference in student outcomes between Pre-test and Post-test
Comparison Group of students who are not in Community Service-Learning Class	Comparison student outcomes measured by Pre-test Survey	Comparison student outcomes measured by Post-test Survey	Differences in comparison student outcomes between Pre-test and Post-test
	Differences before the program	Differences after the program	***Difference of Differences = The estimated impact of the program***

► Measures and Indicators

So how do we know that an outcome has been achieved? One way is to measure the outcome. A measure represents an established standard set of units designed for definitive quantitative comparisons. So, miles, feet, or meters are direct measures of height, length or width. Liters and pints are direct measures of liquid volume. Most often, however, direct measures of outcomes in service-learning are not available. For example, test grades do not "measure" how much a person knows, their intelligence or even how hard a person studied, because many other factors will affect a student's performance on a test and a test cannot measure the full body of knowledge that is stored in someone's mind. Test grades are ***indicators***. An indicator is a type of data that serves as a proxy measure of a particular outcome. An indicator is a type of data that serves to approximate outputs, outcomes, or other results. Indicators can be part of both process and outcomes evaluations. The student surveys in *MKP* include questions that can serve as indicators of social, civic and academic knowledge, skills, attitudes and behavior.

► Some Fancy Words that You May Hear about "Evaluation"

This chapter has provided a brief introduction to some of the key concepts in program evaluation. We realize that it has been full of some 'fancy' terminology. Nevertheless, we hope that this provides you with a resource to approach your program evaluation work thoughtfully and confidently. There are a few additional terms and concepts that you might want to be familiar with.

Triangulation is a five-syllable way of saying that if you reach the same conclusion via two or more different sources of information, then your conclusions are more likely to be valid than if they are only based on one. For example, in assessing the success of your service-learning program, you might want to gather survey data from students, teachers and

What's all the Talk about "Random Assignment Control Experiments"?

Recently government agencies, such as the U.S. Department of Education, and some foundations have made statements in support of a well-respected approach to program evaluation called "random assignment controlled experiments."

Random assignment is based on the scientific method and refers to the process of randomly selecting (i.e.- by lottery) people or groups who will participate in a program (and the study) and not offering such program options to other people or groups who will serve as a "control" group (and who also participate in the data collection for the study). For example, one group of students might be randomly assigned to a service-learning course, while others would be assigned to a study hall or a non-service-learning class. We might expect (hypothesize) to see a greater increase in civic commitment among those that participated in service-learning than those who did not.

Random assignment experiments represent the most rigorous form of research design and have become the most widely accepted way to obtain evidence for the impact of a program. Its great strength is that, since individuals are randomly sorted into or out of the program, it controls for any unmeasured differences between those who would normally have selected to participate or not (for example, because they have a prior interest or are more motivated). At the same time, there is also widespread recognition among professional evaluators that random assignment carries both practical and ethical challenges. Further, random assignment studies still demand proper tools and instruments free of bias and careful monitoring. For these reasons, such studies can be difficult and expensive. We do not recommend that practitioners attempt to create their own random assignment control studies without professional researchers.

Instead, Making Knowledge Productive provides you with the tools to conduct a "quasi-experimental" design that does not include random assignment but is widely consider the "next best" way to prove causality. In future chapters, we present the broad thinking behind these kinds of studies and tools that have been designed to select samples and comparison groups, along with pre and post tests. As we shall see, even these evaluation tasks still have substantial challenges of tracking participants while respecting ethical guidelines such as confidentiality.

For example, test grades do not "measure" how much a person knows, their intelligence or even how hard a person studied, because many other factors will affect a student's performance on a test and a test cannot measure the full body of knowledge that is stored in someone's mind.

Test grades are indicators.

community representatives. Whether you need to triangulate your data depends on the purposes, audiences and approach to your program evaluation. Especially when using qualitative data (for which statistical tests of validity are not available), you may want to use triangulation involving some quantifiable data as well.

Reliability refers the consistency or dependability of your measures or indicators. It refers to the notion that if you were to implement the same evaluation tool multiple times with the same people and under the same circumstances, it would produce the same information and findings each time.

Validity refers to the degree to which you can be sure that the data actually reflects the truth. That is, does your tool measure what it claims to measure? Taken together, reliability and validity are a good way to think about your own tools. Are they as reliable and valid as possible?

Statistical Significance is often used as a test of the validity of a quantitative finding. A statistically significant finding refers to one that is unlikely to have occurred only by random chance and this would only show up if the program were making a difference. As the term is generally used, a ***significant*** finding occurs if the probability of the event taking place randomly occurs fewer than 5 in 100 times ($p < .05$). A ***strong or highly significant*** finding refers to a finding that would occur randomly fewer than 1 in 100 times ($p < .01$). In your own program evaluations, you will likely not need to worry about having to test for statistical significance. Please note, that a statistically 'significant' or even 'highly significant' finding does not mean it is necessarily a more important outcome area– only that the results are less likely to have been the result of chance. We provide the term to you primarily because you may often hear it used –and sometimes misused – by policy makers, practitioners or evaluators.

►►►► While writing *Making Knowledge Productive,*

we have attempted to avoid all of the above fancy terms.

Instead, we have tried to provide an approachable

set of tools for you to use in the evaluation of your programs.

CHAPTER TWO

Preparing and Planning for Your Program Evaluation

This chapter will discuss how to prepare and plan your program evaluation. When is a good time to evaluate a program? When is not? Who is the audience for your evaluation? What types of questions do you want answers to? What are the pros and cons of selecting an external evaluator or of doing it yourself? What are the ethical issues around evaluation? *MKP* has simplified the process of planning and preparing for your program evaluation into three steps.

1. Pull together your Evaluation Team
2. Define the key features of your program
3. Develop an Evaluation Plan

The rest of this chapter and the tools associated with it in the second part of this toolkit will describe each step. We recommend being sure you plan enough time and resources to follow-each step.

► Pull Your Evaluation Team Together

The *MKP* approach to program evaluation begins with developing an "evaluation team" for your project that includes the key stakeholders in the program and the evaluation. This team can help you to define your evaluation questions and begin to plan the steps you need to take for your evaluation. Moreover, the formation of a team provides an opportunity to involve youth in the leadership of your program. By involving others in the evaluation process, you can be more likely to get buy-in from key stakeholders to the evaluation process that you are using. This will help these same stakeholders to accept and appreciate the results of your program evaluation later.

Please note, when you consider who to invite to be part of the evaluation team, you may want to consider both the 'queen bees' from whom you want support, and the 'worker bees' who will help you get the job done.

Your evaluation team may include many of the same people that form your local Community Service-Learning Advisory Committee.

- Youth participants
- Project Funders
- Project Staff and Administrators
- Community Leaders
- Collaborating Agencies
- Others with an Interest in Program Effectiveness

Question:
What Roles can young people play in an evaluation?

Answer:
All of them!

Young people can be involved in:
- Planning and Design Team
- Instrument Design
- Data Collection (Surveys, Interviews, Site Observation, etc.)
- Data Analysis
- Reporting/Presentation/Discussion

Some programs have even used their program evaluation as a service-learning project!

► Define Your Program (Logic Model)

Before you can evaluate your program, you need to clearly define the key parts of your program that you are evaluating. A now common approach to do this involves developing a logic model.

A logic model is a ***picture*** of how your program works – the theory, assumptions, and expectations underlying your program. It generally consists of a ***group of boxes*** that are connected by arrows, one leading to the next leading to the next, on an "if-then basis." There are lots of varieties of logic models but there is ***no one right way*** to create and use them.

► Why Use Logic Models?

1. ***Improved Program Design.*** The process of creating a program logic model helps clarify your thinking about the program, how it was intended to work, and what adaptations need to be made once it is operational.
2. ***A Starting Point for Management Improvement and Evaluation.*** A logic model makes the program design explicit so you can decide more systematically what pieces of the program to study and what outcomes are important to track.
3. ***Understanding Complex Initiatives.*** In complex programs or initiatives, a logic model can lay out interim outcomes, highlight assumptions, and make it easier to identify gaps.
4. ***Partnership Building.*** The process of developing a logic model requires stakeholders to work together, to clarify the rationale for the program and the conditions for success. The model becomes a focal point for discussion and a means of creating ownership among the stakeholders.

SAMPLE LOGIC MODEL (Tool #2)*
Organization's Name

Template and Definitions for Logic Model

For Whom	Assumptions (Theory of Change) Moving From...	Inputs	Strategies/Activities Through to...	Outputs	Outcomes Through to...	Impact
Who are you ultimately doing this for? Who will benefit from your work and the changes you create? Who (described as specifically as possible) should this be in the first year? In two years? Three years? Four years?	Why are you doing this? Why does the program/ organization exist in this local area? Why does change need to happen? What are the social, political and economic conditions or environmental factors that make you want to do it? What identified needs or assets led you to address the issue? What assumptions guide your work? What do you know, think, and believe about why you do what you do?	What resources will your program need to implement its activities?	What supports and opportunities are needed to provide a positive developmental setting for youth? What mix of programs, services, and activities will be needed to achieve the desired outcomes? What systems need to be in place?	What products or services were created and/or delivered by your program? How many?	What outcomes do you want to achieve for youth? What changes for youth do you want to see over the next four years and beyond? How would your community and institutions like to operate differently?	Changes in the lives of the beneficiaries and/ or participants that constitute significant long-term benefits to them.

* *NOTE: There are many approaches to logic models. This approach to logic models is the signature framework for logic models used by the Center for Youth and Communities. For more information on logic models see the W.K. Kellogg Foundation Evaluation Handbook (1998) and the W.K. Kellogg Foundation Logic Model Development Guide (2004). Both are available at* ***http://www.wkkf.org****.*

► How do I develop a logic model for my program?

We recommend developing the logic model with the entire program evaluation team as a way of ensuring buy-in from all of your key stakeholders. You may want to facilitate one or more discussions by using the questions provided in the logic model tool. The Brandeis approach to logic models asks for explicit answers to the following questions:

- Who your program is for?
- What assumptions do you hold about it?
- What are the "inputs" that you need for your program to function?
- What activities or strategies do you engage in?
- What are the immediate "outputs" of these activities or strategies?
- What outcomes do you expect at the end of the program?
- What is the expected eventual long-term impact of your programs?

Once you have developed a logic model for your program (Tool #2), you are ready to begin to plan your evaluation. The approach to your program evaluation will depend on the program, the purpose and audience of the evaluation and the time and resources that you have available.

► Develop an Evaluation Plan

We recommend developing your evaluation plan early in the program with your evaluation team. Your evaluation plan should respond to the following questions:

1. What questions do you want to answer?
2. What information/ data do you need to answer these questions? What data do you already have or do you already collect? What data do you need to collect?
3. What methods can you use to collect this data?
4. From whom or from where will you collect this information? Should you select a sample or collect data from every participant? Should you create a comparison group?
5. How will you analyze this data?
6. How can you use the results?
7. How can you make this do-able?

1. What questions do you want to ask?

Your logic model may help you think about what kinds of questions you want to ask and answer through your program evaluation. Process or implementation evaluations tend to refer to questions about assumptions, strategies, and inputs. You may, for example, want to know if the strategies you've designed are taking place as expected – are service projects meeting your standards of quality; is reflection taking place as expected? Outcomes evaluations tend to respond to questions about the outputs, outcomes or impact of your program – for example, are student attitudes changing in the ways we had planned?

Think about the key questions that you want to answer about your program. The questions you decide to ask will depend on the purpose and audience of the evaluation (who wants to know what and why), on the goals of the program, and on the time and resources that you have available for evaluation. There is no single "right" question – you need to think about what you want and need to know. The "Developing Your Evaluation Questions" tool provides space and a structure for thinking about your evaluation alone and with your evaluation team.

2. What kinds of information can you use to answer your questions? What data should you collect?

First, think about any data that you already collect. This may include participant grades or MCAS scores, tracking tools, forms for keeping track of attendance and participation, samples or portfolios of student work, or records for following participants when they leave the program. Try to match this data to the questions that you asked. Next, ask what additional information you will need to answer your questions.

3. What kinds of methods can you use to answer your questions?

It is important to design your evaluation to be both flexible and responsive – the program should drive the evaluation, not the other way around!

Surveys

Part Two of *MKP* provides several types of surveys for you to use in your evaluation. These include:

a. Participant surveys to measure pre and post program differences. Two versions are available, one for high school/middle school students and one for elementary students.
b. Pre and Post program comparison group survey for non-participants. The questions in these surveys parallel those in the participant surveys so that the results can be compared. Two versions are available for high school/middle school students and elementary school students.
c. A teacher survey, which may be completed at the end of the community service-learning program.
d. A community partner survey which may be completed at the end of the community service-learning program.

Surveys are a common data collection method, but are by no means the only method that you may use. Some other common program evaluation methods include:

- Focus groups: You can ask one or more small groups of people questions or facilitate a discussion about ideas, and collect their responses.
- Interviews: Asking someone to answer questions individually and recording his or her responses (on tape and/or using notes).
- Site visits: Visiting and observing a program on one or more occasion and recording your observations.
- Photographic documentation: Taking photographs of a program or event to be coded and analyzed later according to specific themes or characteristics.

The important thing to remember in using any of these methods is to develop a strategy for using them systematically, so that the data collected is representative and reliable. For focus groups, for example, make sure you have a standard set of questions and a clear plan for selecting who will be in the groups. For site visits, think about what kind of observation guide or rubric you can prepare, so you know ahead of time what you need to look for in each program that you visit.

The decision about what kind of data collection method to use usually involves balancing several considerations: what kind of question you are trying to answer; what kind of evidence your audience is mostly likely to accept; how much time you and/or the program staff have to devote to data collection; and the resources available for collecting and analyzing the data. Often, some kind of trade-off is involved: for example, surveys generally cannot capture information on process or the richness of participants' experience, but are often substantially easier and cheaper to administer on a large scale than focus groups or interviews. Portfolios may provide an excellent way of assessing student work, but your funder may want to see impacts measured by more traditional assessment tools. In the end, it is important to recognize that there is no single "right" answer. The best you can do is to thoughtfully weigh your options as you make your choice.

4. From whom or from where will you collect this information? Should you select a sample? Should you create a comparison?

After you have selected your methods, you will need to identify the people or places that will participate in the evaluation.

Who should you collect data from?

As noted in our discussion of "triangulation" in Chapter One, we recommend that you collect and analyze data from multiple perspectives – don't rely on just one type of source! There are two primary reasons for this. First, each source may have a different perspective on the "true" understanding of your program. No one perspective contains all of the truth. It is only by compiling multiple perspectives that we can be confident in our results. Second, by incorporating multiple stakeholders as sources in the data collection, you increase the likelihood that different stakeholders will accept and be willing to use your results.

Some potential people from whom you should collect data include:

- Program Participants
- Staff/Teachers
- Community or Agency Partners
- Parents

You may notice that many of these stakeholders were also seen as potential members of your evaluation team. By including representatives from multiple sectors in your evaluation team, you increase your likelihood of knowing how to collect and analyze data from those groups.

How many people should you collect data from?

Generally, in a smaller program you will collect data on all the participants for your evaluation. However, sometimes it is not feasible to collect data from all participants or stakeholders. When you select a portion of your participants or stakeholders to collect data from, that portion is a 'sample.' A sample is a smaller set of cases that a researcher selects from a larger pool and uses to make generalizations about the entire pool or population. Ideally, you want to select a sample because it is representative (typical) of all of the participants or stakeholders. A 'representative' sample allows you to draw conclusions that apply to all participants or subgroups within your program.

The size of your sample is very important. In general, the larger the sample size, the more likely it is to be representative and allows you to detect smaller changes as measured by the "margin of error." The margin of error is a way to measure whether your survey is accurate. In the world of gambling a margin of error is sometimes called the "spread." But in research, it is the percent or number that represents how far your numbers can vary. Obviously a smaller margin of error indicates that your survey measures are more accurate. A larger sample size tends to decrease your 'margin of error.' Given this, we suggest, whenever possible, that you try to obtain a sample size of at least 30 respondents for your survey. However, in general a larger sample will give you more useful results. Typically, for example, a sample of 100 participants will provide a margin of error of +/- 10%. This means that the responses could very well vary by ten percent. Often 300 participants or more are needed to provide a margin of error of less than 5%. If you plan to be doing sampling on a large scale, it would be a good idea to consult with a professional researcher or evaluator for guidance.

Do you need a comparison group?

In Chapter One, we discussed the value of having a comparison group. As part of your evaluation planning process, you and your evaluation team should discuss whether and how you will select a comparison group. You and your evaluation team may want to think about which classes or groups seem to be most similar to participants if they had not been involved

in the community service-learning experience. Then you may wish to consider whether it is feasible to ask them to complete pre and post surveys. If you do not think you can make this happen, then you simply should proceed without having such a comparison group. Your study will be less than ideal, but is still highly likely to produce credible, useful information.

Increasing Your Response Rate

When you conduct a survey, it is common that not everyone will respond and return the survey. When you divide the number of surveys returned by the number of surveys that you distributed you obtain a "response rate." Ideally, you want a high response rate. There are several ways to increase your response rate. Some ideas for increasing your response rate include:

1. Provide clear instructions for completion of your surveys.
2. Include a return address envelope with the surveys.
3. Follow-up with reminders before your deadlines approach.
4. Be polite. Remember to say "please" and "thank you".
5. Whenever possible, provide incentives. These can be things like gift certificates for teachers to purchase books or a simple "thank you" lunch for survey respondents.
6. Think of ways of having the surveys done all at once or in a group. Here is another way in which food can provide a great incentive – invite everyone in the program to a pizza party and ask them to complete the survey while they are there. In short, be creative, and think about how best to get the cooperation you need.

The best way to get responses is to have students fill surveys out in class and hand them in before she or he goes out the door.

Think of ways of having the surveys done all at once or in a group. Here is another way in which food can provide a great incentive – invite everyone in the program to a pizza party and ask them to complete the survey while they are there. In short, be creative, and think about how best to get the cooperation you need.

5. How will you analyze the results?

Chapter Four will describe the process of analysis in more detail, but decisions about who will enter, clean and analyze the data that you collect should be written into your evaluation plan. Also, leave enough time to complete the data analysis before you plan to release your results to stakeholders. *MKP* provides a tool (Tool # 11) that will guide you through the steps for data analysis.

6. How are you going to use the results?

Chapter Five of *MKP* looks at how you will use your evaluation, but your plan for using the evaluation needs to be built into your overall evaluation plan. Some uses to consider include:

- Use the Findings
  - — Improving your program
  - — Evaluating the effectiveness of your program
  - — Generating new knowledge
- Use the Evaluation Process
  - — Building shared meaning and understanding
  - — Supporting and enhancing the program
  - — Supporting human and organizational development

7. *How can you make it do-able? What time and resources do you have?*

No plan would be complete without considering what is practical and feasible. The two most common challenges to conducting a program evaluation are time and money. *MKP's* Evaluation Planning Tool includes worksheets for developing an evaluation budget and developing a timeline. You may also want to consider whether there are people with experience or knowledge and willingness to help manage or conduct your program evaluation. Colleges and universities may be able to provide student interns or faculty who might be willing (if asked) to provide advice to your program evaluation process. This is also a place where parents and students can be involved as volunteers, helping to collect and analyze data for you and with you.

► Final Thoughts As You Prepare and Plan your Program Evaluation

- Involve your stakeholders in the process.
- Design evaluation to meet your needs. There is no one "right" approach. Your logic model should help you define your program's unique evaluation needs.
- Make effective use of the resources (people and information) that you have on hand.
- Make evaluation a living, useful process – a "want to" instead of a "have to."

► ► ► ► The two most common challenges to conducting a program evaluation are time and money. *MKP's* Evaluation Planning Tool includes worksheets for developing an evaluation budget and developing a timeline.

CHAPTER THREE

Collecting Data About Your Program

Depending on your evaluation plan, there may be many activities that you will need to engage in. This chapter will discuss two ways to collect data about your programs: existing sources and surveys.

► Existing Data

In today's educational environment, it is impossible to imagine a school that did not already use data to assess student progress and learning. Data about students are abundant in schools and include grades, test scores (i.e., - MCAS proficiencies), attendance and disciplinary records. Different school districts have different rules about the uses and privacy of these types of data. Be sure to check with your building or district administrators if you have any questions about using school record data.

Most classrooms also have an abundance of data on students – class projects, journal entries, directed writing essays, presentations, quizzes and other tools used for grading and assessing classroom progress can also serve as evidence for evaluation, if used consistently and with a clear set of standards for performance. Finally, many service-learning programs use time sheets and other tools for tracking student hours and service activities – all of which provide data that can be used as part of the evaluation process.

By thinking about your evaluation plan, you may find that you already have substantial data on pre and post and comparison groups from existing sources.

► The *Making Knowledge Productive* Student Surveys

If you are a Massachusetts Department of Education Community Service-Learning grant recipient, then you will need to use the student, teacher and community surveys enclosed in this toolkit as one part of your program evaluation. Otherwise, you still may want to use the surveys that we developed to collect and analyze data about the impact that your service-learning project has had on your students.

A first version of the *MKP* surveys were created in conjunction with the practitioner working group and the Massachusetts Department of Education in order to provide Massachusetts service-learning grantees with a standard, easy-to-use-tool for evaluation, and to also help the state gather consistent and reliable data from its many grantees. These surveys sought to measure the impact of community service-learning on students in civic, social, and academic "domains" or categories of questions. The *MKP* surveys were pilot-tested at two schools in Massachusetts and feedback was considered – especially the need to shorten the surveys. The feedback from this pilot test led the Brandeis team to reconsider the survey instrument and to create a new instrument based largely on a survey that the Brandeis team had developed with and for KIDS Consortium, Inc..

The current *MKP* surveys are an adaptation of the surveys used by KIDS Consortium. The domains are:

- ***Academic Domains:*** Education Aspirations, Importance of Understanding History, and School Engagement
- ***Social Domains:*** Belonging, Caring Adult, Confidence/Relationships, Personal Empowerment

- ***Civic Domains:*** Civic Empowerment, Civic Knowledge, Future Civic Involvement and Social Responsibility

This survey has been used among KIDS Consortium grantee sites in Massachusetts, Maine, Rhode Island, Vermont and New Hampshire. Pre and post surveys, comparison surveys and elementary school surveys were then created based on the high school survey template.

It is important to recognize what these surveys are, and what they are not. What they are is a ready-to-use set of tools that address some of the most common outcomes associated with service-learning programs. As such, practitioners can pick them up and use them with some confidence that they are assessing outcomes that are appropriate for a service-learning evaluation. At the same time, they are not customized to individual programs. If there are outcomes that are important to you that are not covered by the surveys, you will need to supplement them with other tools or assessment methods.

► Ethics In Program Evaluation

> ***"Evaluation deals with real people in real programs, often people in serious need of help. The results of the evaluation may have real consequences for these programs and these people. Therefore, evaluation has an obligation to pay even more attention to ethical questions than most other kinds of social science research have to do."***
>
> ***— Carol Weiss, Evaluation***

Even if you are not a professional evaluator, ethics remain an important concern for you as you design and implement your own program evaluation. Evaluators need to be mindful of professional ethics regardless of the methods they use. Evaluators must consider ethical concerns in all phases of the evaluation process: planning, collecting, analyzing and reporting. Key ethical concerns of evaluators are:

- ***Honesty:*** Of course most of us recognize that honesty is an important ethical consideration in any work that we do. It is especially important to tell the truth to the participants and comparison groups.
- ***Informed Consent:*** Informed consent refers to the idea that participating in a survey or evaluation project is voluntary and that people decide whether to participate/have enough information to decide what to do. This is especially important when working with groups of people who may be vulnerable due to power differences in age, physical ability, rank or position, or historical oppression and should always be a consideration when administering surveys of students in schools.
- ***Confidentiality and Anonymity:*** Confidentiality means you will not share people's names or identifying information associated with their responses when you write up the study results. Anonymity means that you will not even collect the names or identifying information from people. Both confidentiality and anonymity are important in order to allow students to report their honest responses, without fear of embarrassment or pressure – and both increase the trustworthiness of your results.
- ***Professionalism and Competence:*** *Making Knowledge Productive* does not presume that you are professional evaluators. Given this, it is important to be honest, respectful and ethical, and to plan and carry out your evaluation as carefully and thoughtfully as possible.
- ***Reciprocity:*** Reciprocity means that people who provide data should also receive something in return, at minimum we should share the results of the study and discuss them with the students.

► Creating a Code to Match Pre and Post Surveys

In the previous two chapters, we discussed the value of pre and post program surveys for measuring change among participants in your community service-learning program. This can be done by simply comparing the average scores before and after the program, but it is more convincing to match the responses of each student's pre-test with their responses on a

post-test.

Given the promise of privacy and the issues that are sometimes involved in securing permission to use students' names, matching pre and post surveys is always a challenge.

One approach is to not do it at all. If you do not match pre and post surveys, then you can still compare the overall responses between the beginning of your program with the end of the program in the aggregate (that is, looking to see if the average "score" for all students combined changes significantly from pre-test to post-test). To illustrate a limitation with this approach, let's use a sample that at the beginning of your program exactly half of your students report that they love school and the other half report that they strongly dislike school. If all of the students who loved school at the beginning of the program, later report that they hate school, and all of the students who hated school later report that they love school – that would be an amazing (and bizarre) result for your program – but you would never know about it by looking at the aggregate scores. The results would appear that there was no change in the average score for students as a whole at all. This is an extreme example, of course, but you will hopefully see the point. By comparing the averages, it is not possible to learn as much as comparing each student's response.

So, another way to match responses is to develop a code that will be consistent over time but also respects privacy. Some ways this can be done:

- Ask students to write their official student ID Numbers on surveys. If students do not know their ID number, you might want to distribute pre-printed labels with student ID numbers to the students.
- Ask students to create their own code using a pseudonym or nickname. This is easier to administer, but depends on students remembering which nickname or pseudonym they used from beginning to end of the program.
- Develop a code. (See the information presented below.) We have included an approach that asks students to develop a code out of their birth date, first and last name and eye color. You can develop any kind of code you want, but try to find something that (a) students will be able to remember or reconstruct at the end of the year, and (b) depends on characteristics or features that do not change (phone numbers, for example, often change over the course of the year).

DEVELOPING A CODE TO MATCH Pre and Post SURVEYS

As the first step in the survey process, we want students to create a personal code. The code will let us keep track of your surveys without knowing student's names. The code has three parts: initials, birth date, and the eye color.

1. What are the first initials of your first name, middle name, and last name?
   First name initial: ___
   Middle name initial: ___ *(leave blank if no middle name)*
   Last name initial: ___

2. What is your date of birth? (Month/date/year): ____/_____/_____

3. What color are your eyes? ☐ Brown ☐ Blue ☐ Hazel/Green ☐ Something Else

Of course, none of these approaches is fool-proof. Any code that can be created can be broken. And, as we previously discussed as part of program evaluation ethics, it is important to try to maintain the privacy of student information.

► Administering the *Making Knowledge Productive* Surveys

Before you administer any survey, please make sure that ***you*** understand what the purpose of each survey is and why we are asking each question in the survey. If you don't understand any part of the survey, please contact the Massachusetts Community Service-Learning Coordinator to discuss these issues. We want to ensure that ***you*** will be able to answer any questions that participants ask you about the surveys.

Please determine if you need to use the parental permission form, and if you do, please discuss it with the leaders of your school. Some school districts require the completion of their own permission forms before you can collect any data from students. The permission forms that we provided in the toolkit are examples. Please feel free to adapt to the specific needs of your school and/or district.

It is important to read the instructions of the cover page of the surveys to the students and let them know that they can skip any questions they want for any reasons they want and let them know that the results will be kept private. Also, please give them a chance to answer any questions about the survey before they start to fill it out, and tell them that the results of the survey should be available to anyone who wants to know within a few months.

> ***The main thing is that the surveys should be completed while the students are present in order to make sure they fill them out. They should NOT be sent home with students who promise to return them at a later time. This approach leaves you open to having too many students forget to fill them out or lose them.***

Ideally, you should give each student an envelope to put her/his completed survey in and then ask them to seal it, to insure privacy, then, collect the envelopes with the enclosed surveys. The surveys are now ready for forwarding to the people who will do the data entry and analysis – ideally not the same people.

► Teacher Surveys and Community Partner Surveys

In addition to the *MKP* student surveys, we have included teacher surveys and community partner surveys (see Tools #9 and #10). Massachusetts Community Service-Learning grant recipients will be asked to complete and submit these as well. Even if you are not reporting on a grant, your evaluation team may decide that you want to complete and collect these as part of your evaluation in order to get multiple perspectives about the effectiveness of your community service-learning program.

CHAPTER FOUR

Analyzing Your Data

We don't believe that there is any one "cookie-cutter" approach to analyzing your data. There are many ways to approach your analysis and comparisons that you could make. This is true of program evaluations in general, and especially true of analysis. Analysis involves making decisions – and these decisions get easier as you get more practice making them. In this way, analysis can be as much of an art as a science.

At the same time, there are some essential things you will need to do – especially if you are working with the surveys provided within *MKP*. You will need to "code" your data, enter your data, "clean" your data and summarize your results. This chapter briefly describes each of these steps. We will then introduces the *MKP* Data Entry and Analysis Tool (Tool #11) - a pre-formatted spreadsheet for entering and analyzing data from the *MKP* surveys. Then, for advanced users, we include an introduction to the technical uses of data analysis tools within a common computer software package, *Microsoft Excel.*

► Coding and Entering the Data

Coding data is the first step in transforming student survey responses into useful information. "Coding data" refers to the process of labeling responses on a survey in a way that they can be easily entered into a computer and analyzed. Coding data is also a way to 'systematically' (remember our definition of program evaluation in Chapter One) turn responses into numbers (or quantitative data) which can be counted and summarized.

MKP includes a "Code Sheet" for all surveys included in Part Two (Tool #11). Student participant and comparison surveys (both pre and post) have been intentionally numbered the same way, and therefore use the same code sheet. As described below, when you enter the data, you will enter the number of the student's response for each question into your spreadsheet.

Entering data can be time consuming and tedious. At the same time, it is an important and necessary step that needs to be done accurately. Also, you will notice that there may be decisions that come up during the data entry process. For instance, what if the student did not answer a question – do you leave it blank or enter a 'zero.' (Note: For the *MKP* surveys leave it blank!) Or, what do you do if someone gave multiple responses? (Note: For the *MKP* surveys you should enter a "9" which indicates invalid or missing data.) You should keep track of these and other questions as you enter your data.

► Clean Your Data

Once your data are entered they need to be "cleaned." This will not involve any liquid soap products. Instead, "cleaning your data" means fixing any problems that may have come as a result of the respondent or your data entry process. It may mean deciding what to do if a respondent did not complete the survey – do you drop the entire survey or keep the parts that were filled out? Are there questions or entire surveys that need to be dropped?

Cleaning your data also means taking a first look at the responses to see if any are "out of range" or raise questions. For example, a classroom teacher may indicate that she or he has 300 students in the program, but only a single class – clearly a typo. By conducting an initial review of the data (by eye, or by generating an initial set of frequencies and averages) you can identify problems or data questions before you start your final analysis and then take steps to correct them.

► Analysis of Pre and Post Surveys

There are two ways to analyze pre and post surveys. Either you can compare the overall change of all your student responses at once (Option A) or you can compare individual changes (Option B). Of course, Option B is a better option, but it is only available if you found a way to match the pre and/or post responses through the use of a code, student IDs or names (see Chapter Three for the pros and cons of each approach).

OPTION A

	Pre	Post	
Student A			
...Students B-Y...			
Student Z			
Total	**Total or average Pre**	**Total or average Post**	**Difference between totals or averages**

OPTION B

	Pre	Post	Difference
Student A			Difference between pre and post for A
...Students B-Y...			Differences between pre and post for B-Y
Student Z			Difference between pre and post for Z
Total			**Average of the differences**

For the *MKP* surveys, you can do this kind of analysis for each individual item or question on the survey (e.g., was there a change in student responses on question 1 from pre to post), or for groups of related items (did the average response for the items related to civic knowledge change from pre to post).

If your program is small (a single classroom, for example) you could do most of this analysis by hand, using a calculator. However, especially when there are more surveys, it is much easier to work with and analyze the data using the computer. There are a number of specialized statistical analysis programs (SPSS and SAS are among the most common). Your best bet for this type of analysis may be to find a graduate student or consultant who can help you run your statistics.

However, you can also do much of the analysis on your own using the tools that are included in a spreadsheet. The next section describes how to use some of the tools in the most popular spreadsheet software, *Microsoft Excel (MS Excel)*.

► Using the *MKP* Data Entry and Analysis Tool

The *MKP* Data Entry and Analysis Tool (Tool #11) provides you with pre-formatted spreadsheets that may be used for data entry and analysis of the *MKP* Surveys. The *MKP* Data Entry and Analysis Tool uses Microsoft Excel software. If you do not have access to Microsoft Excel, you may want to download a free software spreadsheet program ("Calc") found in

"OpenOffice" (http://www.openoffice.org/) that has the same features described below and is compatible with MS Excel.

The file contains the following spreadsheets:

1. ***Read This First*** is a welcome and overview to the *MKP* Data Entry and Analysis Tool.
2. ***Student Survey Data Entry*** is the page for entering the pre and post, elementary school surveys and middle/high school participant and comparison surveys. Although there are columns for all of these, there are some questions asked on the post survey that are not on the pre survey. Also, there are some questions asked of participants that are not asked of comparison students.
3. ***Student Domains Summary*** is a pre-formatted summary of the outcomes measured by the survey domains on a case by case basis.
4. ***Student Survey Data Analysis*** provides a pre-formatted pivot table (See below about pivot tables) that provides data about the outcome domains measured by *MKP* student surveys.
5. ***Community Survey Data Entry*** is the page for entering the results of the community partner surveys. Each row represents a survey. Each column represents a variable.
6. ***Community Survey Analysis*** provides a summary analysis of the community surveys.
7. ***Teacher Survey Data Entry*** is the page for entering the results of the teacher surveys. This is designed for school districts with multiple teachers using service-learning.
8. ***Teacher Survey Data Analysis*** provides a summary analysis of the teacher surveys.

If you are completely unfamiliar with using spreadsheets, you may want to ask someone for a brief introduction to MS Excel until you feel more comfortable navigating the spreadsheet.

► Advanced: Conducting your own Analysis Using MS Excel Data Analysis Tools

The following provides a set of technical instructions on how to use spreadsheet tools that are readily available within most spreadsheet packages, such as MS Excel. These are designed for analysis beyond the summary statistics that are pre-defined in the *MKP* toolkit. We recommend that you save a back-up copy of your original spreadsheet prior to attempting to use these features.

1. Data Entry

If you decide that you want to create your own survey you will need a different spreadsheet from the *MKP* Data Analysis tool in which to systematically analyze that data. MS Excel has a feature (Forms) that lets you set up a quick and easy form to help you enter your data in a spreadsheet. MS Excel data forms can work with a maximum of 32 variables. (A "variable" is a type of information that you are collecting information on. For example, age of students is a variable, whether the student finished the CSL course is a variable, and so forth. The steps are as follows:

a. Open a new spreadsheet, and type the variable names, which correspond to each of the questions across the top of the first row. These will serve as column headings.
b. Select/Highlight the variable names.
c. On the menu at the top, click on Data, then Form.
d. If you get a message asking if you want the first row to be used as labels, click OK.
e. You will see a data entry form with spaces for each variable.
f. Enter the data for each case (or survey), using the TAB key to move from space to space.
g. Hit enter to move to the next case.

h. We recommend saving your work as you go along. After every few surveys are entered, move your cursor to "File" and click "Save As" to save your data.

2. Data Analysis

Excel also has an easy-to-use set of sophisticated data analysis tools, as well as a number of specific data analysis functions that can be used in describing and analyzing data.

A. Data Analysis Tool

Excel has a standard 'add-in' that provides a number of statistical analysis functions, ranging from basic descriptive statistics to more complicated analyses. If you don't have access to a statistical package like SPSS, these Excel tools still let you do much—if not all—of the analyses that you will probably want to carry out.

To set up the data analysis 'add-in' for the first time, click on Tools/Add-ins, then click on the checkbox for 'Analysis Tool Pak.' And click on OK. This will install the Tool Pak.

To use the Analysis Tool Pak, click on Tools, then Data Analysis. You will see a list of statistical operations and can select one. Each then has its own 'wizard' to walk you through the process.

B. Basic Descriptive Statistics in Data Analysis Tool Pak

The analysis you are most likely to use regularly is the Descriptive Statistics tool, which provides the mean, median, mode, sum, standard deviation, number of cases, etc. – that is, most of the descriptive data you will need – for the data in any column that you highlight.

To use the Descriptive Statistics tool:

1. Click on Tools/Data Analysis/Descriptive Statistics. This will open a 'wizard' to help you set up your inquiry.
2. To identify the data you want to analyze, click once in the input range box to set your cursor there, then go to your spreadsheet and highlight the data you want summarized. You should include the column headings in the area you select.
3. Back on the wizard, click on the appropriate button to indicate that your data is organized in columns (that is, each column is a separate variable). Also click the box to indicate that your variable names/labels are in the first row of the spreadsheet.
4. Decide whether you want the output (results) to be placed on the same page as your data, or on a new page or new workbook. A new page is usually the easiest.
5. Then click on the box for Summary Statistics to get the mean, median, etc. for each column of data.
6. Click on OK and Excel will move you to a new page with the results. You may want to adjust the column width on the new page to see the results more easily (Format/Column Width – I'd suggest about 15 characters for a column).
7. Click on "save as" in order to save the file for future reference.

This procedure will provide you with the following information for each variable that you may want to put into your report(s):*

---

* *There are also more esoteric statistics available which you probably will not want to incorporate into your reports.*

ITEM	WHAT IT MEANS	HOW YOU CAN USE IT
Count	The number of people that we have information on for this variable.	"We got usable information from xxx of the xxx students in the course..."
Mean	The arithmetic average.	"The average age of the seniors was 18.5 years..."
Median	The number that would be in the middle if the numbers were listed in a row from the lowest to the highest.	"The average family income for the students in the class was $ 100,000. However, this number is distorted by the two students from families who are millionaires, and so it makes more sense to say that the typical student in the class came from a family with a median income of $ 53,000.
Maximum	The highest value on the list.	The SAT scores ranged from 781 to 1576.
Minimum	The lowest value on the list.	The SAT scores ranged from 781 to 1576.

C. Other Statistical Functions in the Analysis Tool Kit

The other statistical procedures in the Analysis Tool Kit work the same way. Once you select the procedure, you get a wizard that provides a place to select the data you want included and where you want the results displayed. Some of the other procedures include T-Tests (for comparing results from two groups such as pre and post surveys), correlations, analysis of variance and regression. With the exception of the T-tests, it is unlikely that you will want to/need to use any of them.* The T-test will allow you to determine if the difference between the two groups is "statistically significant", i.e., larger than one might expect to frequently occur if there were in fact no differences between the two groups.

---

* *The one type of procedure that the Analysis Tool Kit does not include is that for looking at relationships among certain kinds of data (e.g., Chi square tests, used for looking at categorical data in cross-tabs tables).*

D. Excel Functions

Excel also has individual 'Functions' that can be used to do lots of calculations that you may want/need for your reports. The most commonly used function is the SUM command, which gives you the total (sum) for the numbers in a column. Some other useful functions include:

- → AVERAGE – gives the mean average for the numbers in a specified group of cells. Average ignores empty cells and cells with text, so it gives an accurate result.
- → MEDIAN – gives the median for the numbers in a specified group of cells.
- → MIN and MAX – gives the highest or lowest values in a specified group of cells.
- → COUNT – tells you how many cells have numeric data in them (which can be used to get the 'number of cases' for a file.

To use a function, you can go through the menu (Insert/Function, then select the function you want) or you can type the function into a cell directly following an = sign. For example, to get the average for a group of cells:

1. Place your cursor on an empty cell where you want the result to appear.

2. Type the following: "=AVERAGE (range of cells)". The range of cells is the set of cells you want included in the calculation. In Excel, you type this as: beginning cell:ending cell. For example, to include the data in rows 2 through 8 in column C, you would type C2:C8. So, the whole formula would be: "=AVERAGE (C2:C8)".

OR

1. Click on Insert/Function, and then select Statistical Functions/Average.
2. A 'wizard' will appear. Click in the space for Number 1, then highlight the cells on your spreadsheet that you want included in the calculation (in this example, C2:C8). Then click OK. The wizard will insert the formula on your spreadsheet.

E. Frequencies

If you want to know how many cases fall into individual categories or ranges of values (for example, in a marketing survey, how many people report incomes between $0 and $25,000, $25,000 and $50,000, etc.), you can use Excel's FREQUENCY function.

To use Frequency, you need to provide two types of information: the range of cells where your data is found (called the data array), and the values or ranges of values that you want to count (called a BIN array – that is, you are sorting data into bins…).

To set up the Bin Array, simply go to an empty column and type in the break points or values you want to sort your frequencies by. For example, if you wanted to break incomes into four or five groups, you might go to an empty column and in Cell C2 type 0, C3 type 25000, C4 type 50000, and C5 type 100000. Leave C6 blank. This would let you sort incomes in your file into 5 groups, 0-25000, 25 to 50000, 50 to 100,000, and "Other" (i.e. the blank cell).

To run FREQUENCY take the following steps:

1. Highlight the cells where you want your answers to appear – usually the column next to the Bin array. Be sure to highlight at least one more cell than is in your Bin array (to have a place for the 'other' data to appear).
2. Click on Insert/Function and then click on "statistical" under the category, and "frequency" under the function on the menus.
3. Click once in the 'data array' space to locate your cursor there, and then highlight the cells in your spreadsheet where the data is located.
4. Click once in the 'bins array' space, then highlight the cells where you have put your Bin values (C2:C6).
5. To finish, you must press **SHIFT+CTRL+ENTER** to enter the formula. If you just click 'OK' the function will not work. That is because this is a special 'array' function.
6. The highlighted cells will then show the number of cases that fall into each category in your Bin array. You can use these counts as is (to say how many people fall into each group) or use them to calculate the percentage of the whole in each group.
7. Click on "save as ...." in order to save your file.

F. Pivot Tables

Excel's Pivot Tables function lets you develop simple cross-tabulations or "breakdown tables" that show frequencies or averages broken down into subgroups. You can use Pivot Tables, for example, to show differences in test scores for males and females or for students in different classes as is illustrated below.

SCORE	MALE	FEMALE	TOTAL
High Score	12	8	20
Low Score	4	8	12
TOTAL	**16**	**16**	**32**

To build a Pivot Table, you click on Data, and then select Pivot and Pivot Chart Report from the menu. A Pivot Table wizard will open and you can then:

1. Indicate the ***type of data*** you will be using (usually from an excel spreadsheet).
2. Indicate the ***location of the data*** on the spreadsheet by highlighting the cells you want included. Be sure to include the cells with the column headings or labels in the area you highlight.
3. Indicate ***where you want to put*** the new table (on the same page or a new one).

On the third step of the wizard, you will see a button for 'layout' and one for 'options'. Click on the 'layout' button to set up your new table.

To set up the layout of your table, you will 'drag' variables from the variable list in the wizard to either a row or column heading, or to the data section. The rows and columns sections determine how your data will be displayed, and how it will be broken down. The variables in the 'data' section determine what specific information will be displayed (counts, averages, etc.). You can click twice on the data buttons to change the data type.

CHAPTER FIVE

Using the Results of Your Evaluation

Although this is the shortest chapter in Making Knowledge Productive, it is also the chapter that we consider to be most important. It is here to ensure that you think about how to take what you learned and ***apply*** it to the improvement of your community service-learning program.

► Consider Your Audience(s)

Although you can report all information to all audiences, different groups of stakeholders may have different needs and interests. This is important to consider as you prepare to write and disseminate the results of your program evaluation. The table below provides some ideas for how to report your results to different audiences. (See Tool #12: "How to Use Data.")

AUDIENCE	HOW TO REPORT?
Massachusetts Department of Education Community Service-Learning and other funders	Reports as requested from the Department of Education, specific to your grant, as well as online LASSIE reports and complete excel data sets.
School board, administration and teachers	Summaries, reports and presentations.
Students and parents	Presentations, open houses, written summary reports and clean data sets for student analysis.
The community	Press releases, and newspaper inserts.

► Write It Up!

Although every report is different (we do not recommend 'cookie-cutter' software that will write a report for you), we do suggest including some of the following sections in your write-up of your program evaluation.

1. ***Provide a context.*** Describe what your program does and why it does it. Briefly, what is the history of your program? Who is involved? You may also want to explain why you decided to conduct the program evaluation. You may want to look back at your logic model to inform this section of your report.
2. ***Describe what you did (Your research and evaluation methods).*** Describe your approach to your program evaluation. Who participated and how? What methods did you use? Did you select a sample and if so how? Did you choose a comparison group and if so, how? You may want to look at your Evaluation Plan to inform this section of your report.
3. ***Present the facts.*** At some point in your report, you will want to present a summary of the data that you collected. Often this begins by presenting the frequencies of your results, perhaps in the form of tables or visual charts. In

addition, you may want to include some of the following:

- How many students participated in your program?
- Who were they?
- What percentage responded to the surveys?

4. ***Make Comparisons.*** After you have presented the basic data, if you have the time and need to do so, you may want to compare particular groups within your data. As noted in Chapter Four, one way that this can be done is to use a "pivot table" in Excel. This type of table allows you to display information on things like whether there was a difference in test scores between boys and girls? If you collected data from different ages or grades, were there differences amongst them? If you collected program evaluation data over multiple years, you may want to report the trends in this information over time. This provides more information to help people understand the impact of your programs in new ways.
5. ***Stories and Photos.*** It may be helpful to include more than numbers in your write-up. Stories and photographs may help draw attention and present your results in a more read-able interesting way.
6. ***Drawing Conclusions.*** Here you want to consider whether and how your evaluation results answered the question that you posed in your Evaluation Plan. At the same time, you should recognize the limitations of your study (what you can and cannot say) based on your methods. Every evaluation cannot answer every question. It is important to let the reader know any limitations that you recognize in your own results
7. ***Suggest next Steps.*** Finally, your report should move your group towards some type of action – even if it is only to propose questions that the Evaluation team or School Board might want to consider when you discuss your evaluation with them.

► Focus On Continuous Program Improvement

Community Service-Learning Programs evolve, and the best programs continually look for ways to improve. When the findings and conclusions begin to emerge, we suggest discussing themes with the key stakeholders in your program – including youth. Think about them. When you do, here are some ways that you might frame that discussion.

Consider whether to distribute the results to the key stakeholders or to present them in person. Young people, especially when they were involved in the evaluation data collection, can be great presenters of their information.

Be Skeptical! Begin by asking yourself how much you believe the results. What types of results seem to fit with your experiences? Which seem to contradict your experiences or observations? Can you think about reasons why the results might not be accurate or are biased? If so, we suggest that you acknowledge that or at least be aware of it. It is usually better to encourage self-criticism among your evaluation team members than to be surprised by outside criticism later.

Think about what these results mean. You may want to ask yourself, "What do these results mean for my program? What kinds of things appear to be going well? What kinds of things could be improved? Why?" Then ask yourself, what can you do about it?

► Spread the News!

If your key stakeholders agree, and when you are ready, there are many ways to disseminate the results of your program more broadly. Groups can write press releases, send out mailings to homes, write and distribute newspaper inserts, and/or hold an event or party. You can also attend conferences, such as the Massachusetts Department of Education Community Service-Learning annual conference (usually held in May). You and your evaluation team will know best how and with whom to share the news of your evaluation. Good luck and have fun!

► Conclusion

MKP has been built around several key assumptions. First, just as there is a cycle for student learning through service-learning that involves experience, reflection, planning and further action, so too, there is a cycle for the evaluation process. Further, as with service-learning, we believe the most effective evaluations involve your stakeholders in this process. This makes evaluation an effective use of resources (people and information) that you have at hand and helps your evaluation become a living, useful process that you will 'want to' do, instead of 'having to' do.

► ► ► *Finally, as your evaluation comes alive,*

remember, you should design it to meet your

needs — there is no one "right" approach.

The tools in MKP provide a template,

but you can — and should —

adjust them to your own needs.