From Logic Models to Program and Policy Evaluation

A Workshop Created by the Regional Educational Laboratory-Northeast and Islands

for the Northeast Educator Effectiveness Research Alliance and the Urban School Improvement Alliance

Karen Shakman, Ph.D.

Introduction to Logic Models to Support Program Design, Implementation, and Evaluation Session II: From Logic Models to Program and Policy Evaluation

What is the purpose of this work?

Beginning in 2012 the ten Regional Educational Laboratories have undertaken building and supporting research alliances on topics of importance to states or local school districts. Some examples of these topics include *educator effectiveness* and *supporting all students to become ready for college and careers*. Each of these alliances established a research agenda that includes questions that will guide their work over the next three to five years. The Regional Educational Laboratory Northeast and Islands (REL-NEI) has been involved in work with a set of eight research alliances that are intended to be sustained collaborations among researchers, administrators, policymakers, and practitioners. These alliances focus on a particular priority for the purpose of increasing state and local capacity to use data and research to inform decision-making in that priority area. We developed this workshop to build state and district leaders' capacity to design, implement, and evaluate programs and policies that address some of the most pressing educational issues.

Why the workshop?

This workshop (the workbook and accompanying slide deck) was developed to assist groups, like the alliances as a whole as well as members of the alliances within their own educational contexts, to learn about and build logic models to support effective program designs and evaluations. Based on feedback from alliance members, REL-NEI learned that many of our district- and state-based members would like to build their capacity to develop logic models that may be utilized to both evaluate their own programs as well as to work more effectively with evaluators whom they engage to conduct evaluations on their behalf. This workshop (a two-part series of which this is session 2) is designed to provide a primer on logic modeling, a useful tool for program design, implementation, and evaluation, and provide guidance in how to utilize logic models as a tool for program evaluation.

Table of Contents

From Logic Models to Program and Policy Evaluation (1.5 hours)

Annotated Agenda

7 Minutes	**Overview of Second Session**	Facilitator will review the goals of the session and the agenda.
8 Minutes	**Review Logic Models**	Facilitator will review what was learned about logic models, what they are useful for and what limitations they have. Guidelines for assessing logic models will be provided.
15 Minutes	**Introducing Evaluation**	Facilitator will re-introduce the types of questions that evaluation is designed to answer, the value of implementing evaluation at the onset of program development, and the role that logic models may play in supporting evaluation. Types of evaluation will be presented—formative and summative—and participants will practice generating questions that map onto these overarching categories of evaluation.
13 Minutes	**From Logic Model to Evaluation Questions**	Facilitator will begin this section with more about types of evaluation questions, followed by guidelines for good questions. Facilitator will then introduce the idea of different audiences desiring different information about a program or policy, and therefore asking different questions. Participants will be introduced to a table that delineates different types of audiences, questions, and uses of evaluation. The activity will have participants brainstorm their own audiences, questions, and uses of evaluative information.
25 Minutes	**Generating Indicators** **Identifying the Right Indicators** **Quantitative and Qualitative**	Facilitator will introduce the concept of indicators, and provide an overview of how indicators may be generated from the logic model, specifically from the activities and strategies and outcomes sections of the model. Facilitator will provide an example of this for the College Ready example. Then participants will practice generating indicators, based on one of the cases or their own example. This section closes with a discussion of qualitative and quantitative indicators and the use and value of both types of measures in an evaluation.
15 Minutes	**Building an Evaluation Design**	Facilitator begins this section with the question, "What type of evaluation is right for you?" and

		suggests that determining the purpose/s: formative or summative or hybrid, is critical to building an evaluation. Facilitator then transitions to more discussion about data collection, specifically considering the types of data available to participants. Considerations include what existing data may be available as well as what instruments have already been developed that may be relevant to the evaluation. Types of data, both quantitative and qualitative are reviewed. Then facilitator introduces the data collection framework tool, which outlines the outcomes of interest, data sources, responsible parties and timeline. Participants will practice filling in the data collection framework for their own example. This is followed by discussion of evaluation design, as distinct from a data collection framework, as it focuses on what type of evaluation is planned, who will be studied, how, and (importantly) what comparisons if any may be planned. A table is provided that presents the continuum of rigor of various designs.
6 Minutes	**Putting it all Together**	This final section of the workshop opens with discussion of an evaluation prospectus or overview, and the key questions to consider, when generating this short document, which can serve as the "calling card" for an evaluation, either for potential funders or for potential evaluators. The facilitator will close with presentation of a Gantt Chart as a useful tool for managing an evaluation and considering realistic timelines and deliverables.
2 Minutes	**Closing and Thank You**	Facilitator closes workshop with a thank you and invitation to be in touch with further questions.

Session II Purpose

The purpose of the Second Workshop is to demonstrate how logic models may be used as a tool specific to developing a program or policy evaluation. The workshop will:

- Reintroduce logic models as an effective tool, specifically for evaluation;
- Invite participants to practice using logic models to develop evaluation questions and indicators of success;
- Provide guidance in how to determine the appropriate evaluation for a specific program or policy.

Supporting materials for this section include:

- Slides: 3-4

Pre-Work Assignment

Based on the work in the first session, participants may come to the workshop with a draft logic model for a particular program or policy. If participants do not have their own draft logic model, they should both familiarize themselves with the sample logic models provided in the back of the workbook, as these will be drawn on for examples throughout the workshop, and come to the workshop with an example of a program or policy they would like to evaluate.

Directions: A sample logic model template is provided at the back of the workbook and may be used to generate a simple logic model. Participants will use this logic model to guide their work in the session.

Reviewing Logic Models

In Session I, we discussed the elements and logic of a logic model. Here are a few quick reminders about what a logic model is, and what it isn't. A logic model is:

- A graphic representation of the theory of change driving a program or policy;
- A framework for planning, implementation, and evaluation.

A logic model is not:

- A strategic plan;
- An evaluation design.

While a logic model is not a strategic plan or an evaluation design, it can be useful in developing either of these more detailed resources. The focus of this workshop is on the latter—how does a logic model support the development of an evaluation plan for a program or policy?

Case Examples Revisited

Case Study #1: College Readiness High School Program

College Ready is a school-based college access program for 9th-12th grade students. Students are identified for the program based on Free and Reduced Lunch status, recommendations from school guidance counselors, and/or recommendations from 8th grade English and Math teachers. Students participate in monthly meetings as a group with the College Ready staff, are provided with one-on-one counseling with College Ready staff, are assigned an adult mentor and a peer mentor, and participate in a series of evening and summer workshops. In addition, families make a commitment to the program and attend a series of workshops specifically designed to prepare the whole family for the college application process. The goal of the program is to significantly increase college attendance among the low-income students.

Case Study #2: Redesigning a District's Educator Evaluation Process

A school district wants to review and update the teacher evaluation process they have used for more than 10 years. The new system must reflect the new state guidelines for evaluation, which include a requirement for multiple measures, including a student learning measure. However, much is left to the district to decide about how decisions will be made, what measures to use, who will conduct the evaluations, and how the evaluation process will be managed and supported. The district has determined, in keeping with state guidelines, that the new evaluation will assess teachers' professional practice and their impact on student learning. The district leadership would like the system to be supported by teachers, and they would like it to effectively differentiate among teachers, support teachers' ongoing professional growth, lead to improvements in teacher practice, and ultimately positively influence student learning.

Guidelines for Reviewing a Logic Model

Consider the following questions when evaluating a draft logic model:

- Is the problem statement the "right grain size"?
- Within the strategies and activities, did you identify overarching strategies?
- What assumptions did you uncover?
- What is the timeframe for your outcomes? What are the impacts?
- What was your process for developing the model?
- What requires further explanation or discussion?

You may have created a draft logic model for a program or policy you are engaged in or considering implementing or evaluating. This logic model draft will serve as a template to guide your activities throughout this session.

Supporting materials for this section include:

- Slides: 5-10

Introducing Evaluation

Program and policy evaluation helps to answer important questions that inform our work. At a basic level, evaluation answers the questions: Are we successful? Have we had an impact? What exactly is making the difference?

More specifically, evaluation asks questions such as:

- Is the program or policy effective?
- Is the program or policy working as intended?
- What aspects of the program are working? What aspects are not working?

High-quality evaluation is designed to support your work, inform what you do, and enhance your impact. To do so, evaluation should be considered at the onset of program and policy design, ideally when the logic model is being developed. In other words, as a program or policy is conceived, evaluation of the same program or policy should be a part of the conversation. Questions like:

- How will we know if we're successful?
- What do we anticipate to be the impact of this policy?
- What do we think the most influential aspects of the program will be?

All of these questions suggest directions for evaluation. Do not wait until the program or policy is in the midst of implementation to begin to consider these questions and how to answer them. Invest early in considering these questions and designing an evaluation that will help to answer them. It may also be helpful to involve others, including staff and participants, in helping to plan the evaluation.

Activity I: How will I know?

Consider your own program or policy logic model. How will you know if one or more of your strategies have been successful? Take a moment to write down what might be some ways you will know your efforts have yielded the results you hope to achieve.

Directions: In Adobe, briefly type in some ways you know your program is achieving the desired results. For example: student test scores improve; student attendance goes up; college admissions increases; teachers report increased use of student data to inform instruction, etc.

Supporting materials for this section include:

- Slide: 11-13

Types of Evaluations

The purpose of this section is to provide an overview of the types of evaluations and their different purposes.

- ***Improve*** – These are formative, process, or implementation evaluations.

- ***Prove*** – These are summative, results, or outcome evaluations.

Most evaluation questions emerge out of the strategies and outcomes sections of the logic models. You want to know about the strategies you're trying and how they're going and you want to know about outcomes and impact.

Generally, evaluations that focus on strategies (and outputs) are **process evaluations**, or the evaluations that are designed to help guide changes or improvements to the program or policy. Those evaluations that focus on the outcomes in the logic model are generally **summative evaluations** or those designed to prove the value, merit, or impact of the program or policy.

Evaluations generally fall into four different categories of evaluation (but are not mutually exclusive):

- **Needs assessment**
  This type of evaluation determines what is needed (at the onset) and helps set priorities (e.g., is more money needed to support blended learning?). These types of evaluations are often designed in order to help to create or build a program or policy, so a logic model might be developed after the needs assessment. In fact, the needs assessment might provide information that helps to clarify the problem to which the program or policy is designed to respond.

- **Process/Formative evaluation** (IMPROVE)
  This type of evaluation is one that examines what goes on while a program is in progress. The evaluation assesses what the program is, how it is working, whom it is reaching, and how (e.g., are participants attending as anticipated?).

- **Outcome evaluation** (PROVE)
  This type of evaluation is designed to determine what results from a program, and its consequences, generally for the people most directly affected by the program (e.g., did participants increase their knowledge; change attitudes, behavior, etc.?)

- **Impact evaluation** (PROVE)
  This type of evaluation determines the net causal effects of the program beyond its

immediate results. Impact evaluation often involves a comparison of what appeared after the program with what would have appeared without the program. These evaluations often include comparison groups, interrupted time series, or other designs that allow evaluators to capture what did happen to the target compared to what would have happened without the program (e.g., achievement scores; acceptance rates; etc.).

Activity II: Formative and Summative Evaluation

Consider the different between formative and summative evaluation questions.
Directions: In Adobe, you are provided with the following multiple choice options. Select all answers that apply.

Question I: Which of the following questions would be asked while a college ready program is underway? (select all that apply)

(a) Does the college ready program engage families in the college application process?
(b) Is parent participation holding steady for the monthly workshops?
(c) Are students who attend the College Ready Program enrolling in post-secondary education?
(d) Do students who participate in the program maintain a 3.0 or above GPA in post-secondary education?

Question II: Which of the following questions would be asked to understand outcomes of a teacher evaluation system? (Select all that apply)

(a) Are teachers engaged in developing the new evaluation system?
(b) Does the evaluation system's multiple rating scale (unsatisfactory, needs improvement, satisfactory, exemplary) differentiate among teachers?
(c) Are principals completing all requirements of evaluation (e.g. number of observations, post-observation conference, assessing student learning objectives, etc.)?
(d) Has student achievement improved as a result of the new teacher evaluation system?

Supporting materials for this section include:

- Slides: 14-15

From Logic Models to Evaluation Questions

The purpose of this section is to make the connection between the logic model and development of appropriate evaluation questions, using the logic model as a basis for developing these questions. The first step in making the transition from the logic model to a potential evaluation is to consider the questions that are derived from the model, that you may want answered.

Developing Evaluation Questions

As noted above, in the activity, some questions ask about improvements to the program or policy (Formative/Process/ Implementation/Improve questions), while others ask about the results or impacts (Summative/Outcome/Prove questions). Generally:

- **Formative questions** are asked while the program is operating and are for the purpose of program improvement;
- **Summative questions** are asked at completion or after the program and are for the purpose of determining results and assessing value.

Regardless of the type of questions, there are some guidelines to consider for all evaluation questions.

- Can the question be answered given the program? One of the main reasons for building a logic model as part of program evaluation is to determine what questions are appropriate based on the program. By describing what the program is, the logic model helps determine what is appropriate to evaluate.
- Are the questions high-priority? Try to distinguish between what you need to know and what might merely be nice to know. What are the key, most important questions? For whom? Why?
- Are the questions practical and appropriate to the capacity you have to answer them? Consider time, resources, and the availability of assistance needed to answer the questions. As appropriate, bring stakeholders together and negotiate a practical set of questions. Remember, it is better to answer a few questions thoroughly and well.
- Are the questions clear and jargon free? Apply the "Great Aunt Lucy test." Would someone, like your Aunt Lucy or anyone who is not steeped in the language of your particular field understand the question? Avoid the use of jargon or vague words that can have multiple meanings. Always define key terms so that everyone understands the meaning.

Activity III: Formative and Summative Evaluation Questions

Try it yourself. Come up with a formative and summative evaluation question for a program or policy from your own work.

Formative Example: Does the College Ready Program engage parents in the college application process? OR Are principals completing all requirements of evaluation (e.g. number of observations, post-observation conference, assessing student learning objectives, etc.)?

Formative evaluation:
Topic: _____
Question:

Summative Example: Do students who attend the college ready program maintain a 3.0 in the first two semester of post-secondary education? OR Has student achievement improved as a result of the new teacher evaluation system?

Summative evaluation
Topic: _____
Question:

Slides: 16-18

Considering the Audience

Another key aspect of developing good evaluation questions is to consider different audiences, or the different stakeholders for a program or policy, the different types of questions they might have, and how they would use the answers to these questions (e.g. what decisions would result from answers).

This sample chart outlines some traditional audiences, the types of questions they are likely to have, and how they might apply answers to these questions to make decisions. (Source: Kellogg Foundation Logic Model Handbook)

Audience	Typical Questions	Evaluation Use
Program Staff	Are we reaching our target population (e.g. high school students; low-income families with pre-school age children)? Are participants in the program engaged? Satisfied? Is the program being run well? How can we improve the program?	Day-to-day program operations; changes in program design and delivery
Participants	Did the program help me? Help others? How could the program better serve my needs? How could I get more out of the program?	Decisions about participation/value to them
Public Officials	Who does the program serve? Is it reaching the target population? What difference does the program make? Are participants engaged and satisfied with the program? Is the program cost-effective?	Decisions about support, commitment, funding, scale-up/duplication
Funders	Is the program meeting its goals? Is the program worth the cost?	Decisions about ongoing funding; accountability

Activity: Generating Questions for Different Audiences

Think about your own context and consider:
(1) Audience: Who are the different members of each stakeholder group (staff, participants, etc.)?
(2) Questions: What questions might different stakeholders have about the program or policy?
(3) Use: How might these different stakeholders use the answers to these questions?

Note: We will not complete this activity during the webinar but invite you to complete this table as you begin to consider an evaluation of a program or policy.

Audience	Questions	Evaluation Use
Program Staff:		
Participants:		
Public Officials:		
Funders:		

Supporting materials for this section include:

- Slides: 19-20

Generating Indicators

As we begin to develop evaluation questions, we must consider how we will know we've achieved the goals, and therefore answered our questions of interest. In this section, we demonstrate how the logic model can support generation of good indicators of program or policy success.

If we are asking "Is the program successful?" we need to have some understanding of how we measure "success." In other words, we need to answer the question: "How will we know we're successful?" The logic model provides some support in helping to answer this question.

Starting with the logic model's outputs and outcomes, we can develop indicators that help us to answer the question, "how will we know we're successful?" Indicators are different from the outputs or outcomes you include in your logic model – while the outputs or the outcomes are more general goals for program implementation or outcomes, indicators are specific, measurable targets related to the outcomes of interest.

In short, indicators are:

- Specific, measureable targets;
- Seen, heard, read, felt;
- Connected to strategies, activities, outputs, and outcomes;
- Evidence representing phenomenon of interest (e.g. the outcome).

For example, if the outcome is increased parental engagement, then the indicator is a specific percentage of parents engaged or a specific increase in number of parents engaged. It is these measurable indicators that lead eventually to answer the question, "Is the program successful?"

Indicators can be seen, heard, felt, and/or read. They are tangible results.

Here's an example: How do we know a child has the flu? We feel her forehead for fever, listen to her sniffles, notice her lethargy, recognize a drop in appetite, and eventually take her temperature. All of these are indicators of the flu. They do not mean that the child *absolutely* has the flu but they do provide specific and measurable evidence that suggest the flu.

Similarly, the indicators do not absolutely mean that a policy or program is responsible for the results we measure. To use an example from our College Ready case, a parental engagement program may not be responsible for the rise in college applications among student participants. There might be other factors such as a decline in college costs or a particularly influential teacher at the school encouraging and supporting applications, that leads to an increase in the number of college applications submitted, but this increase in college applications could *reasonably be attributed* to a program that works with the students and their parents to support college readiness.

Supporting materials for this section include:

- Slides: 22-24

Using the Logic Model to Generate Indicators

Just as the logic model follows the basic format from Inputs (resources) to Outputs (strategies & activities, evidence of participation) to Outcomes (from short-term to long-term to impact), we follow this same logic to generate indicators.

As stated above, indicators are related to the logic model categories of resources, strategies, activities, and outcomes/impact. They go a step further, and provide clear numbers or percentages, when appropriate, associated with these resources, activities, outputs, or outcomes. They provide measurable evidence of the phenomenon of interest represented by the outputs or outcomes.

- Indicators related to Inputs are those that provide information about the resources used, the timeliness of the resources, the relevance of the resources (whether tangible or intangible). Indicators related to these inputs may help to answer questions about impediments or facilitators of implementation;
- Indicators related to Outputs are those that capture the numbers or percentages of workshops presented, the numbers of participants, and other data that provide information about whether the program was implemented as intended. Did it do what it set out to do? Did it reach the right people?
- Indicators related to Outcomes or Impacts are those that provide data about the results of participation, such as changes in knowledge, skill, behavior, and attitudes among individuals or groups targeted by the program or policy.

For example, if the strategy or activity is to deliver a parent education class, an indicator related to that activity might be the number of classes delivered or the number of parents who attended (these are process-related indicators). If an outcome is parents' increased understanding of the college application process, an indicator would be the number or percentage of parents reporting increased understanding (outcome indicator).

- Process indicators have to do with the extent to which strategies and activities are implemented as intended.
- Outcome indicators have to do with the extent to which the strategies and activities yield the desired results.

When generating indicators based on the various elements of the logic model (inputs, outputs/activities & strategies, and outcomes), ask yourself these basic questions:

- What would achieving the goal reflected in the outcome look like?
- How would we know it if we achieved it?
- If I were visiting the program, what would I see, hear, or read that would tell me that the program is doing what it intends?

Activity V: From Logic Model to Indicator

Map a path from an activity in your logic model, to an output, to an outcome, to an indicator.

Example: College Access Program Case

Activity	To deliver a series of parent workshops on college readiness.
Output	Six workshops developed and delivered; 100 parents recruited to participate.
Outcome	Parents increase their understanding of college application process
Indicator	*Process indicator*: (1) 70% of parents attend at least 5 of the 6 workshops (2) 85% of parents who attend more than 4 workshops report increased understanding *Outcome indicators:* (1) 80% of students in program complete at least 1 college application by deadline (2) 75% of students in program complete first semester of college and report intention to continue.

Your turn: Educator Evaluation Case

Directions: In Adobe, enter your suggestions for a process and/or an outcome indicator for the example below.

Activity	Develop and implement evaluation training for evaluators
Output	Develop and deliver two summer and two fall workshops
Outcome	New evaluation system piloted with evaluators in 2 elementary schools & 1 high school in 2014-15 school year.
Indicator	*Process indicators:* *Outcome indicators:*

Later, try completing this activity for your own program or policy:

Activity	
Output	
Outcome	
Indicator	*Process indicator*: *Outcome indicators:*

Supporting materials for this section include:

- Slides: 26-33

Identifying the Right Indicators

Different indicators are related to different types of questions.

- For example, if we are interested in whether the program has sufficient resources or funding to operate, we would look at indicators related to the program inputs.

- If we would like to know whether the program was implemented as intended, we would look at indicators related to the activities and strategies.

- Finally, if we are interested in the ultimate value and impact of the program, we would look at the outcome-related indicators.

Some indicators may be more straightforward or easier to measure than others. Sometimes one indicator is all that is needed for a clear explanation. For example, school graduation rate might be the agreed upon indicator for the outcome of decreasing the school dropout rate. In other cases, more than one indicator is necessary to capture a more complex outcome.

For example, if the outcome of interest is improved parental involvement in school, then several indicators may be necessary, such as:

- Attendance at school meetings,
- Participation in parent-school organization,
- Parental calls made to the school,
- Attendance at school functions, and so forth.

Quantitative and Qualitative Indicators

Indicators may be quantitative or qualitative. Given the interest in and demand for measurable outcomes, often our evaluation questions focus only on outcomes and the quantitative measures associated with these outcomes. Remember, however, that to attribute these quantitative outcomes (such as graduation rates or improvements on standardized tests) to your program, you also need to ask questions about the process that contributed to those outcomes. In addition, there are some outcomes that are best measured with a mixture of both quantitative and qualitative data.

It may be helpful to revisit the types of evaluations—those that focus on how to improve a program or policy (formative) versus those that focus on proving the value or impact of the program (summative).

- Quantitative data may be best suited to summative evaluations, e.g. information related to proving the value or impact of the program or policy, and

- Qualitative data may be better suited to formative evaluations, e.g. those that focus on how to improve the program.

However, this is not to suggest that all formative evaluations should be qualitative and all summative evaluations should be quantitative. Often, a mix of measures is the best approach. The qualitative data we collect, via interviews, observations, and other methods, often provides the depth of information we need to help interpret quantitative data such as test scores or graduation rates. Often, we want to know both whether a program or policy had the desired impact and how. Thus, a mix of measures is advisable.

It is often a good idea to collect different types of data (e.g., quantitative and qualitative) from several sources (e.g., surveys, interviews, grades, etc.), and from different groups of stakeholders (e.g., students, parents, mentors, staff, partners such as schools, etc.). While some funders may prefer quantitative data on outcomes, others, such as staff or parents, may prefer qualitative data from students or parents. Utilizing a range of qualitative and quantitative data sources will help to paint a full picture of the program or policy and will provide useful information for different audiences.

For example, in our College Ready case example, if we are interested in knowing whether the program increased student interest in college, then the indicators might include both quantitative and qualitative data to indicate this (# of applications completed AND guidance counselors reporting on student interest).

Activity VII: Identifying Quantitative and Qualitative Indicators

Consider the educator evaluation example. If you could only select one quantitative and one qualitative indicator to measure the question below, what would you choose?

Did the new evaluation system contribute to improvements in teacher accountability for student performance?

(a) X% of teachers report in teacher survey that they feel increased accountability for student performance.
(b) Student performance on standardized assessment improves.
(c) Establishment of data teams in all pilot schools.
(d) X% of student learning objectives are met in all schools.
(e) Interviews with evaluators indicate that teachers are taking more responsibility for student achievement.

Directions: In the poll, please indicate your selection.

Supporting materials for this section include:

- Slides: 34-44

Later, as you think about your own programs, you may think of both quantitative and qualitative indicators of success. Go back to the evaluation questions you generated for different audiences, earlier in this workshop on page 16. Identify two questions and brainstorm both quantitative and qualitative indicators for each of those questions.

Evaluation Question:	
Quantitative Indicator	Qualitative Indicator
Evaluation Question:	
Quantitative Indicator	Qualitative Indicator

Remember…Indicators:

- match the outcomes of interest/questions asked;
- may be singular or several for a given outcome or question;
- may be quantitative or qualitative;
- vary based on the audience.

Supporting materials for this section include:

- Slide 45

Building an Evaluation Design

The purpose of this section is to provide some tools for building an appropriate evaluation design.

Once you have generated a good logic model, come up with the best evaluation questions, and considered what you believe will be the best indicators of program or policy success (either for purposes of program improvement or to make the case for program impact), then you are ready to spend more time building the evaluation design.

What type of evaluation is best for you?

The first question to answer about evaluation design is quite basic: is the **purpose** of the evaluation to examine process elements (e.g. formative; improve) or to examine overall impact (e.g. summative; prove), or is it a hybrid evaluation, with some of each? Answering this question should help to clarify what is in and what is out when it comes to the evaluation you will conduct.

You should also return to the question of **audience** raised above. Who is the audience for your evaluation, what do they want to know, and how will the information be used? Finally, there is the issue of **capacity**: who will conduct the evaluation, using what resources, and within what timeframe? Assuming there are some financial, time, or other capacity constraints, what are your **priorities**?

Identifying Appropriate Sources of Data

When choosing measures for program evaluation, think about the data collection needs as well as data collection capacity.

- Access to pre-existing data. Consider collecting pre-existing sources of data, such as school attendance records or items from surveys that a school district already requires students to take, that will meet your evaluation needs.
- Utilize existing instruments. When the data you need cannot be collected via existing sources, look at existing instruments that measure the same concepts that you are looking to measure. These instruments may be effective as written, you may simply need to make a small tweak, or you may only need to adopt a few items from the entire instrument. It may even be possible to add these instruments or items to existing surveys currently being administered in your school or district.

Consider the types of data sources that might serve as indicators of success, both for process-related questions and for outcome-related questions.

The following are measurement techniques that may be useful for collecting process and/or outcome data:

- Administrative Data: (program documents, activity logs, registration records, etc.)
- Focus groups
- Interviews
- Observations
- Surveys
- Student grades
- Test scores
- Teacher assessments
- And more…

Activity: Consider Data Sources

Consider the data sources available to you in your program? What relevant data sources do you already collect?

Note: We will not complete this activity during the webinar but encourage you to brainstorm your potential data sources as part of your planning for evaluation.

Data source brainstorm:

Supporting materials for this section include:

- Slides 46-48

Creating a Data Collection Framework

In addition to identifying potential data sources, the details of data collection must be ironed out early in planning for the evaluation. It is important to develop a clear sense of what types of data are available, who collects and manages the data, and when it is collected. Developing a data collection framework, linked to the activities and outcomes of interest, will help to guide the evaluation process.

In some cases, the data collection framework may look different for process evaluation and outcome evaluation. However, in many cases, an evaluation will include some process, or formative, and some outcome, or summative components. As such, these may be combined into one framework, like the first table you see below.

Example: College Ready Program

Activity / Strategy	Output or Outcome	Formative	Summative	Indicator	Data Sources	Data Collection Instrument	When Collected	By Whom
Parent Education Strategy	High rate of parent attendance at workshops	x		70% of parents attend at least 5/6 workshops	Administra-tive Data	Attendance log at workshops	At beginning of each session	Program Director
Parent Education Strategy	Increased parent understanding of college application process		x	85% of parents who attend >4 workshops report increased understanding	Parent feedback	Survey; interviews	Beginning of program; End of program	Program Staff

It may be useful to make a distinction among short-term, long-term and impact data when creating an *outcome-specific* data collection table. The table below provides that breakdown. The reason for doing this may only be relevant to some evaluations, depending on the depth and duration of the evaluation plan. However, linking back to the logic model, distinguishing between

the short-term and long-term (and impact) outcomes in the logic model should help to guide the types of data that ought to be collected (and when) to reflect those outcomes. For example, if changes in student test scores are not anticipated until a program has been up and running for 3-5 years, then this data should not be collected (except to serve as a baseline, or point from which change will be measured) until the long-term phase of data collection.

Activity/ Strategy	Outcome	Indicator	Data Sources	Data Collection Instrument	When Collected	By whom
Short –Term						
Parent Education Strategy	Increased parent understanding of college application process	85% of parents who attend >4 workshops report increased understanding	Parent feedback	Survey; interviews	Beginning of program; End of program	Program Staff
Long-Term						
Impact						

These tables are meant to serve as a tool for you and may be combined, modified, and amended to serve your needs.

Activity: Data Collection Framework

Try it yourself. Consider your program or policy and generate a data collection strategy.

Note: We will not complete this activity during the webinar, but encourage you to develop a chart like this as part of your evaluation planning.

Activity/ Strategy	Outcome	Indicator	Data Sources	Data Collection Instrument	When Collected	By whom
Short –Term						
Long-Term						
Impact						

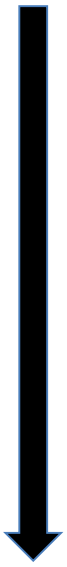
Supporting materials for this section include:

- Slides: 49-53

The Continuum of Rigor in Outcome/Impact Evaluation Designs

An **evaluation design** indicates when data collection will occur, which people or groups will be studied, how they will be selected, and what comparisons to other people or groups, if any, will occur. This is different from the data collection chart above, which is focused solely on the methods (e.g. surveys, interviews, etc.) and the logistics of managing the collection of data.

With regard to outcome evaluations, the evaluation design you select informs the level of confidence you can have in the relationship of your program or policy to the change that you observe. In other words, the design you choose, and the relative rigor of the design, will inform how convinced you and others will be that your program is responsible for the change observed in outcome measures such as test scores or college acceptance rates. The chart below outlines some of the different designs that may be employed in an outcome evaluation. These designs inform both the data collection phase and the analyses of this data.

Evaluation Design	Example	Type of Questions the Evaluation Can Answer	Rigor
Matched comparison group	Schools are selected to implement a new program through some nonrandom process (e.g., they volunteer). Before the program begins, these schools are matched on important background characteristics (e.g., student demographics and average test scores in the prior academic year) to other nonparticipating schools. After the program has been implemented in the participating schools, the outcomes for the two groups of schools are compared to estimate the program's effect.	Did outcomes differ between the matched groups of participating and nonparticipating schools?	**Least Rigorous Design** (Lowest confidence that results can be attributed to program) 
Comparative interrupted time series	Schools are selected to implement a new program, again through a nonrandom process. Before the program begins, these schools are matched to comparison schools with similar ***histories*** of background characteristics and outcomes. After the program has been implemented in the participating schools, ***trends*** in outcomes over time are compared with the outcomes for schools that "just missed" being selected.	Did outcomes in the program schools improve more than would be expected ***given trends*** in similar nonparticipating schools?	
Regression discontinuity *(**NOTE**: NCEE is*	Schools are selected to implement a new program based on a predetermined "cut point" on a well-defined and easily measured criterion	What is the impact of the program on outcomes? Or, are outcomes in program	

not accepting this method for questions of impact.)	(e.g., proficiency rates below 25 percent). The outcomes for participating schools are then compared with the outcomes for schools that "just missed" being selected.	schools different than they would have been absent the program?	**Most Rigorous Design** (Highest confidence that results can be attributed to program)
Random assignment	A set of schools is selected to implement a pilot program based on a random process (e.g., a lottery is used to select 20 pilot schools from among interested volunteers statewide). At the end of the pilot implementation period, the outcomes for pilot schools are compared with the outcomes for the other interested non-participating schools.	What is the impact of the program on outcomes? Or, are outcomes in the pilot schools different than they would have been absent the program?	

From: Perez-Johnson, Irma, Kirk Walters, Michael Puma and others. —Evaluating ARRA Programs and Other Educational Reforms: A Guide for States. Resource document developed jointly by The American Institutes for Research and Mathematica Policy Research, Inc. April 2011.

For an overview of additional designs, and some guidelines for increasing strength of the evaluation, here's a resource: http://www.uwex.edu/ces/pdande/resources/pdf/Tipsheet36.pdf

Supporting materials for this section include:

- Slides: 54

Putting it all together

Now that you have a logic model, a set of measurable indicators, some evaluation questions, a data collection framework, and at least some idea about evaluation design, you are nearly ready to proceed with evaluation. What you still need is the overall evaluation plan, a timeline, and a budget. We provide some tools here to help you do just that.

At this point, you have learned about all the steps necessary to get you to the evaluation. In brief, you have learned about how to:

- Develop a logic model, in collaboration with stakeholders;
- Clarify who the audience is for the evaluation and how it will be used;
- Identify and prioritize evaluation questions, based on the logic model;
- Select appropriate indicators, based on the outcomes of interest;
- Identify data sources and a data collection plan (including considering whether the evaluation is serving a formative/process and/or a summative/outcome goal); and
- Consider evaluation design, with awareness of resources, capacity, and timeline.

Once you have completed all these steps, it is still a good idea to develop an evaluation prospectus that lays out much of this information in clear, narrative form.

Evaluation Overview/Prospectus

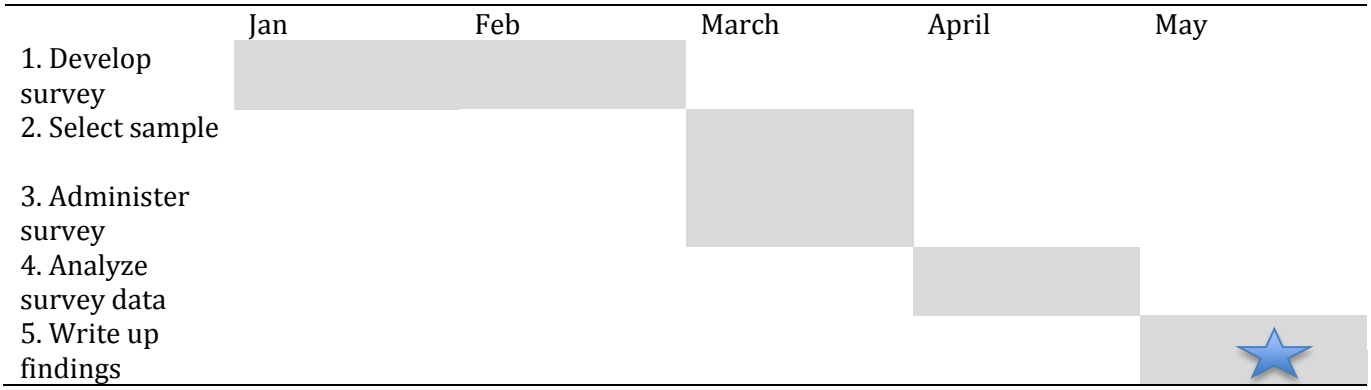
If you are going to look for an external evaluator, or even if you are going to do the evaluation in house, it is wise to have an evaluation overview or prospectus, to accompany all the other materials you are generating. This prospectus should provide a clear and straightforward answer to the following questions:

- What are you going to evaluate?
- What is the purpose of the evaluation?
- How will results of the evaluation be used?
- What specific questions will the evaluation answer?
- What data sources will be necessary to answer these questions?
- How will the data be analyzed (e.g. evaluation design)?
- What resources are needed to conduct this evaluation?
- What is the timeline for the evaluation?
- How will the results be shared/disseminated?
- Who will manage the evaluation?

Example of Gantt Chart

Having done all these things, one more tool may be helpful in implementing the evaluation plan. Often, evaluators work with Gantt charts. These are a kind of timeline, displayed in such a way that readers can immediately see a proportionate, chronological account of the time for each evaluation task.

- The vertical access shows the steps to be completed,
- The horizontal access shows the time scale.
- A horizontal line is drawn for each task to show how long it will take.
- Milestones are denoted with a symbol.



Supporting materials for this section include:

- Slides: 55-57

Closing and Thank You

In closing, if you look back at the beginning of these workshops, when we started with the logic model, we noted that logic models are a very useful tool for program design, implementation, and evaluation. They should be living documents that are referred to throughout the life of the program and the evaluation, and amended as needed. They are also helpful to guide a program as it evolves, and help to ensure that the work of the program remains focused on the key goals and outcomes.

Logic models are useful for program evaluation, especially when evaluation is designed in concert with the logic model. It is much better to consider evaluation at the outset of a program or policy's development rather than as an afterthought or halfway through program implementation.

Finally, engaging key voices, staff, parents, students, funders, and others, in discussions about program design and evaluation will promote the buy-in and ongoing support of these participants as well as increase the authenticity of the model and the evaluation.

Supporting materials for this section include:

- Slides: 58-60

Good luck!

References and Resources

There are many terrific resources available online for logic modeling and program and policy evaluation. Many of these were used in the development of this workshop. Several of the resources below provide links to additional resources, also available online.

Logic Model Resources

University of Wisconsin Extension School: http://www.uwex.edu/ces/lmcourse/#
The University of Wisconsin Extension School has several useful resources related to logic models and evaluation.

W.K. Kellogg Foundation Logic Model Development Guide. Retrieved 3/16/2013 from http://www.wkkf.org/knowledge-center/resources/2006/02/wk-kellogg-foundation-logic-model-development-guide.aspx

Harvard Family Research Project and Coffman, J. (1999). Learning from Logic Models: An Example of a Family-School Partnership Program. Retrieved 3/16/2013 from http://www.hfrp.org/publications-resources/browse-our-publications/learning-from-logic-models-an-example-of-a-family-school-partnership-program

Innovation Network, Inc. (date unknown). Logic Model Workbook. Retrieved 3/18/2013 from *http://www.innonet.org/client_docs/File/logic_model_workbook.pdf*

Evaluation Resources

The Administration for Children and Families, Office of Planning, Research, and Evaluation (2010). The Program Manager's Guide to Evaluation, Second Edition. Retrieved 3/16/2013 from http://www.acf.hhs.gov/programs/opre/resource/the-program-managers-guide-to-evaluation-second-edition

Bond, S., Boyd, S., and Rapp, K. (1997). Taking stock: A practical guide to evaluating your own Programs. Horizon Research, Inc. Retrieved 3/16/2013 from http://www.gao.gov/products/GAO-12-208G http://www.horizon-research.com/reports/1997/taking_stock.php

Centers for Disease Control (August, 2012). Implementing Evaluation: Learning and Growing Through Evaluation. Retrieved 3/16/2013 from http://www.cdc.gov/asthma/program_eval/guide.htm

Chinman, M., Imm, P., and Wandersman, A. (2004). Getting To Outcomes 2004: Promoting Accountability Through Methods and Tools for Planning, Implementation, and Evaluation. Rand Corporation: Santa Monica, CA. Retrieved 3/16/2013 from

http://www.rand.org/pubs/technical_reports/TR101.html

Patton, M. Q. (1997). Utilization-Focused Evaluation: The New Century Text. SAGE Publications: Thousand Oaks, CA.

United States Government Accountability Office (January 2012). Designing Evaluations. Retrieved 3/16/2013 from http://www.gao.gov/products/GAO-12-208G

W.K. Kellogg Foundation Evaluation Handbook. Retrieved 3/16/2013 from http://www.gao.gov/products/GAO-12-208G: http://www.wkkf.org/knowledge-center/resources/2010/w-k-kellogg-foundation-evaluation-handbook.aspx

The CDC has a range of evaluation resources, some produced at the CDC and some from other sources: http://www.cdc.gov/eval/resources/index.htm

These checklists, created by The Evaluation Center at the Western Michigan University may be useful in planning and monitoring evaluation: http://www.wmich.edu/evalctr/checklists/

The author would also like to acknowledge the following people for their expertise and resources in developing these materials:

Katrina Bledsoe, Education Development Center, Inc.
Leslie Goodyear, Education Development Center, Inc.
Brian Lord, Education Development Center, Inc.
Anne Wang, Education Development Center, Inc.

Appendix A: Simple Logic Model Template

Problem Statement:

Resources	**Strategies and Activities**	**Outputs**	**Short-term Outcome**	**Long-term Outcomes**	**Impacts**
What resources are or could reasonably be available?	*What will the activities, events, etc. be?*	*What are the initial products of these activities?*	*What changes are expected in short-term?*	*What changes do you want to after the initial outcomes?*	*What are the hoped for changes over the long haul?*

Assumptions:

Appendix B: College Ready Logic Model

Problem Statement: Low-income high students in selected communities attend college at a lower rate than their middle class peers, leading to more limited opportunities, higher rates of unemployment, and lower earnings.

Resources	**Strategies and Activities**	**Outputs**	**Short-term Outcome**	**Long-term Outcomes**	**Impacts**
What resources are or could reasonably be available?	*What will the activities, events, etc. be?*	*What are the initial products of these activities?*	*What changes are expected in short-term?*	*What changes wanted after initial outcomes?*	*What are hoped for changes over long haul?*
-Partnership with 3 public high schools -Community mentors -Local university space for parent meetings -Volunteer college admissions directors for application workshop -Student volunteers for childcare at parent meetings	-Local college mentorship program -Peer mentors -Student readiness program (workshops) -Parent education (workshops)	-Recruit adequate # of mentors for student cohort -Develop and deliver 12 workshops on college application process; SAT/ACT; FAFSA; college life -Develop and deliver 6 workshops for parents -High interest and attendance at all workshops for parents and students.	-Participating students apply to at least one college on time -Parents report increased understanding of the college application process -Students report increased readiness for college -Participating students complete FAFSA forms on time	-Participating students are accepted to and attend college, remaining enrolled into the 3rd semester of college -Participating students GPAs above 3.0 at college, into the 3rd semester -Increased parental engagement in participating high schools' students education	-Low-income students in participating communities attend college at same rate as middle class peers -Low-income students in participating communities graduate from college at some rate as middle class peers -Participating high schools see increase in parent and student engagement -Participating high schools state test scores increase by x%

Assumptions: College attendance is desired goal for participating communities; high school leaders will remain consistent and support program; parents will show interest and participate in program.

Appendix C: District Educator Evaluation Logic Model

Problem Statement: The district's existing evaluation system is outdated and not compliant with new state guidelines. The current system does not accurately differentiate among teachers nor provide teachers with actionable feedback. The system does not link feedback to appropriate professional support, nor are there any connections between teachers' practice and student outcomes.

Resources	**Strategies and Activities**	**Outputs**	**Short-term Outcome**	**Long-term Outcomes**	**Impacts**
What resources are or could reasonably be available?	*What will the activities, events, etc. be?*	*What are the initial products of these activities?*	*What changes are expected in short-term?*	*What changes wanted after initial outcomes?*	*What are hoped for changes over long haul?*
-Staff meeting time -Race to the Top funds -State evaluation resources -Teacher support -Strong administrator-teacher relations -Current teacher evaluation system -Existing professional climate task force	-Develop and implement evaluation trainings for evaluators -Develop and implement evaluation trainings for teachers -Develop evaluation handbook -Research and select online platform for managing evaluation data -Develop cross-role committee to develop system and monitor implementation.	-Recruit 12 teachers, administrators, school board member, etc. to serve on evaluation committee -Draft new evaluation model by March; reviewed by climate task force and school board by June. -Develop/deliver 2 summer & 2 fall workshops for evaluators. -Develop and deliver 1 summer & 1 fall workshop for teachers, focusing on utilizing student data.	-New evaluation model utilized in 3 pilot schools in 2014-15, low-stakes. -Online platform piloted in 3 schools -Feedback from teachers/evaluators indicate overall positive experience with new model.	-Modifications to new evaluation based on experience of 3 pilot schools. -New evaluation model rolled out to all schools in district in 2015-16. -Teachers and administrators report positive climate in district climate survey -Participating teachers report increased use of student data in evaluation discussions with principal/evaluator.	-All components of new evaluation system, incl. student data, in use in all schools. -Teachers and administrators identify evaluation model as element of successful school climate. -Student achievement on standardized assessments.

Assumptions: Administrator and staff support of the new evaluation system is high and will be sufficient to implement the system as intended. Implementation of a new evaluation system that evaluates teacher practice and student outcomes will lead to more accurate assessments of teachers performance, and will yield positive results for professional practice and climate, and ultimately for student achievement. .