

...But How Do We Know? (Part I)

RILS Day 1, 3:30 - 4:30 p.m. Session
Portsmouth, NH

Outcomes and the Steps to Get There



Outcomes and the Steps to Get There

To catch a mouse...

We know the outcome, but
do we know the steps to get
there?

If we know the steps, do
we know if the steps are
working?



Consider the following Problems

Problem 1


Unfinished Learning/
Gaps in Content
Knowledge

Problem 2

Students are
Disengaged and Feel
Emotionally
Exhausted

Problem 3

Insufficient
Information to Inform
Instructional
Response

- 
- Why is this problem occurring? What are some root causes of it?
 - What might be a potential solution to this problem? How would you implement it?
 - How do you know if you're making progress on this problem?

Attacking the Problem

Please visit the following link: [Problem Exploration Sheet](#)

1. Select a problem from the list (or create your own if you're feeling ambitious!)
2. Specify why this problem is occurring
3. Identify a potential (and plausible) solution as to how you may want to address the problem
4. Identify pieces of evidence or information that help inform whether you're making progress in addressing your problem of practice

We will return to this problem and solution tomorrow...

Consider the following Problems



<https://tinyurl.com/nhecpaxf>



Consider the following Problems

Problem 1

Unfinished Learning/
Gaps in Content
Knowledge

Problem 2

Students are
Disengaged and Feel
Emotionally
Exhausted

Problem 3

Insufficient
Information to Inform
Instructional
Response

- Why is this problem occurring? What are some root causes of it?
- What might be a potential solution to this problem? How would you implement it?
- How do you know if you're making progress on this problem?

[Problem Exploration Sheet](#)

From Problems to Solutions

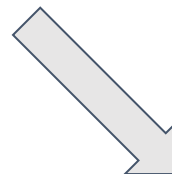
How do you know if a solution is working?

There are two big ideas:



Continuous Improvement

*How do we make it
work better?*



Understanding Outcomes

Did it work?

Setting the Stage for Tomorrow

- Formative vs. summative ***evaluation***? What's the difference and why does it matter?
- Tomorrow's Focus
 - How to examine what works
 - Consider a process to build a useful and robust evaluation plan
 - Group work: Developing the beginnings of an evaluation plan
- But First...

Table Question (Independent Work!)

What is the difference between formative assessment and formative evaluation?

Type your answer and submit it in the jamboard below:

<https://tinyurl.com/6erzkbpj>



Table Question (Independent Work!)

What is the difference between formative assessment and formative evaluation?

Type your answer and submit it in the jamboard below:

<https://tinyurl.com/6erzkbpi>

Do any responses stand out to you?

Are there any with which you might disagree or wish to revise?

A Primer on Evaluation



What is Program Evaluation?



What is Program Evaluation?

Program evaluation is a systematic method for collecting, analyzing, and using information to answer questions about projects, policies and programs, particularly about their effectiveness and efficiency.



What is Program Evaluation?

Program evaluation is a systematic method for collecting, analyzing, and using information to answer questions about projects, policies and programs, particularly about their effectiveness and efficiency.

Three questions to consider...

1 Can we generalize our findings?

2 Are we meant to generalize findings?

3 Do findings serve our intended uses?

Program Evaluation: Evidence of a TOA

Program evaluation is a systematic method for collecting, analyzing, and using information to answer questions about projects, policies and programs, particularly about their effectiveness and efficiency.

Formative Evaluation

Summative Evaluation

What's the distinction?

Program Evaluation: Evidence of a TOA

Program evaluation is a systematic method for collecting, analyzing, and using information to answer questions about projects, policies and programs, particularly about their effectiveness and efficiency.

Formative Evaluation	Summative Evaluation
Evaluation to improve the design, development, or implementation of a program or effort (D'Brot)	Evaluation intended to make a retrospective judgment about a program or effort (D'Brot)

What's the distinction?

Program Evaluation: Evidence of a TOA

Program evaluation is a systematic method for collecting, analyzing, and using information to answer questions about projects, policies and programs, particularly about their effectiveness and efficiency.

How formative and summative evaluation are thought of from the outside (also looking at you, assessment)...



Program Evaluation: Evidence of a TOA

Program evaluation is a systematic method for collecting, analyzing, and using information to answer questions about projects, policies and programs, particularly about their effectiveness and efficiency.

How formative and summative evaluation actually work...



Program Evaluation: Evidence of a TOA

Program evaluation is a systematic method for collecting, analyzing, and using information to answer questions about projects, policies and programs, particularly about their effectiveness and efficiency.

Formative Evaluation	Summative Evaluation
Evaluation to improve the design, development, or implementation of a program or effort (D'Brot)	Evaluation intended to make a retrospective judgment about a program or effort (D'Brot)

The distinction: It's about the
intended use of the information.

Program Evaluation: Evidence of a TOA

Before we adjourn...



[About Us](#) [Our Approach](#) [Resources](#) [CenterLine Blog](#) [Events](#) [Contact](#) [Q](#)

Events

Bringing Educators, Leaders, and Policymakers Together

Events for the passionate changemakers working to improve educational assessment and accountability.

Great work is never done alone. Our events provide an open space for great discussion, idea sharing, collaboration, and inspiration among a wide range of leaders, thinkers, and practitioners in the field – all with a shared commitment to improving outcomes for students.

< > Today Upcoming ~



September 2022

Tue
22

September 22 - September 23

Assessment and Accountability to Monitor and Support Learning Recovery

AC Hotel by Marriott Portsmouth Downtown/Waterfront 299 Vaughan Street, Portsmouth, NH

This year's RILS focuses on how state and district leaders can use assessment and accountability to both monitor and support learning recovery as we emerge from the dark cloud of the pandemic.



Program Evaluation: Evidence of a TOA

Before we adjourn... **3:30** [How Do We Know If Recovery Strategies Are Working?](#)

[Day 1 Evaluation Activity](#)

Juan D'Brot, Center for Assessment


Chris Brandt, Center for Assessment

Scott Marion, Center for Assessment

4:30 Adjourn Day One

Friday, September 23

9:00 Monitoring Progress: But how do you know?

 [Supplemental Reading](#)

[Day 2 Activity Graphic Organizer](#)

Juan D'Brot, Center for Assessment

Chris Brandt, Center for Assessment

Program Evaluation: Evidence of a TOA

Before we adjourn...

How do We Know?

Center for Assessment – RILS 2022

Theory of Action, Logic Model, & Evaluation Design Template

This resource is designed to help states develop theories of action to specify how a program or initiative will promote student-centered learning initiatives.

What is a theory of action?

A theory of action is a hypothesis about how a system produces desired outcomes. At a minimum, a theory of action should have three parts: (1) desired outcomes, (2) resources or inputs, and (3) mechanisms – also known as the activities and outputs – that are expected to produce the desired outputs and outcomes. Each one of these parts can be partitioned further or labeled differently. For example, a theory of action might partition desired outcomes into shorter-term and longer-term outcomes. Or it might make a distinction between “outcomes” and “goals” or between “outcomes” and “impacts.” Some theories of action separate inputs into (human) agents and (material) resources. Similarly, a theory of action focusing on a particular strategy for achieving the desired outcomes might label its mechanisms as “strategies” or “actions.”

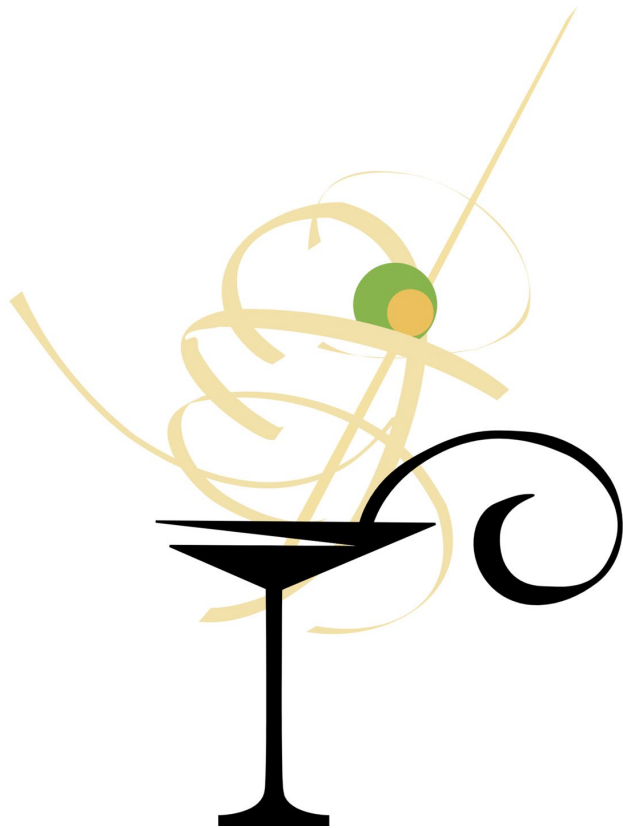
These variations do not change the basic structure of a theory of action. However, they result in varying foci and degrees of specificity in different theories of action.

What is the role of assessment in a theory of action?

Outside the theory of action. Assessment program sponsors often develop theories of action to situate the role of assessment in an educational system. In most systems, assessments play a dual role. The first, most essential role is to ascertain that specific desired outcomes have come about – for example, that students have reached proficiency on academic content standards. In an important sense, this “evidentiary” role places assessment *outside* the theory of action. Here, the assessment is part of an *evidence model* for testing the theory of action. *For example, a state’s end-of-year summative assessment is often used as the primary long-term outcome measure of student proficiency in a theory of action.*

One Last Item!

*Please join us here at 5:30
p.m. for our reception
featuring heavy hors
d'oeuvres and cocktails*



...But How do We Know?

Monitoring Educational Progress (Part II)

RILS Day 2, 9:00 - 10:30 a.m.

Portsmouth, NH

Program Evaluation: Evidence of a TOA

Program evaluation is a systematic method for collecting, analyzing, and using information to answer questions about projects, policies and programs, particularly about their effectiveness and efficiency.

Formative Evaluation	Summative Evaluation
Evaluation to improve the design, development, or implementation of a program or effort (D'Brot)	Evaluation intended to make a retrospective judgment about a program or effort (D'Brot)

The distinction: It's about the
intended use of the information.



Evaluation is Context Dependent

- Why is this relevant for the Center for *Assessment*?
 - Implementing an assessment system or assessment is a akin to implementing a program
 - Context defines the intended purpose or use
 - The intended used defines the evidence needed
- Consider the Program Evaluation Standards (JCSEE, 2014)
 - **Utility:** Are stakeholders' needs met?
 - **Feasibility:** Can it be done effectively and efficiently?
 - **Propriety:** Is it proper, fair, legal, right, and just?
 - **Accuracy:** Are the findings dependable and truthful?

Evaluation is Context Dependent

- Why is this relevant for the Center?
 - Implementing an assessment system or assessment is akin to implementing a program
 - Context defines the intended purpose or use
 - The intended use defines the evidence needed
- Consider the Program Evaluation Standards (JCSEE, 2014)
 - **Utility:** Are stakeholders' needs met?
 - **Feasibility:** Can it be done effectively and efficiently?
 - **Propriety:** Is it proper, fair, legal, right, and just?
 - **Accuracy:** Are the findings dependable and truthful?

***Context
Dependent***

Assessment Depends on Purposes and Uses

- What are some purposes and uses for assessment information?

Assessment Depends on Purposes and Uses

- What are some purposes and uses for assessment information?
 - Diagnosis
 - Formative assessment
 - Progress monitoring
 - Evaluation
 - Prediction

Assessment Depends on Purposes and Uses

- What are some purposes and uses for assessment information?
 - Diagnosis
 - Formative assessment
 - Progress monitoring
 - Evaluation
 - Prediction
- Evaluation, physical therapy, and sports
 - Bad shoulder ? initial diagnosis
 - Physical therapy ← formative assessment
 - End of session ? progress monitoring
 - Can I run my race ? final evaluation
 - Should I change my routine? ? changing the curriculum





Brainstorming Evidence Types

Moving from assessment to evaluation

- Also depends on your purpose
- Also depends on how you want to use it

Recall the following types of evaluation:

- **Formative** - to improve the design, development, or implementation of a program or effort
- **Summative** - to make a retrospective judgment about a program or effort



Brainstorming Evidence Types

Consider the following data elements in the [jamboard](https://tinyurl.com/vwsj6d26): <https://tinyurl.com/vwsj6d26>

Discuss them at your table and categorize them as more summative or more formative.

Rules:

1. Work in tables. Think of it as a team-building activity :)
2. Work fast! You only get to move a post-it **once!**
3. If you disagree with its position, you can change its color, but not its position!
4. We will discuss findings in about 5 minutes.





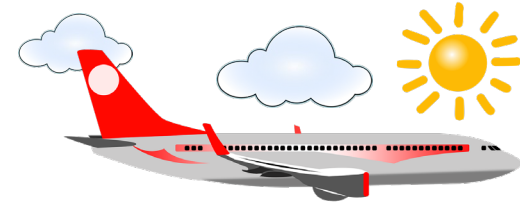
So how *will* we know? Unpacking the steps in the improvement process



Leveraging the Altitude Model (Perspective)

30,000 Feet

**The Why
(Purpose)**



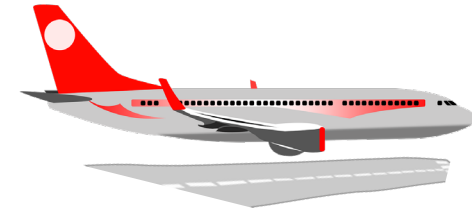
**15,000 Feet
(or 10k, or 14k)**

**The What
(Priorities)**



Ground Level

**The How
(Plan)**





Framing the Improvement Cycle: Programs

Three Core Questions of Continuous Improvement (Bryk, 2015)



1. **Goal:** What specifically are we trying to accomplish?
2. **Program or Theory:** What change might we introduce and why?
3. **Evaluation:** How will we know that a change is actually an improvement?

Evaluation Requires A Clear Problem



Specifying the Goal

Text about goal

Defining the Problem Statement

The problem defines the scope of the program

Identifying the Targets

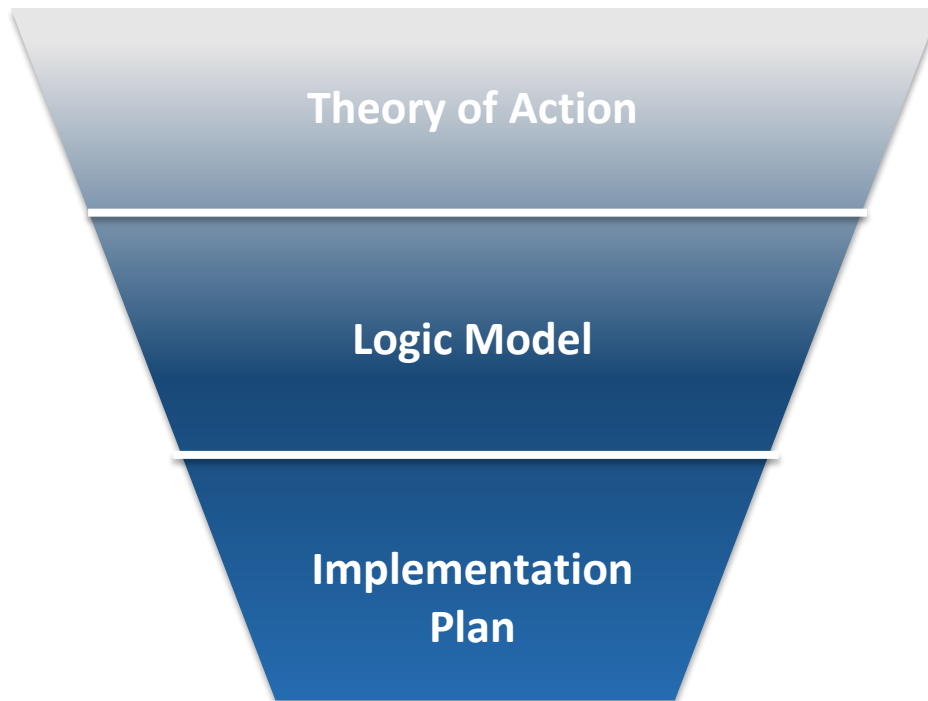
Targets illuminate the population you're interested in

Key Factors to Consider

1. Ultimately, we want to evaluate the outcome.
2. There's a lot that has to happen before we get to the outcome of "accelerating student learning."
3. We need to evaluate of all the steps in between to know if our actions are making a difference (i.e., hitting targets)



Evaluation Plans in the Bigger Picture: The Why



Theory of Action

The **theory of action** casts a wide view of the program by specifying relationships between broader improvement strategies and their expected outputs and outcomes.

Logic Model

The **logic model** draws the logical links and explicit connections between activities, outputs, and outcomes.

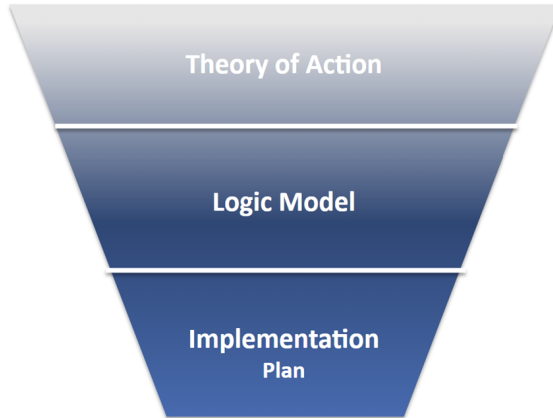
Implementation
Plan

The **implementation plan** builds from the logic model by identifying the actionable components that leverage the knowledge, skill, and will of participants. It can be used to develop assignments, timelines, and who owns what part of the process.



Evaluating the Plan in the Bigger Picture: The Why

EVALUATE



The **theory of action** casts a wide view of the program by specifying relationships between broader improvement strategies and their expected outputs and outcomes.

The **logic model** draws the logical links and explicit connections between activities, outputs, and outcomes.

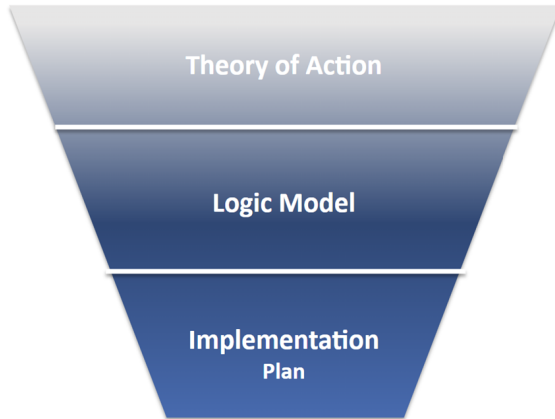
The **implementation plan** builds from the logic model by identifying the actionable components that leverage the knowledge, skill, and will of participants. It can be used to develop assignments, timelines, and who owns what part of the process.



Evaluating the Plan in the Bigger Picture: The Why

We are, in fact, trying to evaluate all of the things in that are included in our bigger picture...

...which is why we need a process.



The **theory of action** casts a wide view of the program by specifying relationships between broader improvement strategies and their expected outputs and outcomes.

The **logic model** draws the logical links and explicit connections between activities, outputs, and outcomes.

The **implementation plan** builds from the logic model by identifying the actionable components that leverage the knowledge, skill, and will of participants. It can be used to develop assignments, timelines, and who owns what part of the process.

EVALUATE





Understanding the Evaluation Plan: The What

01	Strategies used	<ul style="list-style-type: none">• Do you know what strategies were selected?• Do you know what strategies were enacted?
02	Program resources needed	<ul style="list-style-type: none">• What resources are needed?• How these strategies designed to work?
03	Implementation fidelity	<ul style="list-style-type: none">• What evidence do you have that these strategies are actually being implemented as intended?
04	Progress monitoring data	<ul style="list-style-type: none">• Is the information are you collecting relevant to outcomes?• How are you monitoring progress?
05	Determining impact	<ul style="list-style-type: none">• How do you coherently link all of your evidence?• Do you have evidence that you're making an impact?



Building an Evaluation 101: The How

- Developing an evaluation plan is a multi-step, recursive process
- It requires a clear understanding of the
 - Problem
 - Program or initiative
 - Intended outcomes
 - Activities of the program or initiative
 - And how everything is linked together
- Logic models are incredibly beneficial to flesh out a theory of action (see blog [here](#))



Using a Logic Model to Support The How

What is the Activity?	What are the Resources Needed?	What is the (Tangible) Output?	What is the Short-Term Outcome?	What is the Long-Term Outcome?
-----------------------	--------------------------------	--------------------------------	---------------------------------	--------------------------------



Using a Logic Model to Support The How

What is the Activity?	What are the Resources Needed?	What is the (Tangible) Output?	What is the Short-Term Outcome?	What is the Long-Term Outcome?
-----------------------	--------------------------------	--------------------------------	---------------------------------	--------------------------------

Defines the thing we're doing to support the larger initiative



Using a Logic Model to Support The How

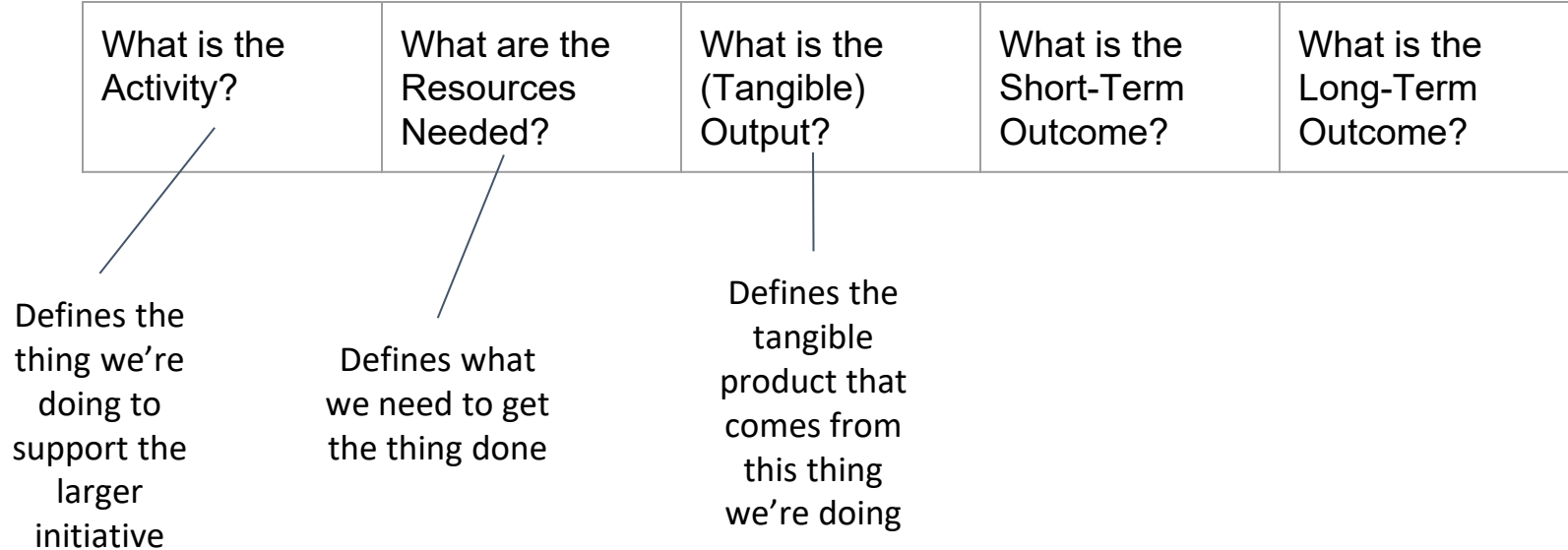
What is the Activity?	What are the Resources Needed?	What is the (Tangible) Output?	What is the Short-Term Outcome?	What is the Long-Term Outcome?
-----------------------	--------------------------------	--------------------------------	---------------------------------	--------------------------------

Defines the thing we're doing to support the larger initiative

Defines what we need to get the thing done

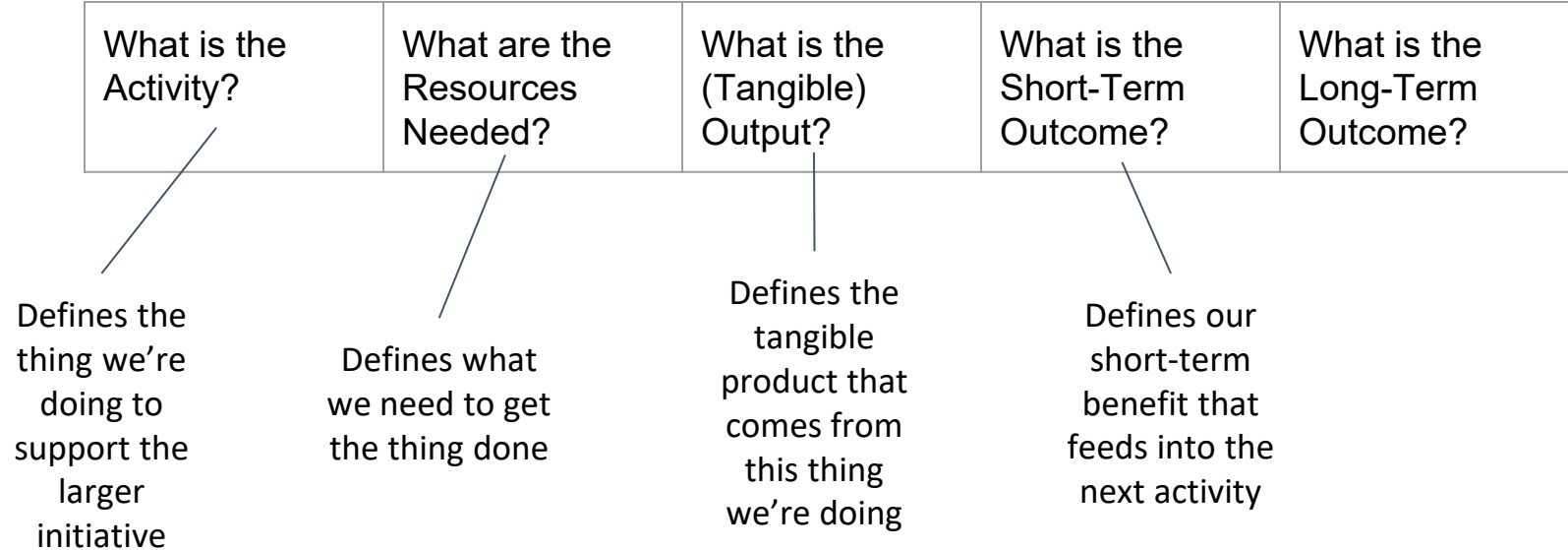


Using a Logic Model to Support The How



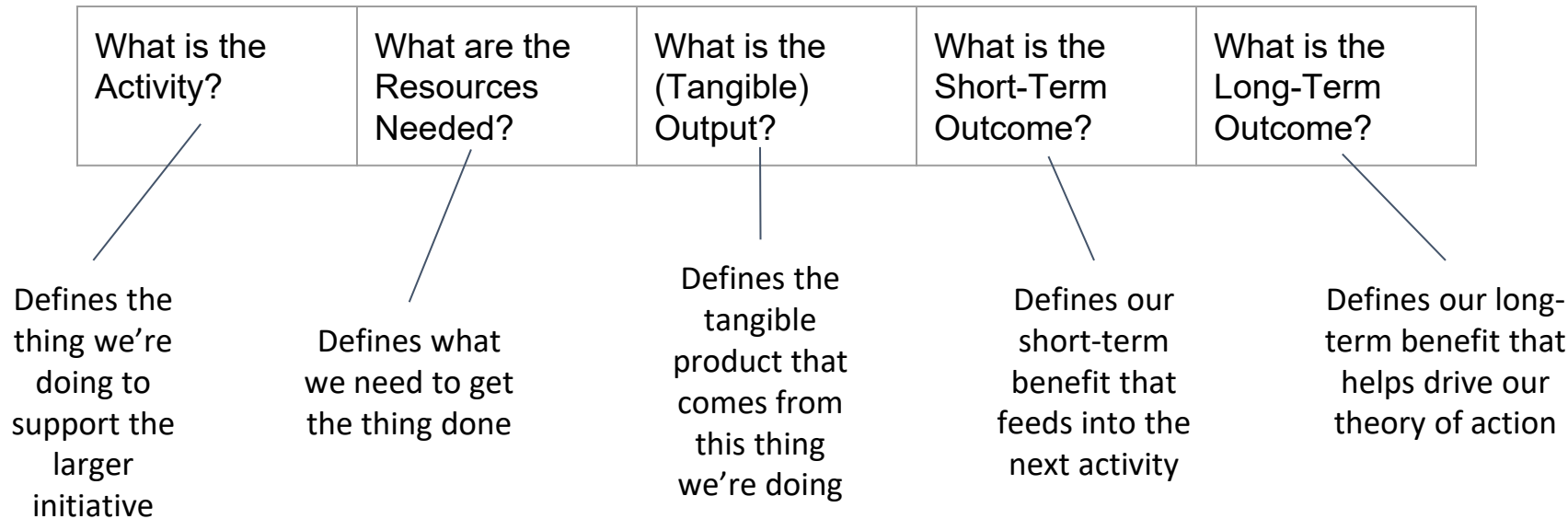


Using a Logic Model to Support The How





Using a Logic Model to Support The How





A Sample Logic Model to Support The How

Theory of Action



A Sample Logic Model to Support The How

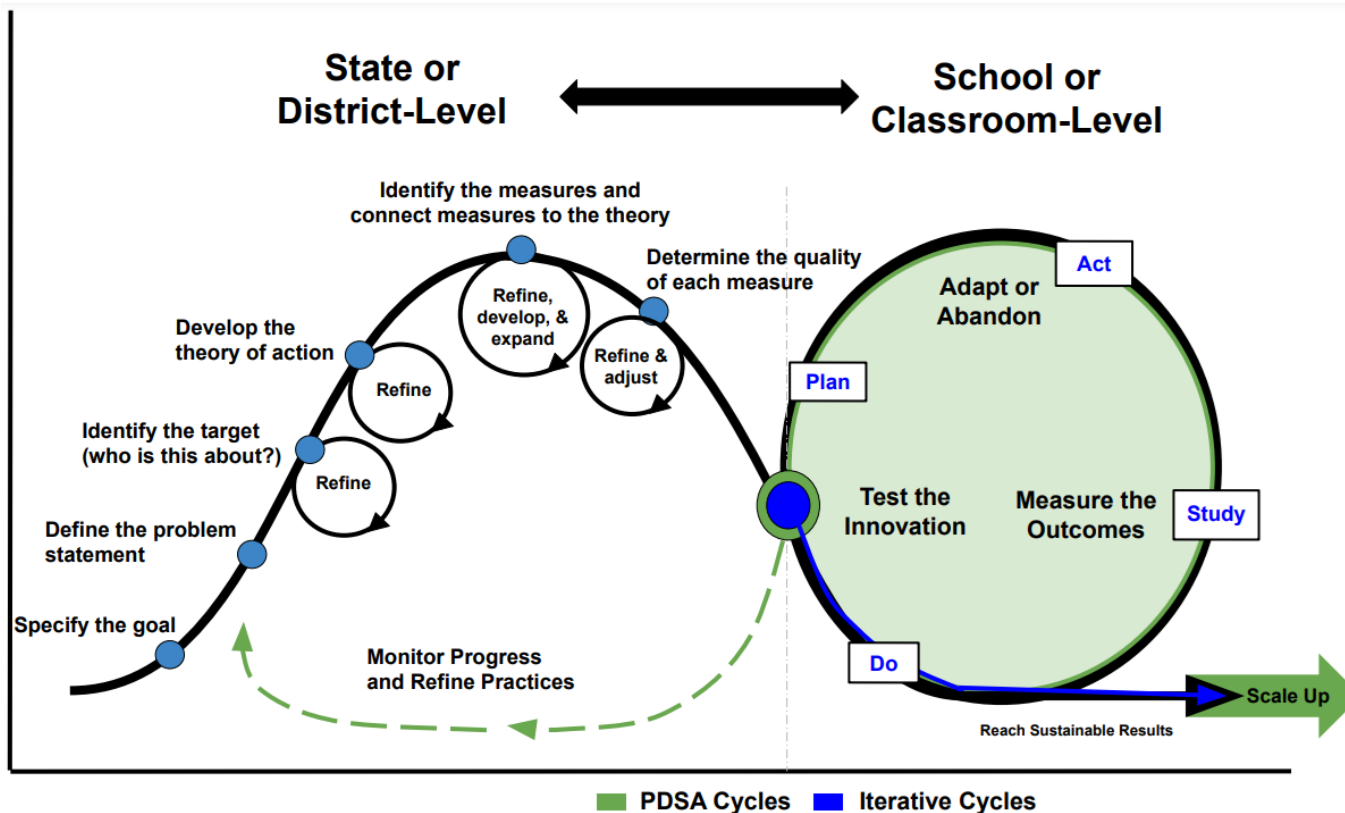
Theory of Action



Sample Logic Model

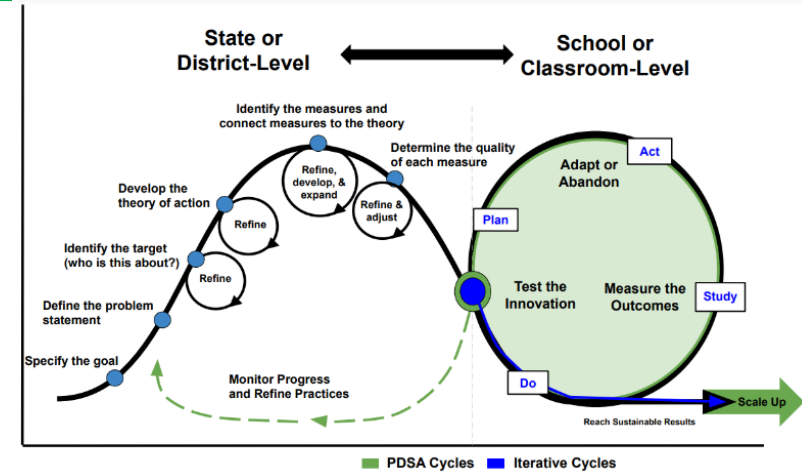
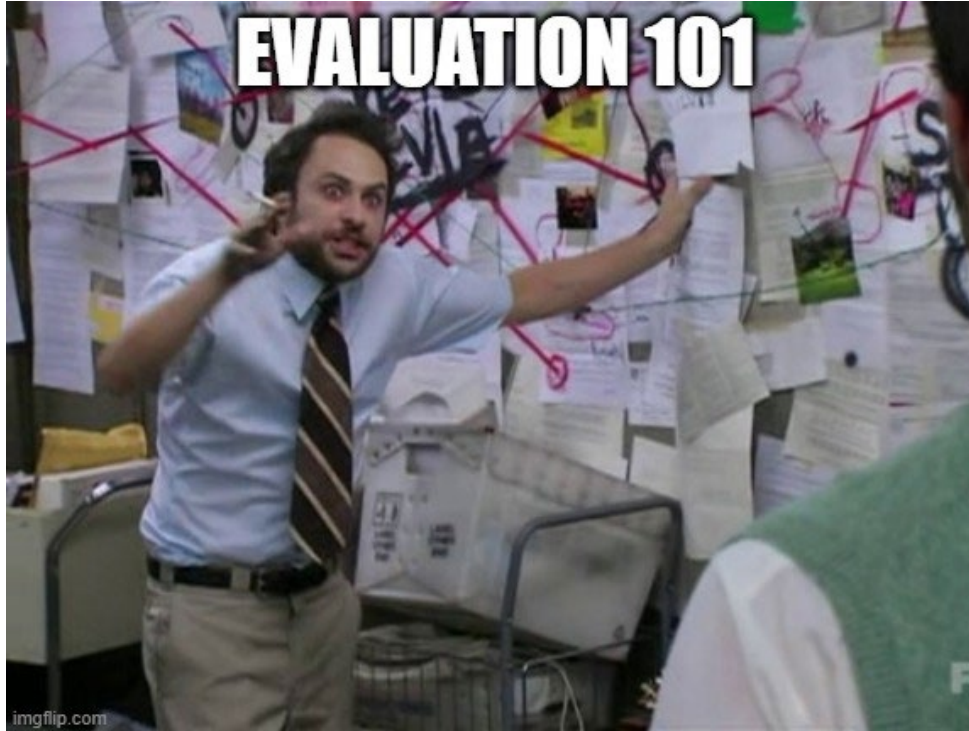
Activity	Resources/Input	Output	Short-Term Outcome*	Long-Term Outcome
Design space	Creative planning; Brain, paper, pencil	A design	Model to align to	Patio design finished
Mark space	Measurement skills, tools, rope, stakes	Stakes and outline	Marked-off space	Patio project foundation complete
Dig and level hole	Shovel, level, rope, a strong back	A hole	A level hole on which to lay	Patio space ready
Lay down paver gravel and sand	Paver gravel, level, rebar, paver sand, 2x4, a sense of balance	Paver gravel and sand on ground	Leveled sand to support pavers on a packed surface	Patio foundation complete
Lay down paver bricks and organize	Paver bricks (lots of bricks), mallet	Paver bricks on sand	A level and well-laid-out patio	Patio structure complete
Fill with polymeric sand and seal stone	Polymeric sand, broom, sealant	Sand and sealant applied	A locked and sealed patio	Sweet patio completed

Continuous Improvement Model





Implementing and Improving a Program



Exactly how do go
about evaluating all of
this stuff?



Evaluation 101: The How

Step	Description
1 Specify the goal and problem statement	It must be attached to the ToA (e.g., evaluate the impact of acceleration strategies).
2 Identify the target	Who is dealing with the problem and how? Specificity is key!
3 Develop the theory of action; build out the program "logic"	Develop a logic model to identify resources, outputs, and outcomes.
4 Connect evidence (measures) to the program logic	How will we measure inputs, outputs, and outcomes? What evidence connects to what outcome(s)?
6 Collect data and determine evidence quality	Are data (1) Complete (2) Consistent (3) Practical (4) Impactful (5) Coherent? (see D'Brot, Landl, Domaleski, & Brant, 2020)
7 Analyze and build a data story	Connect the dots between the activities, evidence, and larger theory of action. Rely on local PDSA cycles and study the variation (why is it occurring?)
8 Tell the story	Document and communicate the results.
9 Make changes	Adjust the program or initiative to improve how you are attacking the problem.



Evaluation 101: The How

Step	Description
1 Specify the goal and problem statement	It must be attached to the ToA (e.g., evaluate the impact of acceleration strategies).
2 Identify the target	Who is dealing with the problem and how? Specificity is key!
3 Develop the theory of action; build out the program "logic"	Develop a logic model to identify resources, outputs, and outcomes.
4 Connect evidence (measures) to the program logic	How will we measure inputs, outputs, and outcomes? What evidence connects to what outcome(s)?
6 Collect data and determine evidence quality	Are data (1) Complete (2) Consistent (3) Practical (4) Impactful (5) Coherent? (see D'Brot, Landl, Domaleski, & Brant, 2020)
7 Analyze and build a data story	Connect the dots between the activities, evidence, and larger theory of action. Rely on local PDSA cycles and study the variation (why is it occurring?)
8 Tell the story	Document and communicate the results.
9 Make changes	Adjust the program or initiative to improve how you are attacking the problem.



Formative Assessment Practices ([Furtak, et. al, 2016](#))

Theory of Action

Teachers Explore
Student Conceptions



Design & Practice
FA Tasks



Enact & Reflect
on FA Tasks



Student Learning
Improves



Formative Assessment Practices ([Furtak, et. al, 2016](#))

Theory of Action

Teachers Explore
Student Conceptions → Design & Practice
FA Tasks → Enact & Reflect
on FA Tasks → Student Learning
Improves

Activities
(some but not all...)



Sample Logic Model

Resources	Activities	Outputs	ST Outcomes	LT Outcome
Certified Facilitator, Learning Progressions	Explore student ideas & teacher understandings about concepts taught.	Clearer representations about how concepts develop	Deeper content expertise	<ul style="list-style-type: none">• Improved engagement• Deeper learning
Design templates	Design FA tasks	Quality of formative assessment task design	Improved lesson plans; pedagogy	
Content experts, teacher teams, tasks	Practice using FA tasks	Quality of questions to elicit student thinking	Improved use of FA practices	
Video, trained observer	Enact FA tasks	Quality interpretation of student ideas and feedback	Improved use of FA practices	
Content experts, teacher teams, tasks	Review video; observation summary. Reflect on & update FA tasks	All of the above	Improved pedagogy, instruction across courses taught	



Sample Measurement Plan

Activities	Measure Activities	Outputs	Measure Outputs
Explore student ideas & teacher understandings about concepts taught.	<ul style="list-style-type: none">• PD attendance records	Clearer representations about how concepts develop	<ul style="list-style-type: none">• Teacher surveys
Design FA tasks	<ul style="list-style-type: none">• PD attendance records	Quality of formative assessment task design	<ul style="list-style-type: none">• FA task ratings
Practice using FA tasks	<ul style="list-style-type: none">• PD attendance records• Team meeting logs	Quality of questions to elicit student thinking	<ul style="list-style-type: none">• Videotaped lesson ratings• Classroom observation ratings<ul style="list-style-type: none">○ Quality of questions○ Quality of instructional FB
Enact FA tasks	<ul style="list-style-type: none">• Teacher logs• Class observations• Teacher surveys	Quality interpretation of student ideas and feedback	<ul style="list-style-type: none">• Sorting task



Sample Measurement Plan

Short Term Outcomes	Measure Short Term Outcomes	Long Term Outcomes	Measure Long Term Outcomes
Deeper content expertise	<ul style="list-style-type: none">• Content-based assessments	Improved student engagement	<ul style="list-style-type: none">• Teacher surveys• Student surveys• Parent surveys
Improved lesson plans; pedagogy	<ul style="list-style-type: none">• Lesson plan ratings	Deeper student learning	<ul style="list-style-type: none">• Interim assessments• Summative assessments• Performance tasks
Improved use of FA practices	<ul style="list-style-type: none">• Class observations• Walkthrough ratings• Teacher surveys, logs		
Improved use of FA practices Improved pedagogy	<ul style="list-style-type: none">• Class observations• Walkthrough ratings• Teacher surveys, logs		



Evaluation 101: The How

Step	Description
1 Specify the goal and problem statement	It must be attached to the ToA (e.g., evaluate the impact of acceleration strategies).
2 Identify the target	Who is dealing with the problem and how? Specificity is key!
3 Develop the theory of action; build out the program "logic"	Develop a logic model to identify resources, outputs, and outcomes.
4 Connect evidence (measures) to the program logic	How will we measure inputs, outputs, and outcomes? What evidence connects to what outcome(s)?
6 Collect data and determine evidence quality	Are data (1) Complete (2) Consistent (3) Practical (4) Impactful (5) Coherent? (see D'Brot, Landl, Domaleski, & Brant, 2020)
7 Analyze and build a data story	Connect the dots between the activities, evidence, and larger theory of action. Rely on local PDSA cycles and study the variation (why is it occurring?)
8 Tell the story	Document and communicate the results.
9 Make changes	Adjust the program or initiative to improve how you are attacking the problem.

Activity Slide

Please visit the “logic model builder”
here: <https://tinyurl.com/yw693xek>

1. Revisit the problem of practice from yesterday
2. Review your work
3. Tackle the three steps named previously using the “logic model builder”



Note: The goal is to not have completed a program evaluation design, but to better understand the need for both formative and summative evaluation evidence.

In Closing

- What are your main takeaways about the last two sessions?
- What outstanding questions would be helpful to write about, create resources for, or revisit in a blog?
- We invite everyone to continue to push our thinking to help us help others.