



# Overview of Evaluation Designs



# Learning objectives



By the end of this presentation, you will be able to:

- Explain evaluation design
- Describe the differences between types of evaluation designs
- Identify the key elements of each type of evaluation design
- Understand the key considerations in selecting a design for conducting an evaluation of your AmeriCorps program

# Overview of presentation

- What is evaluation design?
- CNCS's evaluation continuum
- How to select an appropriate evaluation design for your program
- Key elements of each type of evaluation design
- Evaluation resources and tools

# What is evaluation?



- Evaluation is the use of research methods to assess a program's design, implementation, outcomes, or impacts.
- Evaluation looks at the results of your investment of time, expertise, resources, and energy, and compares those results with what you said you wanted to achieve in your program's logic model.

# Performance measurement and program evaluation

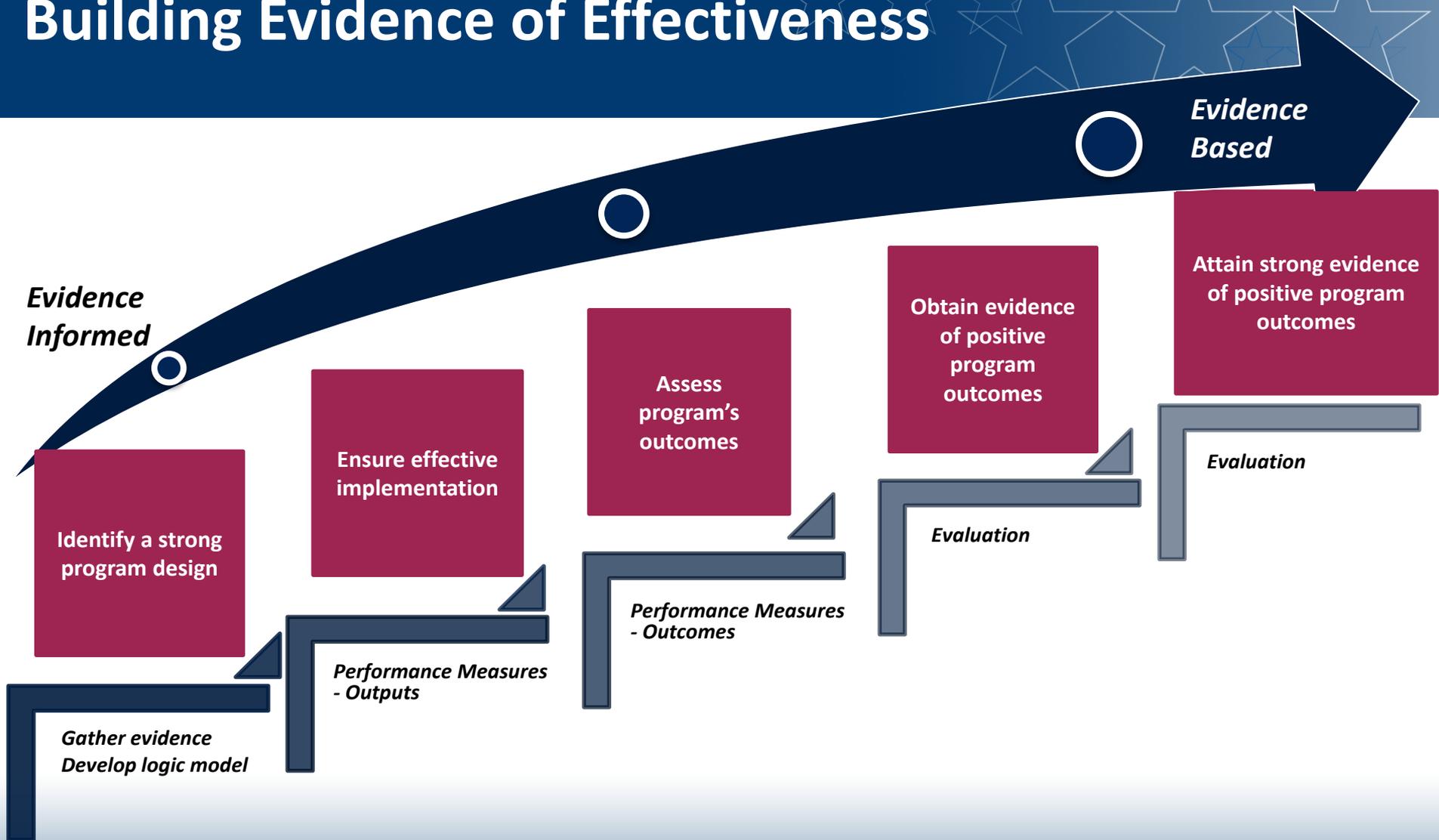
## Performance Measurement

- Ongoing monitoring and reporting of program accomplishments and progress
- Explains what level of performance is achieved by the program

## Program Evaluation

- In-depth research activity conducted periodically or on an ad-hoc basis
- Answers questions or tests hypotheses about program processes and/or outcomes
- Used to assess whether or not a program works as expected and why (e.g., did the program cause the observed changes?)

# Building Evidence of Effectiveness



# What is evaluation design?

- Evaluation design is the structure that provides information to answer questions you have about your program. Evaluation design means thinking about:
  - Why conduct an evaluation
  - What to measure
  - Who to include in the evaluation (e.g. all beneficiaries or a sample)
  - When and how often data will be collected
  - What methods will be used to collect data
  - Whether comparison with another group is appropriate and feasible
- The evaluation design you choose depends on what kinds of questions your evaluation is meant to answer.

# Key considerations in selecting a design

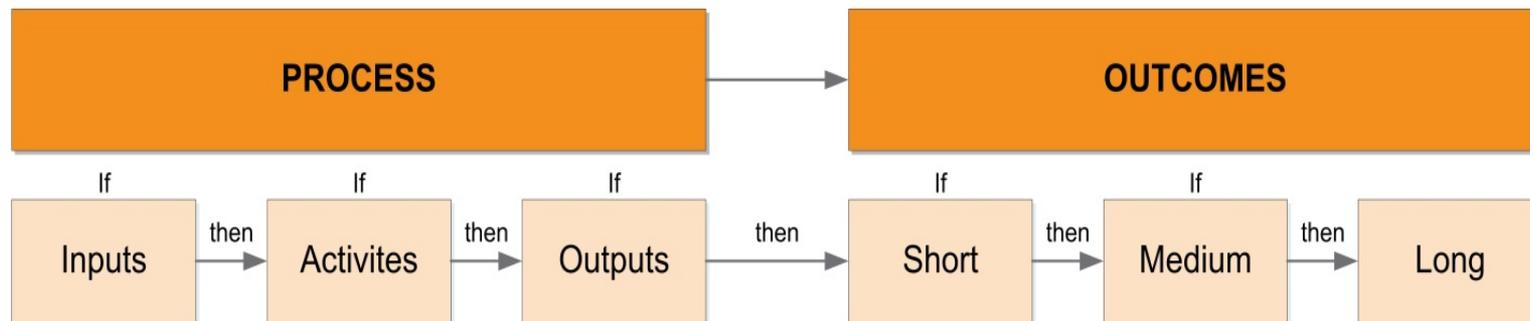
The appropriate design will largely depend upon:

- Your program model
- The primary purpose or goal of the evaluation
- The specific question(s) the evaluation will address
- Resources available for the evaluation
- Funder evaluation requirements

# Program logic model



- A program logic model is a detailed visual representation of a program and its theory of change.
- It communicates how a program works by depicting the intended relationships among program components.



# Example logic model for a literacy program

INPUTS	ACTIVITIES	OUTPUTS	Outcomes		
			Short-Term	Medium-Term	Long-Term
What we invest	What we do	Direct products from program activities	Changes in knowledge, skills, attitudes, opinions	Changes in behavior or action that result from participants' new knowledge	Meaningful changes, often in their condition or status in life
Funding Staff 100 AmeriCorps State and National members 75 non-AmeriCorps volunteers Research	One-on-one tutoring to students below benchmark	Number of students receiving tutoring assistance	Increase in number of students scoring at or above benchmark on literacy assessments Improved student self-efficacy	Increase in number of students reading on grade-level	Students maintain grade-level proficiency in reading

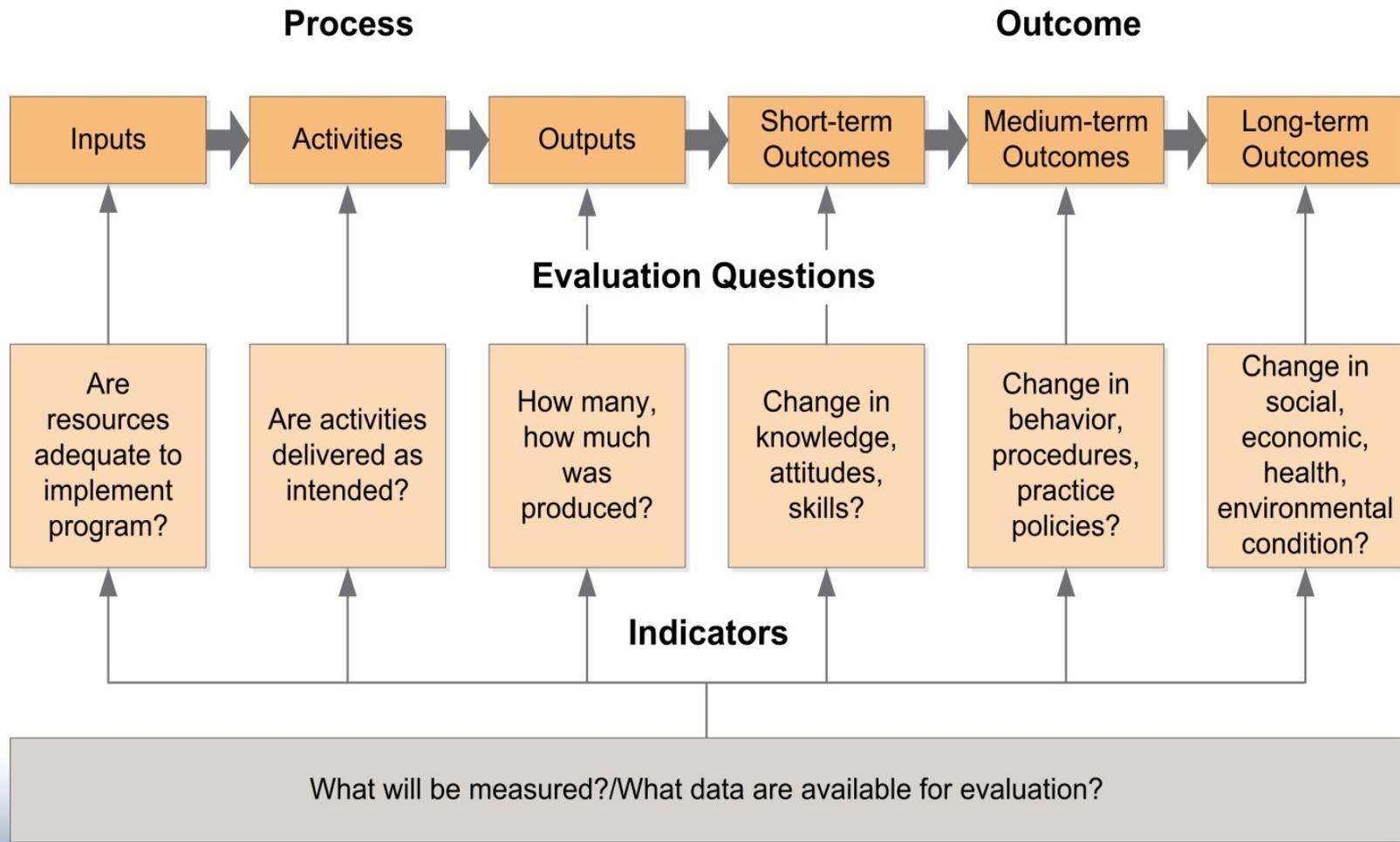
For an overview of logic models, CNCS grantees can refer to the module, "How to Develop a Program Logic Model" located on the Knowledge Network.

# Define purpose and scope

Each evaluation should have a *primary* purpose around which it can be designed and planned.

- Why is the evaluation being done?
- What do you want to learn?
- How will the results be used? By whom?

# Selecting research questions



# Resource considerations

Consider what resources are available to carry out the evaluation:

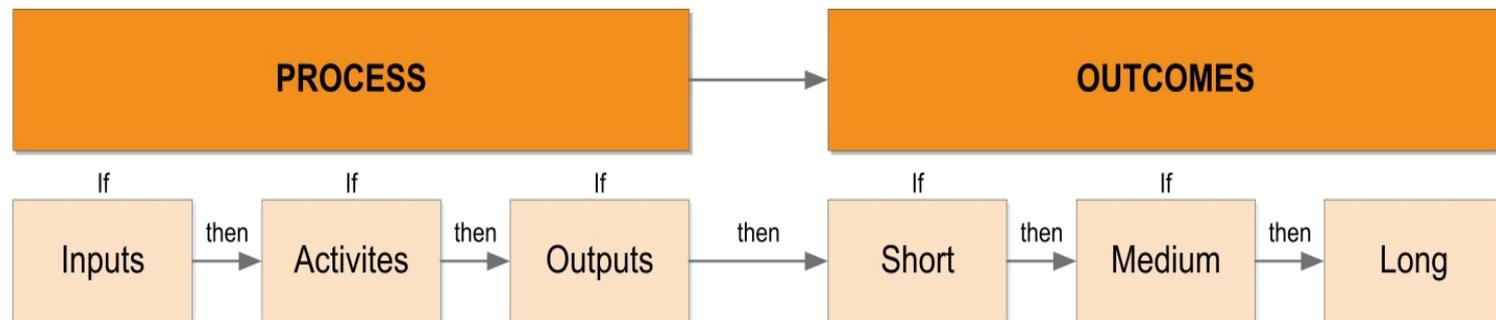
- Staff time
- Funding
- Outside evaluation expertise

It is not necessary to evaluate your entire program.

- Evaluation can be narrow or broad depending on questions to be answered
- Evaluation is not a one-time activity but a series of activities over time that align with the life cycle of your program

# Basic types of evaluation designs

The two “sides” of a program’s logic model align with the two types of evaluation designs: Process and Outcome.



# Process evaluation



## Goals:

- Documents what the program is doing
- Documents to what extent and how consistently the program has been implemented as intended
- Informs changes or improvements in the program's operations

## Common features:

- Does not require a comparison group
- Includes qualitative and quantitative data collection
- Does not require advanced statistical methods

# Examples of process evaluation questions

- Is the program being implemented as designed or planned?
  - Is the program being implemented the same way at each site?
  - Is the program reaching the intended target population with the appropriate services at the planned rate and "dosage"?
- Are there any components of the program that are not working well? Why or why not?
- Are program beneficiaries generally satisfied with the program? Why or why not?
- Are the resources adequate for the successful implementation of the program?

# Examples of methods and data collection tools for process evaluation

## Data Sources:

- Program and school level administrative data
- Site visits to the schools to examine the fidelity of program implementation
  - Observations of literacy intervention with individual students
  - Interviews with school staff and administration
  - Focus groups with teachers and students

## Analysis:

- Thematic identification
- Confirmation of findings across sources

# Group exercise #1: Designing a process evaluation for a literacy program

## Research question:

- Is the literacy program being implemented consistent with the program's logic model and theory of change?

## Design considerations:

- What to measure
- Who to include in the evaluation
- When and how often data will be collected
- What methods will be used to collect data

# Group exercise #1: Designing a process evaluation for a literacy program

**Crosswalk for Process Evaluation of a Literacy Program**

Research question	Indicators	From whom / data sources?	When collected and by whom?	How will you analyze the data?
Is the literacy program being implemented as designed?				

# Example crosswalk for a process evaluation of a literacy program

**Crosswalk for Process Evaluation of a Literacy Program**

Research question	Indicators	From whom / data sources?	When collected and by whom?	How will you analyze the data?
Is the literacy program being implemented as designed?	<ul style="list-style-type: none"> <li>a) Member use of program curriculum during tutoring</li> <li>b) Duration of tutoring</li> <li>c) Student participation rates</li> </ul>	<ul style="list-style-type: none"> <li>a, b, c) Members</li> <li>a, b) Evaluator observes members tutoring students</li> </ul>	<ul style="list-style-type: none"> <li>a, b, c) Members report details of sessions in tutoring logs quarterly</li> <li>a) Quarterly observations by the evaluator(s) using structured observation protocols</li> </ul>	<ul style="list-style-type: none"> <li>a, b, c) Generate frequencies on use of curriculum; average duration of sessions; and average rate of student attendance</li> <li>c) Generate frequencies and averages on quantitative data (e.g., ratings scales, frequency scales) and thematically code and analyze open-ended comments/notes</li> </ul>

# Optional exercise #1: Designing a process evaluation for a literacy program

**Crosswalk for Process Evaluation of a Literacy Program**

Research question	Indicators	From whom / data sources?	When collected and by whom?	How will you analyze the data?
<p>Are program beneficiaries generally satisfied with the program? Why or why not?</p>				

# Example crosswalk for a process evaluation of a literacy program

**Crosswalk for Process Evaluation of a Literacy Program**

Research question	Indicators	From whom / data sources?	When collected and by whom?	How will you analyze the data?
Are program beneficiaries generally satisfied with the program? Why or why not?	<ul style="list-style-type: none"> <li>a) Satisfaction level</li> <li>b) Attendance</li> </ul>	<ul style="list-style-type: none"> <li>a) Parents, teachers, administrators</li> <li>b) Students</li> </ul>	<ul style="list-style-type: none"> <li>a) Parent survey sent home at end of tutoring; teacher focus groups and administrator interviews completed by evaluator</li> <li>b) Members record attendance in tutoring logs quarterly</li> </ul>	<ul style="list-style-type: none"> <li>a) Generate averages on quantitative survey data (e.g., ratings scales, frequency scales) and thematically code and analyze responses from focus groups and interviews</li> <li>b) Generate average rate of student attendance</li> </ul>

# Outcome evaluation



## Goals:

- Identifies the results or effects of a program
- Measures program beneficiaries' changes in knowledge, attitude(s), and/or behavior(s) that result from a program

## Common Features:

- May include a comparison group (impact evaluation)
- Typically requires quantitative data
- Often requires advanced statistical methods

# What is a comparison or control group?

- A group of individuals not participating in the program or receiving the intervention
- Necessary to determine if the program, rather than some other factor, is causing observed changes
- “Comparison group” is associated with a quasi-experimental design and “control group” is associated with an experimental design

# Outcome evaluation questions

- Are there differences in outcomes for program beneficiaries compared to those not in the program?
- Did all types of program beneficiaries benefit from the program or only specific subgroups?
- Did the program change beneficiaries' knowledge, attitude, behavior, or condition?

# Outcome evaluation designs

- Non-Experimental design
  - Single group post design
  - Single group pre-post design
  - Retrospective study designs
- Quasi-Experimental design
- Experimental design (Randomized Controlled Trial)

# Less rigorous outcome evaluation designs

- Single group post design
  - Examines program beneficiaries after they receive program services
- Single group pre-post design
  - Provides a comparison of program beneficiaries before and after they receive program services
- Retrospective study designs
  - Ask previous program beneficiaries to provide opinion on the effects of the program services they received
- Member surveys
  - Survey members on their program experiences and opinions of the results of their service

# Quasi-Experimental and Experimental Designs

## Quasi-experimental design (QED)

- Form comparison group from a similar population of program participants (e.g., similar participants from another program, extra applicants, etc.)

## Experimental design (Randomized Controlled Trial- RCT)

- Randomly assign new eligible applicants to either receive intervention/program or alternative/delayed services (control group)

# Quasi-Experimental and Experimental Designs

Quasi-Experimental	Experimental
<ul style="list-style-type: none"><li>• Can be challenging to identify a similar comparison group</li><li>• Because program and comparison groups are different, results are considered less rigorous</li><li>• Often less labor intensive and expensive than experimental design</li></ul>	<ul style="list-style-type: none"><li>• Most rigorous design option, so results tend to be more highly regarded</li><li>• Often requires increased program recruitment</li><li>• Applicant acceptance is randomly determined</li><li>• Can be more labor intensive and expensive</li></ul>

# Group exercise #2: Designing an outcome evaluation of a literacy program

## Research question:

- What impact does the literacy intervention program have on student reading levels relative to a comparison group of students?

## Design considerations:

- What to measure
- Who to include in the evaluation
- When and how often data will be collected
- What methods will be used to collect data

# Group exercise #2: Designing an outcome evaluation for a literacy program

**Crosswalk for Outcome Evaluation of a Literacy Program**

Research question	What is collected and how?	From whom / data sources?	When collected and by whom?	How will you analyze the data?
What impact does the literacy intervention program have on student reading levels relative to a comparison group of students?				

# Example crosswalk for an outcome evaluation of a literacy program

**Crosswalk for Outcome Evaluation of a Literacy Program**

Research question	What is collected and how?	From whom / data sources?	When collected and by whom?	How will you analyze the data?
What impact does the literacy intervention program have on student reading levels relative to a comparison group of students?	Student reading achievement is measured with literacy assessment tests.	Students participating in the program serve as the intervention group. Students enrolled at a similar school with no program serve as the comparison group.	The evaluator administers the assessments at two time points: <ul style="list-style-type: none"> <li>- At the beginning of the school semester</li> <li>- At the end of the school semester</li> </ul>	Calculate the difference in average outcome in the intervention group minus the difference in average outcome in the comparison group before and after treatment (difference in differences method)

# Optional exercise #2: Designing an outcome evaluation for a literacy program

**Crosswalk for Outcome Evaluation of a Literacy Program**

Research question	What is collected and how?	From whom / data sources?	When collected and by whom?	How will you analyze the data?
<p>What impact does the literacy intervention program have on student self-efficacy relative to a comparison group of students?</p>				

# Example crosswalk for an outcome evaluation of a literacy program

**Crosswalk for Outcome Evaluation of a Literacy Program**

Research question	What is collected and how?	From whom / data sources?	When collected and by whom?	How will you analyze the data?
What impact does the literacy intervention program have on student self-efficacy relative to a comparison group of students?	Student self-efficacy is measured with existing tools, such as the self-efficacy questionnaire for children (SEQ-C).	Students participating in the program serve as the intervention group. Students enrolled at a similar school with no program serve as the comparison group.	The evaluator administers the assessments at two time points: <ul style="list-style-type: none"> <li>- At the beginning of the school semester</li> <li>- At the end of the school semester</li> </ul>	Calculate the difference in average outcome in the intervention group minus the difference in average outcome in the comparison group before and after treatment (difference in differences method)

# Evaluation designs and CNCS's requirements

Evaluation Study Designs	Meet Requirements	
	Large Grantees	Small Grantees/ EAP Programs
Process Design (Non-Experimental Design Studies)	No	Yes
Outcome Design (Non-Experimental Design Studies)	No	Yes
Outcome (Impact) Design (Quasi-Experimental* or Experimental Design Studies)	Yes	Yes

\*Fulfills CNCS evaluation design requirement for large, recompetete grantees if a reasonable comparison group is identified and appropriate matching/propensity scoring is used in the analysis.

# Resources



- **CNCS's Knowledge Network**
  - <https://www.nationalservicerresources.gov/evaluation-ameriacorps>
- **The American Evaluation Association**
  - <http://www.eval.org>
- **The Evaluation Center**
  - <http://www.wmich.edu/evalctr/>
- **Innovation Network's Point K Learning Center**
  - <http://www.innonet.org>
- **Digital Resources for Evaluators**
  - <http://www.resources4evaluators.info/CommunitiesOfEvaluators.html>

# Questions and Answers

