Evaluation Learning Community

Evaluation Basics for Adult Education Administration

NATIONAL REPORTING SYSTEM for Adult Education

U.S. Department of Education, Office of Career, Technical & Adult Education

November 19–21, 2019

Contents

Page

Introduction 1
Evaluation Learning Community 2019 1
Workbook Sections 3
Why Conduct Research and Evaluation 3
Evaluation Planning Model 4
Step 1: Identify Topics and Questions 6
Step 2: Design and Plan 18
Step 3: Analyze, Report, and Use Results 22
Team Planning 27
Additional Notes 27
Appendix: Handouts 28
Handout 1: Icebreaker Activity 29
Monitoring and Evaluation in Everyday Life 29
Handout 2: Threats to Validity Activity 31
Threats to Validity 31

Figures

Figure 1. Evaluation Planning Model 5
Figure 2. Common Logic Model Components 6
Figure 3. Logic Model: Examples of Each Component 8
Figure 4. Sample Logic Model 9
Figure 5. Continuous Improvement Model 23

Introduction

Evaluation Learning Community 2019

Purpose of the ELC Targeted Training

The Evaluation Learning Community (ELC) provides an opportunity for your state team to learn the principles of evaluation and how to conduct valid evaluations to answer critical questions around program practices and procedures.

With additional services, new performance indicators, and partnerships influencing the data collected under the Workforce Innovation and Opportunity Act (WIOA), states need to examine strategic initiatives and other issues of concern by asking different questions and using additional data sources than in years past. The increased emphasis on performance requires that states identify practices, services, and procedures that are most effective. To do this, states must collect quality data in a variety of areas, such as the effectiveness of instructional approaches, enrollment, persistence, and measurable skill gains.

Through participation in the ELC, state teams will engage in a three-stage process: (1) identify research questions, (2) design a research approach, and (3) analyze and interpret results to make decisions about how to make improvements. By using this evaluation process, states can determine whether they are implementing programming as it was intended, identify what is working across program implementation, and decide at the end of the program year or project completion whether an initiative was successful.

Objectives of the Evaluation Learning Community

During the face-to-face training, your state team will engage in activities that will help you:

- Understand the main approaches to research and evaluation
- Develop and refine a logic model and research questions
- Create feasible evaluation and continuous improvement plans
- Implement the evaluation plan
- Begin to implement the continuous improvement plan

Purpose of the Workbook

This workbook is a resource designed to help your state team take an in-depth look at challenges so you can improve service delivery to adult learners.

Completing the activities in this workbook during the face-to-face training will get you started on your evaluation project. Refer to it as you continue your work when you return to your states.

The workbook is designed for team activities, full group discussions, and team planning. The icons on the next page indicate what you will be doing during the face-to-face training.

Team Activities (Your State Only or Mixed Teams)

Share Out/Full Group Discussions

Team Planning

Workbook Sections

Why Conduct Research and Evaluation

You are steeped in the work of your agency every day, managing student needs, federal and state requirements, program performance, partner relationships, and more. Your extensive experience tells you a lot about what is working, areas that could be improved, and promising practices you would like to try. But sometimes it is helpful to take a targeted look at a specific area to determine how effective a process is or whether a different approach could reap better results. You applied to this training to conduct an intensive investigation of an issue or topic that may confirm your assumptions about your programming, or it could uncover something new.

Examples of Areas You Could Explore

- The relationship between the number of hours students attend instruction and their test scores
- The impact of teacher training on student outcomes
- The relationship between participation in Integrated Education and Training (IET) classes and employment outcomes
- The impact of goal setting on Measurable Skill Gains (MSG) attainment

What Areas Do You Want to Investigate?

- In your application and within the work you completed during the online pre-session, you identified topics you would like to explore.

Evaluation Planning Model

The evaluation planning model provides a framework with three steps.

Identify Topics and Questions

- **Determine What You Should Research.** What do you or your stakeholders need to know and hope to gain? Are there federal or state requirements that will direct your efforts? What is a timely topic, versus one that could be examined later? The primary topics your evaluation will focus on include issues or initiatives related to participants, services, and outcomes.

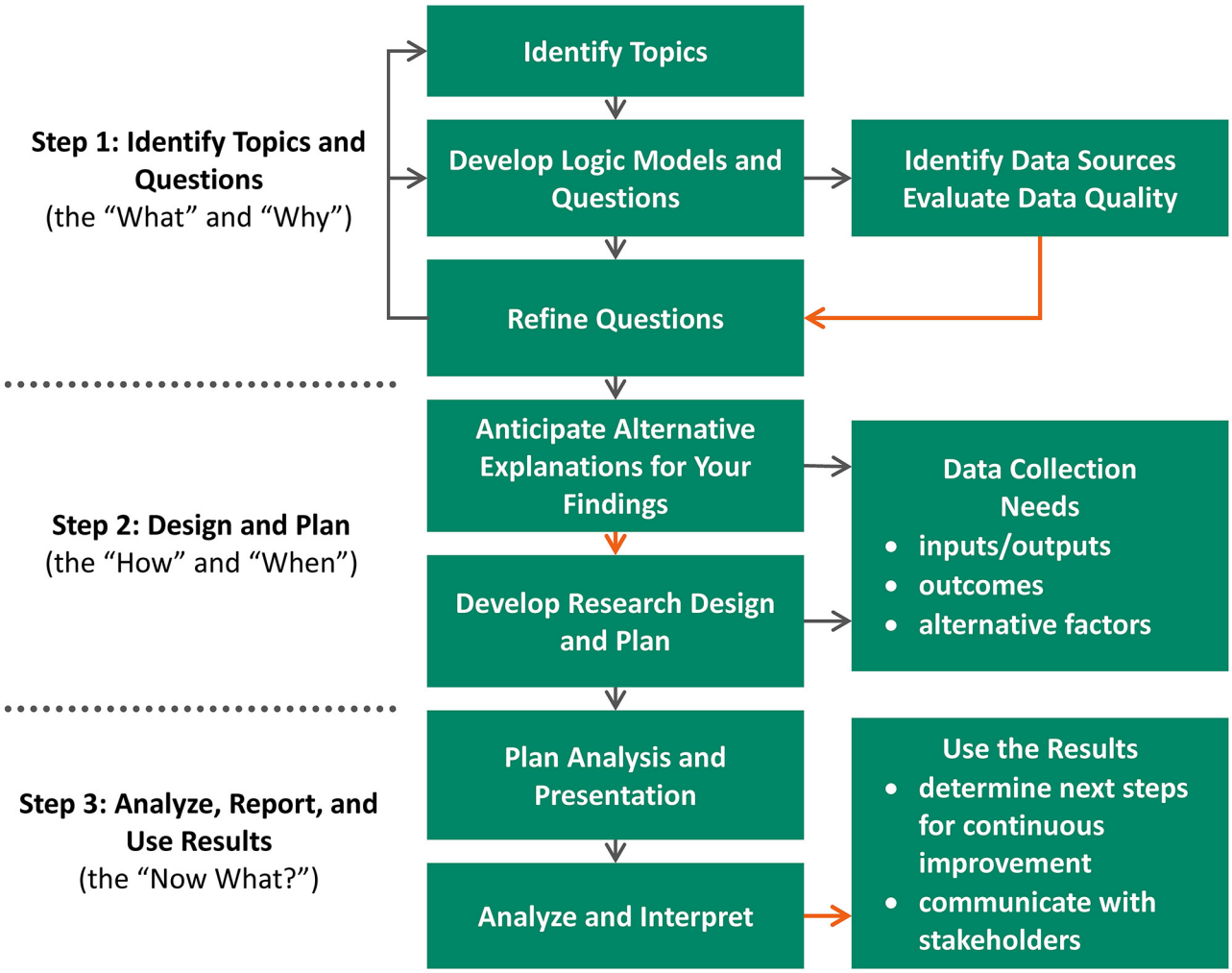
Design and Plan

- **Determine the Type(s) of Studies Needed Based on Your Goals.** Do you want to explore a topic to learn more? Do you want to determine how an activity, program, or practice is being implemented? Do you want to determine the effectiveness of a new initiative on improving student outcomes? The research methods you use to conduct your evaluation will be dictated by the types of answers you are trying to find.

Analyze, Report, and Use Results

- **Determine Your Analysis Methods Based on the Data Available.** What types of data do you have available? Were the data collected in the correct way? Which methods should be used to analyze the types of data you have available? What is the quality level of your data: Is anything missing? Do the results look accurate? Once you have analyzed your data, you can draw conclusions and make decisions about maintaining or changing your practices.

Figure 1. Evaluation Planning Model



Note: Red arrows in the Evaluation Planning Model diagram indicate key decision points.

Step 1: Identify Topics and Questions

Identify Topics: What Are Your Goals?

As leaders in adult education, you are constantly reviewing data and analyzing how your programs and state are performing to improve service delivery to adult learners. There are many topics that can be addressed when evaluating adult education programming and performance. In the chart below are some sample evaluation topics and the specific goals that could be achieved by investigating those topics.

Examples of Evaluation Topics and Goals

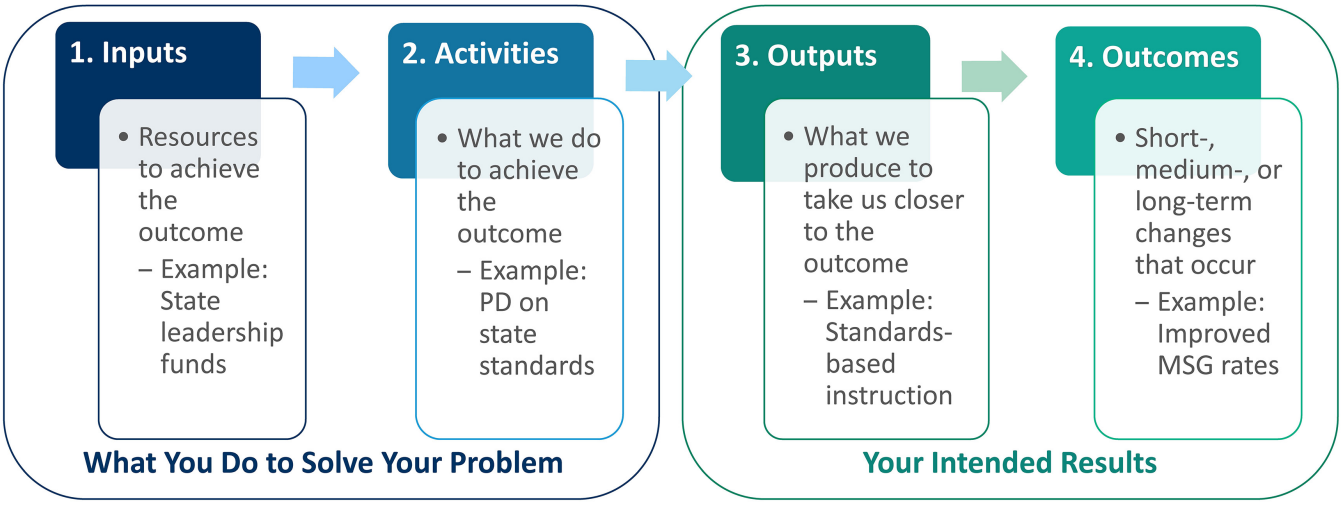
Topic or Problem	Goal
Enrollment of higher-level ABE and ESL participants is lower than the state's targets	Increase enrollment in levels 4–6 of ABE and increase ESL participation
Employment outcomes for participants in IET programming are below our state's target	Increase employment and job retention of participants who attend IET programming

Logic Model

Logic models are powerful tools for understanding the relationships among the activities and services you are providing, how you are providing them and the resources required to do so, and the outcomes they are intended to achieve. The most common uses of logic models are for program design, management, and evaluation, and there are many ways to organize a logic model, depending on its purpose. While all approaches have the same underlying concepts, the model you will use in this training is streamlined so you can easily pinpoint what you want to learn.

In Figure 2, you can see some of the most common components of logic models. Figure 3 provides descriptions and examples of each of the components of a logic model to help you develop your own, and Figure 4 is a completed sample logic model for reference.

Figure 2. Common Logic Model Components



When looking at the topic or problem areas, and thinking about your ultimate goal, you can use the logic model framework to help you think through how your programming and services are set up for delivery. You can see the relationships among the resources you have, the activities you provide, and the outcomes you are seeking to achieve. Using the visual representation allows you, your staff, and providers to see the logical connections between inputs and the outcomes you are trying to achieve, and whether the resources you have in place will allow you to provide the activities and services to the intended participants.

The Components of the Logic Model You Will Use

Logic models for the purpose of the ELC consist of the following components:

- **Topic and Goals**: What you are evaluating and what you want to accomplish.
- **Inputs** (program resources): What the program has. Examples include teachers, staff, and funding to address the topic.
- **Outputs** (activities and audience): What the program does with its inputs and who participates in activities. Examples include training, webinars, and written resources for administrators, teachers, and students.
- **Outcomes** (short term, intermediate, and long term): Results of activities, changes, and impact, immediately and in the future. Examples include the following: short term—skills development and learning gains; intermediate—obtain secondary diploma, enter postsecondary education, or get a job; long term—gain credentials, complete college, or enter a career path with competitive wages.
- **Alternative Explanations or External Factors**: Events, activities, conditions, or situations that, if not controlled, may affect implementation and outcomes in ways that call your evaluation results into question.

Figure 3. Logic Model: Examples of Each Component

Topic or Problem:
Something that you want to change or improve

Goal:
What it will take to change or improve

Inputs	Outputs		Outcomes		
	Activities	Participants	Short-Term	Intermediate	Long-Term
Examples: *program staff, funding, time, external partners, materials*	**Examples:** *events, classes, informational materials, products, workshops, trainings, conferences, exhibits*	**Examples:** *students, teachers, clients, customers, administrators*	**Examples:** *knowledge, attitudes, awareness, opinions, skills*	**Examples:** *knowledge, attitudes, awareness, opinions, skills*	**Examples:** *educational, environmental, quality, or human health improvements*

Alternative Explanations or External Factors
Events, activities, conditions, or situations that, if not controlled, may affect implementation and outcomes in ways that call your evaluation results into question

Figure 4. Sample Logic Model

Topic or Problem:
IET programs are not achieving the targeted outcomes

Goal:
Increase employment and credential outcomes for IET participants

Inputs	Outputs		Outcomes		
	Activities	Participants	Short-Term	Intermediate	Long-Term
• *Funding for IET* • *Teachers trained in IET instruction* • *Providers of IET Instruction* • *Relationships with employers*	• *Number of classes taken* • *Contact hours received* • *Other training services received* • *Professional development on IET, instruction materials use, program design*	• *Participants* • *Teachers and staff*	• *Skills gains* • *Job placement*	• *Job advancement, higher wages* • *Employer satisfaction*	• *Reduced unemployment, higher skilled workforce*

Alternative Explanations or External Factors
1. The unemployment rate in the community
2. Employer availability and willingness to participate

Activity: Draft Your Logic Model

As a team, use your problem statement and the outcomes you are trying to achieve to flesh out your logic model using this template.

Topic or Problem:

Goal:

Inputs	Outputs		Outcomes		
	Activities	Participants	Short-Term	Intermediate	Long-Term

Alternative Explanations or External Factors

Developing Research Questions

Connection Between Your Logic Model and Your Research Questions

Now you have mapped out the problem so you have enough information to formulate the research questions that will be the focus of the rest of your work in the ELC.

Using your logic model, plug in the components to see what you can focus on as you move forward.

- Is **[output]** being implemented as intended across programs? If not, what are some factors affecting implementation?
- Is our **[input or output]** positively associated with **[outcome(s)]**?
- Is **[input or output]** showing promise in improving **[outcome(s)]**?
- Is **[input or output]** effective in improving **[outcome(s)]**?

What other questions could you ask about your topic? Do you have one main question, or do you have multiple questions (e.g., about both implementation *and* outcomes)?

Activity: How to Improve These Questions

In a team with people from other states, improve the questions below. We will then discuss them as a large group.

Poor Question	Good Question	Better Question
Is my program effective for all students?		
How long do students have to be in our programs to be helped?		
What is a good teacher?		

Tips for Formulating Questions

As you formulate your questions, consider the following:

- Is the question focused and specific?
- Is the question narrow enough that the scope is manageable?
- Do we have access to the data we need to answer this question?
- Do we have the resources (staff, time, funding, etc.) necessary to answer this question?

Our Research Question(s)

Using your logic model, and what you know about formulating an effective question, record your evaluation research question(s) in the space below. You may have one or you may have several as you are deciding upon your focus.

Our Research Question(s)
Q1:
Q2:
Q3:

Data Needs

Answering your questions will require quantitative and/or qualitative data. Look at your questions and determine which data you need to address your topic and whether you already have it or you need additional data, and whether you have the time and resources to collect additional data from other sources. There are many ways you currently collect information, such as through the NRS database, other state databases, desk monitoring, site observations, interviews, focus groups, and surveys. The information you are gathering represents factors related to participants, instructors, programming, instruction, outcomes, and more. It is important to think through all the various types of data that can help you answer your research questions.

Activity: Identify Data Needed for Your Logic Model

For the inputs, outputs, and outcomes you outlined in your logic model, map out the data you will need for each. Think about the NRS data you have, other data you have, or data you may need to access and gather to supplement those data, and enter the information in the chart below.

	NRS Data	Other Data
Inputs		
Outputs		

	NRS Data	Other Data
Outcomes		

Data Quality

Your ability to reach valid and reliable conclusions that represent the concept you are studying relies on the quality of your data. Your study can be well designed, but if the data you are collecting have issues with consistency, comprehensiveness, accuracy, or compatibility, you will not reach sound conclusions that answer your research questions.

Issues to consider when looking at existing data:

Definitions and Codes	**Definitions** • Consistent definition and interpretation across programs • Consistent across databases • Remain the same over time • Matches research needs **Coding** • Categories are continuous • Categories match needs
Coverage	• Complete, no missing test data, especially posttests • Consistency in tests used • Data availability • Longitudinal data: How many years do you need? • NRS follow-up measures collected – are they being collected or is a high number missing • Data availability across local programs
Data Collection Procedures, Errors	• Data entry errors • Incorrect administration (tests, survey items) • Missing or incomplete data
Technical Issues	• Different database formats across databases • Database relationships • Changes in database variables, names, codes, definitions, formats, etc., over time • Confidentiality and privacy concerns such as permission and personally identifiable information

Possible Solutions to Data Quality Issues

- Review error checks in data system
- Look at tables of all data to gauge missing data and identify errors
- Recode by combining or creating categories (e.g., age, race/ethnicity)
- Conduct separate analyses for incompatible variables (e.g., look separately at different tests)
- Obtain missing data from another source
- Select classes, programs that have the data already
- Conduct statistical corrections (complicated!)
- Use proxy measures (e.g., student years of education for literacy level) to increase coverage
- Collect new data or revise your research question

Activity: Your Data Quality Challenges and Solutions

In your state teams, discuss your data challenges and how you might resolve them.

Our Data Challenges	How We Can Address Our Data Challenges

Look at the questions you developed in the *Our Research Question(s)* section. Will you have the data you need to answer those questions? Indicate whether you think you have the quantity of data needed, and whether the quality is high enough to use to answer your question. Record this information in the chart below.

Question	Data Necessary to Answer Question	Quantity/ Completeness	Quality

Question	Data Necessary to Answer Question	Quantity/ Completeness	Quality

Based on your determinations regarding the quantity and quality of the data you already have or will need to acquire, do you need to revise your questions? If you don't have the data already, can you acquire them easily? Do you have the resources or staff capacity you'll need? If you already collect data, are you are confident the data are reliable and valid, or are there concerns that data collection is incomplete or inaccurate? After considering these issues, you may want to revise your questions using the space below.

Our Revised Research Question(s) *[Only if necessary]*
Q1:
Q2:
Q3:

Step 2: Design and Plan

Threats to Validity

In designing and planning your evaluation studies, you need to consider factors that may skew the findings. There are a number of influences that threaten the accuracy and legitimacy of your conclusions.

Threat Type	Cautions	Mistake in Action
Confounding Factors	Are you looking at all the variables that could explain your results?	You conclude that the intervention is working, but you have not looked at the impact of other program initiatives.
Nonrepresentativeness	Are you drawing broad conclusions based on limited observations and data?	You conclude that because one site was successful with the intervention that all of your sites will be successful with that intervention, without considering the differences among the various sites.
Hawthorne Effect	Did participants change their behavior because they knew they were being observed?	You conclude that students were doing better because of the design of the intervention, but it could be that students tried harder because their performance was being judged in other ways, e.g., by researchers and not just teachers.
History Effect	Did something external to the study change over time that affected the outcomes?	You draw conclusions about an intervention without considering the fact that there were changes to design or outcome measures.
Maturation	Did the participants mature significantly over the period of the study?	You conclude that the intervention is working when the participant gains could be from factors that naturally take place over time.
Regression to the Mean	Are measures that were high dropping and measures that were low increasing?	You conclude that an original measurement and a second measurement indicate that an intervention had a significant impact, but the increase or decrease could be the result of measurement error.
Participant Mortality	Did certain types of participants drop out of the study?	You conclude that the intervention is working or not working without considering who is still in the study at the end, when the results could be skewed by high performers staying and low performers leaving or vice versa.
Testing Effect	Did you use the same test questions for participants more than once?	You conclude that the intervention is successful when scores are high, when it could be that the participants are mastering the test itself.

There are different ways to control for potential threats to validity statistically—either by including extra variables in your analyses, or by carefully creating matched comparison groups. Controlling for alternative factors statistically when you do your data analysis is not ideal for "effectiveness" research questions, but are more realistic and common practice when randomly assignment is not feasible or desirable. A "Gold standard" experimental impact study is considered the ideal for studying the effect of something (e.g., PD) on student outcomes. This type of study usually randomly assigns students (like in a lottery) to either get whatever intervention or initiative is being studied, or to be in a comparison group. A "quasi" experimental approach is to try and mimic random assignment by creating matched comparison groups after-the-fact.

Activity: Diagnose the Threats to Validity

Refer to Handout #2 in the Appendix to examine a research study scenario and identify the pitfalls of the design and implementation.

Research Methods

There are many types of studies you can conduct based on the goals you have and the questions you are asking. Maybe you just want to know more about a topic as it relates to your program. Or you may want to find out how and/or how well an activity, program or practice is being implemented.

Type of Study	Principles	When to Implement
Exploratory	• Can be based on experience or theories • Can uncover patterns	• When you have broad, more open- ended questions or a general hypothesis you are testing • When you want to learn something that can advance theory or the research base
Descriptive Correlational	• Requires quantitative data • Used to describe and compare • Used to look at relationships	• When you have existing data or survey data • When you are exploring correlations
Qualitative	• For understanding the what, how, when, and where of a topic • Not applicable on a large scale • Results are not generalizable to other groups and settings • Cannot be used to establish causation	• You can use this type of study to complement other research you are doing or others are doing • If you have staff and resources to invest in a labor-intensive process • If you have the capacity to collect data through methods such as: – Case studies – Focus groups – Interviews – Document reviews

Type of Study	Principles	When to Implement
Formative	• Based on a priori hypothesis • Used to informally evaluate how or how well an intervention is implemented, or whether it is or is not showing promise • Results are used to refine the intervention	• When you want to assess how promising an intervention might be for the sake of making improvements and planning • When you want to see how things are going over time and you can incorporate what you learn into making changes; conduct formative evaluation multiple times during intervention development • Types of formative evaluation: – Implementation study/process evaluation – Progress evaluation
Implementation Study	• Usually takes place within a larger study • Used to find out if an intervention being implemented as designed • Increasingly used as part of summative evaluation	• When you want to determine whether changes need to be made to improve the implementation of the intervention • When you want to uncover why an intervention did or didn't work
Process Evaluation	• A major component of a formative evaluation • Provides an indication that the intervention had the intended effect	• When you want to find out whether changes are needed in implementation – Methods are more likely to be quantitative, but can include qualitative components – Often descriptive/correlational – Quasi-experimental
Summative Evaluation	• Findings suggest whether the intervention is effective or not • Based on a priori hypotheses • Used as a more formal test of an intervention	• When you want to test a theory or hypothesis about an intervention
Experimental	• Known as impact study • Establishes causation • Uses random assignment to rule out alternative explanations for an intervention's effect	• When you are conducting a summative study • When you want to learn whether intervention worked by comparing outcomes of students in the intervention group with those not in the intervention group, using a "control" group
Quasi-Experimental	• Establishes causation • Can be used in formative or summative evaluations	• When you want to learn whether an intervention worked by comparing outcomes of students in intervention group to those in a matched "control" group

Activity: Select Your Research Methods

What are your research goals?
What are your research questions?
What data are available to you?
What types of studies are feasible for you to do?

Planning and Conducting Your Studies

You will develop procedures for conducting your study so that there is a guide that each of you can follow for accurate implementation; it will also help others interpret and replicate the study.

Activity: Your Trip of a Lifetime

To plan a trip, there are many steps that happen before you set foot in your desired destination. You think through your desired outcomes, the resources you need to make the trip happen, and how you'll know whether you got everything you wanted out of your travels. In your planning process, you'll collect a lot of different types of data to make it the best vacation it can be. Jot down some of the types of information you will seek out and the tasks you will do in your planning process.

Step 3: Analyze, Report, and Use Results

Key to Keep in Mind When Analyzing Results

Answer Your Question. Look at your logic model, and make sure the data you used are appropriate for what you are trying to learn.

Look for Patterns and Differences. Are you seeing patterns, trends, themes? Are there differences you notice between groups and categories? Are there extreme highs or lows?

Use Appropriate Data and Statistics. Are you using the data that will answer your question? Are the data accurate? Are you processing them correctly? Are you running the analyses that are related to the question you are trying to answer?

Draw Appropriate Conclusions. Be careful in making inferences and drawing conclusions. Avoid being too narrow in considering explanations, or establishing relationships that are not reflected in the data.

Questions to Consider When Using Statistics

- Do the results look feasible and make sense?
- Are there numbers that are too high or too low to be possible?
- Does each statistic (e.g., mean) make sense and fit within the expected range?
- Are the "counts" or number of observations right? How many missing data points are there? How were missing data coded?

Activity: Run Tests

In your teams, you will have the opportunity to use sample data to learn how to use software to process and analyze your data.

Analysis Approach	How This Approach May (or May Not) be Relevant to Our Study
Descriptive Statistics	
Correlational Analysis	
Regression Analysis	

Action Planning

You will find an Excel planning document on the Moodle site with two tools, each on a different tab. The purpose of these tools is to help you think through the various tasks, approaches for completing the work, and the beginning, interim, and final steps to your process. There is some overlap between them, but they provide a view of information in different formats and levels of specificity. You can transfer the work you've done in the activities to the plans so that you have all your ideas in one document.

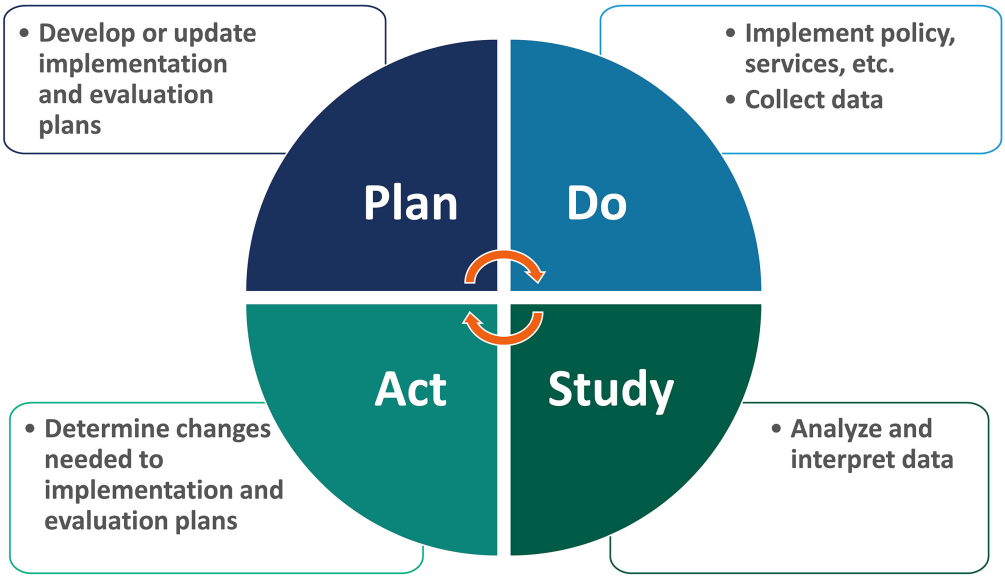
(1) **Final Evaluation Plan**: This tool allows you to map out your entire plan with content, tasks, timeframes, responsibilities, and so on.

(2) **Procedural Plan**: This tool includes scheduling, staffing, format decisions on presentation and communication, and so on.

Continuous Improvement

Throughout this evaluation development process, you have thought through your objectives, what you would like to discover about your program, and how you will collect data. But designing the evaluation is only the beginning of the process. Then you will move into implementation of the plan—collecting data and monitoring the process. Next, you'll analyze and interpret your data and draw conclusions to share with stakeholders. Finally, you'll assess whether you need to make changes to your program implementation and to the evaluation process as you move forward. Figure 5 depicts one continuous improvement model, but there are others you can apply to your process.

Figure 5. Continuous Improvement Model



Activity: Considerations for Continuous Improvement

Your team has successfully designed and implemented the plan and you now have your results. Working in your designated teams, brainstorm and consider responses to the following:

1. Regardless of the findings, how might you use the steps leading to the results and the results themselves to inform:
   - Policy
   - Programs
   - Process
2. Who will you communicate the results to and why?
   - Who are the stakeholders and why are the results important to them?
   - What are at least 3 potential methods of communication and engagement of these stakeholders around the results?
3. If new data becomes available or changes to the data that led to your findings occur, identify at least two strategies or approaches you would take to account for these changes
4. What would be the impact, if any, if you:
   - Do nothing at all with the results
   - Choose not to reflect on the process that led to the results

Use the template below to help guide your responses.

Activity Worksheet: Considerations for Continuous Improvement

	Policy	Program	Process
1. How might you use the evaluation planning steps and findings to inform policy, programs and process?			

2. Who will you communicate the results to and why? List your stakeholders:	Stakeholders	Value/Relevance

List three potential methods of communication or engagement of stakeholders around the results		

3. Identify at least two strategies or approaches you would take to account for new data or changes to data that relate to or directly impact your findings	

	Policy	Program	Process
4. What would be the impact, if any, if you:			
Do nothing at all with the results			
Choose not to reflect on the process that led to the results			

Team Planning

Refer to the Final Evaluation Plan tool and the Procedural Plan tool to continue planning for the next 12 months.

Additional Notes

Appendix: Handouts

Handout 1: Icebreaker Activity

Monitoring and Evaluation in Everyday Life

Think again about the differences between monitoring and evaluation presented in the online pre-session. In section A of the chart below, indicate whether the activity is monitoring or evaluation. One example of each is filled out.

Find a participant who you do not know and introduce yourself. Discuss your responses to section A. Then in section B, identify with your partner at least one activity, unrelated to your work, that can be considered monitoring and one that is evaluation. Be prepared to introduce your partner and discuss your responses with the whole group. Extra space is provided for you to add responses from other participants.

A. Activity	Monitoring	Evaluation
Taking a test drive of a car to see if you want to buy it		✓
Checking the dashboard of your car as you drive	✓	
Following a recipe while you cook		
Asking your family for feedback on a new recipe you made		
Reviewing your retirement plan to see if you can retire		
Reviewing your bank statement to track monthly expenses		
House hunting for your forever home		
Finding lodging for your next vacation		

B. A recent example in everyday life (unrelated to your work)	Monitoring	Evaluation

B. A recent example in everyday life (unrelated to your work)	Monitoring	Evaluation

Handout 2: Threats to Validity Activity

Threats to Validity

Read the scenario describing a research study and identify all the threats to validity from among those listed below. Discuss how the threats can be addressed and other factors you would consider with the design.

Threats to Validity		
• Nonrepresentativeness	• Selection effect	• Participant mortality
• Hawthorne effect	• Maturation effect	• Testing effect
• History effect	• Regression to the mean	

Scenario

Michelle directs a small adult literacy program in Maryland and recently returned from a conference on technology and instruction for adult learners. She was very excited about the possibility of using instructional software in her program's classes to supplement classroom instruction. She thinks the software will be especially helpful for students who cannot attend regularly and found two products she thought would work well with her low-level ESL learners: Quick English and English for All. They both cost about the same and seemed to have the same content, so she decided to try both to see which worked better before purchasing one for her program.

She reported about the conference at a staff meeting and asked if two teachers would try the software so she could compare which helped students better. Two teachers volunteered: Martha, who teaches a morning class and Kirk who teaches a night class. Both classes have about 30 low-level ESL students and meet three days a week for two hours a day. Michelle gives Caroline Quick English and Kirk, English for All and tells each teacher to use the software for one of the instructional days each week. Teachers also to tell their students to use the software at home whenever they want.

The program tests all students when they enter the programs and after three months, each teacher posttests the students who are still in the class (about a third have dropped out from of each class) and asks the students how often they used the software outside of class. After comparing pre-and posttest scores, Michelle finds the students in Caroline's class had higher test gains and used the software more outside of class than students in Kirk's class. Michelle decides to purchase Quick English for her program, based on this study.

What would you tell Michelle about threats to validity and her decision?