

# **Designing Evaluations: 2012 Revision (GAO-12-208G)**

---

**Stephanie Shipman**  
**Center for Evaluation Methods and Issues**  
**Applied Research and Methods Team**  
**April 2012**

# Background

---

- In the 1980s, GAO introduced methodology transfer papers to provide GAO and executive branch agency staff with basic information on commonly used methodologies
- Designing Evaluations was first published in 1984, and revised in 1991.
- Revised in 2012 to:
  - Recognize rise of performance reporting under GPRA
  - Incorporate methods used to evaluate a broad range of programs
  - Incorporate lessons learned from GAO studies of agency evaluations



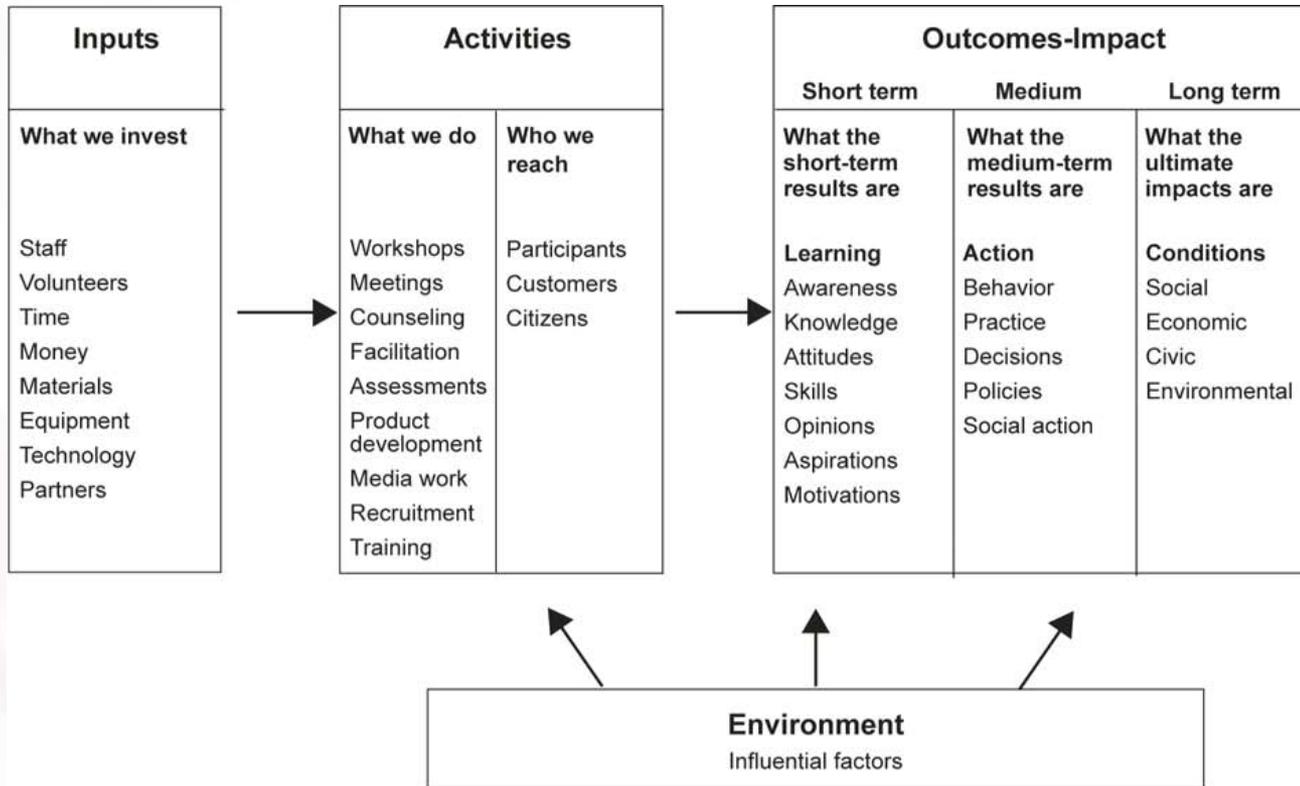
---

# Chapter 1: The Purpose of Evaluation and Importance of Design

---

- Broad definition: “systematic study using research methods to collect and analyze data to assess how well a program is working and why”
  - Different from performance measurement and reporting
  - Used for both accountability and program improvement
  - Studies tailored to address customer questions credibly within available resources
- 5 key steps addressed in guide:
  - Clarify understanding of program’s goals and strategy
  - Develop relevant and useful questions
  - Select an appropriate approach to each question
  - Identify data sources and procedures to obtain credible information
  - Develop plans to analyze data to draw valid conclusions

# Chapter 2: Defining the Evaluation's Scope - Clarify the program's goals and strategy



Source: GAO and University of Wisconsin-Extension, Program Development and Evaluation.



## Chapter 2: Defining the Evaluation's Scope – Develop relevant and useful questions

Program stage	Common evaluation questions	Type of evaluation
Early stage of program or new initiative within a program	<ul style="list-style-type: none"><li>• Is the program being delivered as intended to the targeted recipients?</li><li>• Have any feasibility or management problems emerged?</li><li>• What progress has been made in implementing changes or new provisions?</li></ul>	Process monitoring or process evaluation
Mature, stable program with well-defined program model	<ul style="list-style-type: none"><li>• Are desired program outcomes obtained?</li><li>• What, if any, unintended side effects did the program produce?</li><li>• Do outcomes differ across program approaches, components, providers, or client subgroups?</li></ul>	Outcome monitoring or outcome evaluation
	<ul style="list-style-type: none"><li>• Are program resources being used efficiently?</li><li>• Why is a program no longer obtaining the desired level of outcomes?</li></ul>	Process evaluation
	<ul style="list-style-type: none"><li>• Did the program cause the desired impact?</li><li>• Is one approach more effective than another in obtaining the desired outcomes?</li></ul>	Net impact evaluation



---

## Chapter 3: The Process of Selecting an Evaluation Design

---

- An iterative process to complete BEFORE collecting data
  - Collect background info on program and its evaluations
  - Consider conducting an evaluation synthesis
  - Assess relevance and quality of available data sources and measures
  - Plan original data collection
  - Select evaluative criteria
  - Select a sample of observations
  - Pilot test data collection and analysis procedures
  - Assess study limitations



---

## Chapter 3: The Process of Selecting an Evaluation Design (cont'd)

---

- Criteria for a good design
  - Be appropriate for the evaluation questions and context – clearly state any limitations in scope
  - Adequately address the evaluation question – match the precision, completeness, and conclusiveness needed
  - Fit available time and resources – understand the user's timetable
  - Rely on sufficient, credible data – measures should not be unduly influenced by factors outside the program's control



# Chapter 4: Common Designs – Drawing Causal Inferences about Program Impacts

Evaluation question	Design
Is the program responsible for (effective in) achieving improvements in desired outcomes?	<ul style="list-style-type: none"><li>• Compare (change in) outcomes for a randomly assigned treatment group and a nonparticipating control group (randomized controlled experiment)</li><li>• Compare (change in) outcomes for program participants and a comparison group closely matched to them on key characteristics (comparison group quasi-experiment)</li><li>• Compare (change in) outcomes for participants before and after the intervention, over multiple points in time with statistical controls (single group quasi-experiment)</li></ul>
How does the effectiveness of the program approach compare with other strategies for achieving the same outcomes?	<ul style="list-style-type: none"><li>• Compare (change in) outcomes for randomly assigned treatment groups (randomized controlled experiment)</li><li>• Compare (change in) outcomes for program participants and a comparison group closely matched to them on key characteristics (comparison group quasi-experiment)</li></ul>

## Chapter 5: Lessons Learned about Difficulties in Evaluating Federal Programs

### Challenges and approaches (selected):

- Lack common outcome measures:
  - Develop common reporting standards, recode data, or conduct special survey
- Infrequently observed outcomes:
  - Use proxy measures, assess quality of prevention plans, conduct after-action reviews
- Flexible grant programs:
  - Map variation in approaches, measure aggregate improvements, conduct site-level effectiveness studies

## For more info

---

[www.gao.gov/products/GAO-12-208G](http://www.gao.gov/products/GAO-12-208G)

[ShipmanS@gao.gov](mailto:ShipmanS@gao.gov)

More evaluation resources at [www.fedeval.net](http://www.fedeval.net)