

Issue Brief

**OFFICE OF
RESEARCH
EVALUATION &
PROGRAM SUPPORT**

Program Evaluation, Research, and Informal Inquiry

Pathways to Gathering Information

March 15, 2017

When seeking information to inform issues, programs, policies, or practices we have several modes of inquiry to choose from. Selecting the best approach depends on considerations including the nature of the question and the kinds of statements stakeholders want to make about the results. In the fall of 2016, conversations within CUNY's Senior University Dean's Office prompted a closer look at how program evaluation, academic research, and informal inquiry compare.

In this memo,

1. program evaluation is a systematic approach to gathering information to answer questions about projects, policies, and programs, particularly about their implementation, effectiveness, efficiency, and to inform their development;
2. research is the study of a given subject, field, or problem, undertaken to discover generalizable facts or principles; and
3. informal inquiry consists of activities such as staff meetings to brainstorm problems or informal interviews with stakeholders that are designed generate information about specific internal events or issues.

The table in Appendix A summarizes essential elements across each approach.

Two questions frame this discussion. First, *what is the purpose of the inquiry?* Second, *what kinds of conclusions can be made based on the results from each approach?* The intent here is to encourage conversation about fitting the mode of inquiry to the questions at hand.

Program Evaluation

Purpose. In its most literal sense, program evaluation aims to assess the value of a program as it sheds light on what works for whom, when, how, and under what conditions. Evaluations typically inform public- and private-sector stakeholders who want to know if the programs they are funding, implementing, voting for, receiving, or objecting to are having the intended effect. Equally important are questions such as how the program and its implementation could be improved, whether the program is worthwhile, whether there are better alternatives, if there are unintended outcomes, or whether the program goals are appropriate and useful.

Evaluations generate knowledge about programs, policies, or approaches. Findings may be relevant to a collection of similar programs, but the findings tend to be practical in nature and can be directly applied to answer questions about the particular instance being evaluated. Evaluations can be small or large scale and can use any of the same broad array of methods and measures used in academic research (see Appendix A).

Program evaluations are conducted by trained evaluation researchers and are grounded in formal, systematic research methods. Evaluators may be internal or external to the organization or program under scrutiny; in general, more weight is accorded to evaluations conducted by external evaluators. The Office of Research, Evaluation, and Program Support (REPS) is both internal and external: it is external to the programs it evaluates but is housed within the same overarching entity. This arrangement presents challenges—for example, when evaluation findings are negative—but the benefits are several. REPS evaluators are well-positioned to fully grasp the program context; to engage in close, participatory evaluation projects (i.e., programs are involved in the evaluation process); to ensure designs are responsive when program needs shift; and to communicate findings to university stakeholders.

The first step in any evaluation is to define the research questions that will drive the inquiry. Once evaluators and program stakeholders frame the questions, evaluators identify appropriate measures, methods, sampling procedures, and time required to answer the questions. They carefully assess the appropriate level of confidentiality assurances needed to protect participants.¹ Design and measurement depend on the purpose of the evaluation, whether it is to inform program development, implementation, process, or to assess outcomes. After data analysis the evaluators report back to program staff and help to interpret and share results with stakeholders. In some cases, evaluators contribute to the larger field by disseminating knowledge gleaned about the evaluation process and findings in publications and at conferences. Evaluators contribute their expertise to all phases of the project, from formulating appropriate research questions, adopting strong designs, and framing the findings for stakeholder audiences.

Besides conducting evaluations, evaluators are often trained in methods that support program development such as gathering data to inform the program context, conducting policy analyses, providing information to support grant proposals, and developing internal program documents. For example, logic models and theories of change help programs clarify their assumptions, goals, activities, and expected outcomes.

Evaluation standards and principles. Program evaluation is guided by standards and principles. The American Evaluation Association (AEA), the most prominent professional association for evaluators, promotes ethical practice in all types of evaluation. AEA publishes *Guiding Principles for Evaluators* (2004)² to define ethical practice and, as a member of the Joint Committee on Standards for Educational Evaluation (JSCEE), contributes to setting evaluation standards for evaluation utility, feasibility, propriety, accuracy, and accountability.^{3, 4}

Evaluation, then, is a rigorous approach used by trained professionals to collect systematic information about a program, policy, or approach in order to inform design, implementation, or effectiveness. Whether the target is small or large, and regardless of the evaluation design's level of complexity, the systematic nature of the inquiry is essential to any approach.

¹ How and when to provide confidentiality assurances to participants is critically important in both evaluation and academic research. The appendix to this memo goes into more detail about this related and essential area.

² Available at <http://www.eval.org/p/cm/ld/fid=51>

³ Yarbrough, D. B., Shulha, L. M., Hopson, R. K., and Caruthers, F. A. (2011). *The program evaluation standards: A guide for evaluators and evaluation users* (3rd ed.). Thousand Oaks, CA: Sage.

⁴ In addition to JSCEE standards, guidelines exist for government audits, inspections, and international evaluation. For an overview see <http://www.evaluationcenter.net/pages/standards.aspx>

Academic Research

Purpose. The main difference between program evaluation and traditional academic research is essentially one of purpose. As a rule, academic research seeks to gain insight into underlying processes and to generate enduring insights. Research tests hypotheses and the purpose and methods are determined by the researchers—the project may incorporate evaluation and program partnership, but the inquiry typically extends beyond the immediate program. Whereas evaluators develop studies to meet program stakeholder needs, researchers typically have more autonomy. Research that brings students into a lab to study learning styles is an example of non-evaluation research where all aspects of the study are determined by the investigator.

The distinction between evaluation and research is, to some degree, fluid. For example, MDRC’s random assignment study of CUNY Start (funded by the federal Institute of Education Sciences) is an example of evaluation research, where researcher-program collaboration is important to implementing the research but the methods (such as the design) are up to the researchers.

Informal Inquiry

Purpose. Sometimes, organizations or programs seek information to inform an immediate, internal question. The results of these inquiries are not meant to generalize beyond the specific context or to require the perspective or expertise of a trained evaluator. Whether the inquiry gathers information based on conversations, meetings, or informal surveys about non-sensitive topics (e.g., workshop rating forms), this level of information gathering comprises most of how organizations inform day-to-day issues as well as larger ones on a routine basis.

Informal inquiry provides a quick, efficient means to gather opinions. This approach does not require formal confidentiality assurances (see the appendix) beyond what seems appropriate to the instance. For example, if the subject of the informal inquiry could be interpreted as sensitive to some participants, verbal assurances that information shared would be kept confidential might be appropriate (assuming that confidentiality would indeed be assured). However, if the goal is to gain information beyond an internal matter then the inquiry likely falls under the definition of evaluation: if the content of the inquiry is considered sensitive or if a larger, systematic effort to gain insight is required, then a consultant with content and evaluation experience could be helpful.

Drawing Appropriate Conclusions

The nature of the inquiry determines the type of statement that can be made based on the findings. Results from informal inquiry inform the immediate subject and usually do not generalize to other instances. Conclusions drawn from research and evaluation run the gamut from closely limited to widely generalizable depending on the nature of the questions and the study design. Whereas evaluation provides the opportunity to render robust findings, the ability to do that depends on the evaluation questions and the study design. Evaluations that do not include a comparison group render findings that are descriptive but cannot comment on the program’s “value added” or whether it is a better investment than a similarly focused effort. Random assignment studies, when feasible,

arguably offer the most robust evidence of program effectiveness, but they are not always appropriate to answer evaluation questions.⁵

As a rule, research results are more likely to be generalizable than evaluation findings because the purpose is to understand underlying mechanisms (for example, understanding how cafeteria layout affects student meal choice). Sometimes, evaluation research gains generalizability by, for example, examining programs across multiple settings or investigating mechanisms underlying the program or approach (for example, examining which activities are most effective across program sites).

In evaluation, conversations between program partners and evaluators are essential at the outset of a project to determine the right questions to ask because they determine the kinds of conclusions that can be drawn based on the data. If the question is, “what kinds of students are enrolled in our program and what pathways have they followed?” then then a descriptive approach without a comparison group may suffice. If a program seeks robust evidence to demonstrate the value of the program then a carefully chosen comparison group is essential.

Summary & Conclusion

In sum, careful thought at the outset of an inquiry can help determine the optimal approach. The following exemplify the kinds of questions to ask.

1. Is the intention to gather internal information for an internal audience? *Informal inquiry will probably be sufficient.*
2. Will informal conversations, meetings, or simple session rating forms generate the information we need or do we seek something more rigorous? *If informal, internal information will generate sufficient information to satisfy the desired goals then informal inquiry is appropriate. If more objective data are needed to add weight to the resulting report or findings, then outsourcing to an evaluator may be advisable.*
3. Do we need personal, sensitive, identifiable information? *Even if the inquiry is for an internal purpose, providing basic confidentiality assurances and considering outsourcing to an external evaluator is advisable.*
4. Are we looking for insights into our program policies, practices, or procedures to inform our model, participants, implementation, or outcomes? *A range of evaluation designs are possible depending on the kinds of statements the program wants to make. Consult with an evaluator.*
5. Is our program model robust and are we hoping to scale up across multiple sites to generate generalizable results? *Consult with an evaluator or researcher to guide the process of identifying principal investigators to design a research proposal.*

REPS staff encourages careful thought to fitting the mode of inquiry to the question at hand. We invite requests for consultation if we can help sort through the context to find the right approach.

Please direct questions about this brief to Carol Ripple, carol.ripple@cuny.edu

⁵ Heckman, J., & Smith, J. (1995). Assessing the Case for Social Experiments. *The Journal of Economic Perspectives*, 9 (2), 85-110.

Appendix A: Comparing Program Evaluation, Research, and Informal Inquiry

	Program Evaluation	Research	Informal Inquiry
Purpose			
Nature	Practical, applied	Typically theoretical, but may have practical application	Practical, often immediate application
Type of Insight	Determine performance or outcome as the basis for decision-making	Gain insight into underlying mechanisms	Gain insight into internal issue
Level of Insight	Generate information to reflect on and inform programs, processes, systems, approaches	Generate enduring insights	Generate internal feedback/opinions on internal matters
Aim	Describe program conditions; assess program value relative to criteria; inform future direction	Test research hypotheses	Gather opinions
	Inform program development, implementation, and improvement by examining processes and/or outcomes	Gain insight into underlying mechanisms	Gain insight into a particular issue
Accountability	Stakeholder accountability, program development	Not typically focused on accountability	None, typically
Source of Inquiry	Client-driven inquiry	Researcher-driven inquiry	Staff-driven inquiry
Reporting	Reporting to stakeholders	Reporting in academic journals	Internal reporting
Generalizability	Narrow	Broad	None
Scope			
Questions	Range of questions about a particular program, practice, or policy	Research hypotheses	Specific questions, narrow aims
Tools	Broad range of instrumentation, methods	Broad range of instrumentation, methods	No instrumentation
Design & Measurement			
Methods	Wide array of research methods depending on purpose	Wide array of research methods depending on questions and approach	Informal information gathering, e.g., meetings, conversations, informal interviews
Measures	Measures fit the evaluation questions	Measures suitable to test research hypotheses	No measures

Appendix B: Confidentiality Assurances

How and when to provide participants confidentiality assurances is critically important in evaluation and academic research. Any formal inquiry requires some level of assurance, from a minimal verbal statement to a signed consent procedure. This next section describes the fundamentals as they apply to systematic inquiry. Throughout, the principles apply to both evaluation and research.

Assurances of confidentiality for evaluation and research participants establish *what the participant can expect* and *what the evaluator/researcher commits to do* to uphold those assurances. Regardless of the purpose of the inquiry or the nature of the questions asked, evaluators must ensure they have willing participants. In short, participants must understand their fundamental rights to:

- Choose whether or not they want to participate without penalties (e.g., participating is not required to receiving services or positive regard).
- Withdraw from the project at any time, even if they previously agreed to participate.
- Refuse to complete any part of the project, including refusing to answer any questions.
- Understand what will be done with the information they provide, including the level of confidentiality they can expect.

Some types of evaluation require formal assurances whereas others may not. The most formal method of assuring confidentiality and obtaining informed consent is having participants sign a consent form before any information is gathered. In less formal inquiries where consent forms are not required, assurances may be provided verbally, for example at the start of a focus group. Examples of evaluations that may not require assurances of confidentiality beyond the essential rights above are:

- Strictly internal use of the findings where no personal identifying information is collected.
- Information collected is not personal, sensitive, or identifiable.
- Evaluations of routine education practice.
- Inquiries do not pose significant risk.

Each of these situations presumes research participants are adults; any research or evaluation with children comes with its own set of requirements.

Institutional Review Board approval. Some evaluations—particularly those seeking to generate generalizable information—require Institutional Review Board (IRB) approval. Anyone at CUNY involved in research with human subjects is required to complete the online Collaborative Institutional Training Initiative (CITI) units on research compliance, which familiarizes researchers with the responsibilities associated with protecting the rights of research participants. CUNY’s procedures are covered in detail through its own IRB.⁶ To require IRB approval, an inquiry must meet the definitions of research *and* of human subjects.

⁶ See CUNY’s Human Research Protection Program (HRPP) Policies and Procedures, available at <http://www2.cuny.edu/research/research-compliance/human-research-protection-program-hrpp/hrpp-policies-procedures/>

To meet the definition of research with human subjects, one or both of the following must be true:

1. The project involves conducting a pilot study, a preliminary study, or other preliminary research.
2. The study is designed to collect information in a systematic way with the intention of contributing to a field of knowledge.

And the evaluator/researcher must be:

1. Interacting with living human beings in order to gather data about them, using methods such as interviews, focus groups, questionnaires, and participant observation, or
2. Conducting interventions with living human beings such as experiments and manipulations of subjects or subjects' environments, or
3. Observing or recording private behavior (behavior that individuals have a reasonable expectation will not be observed and recorded), or
4. Obtaining private identifiable information that has been collected about or provided by individuals, such as a school record or identifiable information collected by another researcher or organization.

To meet the definition of research with human subjects - thereby triggering IRB approval - the project must involve research *and* obtaining information from human subjects.