Evaluation Planning; Holden & Zimmerman Approach; Chen and Program Rationales; Evaluation Capacity-building

Mario A. Rivera
PADM 522
School of Public Administration
Defining Evaluation Phases and Types

- **Design-phase or Developmental Evaluation**: Conducted before or early in program implementation, testing an evaluation plan, a change model (theory of change), action model, etc.

- **Formative Evaluation**: designed to determine (particularly during later developmental phases of an intervention): (1) the feasibility of program implementation; (2) the aptness of change and action models; and (3) the short-term social, behavioral, or other impacts of an intervention. Focused on *program improvement*.

- **Process Evaluation**: Designed to ascertain the degree to which replicable program procedures were implemented with fidelity by trained staff according to an articulated program plan; “black box” evaluation. Examines and assesses program processes. *If also meant to inform and improve these processes, it may properly be called a formative evaluation*. However, “formative” and “process” evaluation are often used interchangeably.

- **Outcome or Impact (Summative) Evaluation**: Intended to assess the feasibility, efficacy, and cost-effectiveness of a program intervention in producing significant, long-term benefits for a well-defined population. Results-oriented evaluation.
Chen’s Evaluation Strategies classification; Holden & Zimmerman on evaluation planning

- Chen proposes four evaluation strategies: (1) assessment strategies (judging the performance of an intervention effort); (2) developmental strategies (judging the planning and early implementation of the intervention); (3) theory-elucidation or “enlightenment” strategies, which aim to make explicit the underlying assumptions, change models, and action models of the programmatic intervention; and (4) partnership strategies (ways of involving stakeholders, as well as other organizations in strategic and operational collaboration).

- The distinction is based on the purpose or objectives of the evaluation, and what aspect of a program is under scrutiny. More than one of these efforts could be undertaken in one evaluation, probably at different stages of the evaluation.

- Both Chen and Holden/Zimmerman stress evaluation planning, the projective process that occurs prior to implementation of an evaluation.
Holden & Zimmerman

Planning for evaluation involves:
1. Stating the purpose of the evaluation
2. Understanding the program’s organizational & political context
3. Determining the uses of the prospective evaluation
4. Working with stakeholders to identify primary and secondary evaluation questions
5. Ensuring stakeholder’s buy-in for the evaluation

- Holden and Zimmerman developed a model called Evaluation Planning Incorporating Context, or EPIC, which aims to engage stakeholders, describe the program, and focus the evaluation plan. It provides a way to address issues in the pre-implementation phase of program evaluation. There are five steps in the model, namely assessing context, understanding the organizational and political environment, defining relationships, determining level of evaluation, gathering reconnaissance, specifying evaluation uses, and validating evaluative perspectives.
Building Evaluation Capacity

- Evaluation capacity building is the intentional effort to continuously create and sustain overall organizational processes that make quality evaluation and its uses routine. (Stockdill, Baizerman, & Compton, 2002)
- It is a way of mainstreaming evaluation, of making it an integral part of organizational functioning.
- An organization’s capacity is its potential to perform – its ability to successfully apply its skills and resources toward accomplishment of its goals and the satisfaction of its stakeholders’ expectations. The aim of capacity development is to improve the organization’s performance by increasing its potential in terms of its resources and management. (www.capacity.org, 2003, p. 3)
Evaluation and Organizational Learning Capacity

- *What* assumptions about evaluation guide capacity building efforts?
- *Who* is involved in building capacity?
- *Where* does capacity building happen?
- *When* can evaluation support organizational and programmatic capacity building?
- *How* do we build evaluation and organizational capacity for learning, growth, change?
- *Why* should Organizational Learning, programmatic, and evaluation capacity building efforts be supported?
What definition of evaluation is most useful?

- Program evaluation is the systematic collection of information about the activities, characteristics, and outcomes of programs to make judgments about the program, improve program effectiveness, and/or inform decisions about future programming. (Patton, 1997, p. 23)

- Evaluative inquiry is an ongoing process for investigating and understanding critical organization issues. It is an approach to learning that is fully integrated with an organization’s work practices. (Preskill & Torres, 1999, p. 1-2)
What assumptions should guide evaluation work?

- Evaluation needs to be framed, first of all, from a systems, or complex systems, perspective.
- Evaluation is (or can be, or should be) a **catalyst for individual, group, & organizational learning**.
- Evaluation is an **intervention** – change oriented series of activities.
- Evaluations should only be conducted when there is **intention to use the findings**.
- Evaluation should contribute to the **decisionmaking** process of organization members.
- **Process use** is as important as the use of findings.
- Evaluation is **everyone’s responsibility**.
What assumptions should guide evaluation work?

- Evaluation is most effective when it is collaborative, participatory and learning-oriented.
- Involving stakeholders in evaluation processes builds evaluation capacity.
- The use of dialogic processes (engagement, consultation) enhances stakeholders’ learning throughout the evaluation.
- The most effective evaluator role is that of facilitator, guide, educator, mentor, critical friend, and consultant. (Preskill, 2004)
When does evaluation and organizational capacity coincide?

- Leadership support (an issue with the Lab team)
- Broad stakeholder involvement
- Spirit of inquiry
- Culture of collaboration
- Learning systems and structures
- Effective communication systems
- Open attitudes towards evaluation
- Perceived need or demand for evaluation
- Members have evaluation skills
- Resources to support ECB
When do evaluation and organizational capacity building coincide?

If...
Evaluation is a systematic and purposeful process of inquiry that addresses important organizational issues and questions.

And, it emphasizes...
- Stakeholder involvement
- Learning
- Communicating & reporting
- Evaluation use
- Decision-making
- Action

While considering...
The organization’s Culture, Leadership Systems & Structures, Communication Channels, Politics, And external demands & requirements

Then...
Individuals, teams and organizations learn, and build their own *internal evaluation capacity for ongoing evaluations*
How can we build organizational and evaluation capacity?

1. Technical Assistance
2. Training
3. Mentoring
4. Technology
5. Written Materials
6. Apprenticeship
7. Communities of Practice
8. Appreciative Inquiry
Evaluation Phases, Purposes & Types

**Program Development Phase**

**Design-phase or Developmental Evaluation**
Helps ensure that programs are well conceived, well designed; considers organizational capacity across the board

**Formative (often also Process) Forms of Evaluation**
Help improve the implementation and management of programs, build programmatic & evaluation capacity

**Program Implementation Phase**

**Summative or Outcome or Impact Evaluation**
Helps determine whether or to what extent a program worked; considers as outcomes organizational & evaluation capacity gains

**Program Effect/Impact Phase**
Multidisciplinary Systemic/Diffusion Model of ECB

Organizational Learning Capacity
- Leadership
- Culture

Sustainable Evaluation Practice
- Evaluation Policies & Procedures
- Evaluation Frameworks & Processes
- Strategic Plan for Evaluation
- Continuous Learning about Evaluation
- Resources Dedicated to Evaluation
- Use of Evaluation Findings

Evaluating Knowledge, Skills, & Attitudes
- Design
- Assumptions
- Expectations

Motivations
- Internship
- Written Materials
- Meetings

Coaching
- Technology
- Involvement in Evaluation

Training
- Appreciative Inquiry
- Communities of Practice

Implementation
- Evaluation
- Systems & Structures
- Communication

Diffusion
Program Models

- Models can not substitute for reality—not logic models, change or action models, nor any other program modeling. It’s about delineation of the essentials of a program, pattern matching models to program reality.

- Models can support:
  - Assessment of “evaluability”
  - Client needs identification
  - Program development, refinement, etc.
  - Monitoring and evaluation
Lessons Learned About Building Capacity

- Capacity building efforts require guidance and ongoing support.
- The success of evaluation capacity building is indicated by the use of evaluation processes and findings.
- Trust is the foundation for shared learning.
- People & organizations learn about evaluation by doing it.
- Concerns about evaluation must be addressed as they arise.
- Gains in evaluation capacity are fragile - change is gradual.
- Capacity building involves using wide variety of approaches.
- Capacity building requires organization development skills.
- Adequate resources are required for sustainability.
References and Selected ECB Resources

- Appreciative Inquiry Commons - [http://appreciativeinquiry.cwru.edu](http://appreciativeinquiry.cwru.edu)
Chen: Program Theory and Fidelity

- Theory-driven program evaluation
  - All programs have implicit theories (change+action models, aka program rationales+implementation plans)
  - Program modeling (e.g., via logic models) helps make implicit (or tacit) theory more explicit and therefore subject to scrutiny

- Implementation fidelity
  - Preserving causal mechanisms in implementation
  - Scaling up
  - Staying close to projected/intended outcomes (judgment call—what of positive unintended outcomes? Or negative unintended consequences of projected and attained outcomes?)
Determinants of Success (Mediating & Moderating Variables, or Mediators and Moderators): Mediators are intervening variables (intervening between the intervention effort and program outcomes), while moderators are contextual factors that constrain or enable those outcomes—Chen, page 91
Chen pp.240-241; Action Model for HIV/AIDS education

**Action Model** (which along with the Change Model = Program Theory)

Implementation (intervention → determinants → program outcomes)

**Mediating Variables**

*Instrumental* variables inherent in program design. E.g., openness to learning and change regarding sexual behaviors may well be either presumed or actively fostered by the program, since this cognitive factor would be considered a key variable intervening between program intervention(s) and behavioral change(s) among program subjects. Also, partner involvement (a moderating, contextual factor mainly) can be fostered instrumentally.

**Moderating Variables**

Often, less than positive: e.g., lack of partner support, social and economic variables such as poverty, education, prejudice. However, may be positive: e.g., the net incidence of help from supportive networks—support groups, supportive partners and family and friends, reinforcing social messages, and institutional and cultural supports.

Impacts on individual subject(s) of the intervention, with “impacts’ defined as the aggregate of *comparative net outcomes*
Logic Model & implicit/explicit program theory

- A good logic model clearly identifies Program Goals, Objectives, Inputs, Activities, Outputs, Desired Outcomes, and Eventual Impacts, in their interrelation.

- More or less explicit program theory specifies the relationship between program efforts and expected results (cf. theory-driven and utilization-focused evaluation—Chen). Some logic models specify the connections among program elements with reference to a specific theory or theories of change and of action; others just include if-then linkages.

- A logic model helps specify what to measure in an evaluation, guides assessment of underlying assumptions, and allows for stakeholder consultation and corrective action, for telling a program’s “performance story.”

- Networked, partnered, collaborative programs involving a number of agencies or organizations have more complex causal chains; it is a challenge to capture and assess these in evaluation, as indicated in the following slide.
Multi-agency Monitoring & Evaluation Logic Model

Adapted from Milstein & Kreuter. A Summary Outline of Logic Models: What are They and What Can They Do for Planning and Evaluation? CDC 2000
Networks can increase program design and implementation capacity, but they may also become insular, losing touch with beneficiary groups and communities (Rivera/Rogers).

Using approaches such as Community-based Participatory Research (CBPR), the evaluator must engage stakeholders fully with regard to any important research design issues.

Most evaluations involve multiple research designs or combination of methods—hybrid evaluation designs. It is important to discuss early to see whether (for instance):

- Focus groups are feasible, and whether intensive interviewing and/or survey methods, document reviews, etc., are feasible.
- What kinds of community approvals are needed (e.g., IRBs).
- Is the design “doable?”
Identifying Appropriate Information Sources

- Once information requirements are agreed upon, the sources of the information must be specified; the following questions are key:
  - Who will have information or access to it?
  - Who will be able to collect those data?

- Using existing data as information source
  - Does necessary information already exist in a readily available form? Preferable to use it

- Commonly used information sources
  - Program recipients, deliverers, persons who have knowledge of the program recipients, public documents/databases, file data, reports, position papers, grant proposals

- Policies that restrict information sources
  - Are there policies about collecting data from clients or program files?
  - Confidentiality, anonymity, privacy, IRB protocols, data ownership
  - Give credit to community members contributing to published reports
Identifying Appropriate Information Sources

Client and stakeholder involvement in identifying sources

- Evaluator, by training and experience, often can identify key sources of information
- Client groups will be able to identify sources of information that may be missed by the evaluator
- This is one area where evaluator-client and evaluator-stakeholder collaboration yields helpful answers and makes for a sense of shared ownership of the evaluation process
Identifying Data Collection Methods, Instruments

- Data collected **directly from individuals** identified as sources of information
  - Self-reports
    - interviews, surveys, rating scales, focus groups, logs/journals
  - Personal Products:
    - Tests, narratives, survey responses
- Data collected by **independent observer**
  - Narrative accounts
  - Observation forms (rating scales, checklists)
  - Unobtrusive measures; participant observation
Data collected from existing information

- Public documents
  - federal, state, local, databases, Census data, etc.

- Review of organizational documents
  - client files, notes of employees/directors, audits, minutes, publications, reports, proposals

- Program files
  - Original grant proposal
  - Position papers
  - Program planning documents
  - Correspondence, e-mails, etc.
After identifying for methods, it is important to review adequacy of techniques

- Will the information collected provide a comprehensive picture?
- Are the methods both legal and ethical?
- Is the cost of data collection worthwhile?
- Can data be collected without undue disruption?
- Can data be collected within time constraints?
- Will the information be reliable and valid for the purposes of the evaluation?
Determining Appropriate Conditions for Collecting Information

- **Examples of issues around data collection:**
  - Will sampling be used?
  - How will data actually be collected?
  - When will data be collected?

- **Specifying sampling procedures to be employed**
  - Sampling helps researcher draw inferences about the population in the study
  - Sampling is useful when it will not diminish the confidence of results
  - Sample size must be appropriate; too small a sample is of limited value, and over-large, unfeasible
Specifying how/when information will be collected

- Who will collect data?
- For interviews, focus groups: Might characteristics of the evaluator or evaluators influence data collection? For instance, cultural distance.
- What training should be given to people collecting the data? Striving for consistency across applications.
- In what setting should data collection take place?
- Is confidentiality protected?
- Are special equipment, materials needed?
- When will the information be needed? Available?
- When can the information conveniently be collected?
Determining Appropriate Methods to Organize, Analyze, and Interpret Information

- Develop a system to code, organize, store, and retrieve data
- For each evaluation question, specify how collected information will be analyzed
  - Identify statistical and other analytical techniques
  - Designate some means for conducting the analysis
- Interpreting results
  - Share information with clients to gain perspective on potential interpretations of the data, and to ensure completeness and correctness of the data collected
  - The evaluation plan should allow for the generation and recording of multiple or conflicting interpretations
  - Interpretations should consider multiple perspectives
Determining Appropriate Ways to Report Evaluation Findings

- What is the appropriate way to report findings?
  - Audience, content, format, date, context of presentation

- Suggested Questions (Chen)
  - Are reporting audiences defined?
  - Are report formats and content appropriate for audience needs?
  - Will the evaluation report balanced information?
  - Will reports be timely and effective? Purposes?
  - Is the report plan responsive to the rights to information and data ownership of the audiences?
The following is a checklist of the primary components of a typical evaluation plan; plans should be tailored to specific requirements, beyond this checklist:

- **Introduction and Background**
  - A description of the project, strategy or activity that you are evaluating

- **Research Questions**
  - Questions that you think need answers in order to understand the impact of your work and to improve the evaluation effort
Checklist—outline

Program Outcomes and Measures

- The desired outcomes of the project or program effort about to be undertaken or already underway, and the measures that you will use to indicate that you are progressing toward those outcomes. The evaluation plan often articulates desired program outcomes (objectives) more fully and clearly than program documents. This is one way that evaluations can play a formative role.

- Methodology and Approach

  - Methodology or techniques (e.g., surveys, use of agency records, focus groups, key informant interviews, pre- and post-tests, etc.) that you will be using to collect the measurement data
Data Collection Management and Work-plan

- The data sources (e.g. administrative data sources, respondent groups) that will be used, how data will be managed, and who will be responsible for data collection, data “clean-up,” quality-control of data collection, and eventual “ownership” of the data. These controls and disposition of ownership question were major concerns with the NARCH program.

Proposed Products

- An evaluation report or several reports, an executive summary, a PowerPoint presentation to program principals, grant proposals, handouts, press releases? Who will receive them—intended audiences? (The contractor and funding agency and other key actors may wish to have distinct reports). How will these products be used? Are various uses to be sequenced in particular ways?

Checklist—outline
Evaluation Research Questions

Most evaluation plans are prepared annually for multi-year programs; the following retrospective and prospective questions often arise:

1. **Planning and Implementation:**
   Was program planning adequate? Was the implementation carried out as planned? How well? Were there process or implementation barriers?

2. **Opportunities:** What anticipated and unanticipated opportunities for the generation of information obtained? Did advisory groups, IRBs, focus groups, and other key respondents function as expected? Were information and resources provided as planned—as to types, quantity, and timing? For instance, in the Native American Research Centers for Health Program, data collection went well, but collection of data from six different sets of project Principal Investigators was often delayed or not available in the right format or containing the information expected.

3. **Participation and Utilization:** Who and how many key stakeholders participated? Were there unexpected barriers to participation?

4. **Developmental/consultative role for the evaluator:** If there are serious shortcomings in any of these areas, should the evaluator become involved in redressing them? Questions about the proper evaluator role.
Evaluation Research Questions

5. **Satisfaction**: Are/Were participants satisfied? Why? Why not?

6. **Awareness**: What is the level of awareness of the subject in the target community? Has awareness increased?

7. **Attitudes, norms**: What is the perception of an activity or service (example: cancer screening)? Have perceptions changed?

8. **Knowledge**: What does the target population know about an issue or service (example: substance abuse awareness)? Do they now know more about it? Are they more engaged? For example, in the NARCH Program, parent-facilitators were trained in two communities to develop and implement a family-based curriculum for their late-elementary-school children, and depth semi-structured interviews indicated a very significant increase in awareness and buy-in on their part.

9. **Behavior**: What do people do (example: display a willingness to undergo cancer screening)?

10. **Capacity**: Has community institutional capacity increased? E.g., in the NARCH program, development of IRB capability.
Outcomes and Measures

- What are the stated goals and objectives of the program? For NARCH they were drawn from NIH: (1) Reducing historic mistrust between tribal communities and university researchers, (2) reducing health disparities between Native communities and the American population at large, and (3) reducing under-representation of AI/AN in the health professions.

- How do goals and objectives connect to one another?

- What are the specific program intervention strategies to attain these goals, objectives? You may need to have a strategic planning retreat or two with clients to define these.

- How do goals and objectives connect to strategies. Examine assumptions as you link the two.

- How will progress toward goal attainment be assessed – what indicators or measures will tell you how you are doing? Short term and long term and the time-frames for each. Define indicators and measures in dialogue with clients and beneficiaries, stakeholders.
Data Collection Management and Work-plan

Projecting—in the Evaluation Plan—time Involved in Data Collection:

- A day of data collection often requires a day of analysis; as a rule of thumb, at least two hours of analysis for a two-hour focus group
- Build time for review and feedback from other stakeholders into the preparation of products phase
- Allow for flexibility in your projection of time involved for general project management, meetings to discuss the data collection and analysis, and unintended or unexpected events

Projecting the Cost Involved:

- Your costs—time invested, for instance—as a consultant. It’s not uncommon for contracts to be priced per deliverables in their totality and not per-hour, while still allowing for separate billing for travel and other extraordinary costs; this has significant advantages for evaluators.
- Direct costs (mailing, copying, telephone use for data collection)
- Incentives for participation (NARCH used $20 gift cards);
- Costs for conducting focus groups (food, space, transportation)
- Presentation materials—usually the evaluator’s responsibility as pertains to her or his own presentations.
Evaluation Products and Deliverables

These are to be specified in the Evaluation Plan:

Evaluation Reports – principal product; annual, semi-annual, end-of-project

Other Products: quarterly reports, reports to boards, separate reports to client and funding source, press releases, position papers, etc.

Audience/Purpose: Specify the various audiences corresponding to each of these

These presuppose an articulation early in the report of key Evaluation Questions, which connect back to program goals and objectives, indicators and measures, type of evaluation (design or developmental, process or formative, impact or summative). They also connect with key stakeholder and program principal questions. For example: What would I want to know as a program manager, funder, board member, community member? What would I want to read to be able to understand the issue? What would I need to know in order to take action? In what ways does the evaluation address program accountability and responsibility? In order to carry all of this out well, you need to engage principals and stakeholders, hold community forums (also opportunities for satisfaction surveys), circulate drafts.
The Evaluation Management Plan (Work Plan)

- The final task in planning the evaluation is describing how it will be carried out.
- An Evaluation Management Plan or Work Plan is often essential to help with implementation and oversight of the project.
  - Who will do the evaluation?
  - How much will it cost?
  - Will it be within budget?
  - How does the sequencing of tasks define these issues?
  - Can evaluators count on continued support from top management, for instance in mandating timely and usable reporting of data by program principles?
Agreements and contracts

- Potential problems that arise during the evaluation can be more easily resolved if client and evaluator share a firm understanding.

- A well-documented agreement prior to launching the evaluation study concerning important research procedures (and caveats, e.g., as to data availability) is essential.

Contract and form samples may be found at:

http://www.wmich.edu/evalctr/checklists/contracts.pdf
The EPIC model provides a *heuristic* (or set of rules) for evaluation planning rather than a specified set of steps that are required for all evaluations. Some parts of the model may be more or less applicable depending on such issues as the type of evaluation, the setting of the evaluation, the outcomes of interest, and the sponsor's interests. Therefore, the EPIC model can be used as a kind of instruction guide to prepare for a program evaluation.
The Holden text incorporates the CDC Program Evaluation Framework; we will add the CDC *Partnership Evaluation* model and some network evaluation models as well.

![Diagram](image-url)