Federal Evaluation Policy

Margaret Johnson, MPA
Graduate Research Assistant
Cornell University
maj35@cornell.edu

Monica Hargraves, PhD
Manager for Evaluation in Extension
Cornell University
mjh51@cornell.edu

Bill Trochim, PhD
Director
Cornell Office for Research on Evaluation
Cornell University
wmt1@cornell.edu
This presentation

• Building a ground-up taxonomy using concept mapping (preliminary results)

• How AEA members rated policy ideas
“To … strengthen program evaluation, OMB will launch the following government-wide efforts as part of the Fiscal Year 2011 Budget process:

- OMB will work with agencies to make information readily available online about all Federal evaluations focused on program impacts that are planned or already underway.

- Together with the Domestic Policy Council, National Economic Council and the Council of Economic Advisors, OMB will establish a new inter-agency working group to promote stronger evaluation across the Federal government.

- As part of the Fiscal Year 2011 Budget process, OMB will allocate a limited amount of funding for agencies which, on a voluntary basis:
  - Show how their Fiscal Year 2011 funding priorities are evidence-based or otherwise subject to rigorous evaluation;
  - Assess their own capacity to support evaluation and suggest pathways for strengthening that capacity;
  - Propose new evaluations that could improve government programs in the future; and
  - Identify impediments to rigorous program evaluation in their statutes or regulations.”

Peter Orzag, Director of OMB, October 7, 2009
So what is evaluation policy?

“An evaluation policy is any rule or principle that a group or organization uses to guide its decisions and actions when doing evaluation.”

(Trochim 2008)

Policy can be formal or informal, explicit or implicit.
Why a taxonomy of evaluation policies?

- allows advocates to quickly grasp important interrelationships among different policies for more timely and effective advocacy

- makes it easier for learning and communication about specific policy areas, so knowledge can accumulate (Trochim 2009)

- organizes efforts to develop well-tailored criteria for good policy in a way that is appropriately specific to particular policy areas

- helps trace the transformation from unwritten, specific policies to written evaluation guidance (Datta 2009)
Some ways of grouping evaluation policy

AEA Evaluation Policy Task Force (2007): Definitions, Requirements, Methods, Human resources, Budgets, Implementation, Ethics

Datta (2009): Who is involved, Who gets resources


AEA Roadmap (2009): Relationship of evaluation to policy making; Scope and coverage; Analytic approaches and method; Management; Quality and Independence; Transparency; Resources; Professional competence; Evaluation plans; Dissemination of results; Agency evaluation policy and procedures; Independence; Institutional structures; Legislated policy; Policies for inter-branch collaboration on evaluation
Why ask random AEA members to “taxonomize”?

-policies are best developed democratically, using broadly participatory methods (Trochim 2009)

-new approaches might identify other, high-level categories not previously identified (Cooksy 2009)
Concept Mapping Process

**Planning:** Planners and key issue advisors develop a focus prompt and identify participants.

**Idea Generation:** Communities of interest and expertise are identified, and respond with brainstormed ideas.

**Structuring:** Communities of interest and expertise sort and rate the results of the idea development, authoring the structure and value domain of the issue.

**Representation:** Compute the maps, pattern matches and “go zones,” and prepare them for interpretation by communities of interest.

**Interpretation and Use:** Strategies and tactics for action can follow from the interpretation of the results. Pattern matches and go zones can help build consensus on action. These results are being used to build a taxonomy of federal evaluation policy.

---

**This Project:**

**Focus Prompt:** “In a comprehensive set of U.S. federal evaluation policies, one policy that should be included is …”

**Brainstorming:** 1003 ideas from 554 people, obtained online, pared down to a final set of 100 idea statements.

**Sorting:** 14 so far

**Rating:** 43 (Merit) and 31 (Feasibility)

Participation is OPEN UNTIL NOV. 30.

**Contributors:** Members of the American Evaluation Association, randomly selected.
Brainstorming policy ideas

Participants completed the sentence:

“In a comprehensive set of U.S. federal evaluation policies, one policy that should be included is …”

Researcher synthesized 1003 statements down to 100 -roughly categorized statements -chose one to represent each idea group -aimed for variety and breadth in final set
SORT
“Please place these evaluation policy statements in categories according to meaning or theme, in a way that makes sense to you.”

RATE
MERIT: “For each policy idea, please indicate whether you agree or disagree that it should be implemented.”
(strongly disagree, disagree, neither, agree, strongly agree)

FEASIBILITY: “For each policy idea, please indicate whether you think it is feasible to implement.”
(very infeasible, infeasible, neither, feasible, very feasible)
To compute concept maps

1. Create a similarity matrix from sort data.
   - A square symmetric matrix that shows the number who sorted each pair of statements together.

2. Do multidimensional scaling (MDS) of the similarity matrix
   - Locates each statement as a separate point on a two-dimensional (x,y) map.

3. Do hierarchical cluster analysis of the MDS (x,y) coordinates
   - Partitions the points on the map into groups.
“In a comprehensive set of U.S. federal evaluation policies, one policy that should be included is …”
(3) Require that evaluators be certified to perform evaluations of publicly funded programs.

(12) Establish a "better business"-type consumer protection rating system for all evaluation companies, groups, and individuals working with publicly funded evaluations.

(55) Make outcomes the primary focus of all evaluation plans.
Clusters show how the ideas might be grouped based on sorters’ sense of “nearness”
Feasibility ratings + sorted clusters shows relative rankings

1. Perspectives
2. Principles
3. Definitions and Understandings
4. Methods and Design
5. Planning Evaluations
6. Evaluator Qualifications
7. Evaluation Culture
8. Governance
9. Funding of Projects
Some ideas from most feasible clusters

Evaluation culture
*Use evaluation continuously during the planning and management of federal government policies and activities, not just at budget time.

Methods and design
*Evaluations shall employ methodologies that fit with program duration, intensity, lifecycle stage, context, available capacity, stakeholder needs.

Funding of projects
*A cost analysis (benefit, effectiveness, utility, and/or feasibility) shall be conducted for all federal programs.

Evaluator qualifications
*Require that evaluators be certified in order to perform evaluations of publicly funded programs.
Merit ratings + sorted clusters shows relative rankings

1. Perspectives
2. Principles
3. Definitions and Understandings
4. Methods and Design
5. Planning Evaluations
6. Evaluator Qualifications
7. Evaluation Culture
8. Governance
9. Funding of Projects

- Red circle indicates a focus on Evaluation Culture.
Some ideas from the Principles cluster

- Evaluate to learn not just what is happening but why.
- The perspectives of diverse stakeholders shall be considered and included in evaluation design, implementation, analysis and reporting.
- All evaluations shall be informed by the most current knowledge about evaluation theory and practice.
Some new ways to frame evaluation policy?

1. Importance of Perspectives
2. Principles
3. Definitions and understandings
4. Methods and design
5. Evaluation planning
6. Evaluator qualifications
7. Evaluation culture
8. Governance
9. Funding of projects
Relative cluster rankings
Merit and Feasibility

Merit:
- Evaluation Culture: 4.08
- Methods and Design
- Principles
- Definitions and Understandings
- Bias
- Governance
- Evaluator Qualifications
- Planning Evaluations
- Funding of Projects: 3.42

Feasibility:
- Evaluation Culture: 3.62
- Evaluator Qualifications
- Methods and Design
- Funding of Projects
- Bias
- Definitions and Understandings
- Principles
- Governance
- Planning Evaluations
- Funding of Projects: 3.08
Bringing Merit and Feasibility together

more feasible

Feasibility

Merit

“Go-Zone”

less feasible

less merit

more merit
Top 10 “go zone” ideas

Creating the conditions for good evaluation:
1. Shared understandings of program goals and mechanisms shall be re-assessed periodically.
71. Allow programs an appropriate period of time in operation before expecting them to achieve and report long-term, sustainable change.
97. Provide technical assistance to foster evaluation capacity building within organizations and agencies.
87. Evaluators reporting evidence of coercion through the exercise of power, influence, or resources shall be protected under Federal Whistle Blowers Laws.

How to do good evaluation:
16. Evaluations shall use many different ways to examine a question, including quantitative and qualitative, broad and narrow, shallow and deep
39. Evaluations shall employ methods and procedures that fit the focus of the evaluation.
42. Evaluations shall employ measures that are reliable and valid for the respondents.
73. Evaluation findings shall explicitly address threats to validity.
91. Every evaluation plan shall indicate how the results are to be used and communicated.
61. Evaluators must adhere to American Evaluation Association (AEA) Guiding Principles for Evaluators on systematic inquiry, competence, integrity/honesty, respect for people, and responsibilities for general and public welfare.
Thank you!

For more information about the member-driven taxonomy project, please contact me at: maj35@cornell.edu