



The Evaluator's Methodological Toolkit



Practical additions from
evaluation-specific methodology

Keynote presentation, Arizona Evaluation Network
Conference, Phoenix, AZ, April 10, 2003

E. Jane Davidson, Ph.D.
The Evaluation Center, Western Michigan University
<http://homepages.wmich.edu/~jdauidso>



Overview



- What makes evaluation tricky/different?
 - The "values" part, which requires ...
 - Evaluation-specific methodologies (ESMs)

- A menu of methodological options for:
 - Determining whether a finding is "excellent" vs. "good" vs. "mediocre" (etc.)
 - Working out how the pros and cons of a program should be weighted relative to each other



What makes evaluation tricky?

- Evaluation is the systematic determination of the quality, value, or importance of something
- e-**VALU**-ation has two components
Descriptive Facts (what's so) + **Values** = Evaluative Conclusions (so what)
- Values are what get us from "What's so?" to "So what?" - and they are the trickiest part of an evaluation!



Evaluation-specific methodology

- What makes evaluation unique is that it involves saying something explicit about the *quality or value* of [a program], not just describing how it is
- Social science methodologies (qual & quant data collection and analysis methods) help us gather and interpret *descriptive* information (the "what's so")
- Evaluation-specific methodologies = those methodological tools and techniques that help us address the *quality/value* question (the "so what")



Where do we need ESMs?

- Putting together a solid evaluation involves several steps, not all of which require evaluation-specific methodology
- So ... what are the main methodological steps, and for which ones do we need ESM?



The methodological steps: Part 1

1. What are the evaluation questions?
 - Who needs to know what about program quality/value, by when, to what degree of accuracy, and for what purposes?
 - * 2. In order to answer the evaluation questions, what *aspects* of the program should we examine?
 - Given the context and the needs, what should distinguish an excellent/exemplary program from a mediocre or poor one?
 3. Where should we look for evidence of how well the program is performing on those aspects?
 - What specific information/data (qualitative & quantitative) will we need and from whom/where (informants/sources)?
- * = *requires evaluation-specific methodology*



The methodological steps: Part 2

- * 4. How should we draw conclusions about performance on some aspect of the program, based on a mix of qualitative and quantitative data?
 - What should "excellent" vs. "good" vs. "mediocre" performance look like on Outcome X?
 - * 5. Which aspects of the program are the most critical to its overall value/success in this context?
 - How should pros and cons be weighted?
 - * 6. How should we add up multiple pros and cons?
 - What methods do we have for synthesizing all of the above to answer the original evaluation questions?
- * = *requires evaluation-specific methodology*

April 10, 2003

Keynote, AZ Evaluation Network

7



Four Important ESMs

Explicitly Evaluative Task

ESM

- | | |
|--|----------------------------|
| ■ Identifying all relevant aspects of the program to investigate | → Needs/value assessment |
| ■ Defining what mix of evidence should constitute "excellent" (etc.) | → Merit determination |
| ■ "Weighting" the different aspects of program performance | → Importance determination |
| ■ Adding up the pros and cons to get answers to evaluation questions | → Synthesis |

April 10, 2003

Keynote, AZ Evaluation Network

8



Merit Determination

Descriptive Facts (what's so) + "Values" = Evaluative Conclusions (so what)

- Merit determination = the process of combining descriptive facts with relevant values to say something explicit about the merit/value of certain performances
- Merit determination methodology = a systematic set of procedures for merit determination



Merit Determination Methodologies

- There are several strategies available for determining merit
- Each one comes with its own
 - Set of assumptions
 - Information requirements
 - Strengths and weaknesses
 - Price tag
- One size does not fit all — merit determination methods should be chosen carefully based on the specific task at hand



Merit Determination Strategies

1. Evaluate relative to **goals** (usually a weak option)
2. Have stakeholders **"vote"** on standards
3. Evaluate individual indicators relative to some standards (e.g., fixed **quality standards**)
4. **"Benchmark"** against (compare with results achieved by) other efforts or alternatives
5. **Use your needs assessment:** Interview/probe the real impact on recipients' needs, e.g., for Outcome X:
How *substantial* is its impact in people's lives?
6. Use **rubrics** that combine several of the above



A Basic Rubric for Determining Merit

Excellent	Clear example of best practice in this domain; no weaknesses.
Very Good	Very good or excellent performance on virtually all aspects; no real weaknesses.
Good	Reasonably good performance overall; may have a few slight weaknesses.
Satisfactory	Fair performance; some serious (but non-fatal) weaknesses on a few aspects.
Poor	Clear evidence of impaired functioning on this dimension; very serious weaknesses.



Sample Rubric for Qualitative Data

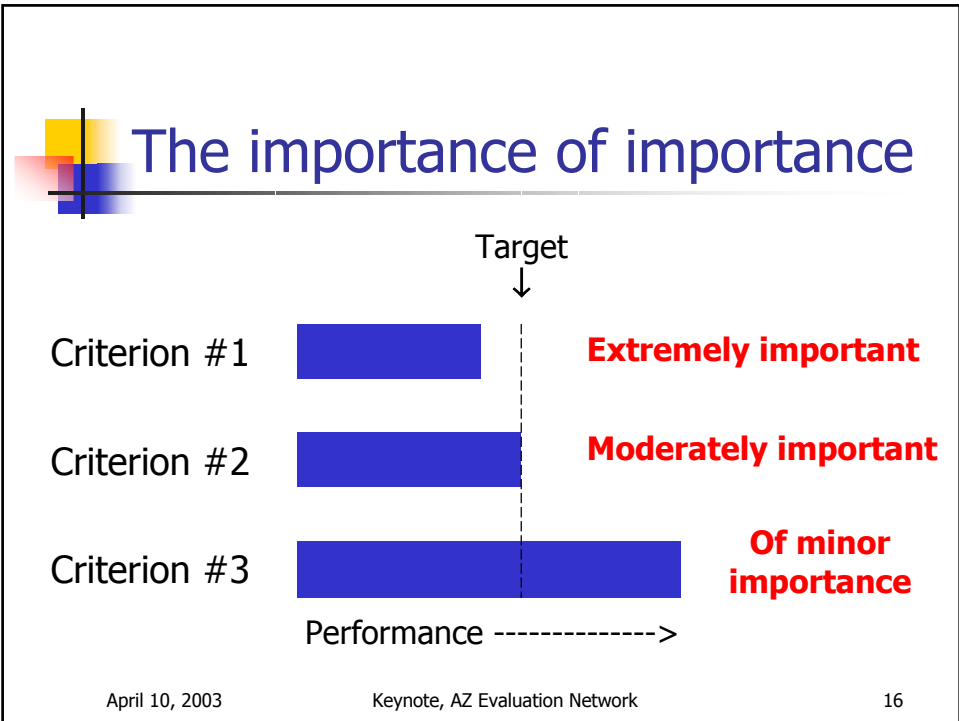
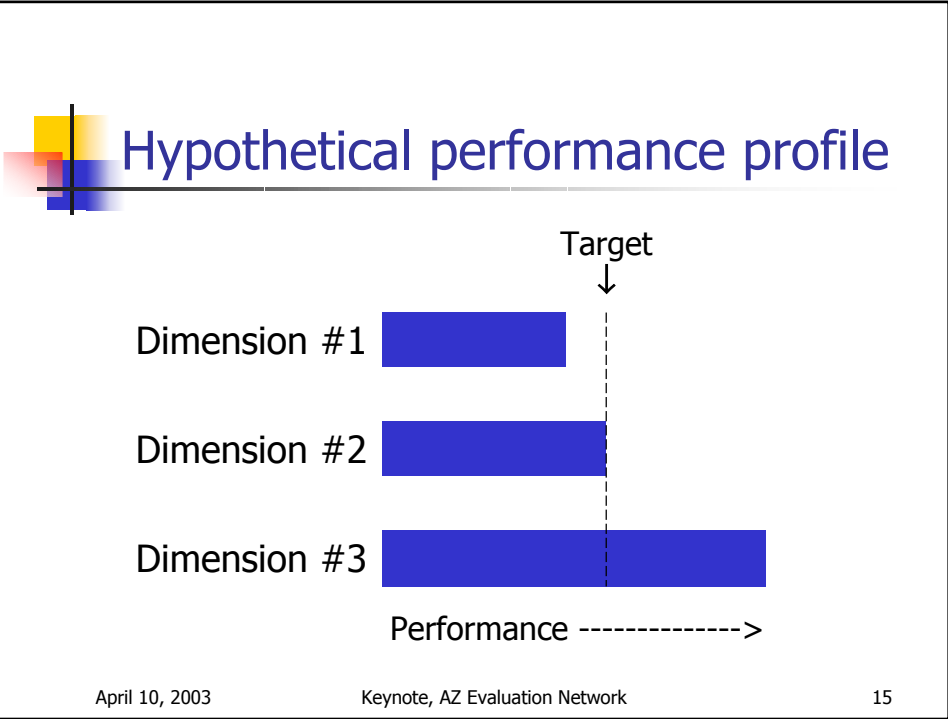
Excellent	Evidence of a strong positive impact: Very positive comments, with a substantial number that indicated a very strong impact; few if any neutral or negative comments.
Good	Evidence of a noticeable positive impact: A good number of positive comments (few neutral or negative) clearly showed that the program had made a noticeable positive effect on recipients.
Satisfactory	Evidence of some positive impact: A mix of positive and negative comments, skewed somewhat toward the positive. Evidence points in the right direction, but not to a very noticeable impact.
Marginal	Little or no impact either way: A real mix of comments; no clear trend in either the positive or the negative direction.
Poor	Evidence of some negative impact: A mix of positive and negative comments, skewed somewhat toward the negative. Not enough evidence to call this a really noticeable negative impact.



Using Rubrics for Merit Determination

- Rubrics are equally useful for:
 - Independent evaluators
 - Participatory evaluations

- They function to:
 - Help people (especially participating stakeholders) clarify their thinking about “how good is good”
 - Provide a clear and transparent record of your reasoning for all interested stakeholders to see (and to remind you what you were thinking!)





Importance Determination

- Importance determination = the process of assigning labels to dimensions (aspects of the program) or components (parts of the program) to indicate their importance

April 10, 2003

Keynote, AZ Evaluation Network

17



Importance Determination

- Use expert judgment
- Have stakeholders “vote”
 - Whose views should you poll?
 - How should you ask them? *
- Investigate for yourself:
 - Probing interviews of key informants → judgment *
 - Evidence from the literature
 - Causal linkages



April 10, 2003

Keynote, AZ Evaluation Network

18



Having stakeholders “vote”

- Survey
 - Rate importance; average responses
 - Paired comparisons → ranked importance
- Group process
 - Rate importance, pool ratings, discuss differences, negotiate agreement
- Key issues:
 - When would opinion NOT be the right way to determine importance?
 - Who should you ask/NOT ask for input?




Probing interviews + judgment

Focus: Which outcome criteria are the most important here?


Probe: Actual or probable impact:

- How beneficial would it be overall if performance on this criterion were really excellent?
- How detrimental would it be overall if performance on this criterion were really poor?
- How do you know? Based on what evidence/experience?



		What evidence exists that it would be <u>detrimental</u> overall if the program did <i>very poorly</i> on this component or criterion?		
		Little/no evidence of potential detrimental impact	Clear evidence that some detrimental impact would be possible	Evidence that <i>unacceptably</i> detrimental impact possible
What evidence exists that it would be <u>beneficial</u> overall if the program did <i>very well</i> on this component or criterion?	Clear evidence that some beneficial impact would be possible	Somewhat important	Important	Important (and set a bar)
	Clear evidence of substantial beneficial impact	Important	Very important	Very important (and set a bar)
	Consistently a major determinant of evaluand quality/value	Very important	Extremely important	Extremely important (and set a bar)

April 10, 2003 Keynote, AZ Evaluation Network 21



Recap: The Methodological Toolkit

1. Identification of all relevant criteria
 - Multi-level needs assessment
 - Search for other relevant values
2. Determining merit of performance
3. Determining importance of criteria
4. Synthesis methodology
 - Pulling together the pros and cons to draw overall conclusions

April 10, 2003 Keynote, AZ Evaluation Network 22



Key Points to Remember

- For all evaluation-specific methodologies:
 - There are multiple options/tools to use, each with its own set of assumptions
 - Mix and match according to the needs of the situation (NOT just to your preferences or skills – get help if you need it)
- This is one of the least developed areas in evaluation, so there is plenty of room for new ideas!



For More Information

- Davidson, E. J.
(forthcoming, 2004).
The Multipurpose Evaluation Guidebook: The nuts and bolts of putting together a solid evaluation.
Sage Publications
- From “Where do I start?” to “How do I pull all this information together into a report for my client?”
 - Step-by-step guide, including checklists, rubrics, and rules of thumb for doing a real evaluation.