World-class evaluation & organisational consulting

Improving evaluation questions and answers: Getting actionable answers for real-world decision makers

E. Jane Davidson, Ph.D.

American Evaluation Association conference
Orlando, FL – 12 November 2009

1

Overview

- What makes evaluations useless for action?
  - Four far-too-common flaws

- What makes evaluations actionable?
  - A good set of 'big picture' evaluation questions
  - Incisive, evaluative answers
  - A radically different evaluation report structure

2

What makes evaluations useless?

- Unbelievably, NO evaluation questions at all! Hello??

- Having evaluation questions, but never actually answering them!

- Having "research questions" instead of "evaluation questions"

- No clear sense of what needs to be understood <u>across</u> a set of evaluations

3

The "no questions" evaluation

- No questions → no answers!

- So, what did these "evaluations" do?
  - Stated the broad purpose of the evaluation
  - Described the evaluand
  - Identified and documented "outcomes of interest" and/or "interesting stories"
  - Found some areas for improvement

- The end product?

Lots of tangential, barely related, wouldn't-it-be-nice-to-know stuff that seemed to be more about the evaluator's interests than anything the client might need to know.

You could read through these reports, they looked very plausible on the surface (to the untrained eye), but after wading through ALL those data, you'd still be none the wiser whether the whole program was a waste of time or not, or how serious the weaknesses were that were found.

The "no answers" evaluation

- Questions → but no answers??

- What did these "evaluations" do?
  - Trotted out the questions at the beginning
  - Leapt straight to measurement, interviews, etc
  - Presented results separately by source & data type
  - Absolutely no attempt to weave evidence
  - Absolutely no attempt to answer the questions!

- The end product?

The "research questions" evaluation

- Research questions → Descriptive/causal
                                            answers

- What did these "evaluations" do?
    - Asked descriptive or causal questions, usually at the micro (variable) level
    - Provided answers that described (e.g.) the size of outcome shift, <u>but not</u> whether that was substantial or worthwhile or problematic …
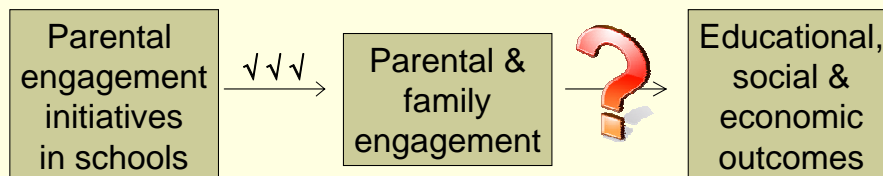
- The end product?

End product – pretty, interesting, scientific/rigorous-looking, but not very actionable

The "better evaluate it" evaluation

- Client gets programs evaluated, and relevant questions are covered each time, BUT …
- A lack of strategic thinking about what needs to be learned across a <u>set</u> of evaluations, e.g.

| Parental engagement initiatives in schools | √ √ √ → | Parental & family engagement | → ? | Educational, social & economic outcomes |

- The end product:
  - Many evaluations, but collectively they have missed important questions and answers

7

Loads of evaluations addressing that front end question: "What programs and strategies help increase parental engagement in the children's education?" but NONE that could answer the more important question of "Why invest in parental engagement?" Really critical in a recessionary environment when an entire agency division is using engagement as a fundamental strategy across programs but is unable to make an evidence-based case for why such initiatives should be invested in.

Also, this may require NOT evaluating ALL programs, but devoting more serious evaluation resources to particular evaluations so that the important investment question can be asked.

What makes evaluations "actionable"?

- Back to the strategic level – making sure the questions cover the 'big picture' issues

- Making the questions evaluative
  - What's so? → So what? → Now what?

- Making the answers <u>direct</u> and <u>explicitly evaluative</u>

8

Don't get lost in the details – step back every now and then, look across multiple evaluations and think what they tell you, what they don't tell you, what you need to know right now

It's not enough to ask whether parental involvement is related to educational and other outcomes (a research question); you need to know how strong the link is, whether it's a causal relationship, AND whether the amount of difference made is substantial enough to warrant the investment – that's the evaluative bit. THAT's the bit you can DO something with; that's the actionable answer.

A set of 'big picture' evaluation questions

- A list of 7 +/- 2 'big picture' questions
- All questions should be explicitly evaluative
- All questions should be roughly the same 'size'/level
- Questions should cover most or all of the following:
  - <u>Process</u> evaluation (quality of content, design, and delivery/implementation)
  - <u>Outcome</u> evaluation (value of outcomes)
  - <u>Learnings</u> (barriers and enablers, nuances, surprises, causal explanations / mechanisms / feedback loops)
  - Overarching questions about <u>value/worth</u>
  - <u>Forward/outward</u> focused evaluation questions (e.g. exportability, sustainability, threats, opportunities)

9

My evaluation questions "cheat sheet"

1. What was the <u>quality</u> of the program's content/design and how well was it implemented?

2. How <u>valuable</u> were the outcomes to participants? To the organization, the community, the economy?

3. What were the barriers and enablers that made the difference between <u>successful</u> and <u>disappointing</u> implementation and outcomes?

4. What else was learned (about how or why the effects were caused/prevented, <u>what went right/wrong</u>, lessons for next time)?

Note the explicitly evaluative words outlined.

More questions from the "cheat sheet"

5. Was the program <u>worth</u> implementing? Did the <u>value</u> of the outcomes <u>outweigh the value</u> of the resources used to obtain them?

6. To what extent did the program represent the <u>best possible use</u> of available resources to achieve outcomes of the greatest possible <u>value</u> to participants and the community?

7. To what extent is the program, or aspects of its content, design or delivery, likely to be <u>valuable</u> in other settings? How exportable is it?

8. How <u>strong</u> is the program's sustainability? Can it survive/grow in the future with limited additional resources?

11

Incisive, evaluative answers

- What's the antidote to …
  - Bureaucratic "weasel words"
  - "Academese"

- Some ideas:
  - Rehearse/discuss/create a guide for evidence interpretation BEFORE collecting the data
  - Use direct, evaluative terms in your answers – like "excellent," "minimally effective," "inadequate"
  - Structure the evaluation report using the questions as the framework

12

How can we solve the problem of reports that leave us none the wiser about whether an outcome was pathetic or brilliant, or whether a program was a complete waste of time or not?

How can we weed out the fuzzy talk and give direct answers that people can DO something with?

One useful method is to rehearse evidence interpretation BEFORE data are collected. What would constitute an "excellent" outcome here, as opposed to "mediocre", etc?

Evidence Interpretation Rubric: *Presence, Participation & Engagement for Students with Special Needs*

Rating	Description
Excellent	Student is positively engaged in interpersonal relationships, social activities and shared learning experiences and is increasingly self-determining. Student is a valued member of the school community and may have some friendships.
Very Good	Student is frequently included in shared learning experiences and social activities. There is evidence of emerging positive interpersonal relationships.
Good	Student is participating and encouraged in appropriate classroom and playground opportunities and interactions.
Just Adequate	Student is attending school and participating in some appropriate classroom and playground activities and interactions.
Inadequate	Student is restricted from attending or participating in school activities or unacceptably high levels of unexcused absences or levels of participation are far lower than the child would be capable of.

13

Evaluative Rubrics - Scope

- May be used to:
    - Interpret outputs, outcomes, process (content, design, implementation)
    - Provide incisive, evaluative answers to broader evaluation questions based on a range of evidence
    - Generate an overall rating of a program's performance or capability

- For good examples of the latter two, see Session #625 – An evaluative approach to quality assurance in higher education (Friday 4.30pm)

14

A radically different report

- Executive Summary
  - 2 pages
  - 7 +/- 2 headings
  - Each heading is one of your key evaluation questions
  - 1-2 paragraphs under each heading
  - These 1-2 paragraphs directly and succinctly answer the question!

- Introduction (what, why)
- Methodology (& why)
- Values & Context
- Key Evaluation Questions
- Findings
  - 7 +/- 2 sections
  - Each section heading is one of your questions
  - Succinct answer, followed by the evidence

15

Some useful references

- Davidson, E. J. (2007). Unlearning some of our social scientist habits. *Journal of Multidisciplinary Evaluation, 4*(8), iii-vi*. http://jmde.com

- Davidson, E. J. (2004). *Evaluation methodology basics: The nuts and bolts of sound evaluation.* Thousand Oaks, CA: Sage.

16