

evaluation-specific methodology



Why it must be in your toolkit

Jane Davidson

Michael Scriven

Evaluation 2013



evaluation

merit

worth

significance

evaluation-specific methodology

not just finding out what people's values are

evaluation-specific methodology

not just determining the "bare facts"

the fact/value distinction is a false dichotomy

the argument that social scientists can't get into values is flawed

"it's too vague to be scientific"

good/bad, right/wrong is useful evaluation even if it's imprecise

evaluation-specific methodology

not social science
not positivism

the ghost of
positivism



geography

10 significant active
subdivisions
3-4 ancient ones
2-3 new ones

a discipline

professional associations in
100 countries
many journals
graduate degrees

a transdiscipline

a key “tool discipline” in
every other discipline
(cf. statistics, logic)

the alpha discipline

the quality control
system that legitimates
any potential discipline

history



evaluation

merit
worth
significance

evaluation-specific methodology

the methodologies for determining merit, worth, or significance

evaluation-specific

needs & values assessment
merit determination
importance weighting
evaluative synthesis
value-for-money analysis

not evaluation-specific

RCTs
statistics
interviews
surveys
content analysis
causal inference methods

not using evaluation-specific methodology?

→ not doing real evaluation

yes, other disciplines use ESM

it's still evaluation-specific

evaluative tasks

1. Identify & define relevant values
2. Weight the values
3. Validate values & weights
4. Gather evidence
5. Synthesis

evaluative tasks

1. Critical description of evaluand
2. The point of view (POV)
3. Dimensions of merit
4. Field work
5. Convert, synthesize via rubrics

rubrics

Evaluative rubrics paint a picture of what the mix of evidence should look like at different levels of performance.

ex. 1: NZQA

How well do learners achieve?

How well does the program match the needs of learners and other stakeholders?

How good is the teaching?

Rating	Performance Descriptors for Answering Key Evaluation Questions
Excellent	Performance is clearly exemplary in relation to the question. Very few or no gaps or weaknesses. Any gaps or weaknesses have no significant impact and are managed effectively.
Good	Performance is generally strong in relation to the question. Few gaps or weaknesses. Gaps or weaknesses have some impact but are mostly managed effectively.
Adequate	Performance is inconsistent in relation to the question. Some gaps or weaknesses have impact, and are not managed effectively. Meets minimum expectations/requirements, as far as can be determined
Poor	Performance is unacceptably weak in relation to the question. Significant gaps or weaknesses are not managed effectively. Does not meet minimum expectations/requirements

Source: NZQA's External Evaluation & Review framework

e.g. when rating "Good"		you need to show ...
Key points from performance descriptors	Performance is generally strong in relation to the question.	Specifically, what evidence led you to believe performance was "generally strong" – as opposed to "clearly exemplary" (excellent) or "inconsistent" (adequate)?
	Few gaps or weaknesses.	What were the gaps or weaknesses, and why should they be considered "few" in number?
	Gaps or weaknesses have some impact ... but are mostly managed effectively.	What impact do the gaps and weaknesses have? What, specifically, is the TEO doing to manage these gaps and weaknesses, and is this "effective management"

example

How good are our outcomes on **violent behavior of adolescents?**



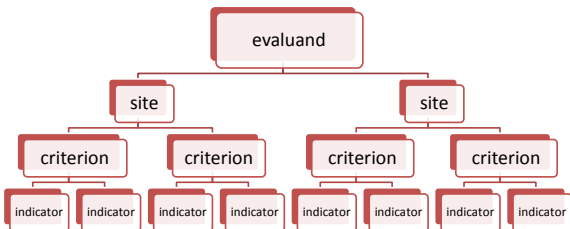
outcome: adolescent violent behavior

If outcome is:	... the evidence will look like this:
Excellent	In school and no trouble with the law; good skills and support to identify triggers and prevent violent outbursts. No serious trouble at home evident. Proactive about sharing skills and helping friends stay out of trouble.
Very Good	Starting to show signs of outbursts but using coping strategies to defuse the situation before it got worse (usually). Starting to share their new skills with classmates, friends and families.
Good	Recognition that outbursts are unacceptable; outbursts are now quite infrequent; they show remorse.
Just OK	Some outbursts at home or school and elsewhere, but none very serious or causing serious damage or injury.
Unacceptable	Any one or more of the following: <ul style="list-style-type: none"> • Serious and frequent outbursts causing injury/serious damage • Violence-related brushes with the law, especially arrests • Stood down from school for violence

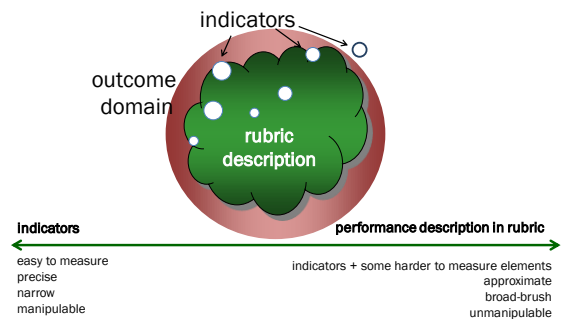
- individual level
- site level
- program level
- policy level



synthesis



indicators



What's So?

non-evaluative facts
(qualitative, quantitative, & mixed method evidence)

+

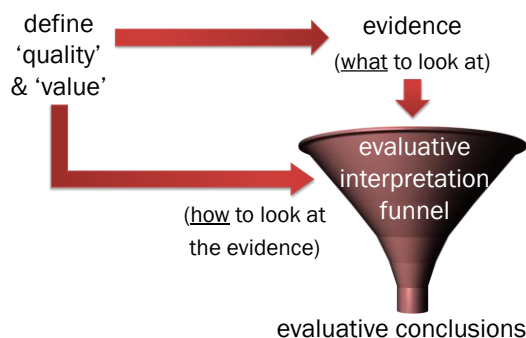
definitions of "quality" & "value"
(also facts, evidence-based e.g. needs, potential, aspirations)

→

So What?

evaluative conclusions
(saying something explicit about quality, value)

evaluative reasoning



your mission

what will YOU do differently now?

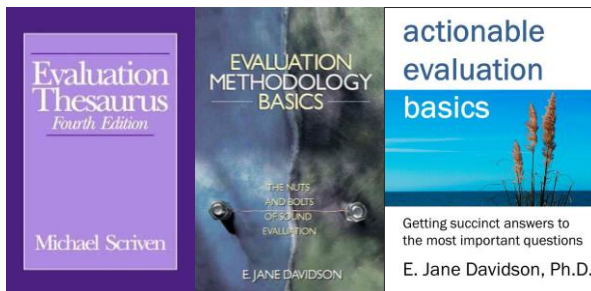
more information

presentations at AEA
books
websites

presentations

- Evaluation-Specific Methodology: Why it must be in your toolkit (Thu 11am, 196)
- Rubrics: What, why, and how (Fri 8am, 444)
- Defining Which Side of the Elephant We Are On: Using Rubrics on Non-Profit & Govt Program Evaluation (Fri 11am, 517)

ESM books



Michael Scriven
<http://michaelscriven.info>

Home
 Papers and Publications
 Faster Forward Fund
 Bio & C.V.
 Presentation: WMI/ Evaluation Cafe 9/11/12
 CEC's Free Webinar Series: Spring 2013

Home
 Welcome and thanks for visiting (or revisiting) my website!
 On this site I've posted a variety of items (about evaluation, philosophy of science and computers) that range from recent and completely unpublished work to newly revised earlier publications. It's a work in progress because I'm hoping to get reactions from readers, from the merely critical--always appreciated--to constructive suggestions for additions/explications, all the way from typos to fundamental errors. Send them to me at mjs@MichaelScriven.info, and I'll respond as usual behind my protective "key" if I think of

home develop learn how read about <http://RealEvaluation.com>

Real Evaluation
 Actionable evaluation for the real world

develop
 in-house capability
 tools & systems
 insight & ideas
 actionable answers

read
 jane's book
 jane's e-book
 publications
 genuine blog

learn how
 tips & guides
 workshops
 seminars
 e-learning

about
 real evaluation
 jane
 join
 contact

actionable answers to the most important questions

email updates
 Email Address

 or specify topics of interest when you sign up

twitter feed
 @janedavidson
 @janedavidson
 like to share the outcomes and impacts are basically the same things, changes that happen to people or communities, or the impact.
 @janedavidson
 The Friday Funny: World's First Placebo Hospital
 How NOT to evaluate proposals

genuine blog
 The Friday Funny: World's First Placebo Hospital
 How NOT to evaluate proposals

Genuine Evaluation
 Patricia Rogers and E. Jane Davidson blog about real, genuine, authentic, practical evaluation

Evaluation-Specific Methodology: The methodologies that are distinctive to evaluation
 Posted by Jane Davidson on March 22nd 2012
 I started a few syndrims last week when I mentioned the idea of Evaluation-Specific Methodology (ESM) as being an essential part of what defines us as a discipline. Of course, a large proportion of people who identify as evaluators consider that evaluation is merely the application of social science research methods to support decisions.
 Read the whole post -> Evaluation-Specific Methodology: the methodologies that are distinctive to evaluation

What is evaluation? Getting clarity about who we are as a profession, and a discipline
 Posted by Jane Davidson on March 22nd 2012
 Getting the definition of evaluation right is not simply a matter of having a popularity vote about it.

Search
 Recent Posts
 Time for evaluation to take the lead of the social sciences - Scriven's battle plan
 The Friday Funny: World's First Placebo Hospital
 How NOT to evaluate proposals
 Recent Comments
 Sheila B. Robinson on Time for evaluation to take the lead of the social sciences - Scriven's battle plan

BetterEvaluation
 Beta
 Select Language | Join | Login

Sharing Information to Improve Evaluation

HOME START HERE FIND OPTIONS APPROACHES THEMES PAGES ABOUT US CONTACT US

Features Blog Members New Studies Events

Synthesize
 Contribute Content

Bringing together data into an overall conclusion and judgement is important for individual evaluations and also when summarizing evidence from multiple evaluations.

Download a summary of the tasks, options, and approaches associated with synthesizing data from one or more evaluations.

Share
 RSS
 Print version
 Add comment

Manage Evaluation
 Define
 Frame
 Describe
 Understand Cases
 Synthesize
 Report And Support Use

Ask a Question
 Ask the Forum

Tasks
 1. Synthesize data from a single evaluation
 An evaluation needs to produce an overall judgement of merit or worth, bringing together data in terms of the agreed evaluator criteria and standards.
 2. Synthesize data across evaluations
 Data from multiple evaluations can also be synthesized to produce an overall judgement about 'what works' or 'what works for whom in what circumstances'.