

# Evidence Battles in Evaluation: How can we do better?

Mel Mark  
Penn State University

# DES Conference 2013

## Three tracks:

1. Evaluation as a force for change
2. New & old roads in impact evaluation
3. Evaluation as a forward-looking perspective

This talk in relation to the  
three conference tracks:

1. Evaluation as a force for change

**2. New & old roads in impact  
evaluation**

3. Evaluation as a forward-looking perspective

# First, apologies

- Language
- Examples
- “Humor”
- And more

# Partial and selective history of evidence (RCT) battles

- Earlier, the paradigm wars
- RCTs: quick overview
- Use of RCTs varied by field
- In US, Department of Education
- International development evaluation
- Pushback, debate, association statements

# The debate

- Generally unproductive, e.g.
  - Talking past each other
  - Critiques, not even-handed
- *On the surface*, about methods
- Instead, probably about other issues, e.g.
  - Role of impact evaluation
  - Relative merits of RCTs for impact evaluation
  - More on these to come
- Aside: “Strange bedfellows”

Instead of method debate,  
consider 'deeper' issues

# Should an impact evaluation be done?

- For early figures, e.g. Campbell
- Assumes “fork in the road”
- But other purposes of evaluation exist:





# **Many evaluation theories, emphasizing different evaluation purposes, e.g.**

- Impact evaluations for selection from among options
  - Info needs of program managers; program improvement
  - Social justice
  - Empowerment of individuals
  - Creating forum for democratic deliberation
  - Development of learning organizations
  - Ongoing construction of an initiative
  - And on and on
- 
- Aside: Knowledge of associated theories as part of content knowledge of an evaluator

# Beyond the many evaluation theories, multiple questions for evaluation, e.g.

- Feasibility of implementing a new program type
- Quality of implementation
- Compliance with regulations, e.g. about client eligibility
- Cost
- Client compliance, retention, perceptions
- Ability to scale up
- **Question:** Is impact the right question, for a given evaluation?

# RCT advocates vs critics: Each side's view of the role of impact evaluation

- Guess.
- Aside: Advocacy of RCTs, and 'gold standard' language, may be an effort to make impact evaluation more salient among policy makers, evaluation funders?

# ***IF*** impact is right question, is RCT useful relative to other methods?

- Needed?
- Practical?
- Ethical?
- Overkill?
  
- Compared to alternative methods,
- And with what method ancillaries for other questions?

# Alternative methods

- Long list (including regression-discontinuity, time series, various quasi-experiments, comparative case studies, participant statements, ....)
- Circumstances may favor or prohibit alternative methods

# RCT advocates vs critics: Each side's view of RCT's comparative advantage

- Guess

# Issues of trade-offs: Estimating effects vs generalizing



# Where to now? 1

- Regarding debates (this and future)
  - Try to find deeper sources of disagreement
    - E.g., role of impact evaluation; whether RCTs are generally preferable for impact evaluation
  - Try to understand other's assumptions, try not to talk past each other
  - Even-handed assessments of one's preferred and not preferred options
  - Less heat, more light

# Where to now? 2

- Evidence hierarchy, not ideal
- Evidence typology, or contingency tree, an option, but may:
  - Ignore specifics
  - Be cumbersome, or incomplete, or both
  - Stifle innovation
  - Ignore quality of information needed
- May still suggest better vs worse options

# An alternative:

- Informed process for selecting evaluation method (given evaluation question, context, etc).
- Leads to questions, e.g.,
  - Evaluation policy that describes
  - The location, organization, independence of evaluation unit
  - Advisory and/or review processes
- “Frameworks as an aid...”

# And keep in mind

- The ‘guiding star’ is not method choice per se
- It’s the potential for evaluation to make a difference, to have positive consequences, to contribute to social betterment
  - Think of evaluation as an intervention
  - Consider the equivalent of “program theory”

Q&A.

Closing