Current Changes and Challenges in the Field of STI Policy-Making and the Effects These Might Have on STI Evaluation Theory and Practice

Irwin Feller, Professor Emeritus, Economics,
The Pennsylvania State University

Presentation to Conference, Evaluation of STI Policies, Instruments and Organizations: New Horizons and New Challenges

Vienna, Austria 14-15, November 2013

Season of (Movie) Sequels

- Iron Man 3
- Fast and Furious 6
- Despicable Me 2
- R.E.D. 2
  (Retired but Extremely Dangerous)

3 Topics

- Shift from Paradigm Wars to Evidence Wars
- Responses to the Marburger Questions
- Big Data

Organization

New Horizons

- Demand & Supply of Evaluation-based Evidence
- Science of Science Policy
- Big Data Revolution

Challenges

- What Constitutes Acceptable Evidence?
- Mapping of Questions & Answers
- Measurement without Theory (as a Guide to Policy)

Observations of a Reflexive Practitioner

- Evaluator; Policy Researcher; Social Scientist
- International experiences
- Examples from US experiences

Four Purposes of Evaluation

1. **Assessment of merit and worth:** the development of warranted judgments, at the individual and societal level, of the value of a policy or program

2. **Program and organizational improvement:** the effort to use information to directly modify and enhance program operations

3. **Oversight and compliance:** the assessment of the extent to which a program follows the directives of statues, regulations, rules, mandated standards or any other formal expectations

4. **Knowledge development:** the discovery or testing of general theories, propositions, and hypotheses in the context of policies and programs

Mark, Henry and Jules, Evaluation, p. 31

The End to (Truce in) Paradigm Wars?

"The origins of this volume can be found in the long-standing antagonisms between qualitative and quantitative researchers in evaluation"
(Reichardt and Rallis, The Qualitative-Quantitative Debate, 1994)

New Horizon: The Pursuit of Evidence

"Based on our rough calculations, less than $1 out of every $100 of government spending is backed by even the most basic evidence that the money is being spent wisely".

Bridgeland and Orzag, The Atlantic,2013

Excellence & Evaluation

"Rigorous, independent program evaluations can be a key resource in determining whether government programs are achieving their intended outcomes as well as possible and at the lowest possible cost…Federal programs have rarely evaluated multiple approaches to the same problem with the goal of identifying which ones are the most effective"

(U.S. Office of Management and Budget, 2009; emphasis added)

Challenges: What Constitutes Acceptable Evidence?

- Apotheosis of randomized control tests
- Increased use of formal meta-analysis to sift multiple, often conflicting, studies (Cochrane Collaboration; Campbell Collaboration; Coalition for Evidence-based Policy)
- Limited efforts at cross-program evaluations

Opportunity Costs of Not Comparing Alternatives: US SBIR Program

- NRC formative evaluation used to advance summative recommendation about increased funding
- No consideration of impact of recommendation on funding levels upon funds available for functionally comparable programs in other Federal agencies (e.g.NSF's university-industry centers programs)
- Incorrect "Lessons Learned" (OECD)

OECD Review of US SBIR Program

"The (SBIR) programme has no budget line and thus requires no new funds; it is therefore politically viable and independent of the budget process" (OECD, 2010, p.3)

The Risks of Policy Transfer (Evaluation Findings) Across Borders

"Once facts leave home, it is more difficult to keep them safe"

(Morgan, How Well Do Facts Travel?, p.6)

The Marburger Questions

- How much should a nation spend on science?
- What kind of science?
- How much from private versus public sectors?
- Does demand for funding by potential performers imply a shortage of funding or a surfeit of performers?”

New Horizons

- Revitalization of field of S&T research
- Increased funding for non-programmatic (evaluation) research
- Building stream of new findings that have important policy implications (e.g. “Growing Stem Cells: The Impact of Federal Funding Policy on the U.S. Scientific Frontier”, Furman, Murray, Stern)

Major Research Thrusts

- Estimating Returns from Public R&D Investments
- Changing Organization of Science
- Determinants of Adoption-Innovation Processes

- Impacts of Organizational Forms and Incentives on Scientific Productivity (Azoulay, Ziff, & Manso)
- Structure ("Backbone") of Science (Borner; Cohen)
- Predicting Scientific Impacts (Evans)

Challenges

- Research portfolio exhibits an admixture of congruent, tangential, orthogonal, and little connection to Marburger's policy-oriented, resource allocation questions

- Management of (Programmatic) Expectations : "limits to knowledge" as guide to future S&T investments/policies

- Willingness to tackle higher risk-higher reward research questions?

What Has a Science of Science Policy Wrought (To Date)?

"The prospects of answering the resource allocation questions that Dr. Marburger used to motivate his call for a Science of Science and Innovation Policy seem much less promising to me. As a practical matter I'm not sure this matters very much. Even with a better understanding of the answers to these questions, it would be hard to influence the political decision-making process that shapes the nation's scientific investments
(Rosenbloom, 2013; p. 22).

Mapping the Frontiers of Science Policy Research

"No theory of scientific progress exists, or is on the horizon, that allows prediction of the future development of new scientific idea or specifies how the different types of scientific progress influence each other-although they clearly are independent…"
(National Research Council, A Strategy for Assessing Science 2006, p. 73)

New Horizons: Big Data

- "Revolution" in the conduct of social science/evaluation research
- Assemble and integrate large(r) data sets
- Essential building blocks for the more precise testing of existing theories/policies and for the development of new theories/policies

Big Data Claims

"No longer do we necessarily require a valid substantive hypotheses about a phenomenon to begin to understand the world…In place of the hypothesis-driven approach, we can use a data-driven one. Our results may be less biased and more accurate, and we will certainly get them much faster"
(Mayer-Schonberger and Cukier, 2013, p. 55)

The Challenge Race on the Use of Big Data in Program Evaluation

1. Use Big Data to access, assemble, and sift large formerly discrete data sets to test implicit theories (logic models) of existing programs and to provide evidence for revised/new policies/programs

2. Search for patterns within the data "…dispenses with the need for hypotheses about how the world functions: "

Risks of Big Data Claims

- “Patterns’ likely consistent with multiple, competing interpretations and policies (absent any hypotheses)

- Conversion into simplistic performance measures and performance assessments

Lessons from Big Data?

Run or Pass?

Challenges

- Measurement without Theory
- Premature, bureaucratic and potentially dysfunctional use in performance measurement systems
- Gaming; Swarming