

# Evaluating Government Funded Research

Karen Sollins  
DOE ASCAC meeting  
May 1-2, 2002

# Sources

- NSF
- DARPA
- NIST ATP
- AAAS panel session
- OMB “Governing with Accountability”

# NSF

- Two primary review criteria
  - Intellectual Merit
  - Broader Impact
- Regular reporting - on approval of cognizant program officer
- GPRA
  - NSF got approval to report in text rather than numerically
  - The effects of research by “nuggets”

# DARPA

- Much more informal
- Criteria published in BAAs
- An informal evaluation of tech transfer, both intended and unintended
- Heilmeier's criteria
  - Director of DARPA in the mid 1970's

# Heilmeier's criteria (used informally)

- What are you trying to do? Articulate your objectives using absolutely no jargon.
- How is it done today, and what are the limits of current practice?
- What's new in your approach and why do you think it will be successful?
- Who cares? If you're successful, what difference will it make?
- What are the risks and the payoffs?
- How much will it cost? How long will it take?
- What are the midterm and final "exams" to check for success?

# NIST ATP (Advanced Tech. Prog)

- “... bridges the gap between the research lab and the market place, stimulating prosperity through innovation.”
- Peer review of projects
- Includes both internal and external contracted review (National Bureau of Economic Research)
- NRC study: assessing the outcomes (“An exceptional assessment effort”)

# NRC conclusions

1. Effective federal partnership program
2. Peer review of both technical feasibility and commercial potential critical to success of program
3. The high quality of both internal and external assessment lend credence to evaluation of accomplishments
4. Assessment indicate achievement of core objectives
5. Some recommendations for specific procedural improvements

# AAAS panel: Governing Science and Science in Government

- The evaluation community and me
  - Professional evaluators - private sector consultants (non US), US federal people
  - Theoreticians (academics): hold to belief that can't do good evaluation without a good theory of evaluation
- Run workshops and study groups, publish books and reports, etc.

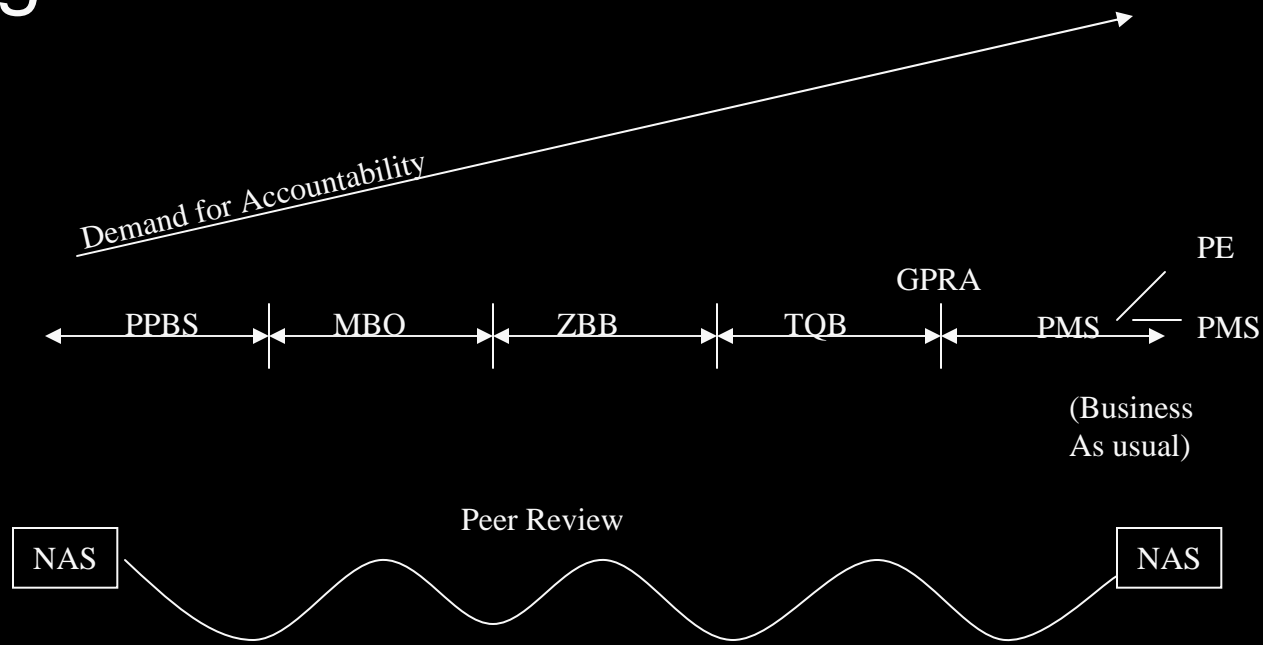


# The Professional Evaluators

- Evaluation important
- Rarely include peer review, although sometimes include expert review (not always)
- Never mentioned impact of evaluation process on work being evaluated
- One speaker reported less than stellar impact in most European venues
- Think highly of GPRA, but it was their jobs to do so
- May reflect significant percentage cost of program

# The Academics

- Evaluation programs not new
- GPRA longer lifetime because legislative



## The Academics (part 2)

- “utilization based evaluation”
  - Goal displacement
  - Decisions precede evaluation
  - Organizational burden
  - The problem of failures
  - At best loosely connected to the political process of agency funding decisions
  - Generally provides little insight into agency priority setting
  - Little connection to larger question of national priority setting

# “Governing with Accountability” (White House/OMB)

- Attract stronger employees to the gov't
- Expose tasks to competition to improve “customer” service while controlling cost
- Improve financial management
- Harness the power of the Internet to improve government
- Start the process of linking resource decisions to results

*Note: This does not address the question of how to set priorities among agencies and programs, nor how to evaluate research.*

# In the document

- The Scorecard (see transparency)
  - To be repeated every 6 months
  - Evaluation left to outside contractor - choice made by White House.
- Legislation to allow for more management freedom within agencies and offices
- Permanent reorganization authority
- Use of scorecard to move programs from one organization to another

*Note: These can be viewed positively or negatively. There was concern in the AAAS mtg that moving programs might be done as punishment for agencies not meeting with approval.*

# Conclusions

- Metrics and evaluation extremely challenging
- Often don't match decision-making process
- Decision-making may be very (completely) political

# Some challenging questions

- Why are measurement and evaluation being done? What will be learned by them?
- What is being measured and how does evaluation proceed from that?
- What decisions might be influenced by the results of the process?
- What are the impacts, elsewhere besides on the decision-making process, e.g. on the agenda itself or the rest of the peer community?
- For the Office of ASC, should there be separate measurement and evaluation procedures for the research in infrastructural components of the agenda?
- How, if at all, does all this relate to strategic planning?