

# The Purposes and Values of Peer Review



Ed Hackett  
Daryl Chubin



# Our aims in this talk:

- Context of peer review
- Institutional analysis of peer review
  - What does it do?
    - Intended and unintended consequences
  - What values or principles guide it?
  - Values compete: ambivalence
- Choices and challenges for IES



# Ways To Allocate Funds For Science

- Congress: Earmarking & Pork Barrelling
  - Democratic
  - Legitimate
  - Distributional fairness
  - Political
  - Inexpert
  - Culturally corrosive
  - Around \$1.8B in FY02



# Another Way To Allocate

- Strong Manager (DARPA)
  - Flexible and responsive
  - Assumes clear objectives and standards
  - Requires outcome accountability
  - May not work for all fields
  - Projects **end**, programs **sustain** fields
  - May not scale up (to, say, NIH's \$25B)
  - Must accept failure and cut losses



# One Final Option...

- Formula funding
  - To states or institutions or departments?
    - Then peer review? **Another formula?**
  - Who writes the formula, and how?
  - Risk-taking and responsiveness?
  - Caring for the young and putting the elderly to pasture
  - Gaming and unintended outcomes



# Two ideas to hold on to

**Boundary objects** (Star and Griesemer, 1987),  
boundary organizations (Guston, 2000),  
boundary processes (us, now)

Peer review spans science and policy,  
academe and government, knowledge  
production and professional practice

**Legitimacy:** what makes peer review  
acceptable *to diverse populations* as a way to  
slice the melon?



# More than a trendy trope...

- In peer review there is a mix of
  - Communities
  - Purposes
  - Evidential standards
  - Argumentative procedures
  - Ethical standards
  - Epistemic cultures



# What is peer review?

- A process for “grading the grain” and allocating scarce resources, of course.
- But also much more...



# A Source of Expert Advice

- To the proposer and to the agency
- Aims to improve science by advice and by advising wise allocations
- Cumulatively, it shapes the research area and the agency research program



# A Flywheel

- Lends stability
- Embodies the “essential tension” between tradition and innovation
- Challenges new ideas to be truly new
  - And insists on connection to tradition
- Helps researchers “stay the course” through obstacles of research



# Mode of Scholarly Communication

- Ideas circulate among influentials in various forms, benefiting from expert advice while preparing the field to accept them
- May speed acceptance of original ideas and results
  - Compare with findings springing from formula-funded research groups



# Entry Point for Social Considerations

- Currently formalized in GPRA and new review criteria at NIH and NSF—societal benefit formally equal to scientific merit
- Program officer balancing portfolio (gender, ethnicity, geography, undergrad institutions)
- NIH Advisory Councils and “specials”
- Is lay participation a rising trend?



# Enactment of Professional Authority

- A buffer or boundary that separates science from other spheres (we don't use peer review to fund highways!)
- Symbolic importance
- Practical benefit of balancing citizen influence and expert autonomy



# Competing Values

- Initially social values were thought to be unequivocal, the standards of goodness, truth, beauty and such that a society shared.
- Robert Merton (again) offered the notion of “sociological ambivalence” to capture idea of values in tension.
- Erikson pursued it further with cultural axes
- When you think about it, peer review serves competing values, honor some and compromise others



# Openness-Secrecy

- Procedures, criteria, rating scales and such are knowable, often published, and transparent.
- Criteria are applied in the same way to all.



# But peer review is secret, too

- Confidentiality of proposals and reviews
- NSF panelists are concealed and panel minutes are minimal
- Anonymity of proposers and reviewers
- Even process of choosing reviewers and panelists may seem secret to some




# Effectiveness-Efficiency

- We ask peer review to accomplish a lot—see purposes, above—and to do all that quite well.
- NAPA report on NSF peer review complains that reviewers don't apply review criteria fully and fairly, which amounts to a complaint about effectiveness.



# But we demand efficiency

- Reviewers are not paid and panelists generally receive modest honoraria (\$280/day for *meeting time*)
- Complaints occasionally arise about how much peer review costs to administer (NSF staff and travel).
- And complaints also arise about the time it takes to review all this stuff.



# Sensitivity-Selectivity

- A sensitive selection process detects merit wherever it is found.
- A selective process filters out work of dubious quality or significance.
- **But** sensitivity is gained at the cost of selectivity, and selectivity compromises sensitivity.
- Each suits different fields, different days



# Responsive-Inertial

- Kuhn's "essential tension" again
- Peer review should respond to every new theory, method, topic, need
- Peer review should also guard the body of knowledge, accepted methods, agreed priorities, and the like...the tradition of the field.
- And peer review is rapped for both.




# Meritocratic-Fair

- Merit: Evaluate all the science, only the science, and nothing but the science
  - separate from other stuff in a proposal
- Fair: Do so in ways that are considered fair in this society, which may mean apply substantive or gender or ethnic or geographic preferences.



# Reliability-Validity

- A reliable measure has little random error: repeated measures agree, different measurer agree.
- A valid measure measures what it is supposed to and not much else.
- Measurement theory tells us that reliability sets an upper bound on validity



# So is inconsistency (unreliability) a problem?

- Scientific quality has many elements, perhaps even many dimensions
- They are unequal in importance
- They interact
- A **valid** review addresses all of these elements and tries to form a weighted composite, with some thought to interactions.



# It gets worse...

- Reviewers are not equally competent to evaluate all aspects of a proposal, nor would fully competent reviewers necessarily agree about weighting.
- So a **valid** review is patched together of diverse reviewers assessing diverse facets of a proposal.
- Parenthetically, in all this don't mistake measures of *validity* for validity itself—an empiricist fallacy.



# So reliability suffers....

- This patchwork—or think of a matrix of reviewer skills by proposal elements—is a recipe for low inter-rater agreement and **low reliability**.
- In fact, if a program officer knew a set of reviewers would agree in their ratings of a proposal, she would only assign one of them!



# Therefore ...

- The pursuit of reliability would compromise validity, and the pursuit of validity jeopardizes reliability.
- Which is more important to you?
- Perhaps I say this only to tweak Hal Arkes, but

It seems fair to say that validity sets an upper limit on reliability!



# Choices and Challenges

- Choose which purposes and values are most important to IES.
- Some preliminary choices confront us now.
- Settled choices will emerge with time and experience.



# Boundaries, Performance, Learning, and Legitimacy

- Policies and procedures are a score without an orchestra:
  - Build a community and culture for reviewers and program officers
  - Process will take time and will be shaped by actions (not just scripts)
  - Boundary process=diverse actors
  - IES must be a robust *learning* organization
  - Legitimacy and unintended consequences



# The Szilard Point

- At a success rate of 10% or so, Leo Szilard noted, one would have to write proposals full-time to remain funded.
- Some education funding rates approach that level...
  - Challenges legitimacy: “why bother?”
  - Exceeds the “resolving power” of peer review, which erodes legitimacy and culture
  - Limits participation, weakening community-building effort



# Bridging policy and practice

- Education is ideal for this! Better teaching and learning is a societal goal that research can serve.
- Involving practitioners in peer review would add this dimension (which many researchers lack in their proposals & reviews)
- **BUT** delicate sensibilities of professionalism may be at stake (cf., ecology and environmentalism)



# Pluralist possibilities?

- With mixed objectives of fundamental knowledge and practical gains, accommodating diverse constituencies and building a research community, IES may be well served by mixing peer review with pockets of strongly managed projects.
- What to do which way, and how to combine them?