

Risks of Quantitative Measurements When Abused

Presented at

*2006 American Evaluation Association Conference
“The Consequences of Evaluation”*

*Portland, Oregon
November 2, 2006*

Yongsuk Jang, Ph.D.
Research Scientist
Center for International Science and Technology Policy
The George Washington University

Email: jang@gwu.edu



Core Messages

- **NOT** Discouraging Efforts toward Quantitative Measurements for S&T
- **BUT** Making Precautionary Signal to Users of Quantitative Measurements for S&T

Background I : Peer Review

- **The Most Prevailing Methodology for Evaluating Science, Technology, and Innovation**
- **Best Assessed by Peers – Experts in their own fields**
- **Expert Review, Committee Review, Council, etc.**
- **Because of Difficulties of Evaluation of S&T**
- **Qualitative Measurements, Basically**
- **Limited Usefulness**

Background II : Quantitative Measurement

- **Strong Demands on Quantitative Indicators**
- **Eager to Seek or Develop Better or New Quantitative Measurements / Methodologies**
- **Worldwide Trend : in All Advanced Countries as well as in the U.S.**
- **Performance-based Program Management & Budgeting**
- **Accountability and Efficiency**

Background III : Dr. Marburger's Initiative

- **Science on Science Policy**
- **Limitations of Prior and Current Data in Understanding Modern Innovation**
- **Need of Better and Rich Information for Robust Science Policies**
- **Development of New Innovation Models**
- **Focus Mainly on Quantitative Indicators**

Why Quantitative Measurements?

- **Objectivity**

- Numbers are Neutral & Universal
- Can Avoid Controversial Situations
- Transparent to the Public as well as Stakeholders

- **Comparability**

- Can Compare across Fields
- International Benchmarks
- Comparable at Aggregate Levels (National, Agency, Local, etc.)

- **Easy to Use**

- Numbers are Clear to Understand
- Can Easily Compare between Losers and Winners
- Useful in Persuading Losers

Where Are We with Quantitative Indicators ?

- **Primarily, Input & Output Measures**
 - R&D Expenditure, Number of S&E Researchers, etc.
 - Number of Publications & Patents
 - Some Econometric Methodologies (e.g., Social Rate of Return, etc.) – but many Problems such as attribution, time lag, etc.
- **Obsolete Innovation Models**
 - Simple Linear Model of Innovation
 - Innovation Processes remained largely as a Black Box
 - Induce to focus primarily on dollar values of inputs & outputs
- **Simple Conversion of Subjective Judgment to Numbers**
 - Majority of Quantitative Indicators are simple conversions of Qualitative Measurements
 - **Pertain all limitations of Qualitative Measurements**



When We Mean “Abused” ?

- **Over-utilized or Over-emphasized**
 - Solely or Heavily depend on single or a few number of Quantitative Indicators
 - Countable Inputs & Outputs are only Small Parts of Innovation - e.g., R&D expenditure, number of publications & patents
- **Wrong Indicators in Wrong Places**
 - Intentionally Manipulating Indicators or Unawarely Adopting Wrong Indicators
- **Interpret Numbers as Absolute Values**
 - Majority of Numbers are Meaningful only in a Context
 - National, Sectoral, Regional Differences change Meanings of Numbers
- **Ignore Time Lag**
 - R&D realized in a long-term
 - Not Linearly Progressive

Risks of Quantitative Measurement When Abused

- **Distort Program Direction or Purposes**
 - Significantly Affect on Program Mgt. & Operation
 - Frequently Numbers themselves become the Goals (e.g., SCI Papers)
- **Change the Behaviors of Researchers**
 - Significantly Affect on Researchers Behaviors
 - Most Likely to make them Short-Sighted
- **Neglect Other Important Elements**
 - Recently, Behavioral Additionalities
- **Misguide Program Management or Science Policy**
 - Provide Misleading Data
 - Annual Evaluation with Numbers may Kill Wrong ones

Case I : Pre-Budget Review in Korea

Expert Review

- 1 Steering Committee
- 3 Sub-Committees by Program groups
- 12 Working-Committees by specific areas
- 255 experts from Industry (20), Univ. (95), GRIs (140)

Evaluation Criteria

- Evaluation Results of Public R&D Programs
- Relevance to the Program Goals and Scopes
- Appropriateness of Government Agencies for the Programs

Forced Distribution in Weighted Scoring

- A (Upper 10%)
- B (10 – 30%)
- C (30 – 70%)
- D (70 – 90%)
- E (Lower 10%)

Descriptive Recommendations

- Re-Planning
- Redirecting its focus on Special Area
- Linking or Merging with Other Program
- Increasing or Reducing Budget

Case II :Program Assessment Rating Tool (PART) in the US

- **OMB-Initiated Assessment Tool**
- **Assessing the Program Performances**
- **Under Initiative of
“Budget and Performance Integration”**
- **Cover 20% of Programs Every Year**
- **Cover all Programs by FY 2006 Request**
- **Eight Program Types**

Four Sections of Questionnaire for PART

**OMB Guidelines for assessing all Fed. Programs (including R&D)
in following four criteria:**

(Will apply to all Federal Programs by FY 2006)

- Program Purpose/Relevance/Federal Role (20%)**
- Strategic Planning (10%)**
- Program Management (20%)**
- Program Results (50%)**

A Typical PART Assessment at OMB

Program: Basic Energy Sciences

Agency: Department of Energy

Bureau: Office of Science

Rating: Effective

Program Type: Research and Development Competitive Grant, Capital Assets and

Last Assessed: 1 year ago

Key Performance Measures from Latest PART	Year	Target	Actual	Recommended Follow-up Actions	Status
Long-term Measure: Progress in designing, modeling, fabricating, characterizing, analyzing, assembling, and using a variety of new materials and structures, including metals, alloys, ceramics, polymers, biomaterials and more--particularly at the nanoscale--for energy-related applications. An independent expert panel will conduct a review and rate progress (excellent, adequate, poor) on a triennial basis.	2006	Excellent		The Department will continue to improve performance reporting and centralize management and planning of operations at its user facilities.	Action taken, but not completed
	2009	Excellent			
	2012	Excellent			
	2015	Excellent			
Annual Efficiency Measure: Average achieved operation time of the scientific user facilities as a percentage of the total scheduled annual operation time.	2003	>90%	91%	The Department will work with its advisory committee to develop research milestones [by September, 2004] against which future outside panels may judge interim progress toward achieving the long-term goals of the program.	Action taken, but not completed
	2004	>90%	92%		
	2005	>90%			
	2006	>90%			
Annual Measure: Improve Spatial Resolution: Demonstrated spatial resolutions for imaging in the hard and soft x-ray regions, and spatial information limit for an electron microscope (measured in nanometers).	2003		130, 20, 0.09		
	2004	<115, <19, <0.08	100, 19, 0.08		
	2005	<100, <18, <0.08			
	2006	<100, <18, <0.08			

Update on Follow-up Actions:

(1) BES has responded to the DOE IG report on performance report at two of its light sources, and is in the process of centralizing the management, planning, and condition and utilization metric reporting for the beamlines at its user facilities. (2) Long-term goals are now included in grant solicitations, but performance reporting at the grantee/contractor level for the entire Office of Science is not yet transparent and readily accessible. (3) The program's research milestones--as expressed in the new DOE program plans--were produced and reflect the strategic goals of the program, but the BES advisory committee has yet to formally comment on the milestones.

Program Funding Level (in millions of dollars)

<u>2004 Actual</u>	<u>2005 Estimate</u>	<u>2006 Estimate</u>
1,011	1,105	1,146



What Can We Do ?

- **Educate Users (Policy Makers / Program Managers)**
 - To clearly recognize the Risks of Quantitative Indicators When Abusively Adopted
 - To Be very Cautious in Adopting Quantitative Indicators for Program Evaluation
 - To Estimate in advance the Impacts on Program Direction and/or Researchers' Behaviors
 - Bottom Line: How Wisely Utilize the Quantitative Measurements?
- **Supplement with Additional Information**
 - Qualitative Measurements can provide useful information
 - Encourage Use of Multiple Indicators
- **Innovate Evaluation with New Innovation Models**
 - Right Indicators are possible from Right Theories