

## **SCIENCE AND TECHNOLOGY TRANSITION METRICS**

BY

Dr. Ronald N. Kostoff  
Office of Naval Research  
800 N. Quincy St.  
Arlington, VA 22217  
Phone: 703-696-4198  
Fax: 703-696-4274  
Internet: [kostofr@onr.navy.mil](mailto:kostofr@onr.navy.mil)

*(THE VIEWS PRESENTED IN THIS PAPER ARE SOLELY THOSE OF THE AUTHOR, AND DO NOT REPRESENT THE VIEWS OF THE DEPARTMENT OF THE NAVY OR ANY OF ITS COMPONENTS)*

KEYWORDS: Metrics; Transitions; Science and Technology; Outputs; Impacts; Outcomes; Productivity.

### I. OVERVIEW

On 27 October 1998, a workshop was convened by the National Institute for Occupational Safety and Health (NIOSH) to identify key metrics for NIOSH's Strategic Goals. The first NIOSH Strategic Goal (Conduct a targeted program of research to reduce morbidity, injuries, and mortality among workers in high-priority areas and high-risk sectors) was the major focus of the workshop. Its two related Objectives addressed 1) the success in implementing a research program based on its 1996 National Occupational Research Agenda (NORA) priorities (NORA is a framework to guide occupational safety and health research into the next decade, and resulted in the establishment of a list of the top 21 research priorities) and 2) success in measuring its safety and health outcomes.

The author was invited to participate as a member of the panel. This report generalizes a document that the author prepared for the NIOSH workshop, and was further refined during preparation for a DOE-sponsored workshop on S&T benefits, 4-5 March 2002. The paper focuses on key metrics for evaluating progress in a mission-oriented research program. The results and conclusions of the analyses are sufficiently generic for applicability to any science and technology (S&T) sponsoring organization.

# Report Documentation Page

*Form Approved*  
*OMB No. 0704-0188*

Public reporting burden for the collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to a penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.

1. REPORT DATE <b>01 MAR 2004</b>	2. REPORT TYPE	3. DATES COVERED -	
4. TITLE AND SUBTITLE <b>SCIENCE AND TECHNOLOGY TRANSITION METRICS</b>		5a. CONTRACT NUMBER	
		5b. GRANT NUMBER	
		5c. PROGRAM ELEMENT NUMBER	
6. AUTHOR(S) <b>Ronald Kostoff</b>		5d. PROJECT NUMBER	
		5e. TASK NUMBER	
		5f. WORK UNIT NUMBER	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) <b>Office of Naval Research, 800 N. Quincy St., Arlington, VA, 22217</b>		8. PERFORMING ORGANIZATION REPORT NUMBER	
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)		10. SPONSOR/MONITOR'S ACRONYM(S)	
		11. SPONSOR/MONITOR'S REPORT NUMBER(S)	
12. DISTRIBUTION/AVAILABILITY STATEMENT <b>Approved for public release; distribution unlimited</b>			
13. SUPPLEMENTARY NOTES			
14. ABSTRACT <b>On 27 October 1998, a workshop was convened by the National Institute for Occupational Safety and Health (NIOSH) to identify key metrics for NIOSH's Strategic Goals. The first NIOSH Strategic Goal (Conduct a targeted program of research to reduce morbidity, injuries, and mortality among workers in high-priority areas and high-risk sectors) was the major focus of the workshop. Its two related Objectives addressed 1) the success in implementing a research program based on its 1996 National Occupational Research Agenda (NORA) priorities (NORA is a framework to guide occupational safety and health research into the next decade, and resulted in the establishment of a list of the top 21 research priorities) and 2) success in measuring its safety and health outcomes. The author was invited to participate as a member of the panel. This report generalizes a document that the author prepared for the NIOSH workshop, and was further refined during preparation for a DOE-sponsored workshop on S&amp;T benefits, 4-5 March 2002. The paper focuses on key metrics for evaluating progress in a mission-oriented research program. The results and conclusions of the analyses are sufficiently generic for applicability to any science and technology (S&amp;T) sponsoring organization.</b>			
15. SUBJECT TERMS			
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT
a. REPORT <b>unclassified</b>	b. ABSTRACT <b>unclassified</b>	c. THIS PAGE <b>unclassified</b>	
			18. NUMBER OF PAGES <b>14</b>
			19a. NAME OF RESPONSIBLE PERSON

## II. BACKGROUND

The implementation of the Government Performance and Results Act of 1993 (GPRA) signaled the codification of the use of quantitative metrics to monitor the progress of government-sponsored S&T. An open question since that time has revolved around the appropriate quantities to measure, and the appropriate metrics to use.

Typically, a major event in the life of an S&T project is its transition from one level of development (e.g., basic research) to another level of development (e.g., applied research, or technology development). Could such transitions be quantified, and used to populate performance metrics? Before this question can be addressed, different types of S&T transitions need to be identified and discussed. The following paragraphs describe transitions in the context of mission-oriented government S&T-sponsoring organizations,

Mission-oriented-government S&T sponsors have the generic mission of providing S&T information to 1) the engineering development and operational/ acquisition components of their parent organizations and/ or to 2) the engineering development components of the commercial sector, depending on their organizational structure and mission. These post-S&T developers and implementers will be referred to as the customer.

S&T information can be provided to the customer through two paths: 1) development sponsored directly by the government S&T organization, or 2) development sponsored by some other S&T organization(s). Resources expended by other S&T sponsoring organizations in a given technical interest area can be much larger cumulatively than resources available to any single S&T sponsor. Therefore, leveraging of these external resources by the customer/ S&T sponsor could have cost impacts far in excess of those resulting from directly sponsored S&T.

However, advanced technical understanding is required to identify the significance of technical advances made by other organizations. S&T sponsoring organizations tend to have the largest concentration of advanced technical personnel within the customer's management purview, and are in the best position to make the customer aware of significant technical developments globally.

Therefore, S&T sponsoring organizations have a dual role in providing S&T information to their customers: direct sponsorship of S&T targeted toward obtaining this information, and making their customers aware of significant technical advances worldwide. Given these two major missions and objectives for the S&T sponsoring organizations, management performance and metrics should focus on progress made for each of these two major roles.

### III. INTRODUCTION

There are four major classes of metrics available for consideration as transition metrics:

- 1) Activity - measures resource expenditures (e.g., people employed, operating budgets, etc), under management control (after resources received).
- 2) Output – tangible products under control of management (e.g., reports produced, components built)
- 3) Impact – measures effects on science and technology, and typically based on external judgements (e.g., transitions, citations, awards). Typically not under management control.
- 4) Outcome – long-term impacts on larger societal goals (e.g., health improvement, environmental remediation, etc)

Activity metrics are used mainly to normalize productivity and impact metrics. Most output metrics are used for superficial reporting purposes by S&T sponsors. Output metrics are rarely used in practice to impact major sponsor or performer management decisions, except in isolated cases like faculty tenure evaluation. They are sometimes used for research performer bonus considerations.

Outcome metrics are useful for long-term program auditing, for retrospective studies to identify critical parameters for fostering quality S&T, and for general documenting and archival purposes. Outcome metrics become operational too far into the future to impact management decisions and performance evaluation. Government military, civilian, and commercial civilian organizations have relatively rapid turnover of their highest level management. Especially in commercial organizations, portable pension plans have increased mobility, and continual de-regulation has enhanced the role of short-term market performance in driving management decisions. Motivation of government or commercial organizational management is to show progress within time frame of highest management cognizance. Management decisions are mainly governed by this time scale.

For S&T sponsors, major metrics used operationally for management decision-making and performance evaluation are transitions from one development level to another. These are metrics that incorporate:

- The number of transitions across development levels per unit of time
- The potential impact or benefit eventually resulting from these transitions
- The probability that each transition will eventually achieve the potential impact

The remainder of this paper will address the impact metric of transitions.

Transitions have two components, one under control of the S&T sponsor, and the other not under sponsor control. The first component is developing S&T to the

point where it has 'positive transitionability characteristics' (e.g., potential for affordability, increased performance, lighter weight, smaller, etc). The second component is the decision by the downstream developer/ user to advance development externally based on a number of exogenous parameters (e.g., geopolitical, legal, financial, etc). To some degree, whatever transition metrics are developed and implemented should reflect this division of responsibility between S&T sponsor and customer.

The transition metrics used presently for S&T sponsor performance and evaluation do not reflect this division of responsibility. Further, they do not reflect the dual role responsibility of S&T sponsors, namely, direct S&T sponsorship and increasing customer awareness of external S&T advancements. This limited scope of present day transition metrics reflects the limited scope of strategic objectives and organizational responsibilities of S&T sponsors. In addition, transitions used presently as S&T sponsor performance and evaluation metrics are not normalized to target productivity levels, and transition efficiency can not be evaluated.

This paper proposes transition metrics be re-defined to 1) reflect transition efficiency, similar to Carnot efficiency for thermodynamic systems; 2) reflect dual responsibilities of direct science and technology sponsorship and enhanced customer awareness; 3) reflect in part shared responsibility of sponsor and customer for effecting transitions successfully. This paper shows how use of these re-defined transition metrics will enhance productivity and the role of S&T sponsors in the full product development cycle. The Appendix provides supplementary information on high quality metrics.

#### IV. ANALYSIS

The approach taken here to re-define appropriate transition metrics is analogous to an approach used for citations [Kostoff, 1998a]. The fundamental principle is to measure the efficiency and effectiveness with which the S&T sponsor is accomplishing its broader mission. The basic objective function that contains these efficiency and effectiveness measures is the ratio of: 1) the impact (benefits) of all actual transitions enabled by the S&T sponsor to 2) the research transitions that would have maximized impacts (benefits) for the American public, given the level of global S&T funding in the topical areas being examined. The term 'enabled' is used in the ratio definition to include the dual role of the S&T sponsor discussed previously.

The objective function can be written in equation form as:

$$R = \frac{\sum_{i=1}^Z (T_i * I_i)}{\sum_{i=1}^Z (T_i * I_i)} \quad (1)$$

where:

R is the objective function,

SUM is the summation operator,

i is the dummy variable that ranges between the limits shown,

T<sub>i</sub> is the 'i'th transition from research to application,

l<sub>i</sub> is the magnitude of the impact (benefit) resulting from the 'i'th transition,

n is the actual number of transitions enabled from all sources, and

Z is the potential maximum number of high impact transitions resulting from a perfect investment strategy applied to the global funding that was expended on the topical area's S&T.

l<sub>i</sub> is the product of the potential benefit (resulting from the 'i'th transition) times the probability that the 'i'th transition will actually realize that benefit, and therefore l<sub>i</sub> should be viewed as the expected benefit.

The stage in time at which the objective function is evaluated determines the credibility of the data. If the evaluation time is far in advance of the transition time frames, then the quantities evaluated are estimates, with all the associated uncertainties. If the evaluation time is far after the transition time frames, then the quantities evaluated are much more credible, but are now outcome metrics, and lose their operational impact for the reasons discussed previously. Thus, the sum of utility and credibility for this metric is probably optimal somewhere in the time frame of the transitions being evaluated.

Obtaining credible data to evaluate the complete objective function is very difficult. In particular, Z is a hypothetical quantity based on a perfect investment strategy. It is included in the fundamental objective function statement to counteract the case where the S&T sponsor could conceivably be investing in very low-risk low-impact safe technologies, could have a high transition efficiency, and yet be ineffective relative to what could have been accomplished with a better investment strategy.

Equation 1 can be re-written to reflect more clearly those transitions resulting from the direct sponsorship of S&T and those transitions resulting from enhanced global data awareness.

$$\begin{aligned} & \dots i=N1 \dots \dots \dots j=N2 \dots \dots \dots i=Z1 \dots j=Z2 \\ R' = & (\text{SUM}(T_i * l_i) + \text{SUM}(T_j * l_j)) / (\text{SUM}(T_i * l_i) + \text{SUM}(T_j * l_j)) \quad (2) \\ & \dots i=1 \dots \dots \dots j=1 \dots \dots \dots i=1 \dots j=1 \end{aligned}$$

where N1 is the transitions resulting from directly sponsored S&T of the organization being evaluated, and N2 is the transitions from other globally-sponsored S&T enabled by the awareness of the technical experts in the organization being evaluated. Z1 and Z2 are the analogous numbers for ideal investment strategy and awareness.

The following section addresses different levels of approximation to the objective function, and includes comments on the strengths and weaknesses of each level.

1) Zeroth order approximation

$$\begin{aligned} & \dots\dots\dots i=N1 \\ R0 & =(\text{SUM}(Ti)) && (3) \\ & \dots\dots\dots i=1 \end{aligned}$$

This approximation applies to the S&T sponsor's projects only. Here, the number of transitions from the sponsor's S&T is the metric. This is the easiest metric for which data can be obtained, but is essentially useless for addressing the accountability components defined above. Unfortunately, this metric is used all too commonly in many organizations. It provides no indication of impact, and no indication of how efficiently the agency is performing its function. Further, it can be 'gamed', where the organization funds a large number of low-risk modest-payoff projects to inflate the transition numbers. The S&T sponsor could then be transitioning a high fraction of its potentially transitionable projects, but collectively these transitions will have low impact relative to what was possible with a better investment strategy.

2) First order approximation

$$\begin{aligned} & \dots\dots\dots i=N1 \\ R1 & =(\text{SUM}(Ti*li)) && (4) \\ & \dots\dots\dots i=1 \end{aligned}$$

Here, the product of number of transitions from the directly-sponsored S&T times expected impact per transition is the metric. It provides an indication of actual impact, but no indication of transition efficiency. Obtaining credible data for impacts and benefits is significantly more complicated than for the zeroth order metric, but much more insight is provided. Further, this metric overcomes the 'gaming' aspect of the previous metric to some degree, since level of payoff is included in the objective function.

### 3) Second order approximation

$$\begin{aligned} & \dots\dots\dots i=N1 \dots\dots\dots i=Z1 \\ R2 & =(\text{SUM}(T_i * I_i)) / (\text{SUM}(T_i * I_i)) \qquad (5) \\ & \dots\dots\dots i=1 \dots\dots\dots i=1 \end{aligned}$$

In this approximation, it is assumed that a panel of experts was convened, and identified the transitions that would have occurred from the directly sponsored S&T if an ideal investment strategy had been followed and executed. These ideal transitions are reflected in the denominator. The complexity of evaluating this metric increases considerably over the first order approximation, since judgements are now required as to how many of the sponsor's projects could have transitioned. However, this metric does offer indication of efficiency, as well as impact.

### 4) Third order approximation

$$\begin{aligned} & \dots\dots\dots i=N1 \dots\dots\dots j=N2 \\ R3 & =(\text{SUM}(T_i) + \text{SUM}(T_j)) \qquad (6) \\ & \dots\dots\dots i=1 \dots\dots\dots j=1 \end{aligned}$$

This approximation sums the number of transitions resulting from directly-sponsored S&T and the number of transitions from global S&T enabled by the global S&T awareness of the S&T sponsor. While it suffers from the types of deficiencies noted in the zeroth order approximation, it nevertheless represents a step forward through the inclusion of enabled transitions from global S&T. This metric, while still primitive, provides some indication of how well the S&T sponsor is performing its knowledge awareness function, in addition to its S&T sponsoring function. However, without impact or benefit level numbers incorporated into the objective function, this metric is subject to 'gaming'.

### 5) Fourth order approximation

$$\begin{aligned} & \dots\dots\dots i=N1 \dots\dots\dots j=N2 \\ R4 & =(\text{SUM}(T_i * I_i) + \text{SUM}(T_j * I_j)) \qquad (7) \\ & \dots\dots\dots i=1 \dots\dots\dots j=1 \end{aligned}$$

Here, the product of number of transitions from the directly-sponsored S&T and enabled S&T times impact per transition is the metric. It provides an indication of actual impact, but no indication of transition efficiency. Obtaining credible data for impacts and benefits is significantly more complicated than for the third order metric, but much more insight is provided. Further, this metric overcomes

the 'gaming' aspect described previously to some degree, since level of payoff is included in the objective function.

6) Fifth order approximation

$$\begin{aligned}
 & \dots\dots\dots i=N1 \dots\dots\dots j=N2 \dots\dots\dots i=Z1 \dots\dots\dots j=Z2 \\
 R5 = & (\text{SUM}(T_i * I_i) + \text{SUM}(T_j * I_j)) / (\text{SUM}(T_i * I_i) + \text{SUM}(T_j * I_j)) \quad (8) \\
 & \dots\dots\dots i=1 \dots\dots\dots j=1 \dots\dots\dots i=1 \dots\dots\dots j=1
 \end{aligned}$$

In this approximation, it is assumed that a panel of experts was convened. They identified the transitions that would have occurred from a) the directly sponsored S&T if an ideal investment strategy had been followed and executed, and b) the globally enabled S&T if the technical experts had been fully aware of the relevant global S&T sponsored and the relationship of the relevant global S&T to the needs of the parent organization. These ideal transitions are reflected in the denominator. The complexity of evaluating this metric increases considerably over the fourth order approximation. Judgements are now required as to how many of the sponsor's projects could have transitioned as well as the number of other global S&T projects that could have been exploited by the S&T sponsor's parent organization. However, this metric does offer indication of efficiency, as well as impact.

## V. SUMMARY AND CONCLUSIONS

Transition metrics have been defined to different levels of approximation. They are based on the rate of flow of expected benefit across a transition barrier. They range in complexity from the rate of flow of numbers of transitions to the normalized rate of flow of actual expected or realized benefits. They take into account transitions resulting from the sponsor's S&T development efforts as well as transitions enabled by the S&T sponsor's awareness of S&T performed globally.

## VI. SUGGESTIONS FOR FURTHER READING

Geisler, E., "The Metrics of Science and Technology", Quorum Books, Westport, CT, 2000.

Kostoff, R. N., "Peer Review: The Appropriate GPRA Metric for Research", Science, Volume 277, 1 August 1997.

Kostoff, R. N., "The Use and Misuse of Citation Analysis in Research Evaluation", Scientometrics, 43:1, September 1998a.

Kostoff, R. N., Science and Technology Metrics,  
[http://www.onr.navy.mil/sci\\_tech/special/technowatch/](http://www.onr.navy.mil/sci_tech/special/technowatch/) , 1998b.

Kostoff, R. N. "The Metrics of Science and Technology". *Scientometrics*. 50:2. 353-361. February 2001.

Kostoff, R. N., and Schaller, R. R. "Science and Technology Roadmaps". *IEEE Transactions on Engineering Management*. 48:2. 132-143. May 2001.

Kostoff, R. N. "Normalization for Citation Analysis". *Cortex*. 37. 604-606. September 2001.

Kostoff, R. N. "Citation Analysis for Research Performer Quality". *Scientometrics*. 53:1. 49-71. 2002.

## VII. APPENDIX

### VII-A. METRICS FOR STRATEGIC PLANNING AND S&T EVALUATION

In today's highly competitive global marketplace, strategic planning has become an integral part of the corporate management process. It is essential that the goals and targets within the plan are well understood by all parties associated with the firm, including investors, managers, and employees. A critical element of a well-designed strategic plan is the quantification of the strategic goals. Use of metrics operationally requires a higher precision in the definition of the goals, allows the goals to be expressed in a more commonly understood language, and allows the progress toward the goals to be monitored continuously. Temporal metrics monitoring can identify when progress diverges from the targets, and allows midcourse corrections to be made. (If metrics are established for every aspect of an organization's operation, and monitoring systems are established to automatically measure deviations from targeted goals, then the physical and biological world's self-regulating system of condition-based maintenance can be readily applied to organizational behavior.) However, metrics should not be used as stand-alone targets or measures of progress. Metrics should always be imbedded in a more comprehensive evaluation process in which a diversity of experts can assess and interpret the metrics' significance, and make recommendations based on these evaluations.

General principles for implementing metrics in strategic plans, and especially for conducting metrics-based science and technology evaluations, have been developed by the author. The remainder of this article summarizes these principles. A much more comprehensive treatment of S&T metrics, and examination of their many facets and dimensions, can be found on the author's web site [Kostoff, 1998b] and in Geisler's book on S&T metrics [Geisler, 2000].

## VII-B. PRINCIPLES OF HIGH QUALITY METRICS-BASED S&T EVALUATIONS

### VII-B-1) Senior Management Commitment

The most important factor in a high-quality metrics-based S&T evaluation is the serious commitment of the evaluating organization's senior management to high-quality metrics-based S&T evaluations, and the associated emplacement of rewards and incentives to encourage such evaluations.

### VII-B-2) Assessment Manager Motivation

The second most important factor is the assessment manager's motivation to perform a technically credible assessment. The manager:

VII-B-2-a. sets the boundary conditions and constraints on the assessment's scope;

VII-B-2-b. selects the final metrics used from a myriad of potential choices;

VII-B-2-c. selects the methodologies for how these metrics will be combined/integrated/interpreted, and

VII-B-2-d. selects the experts who will perform the interpretation.

In particular, if the evaluation manager does not follow, either consciously or subconsciously, the highest standards in selecting these experts, the evaluation's final conclusions could be substantially determined even before the evaluation process begins.

### VII-B-3. Statement of Objectives

The third most important factor is the transmission of a clear, unambiguous statement of the metrics-based evaluations objectives (and conduct) and potential impact/consequences to all participants at the initiation of the process. Participants are usually more motivated to contribute when they understand the importance of the evaluation to the achievement of the organizations goals, and understand in particular how they and the organization will be potentially impacted by the evaluations outcome.

Clear objectives and goals tend to derive from the seamless integration of evaluation processes in general into the organization's business operations. Evaluation processes should not be incorporated in the management tools as an afterthought, as is the case in practice today, but should be part of the organization's front-end design. This allows optimal matching between data generating/ gathering and evaluation requirements, not the present procedure of force fitting evaluation criteria and processes to whatever data is produced from non-evaluation requirements. When the evaluation processes are integrated with the organizations strategic management, the objectives drive the metrics which in turn determine what data should be gathered. Ad hoc evaluation processes tend to let the available data drive the metrics and the quantifiable goals.

#### VII-B-4. Competency of Technical Evaluators

The fourth most important factor is the role and competency of technical experts in a metrics-based S&T evaluation. Metrics should not be used as a stand-alone diagnostic instrument. Analogous to a medical exam, even quantitative metric results from suites of instruments require expert interpretation to be placed into proper context and gain credibility. The metrics results should contribute to, and be subordinate to, an effective peer review of the technical area being examined.

Thus, this fourth critical factor consists of the evaluation experts' competence and objectivity. Each expert should be technically competent in his subject area, and the competence of the total evaluation team should cover the multiple research and technology areas critically related to the science or technology area of present interest. In addition, the team's focus should not be limited to disciplines related only to the present technology area (which tends to reinforce the status quo and provide conclusions along very narrow lines), but should be broadened to disciplines and technologies which have the potential to impact the overall evaluation's highest-level objectives (that would be more likely to provide equitable consideration to revolutionary new paradigms).

#### VII-B-5. Criteria for Metric Selection

The fifth most important factor is criteria for metric selection. These criteria and the resultant metrics will depend on the interests of the audience for the evaluation, the nature of the benefits and impacts, the availability and quality of the underlying data, the accuracy and quality of results desired, the complementary metrics available and suites of metrics desired for the complete analysis, the status of algorithms and analysis techniques, and the capabilities of the evaluation team.

#### VII-B-6. Relevance of Metric to Future Action

The sixth most important factor is one that has been violated in every metrics briefing the author has attended, spanning many government agencies, industrial organizations, and academic institutions.

EVERY S&T METRIC, AND ASSOCIATED DATA, PRESENTED IN A STUDY OR BRIEFING SHOULD HAVE A DECISION FOCUS; IT SHOULD CONTRIBUTE TO THE ANSWER OF A QUESTION THAT IN TURN WOULD BE THE BASIS OF A RECOMMENDATION FOR FUTURE ACTION.

Metrics and associated data that do not perform this function become an end in themselves, offer no insight to the central focus of the study or briefing, and provide no contribution to decision-making. They dilute the theme of the study, and, over time, tend to devalue the worth of metrics in credible S&T evaluations. Because of the political popularity and subsequent proliferation of S&T metrics, the widespread availability of data, and the ease with which this data can be electronically gathered/aggregated/displayed, most S&T metrics briefings and studies are immersed in data geared to impress rather than inform.

#### VII-B-7. Reliability of Evaluation

The seventh most important factor is reliability or repeatability. To what degree would a metrics-based evaluation be replicated if a completely different team were involved in selection, analysis, and interpretation of the metrics data? If each evaluation team were to generate different metrics, and particularly far different interpretations of metrics, for the same topic, then what meaning or credibility or value can be assigned to any metrics-based evaluation? To minimize repeatability problems, a diverse segment of the competent technical community should be involved in the construction and execution of the evaluation.

#### VII-B-8. Metrics Integration

The eighth most important factor is the seamless integration of metrics in particular, and evaluation processes in general, into the organization's business operations. Evaluation processes should not be incorporated in the management tools as an afterthought, as is the case in practice today, but should be part of the organization's front-end design. This allows optimal matching between data generating/ gathering and evaluation requirements, not the present procedure of force fitting metrics and evaluation processes to whatever data is produced from non-evaluation requirements.

#### VII-B-9. Normalization Across Technical Disciplines

For evaluations that will involve comparison of science and technology programs or projects, the ninth most important factor is normalization and standardization across different science and technology areas. For science and technology areas that have some similarity, use of common experts (on the evaluation teams) with broad backgrounds that overlap the disciplines can provide some degree of standardization. For very disparate science and technology areas, some allowances need to be made for the relative strategic value of each discipline to the organization, and arbitrary corrections applied for benefit estimation differences and biases. Even in this case of disparate disciplines, some normalization is possible by having some common team members with broad backgrounds contributing to the evaluations for diverse programs and projects. However, normalization of the metrics for each science or technology area's unique characteristics is a fundamental requirement. Because credible normalization requires substantial time and judgement, it tends to be an operational area where quality is sacrificed for expediency.

#### VII-B-10. Global Data Awareness

The tenth most important factor is global data awareness. What S&T projects, developed systems or operations, or events, that exist globally are in any way supportive of, related to, or impacted by, the S&T programs undergoing a metrics-based evaluation? This factor is foundational to S&T investment strategy, and how a program or body of S&T is planned, selected, managed, coordinated, integrated, and transitioned. It is imperative that the latest information technology resources in tandem with human investigative efforts be used to the greatest extent possible during the complete evaluation process to insure that global S&T resources are being exploited maximally.

#### VII-B-11. Cost of Metrics-based Evaluations

The eleventh critical factor for quality metrics-based evaluations is cost. The true total costs of developing a high quality evaluation using credible suites of metrics, sophisticated normalization techniques, and diverse experts for analyses and interpretation can be considerable, but tend to be understated. For high quality evaluations, where sufficient expertise is represented on the evaluation team, the major contributor to total costs is the time of all the individuals involved in normalizing and interpreting the data. With high quality personnel involved in the evaluation process, time costs are high, and the total evaluation costs can be non-negligible. Especially when a metrics-based evaluation is performed in tandem to a qualitative peer-review process [Kostoff, 1997], the real costs of these experts could be substantial. Costs should not be neglected in designing a high quality metrics-based S&T evaluation process.

#### VII-B-12. Maintenance of High Ethical Standards

The final critical factor, and perhaps the foundational factor, in high quality metrics-based evaluations is the maintenance of high ethical standards throughout the process. There is a plethora of potential ethical issues, including technical fraud, technical misconduct, betraying confidential information, and unduly profiting from access to privileged information, because there is an inherent bias/ conflict of interest in the process when real experts are desired to design, analyze, and interpret a metrics-based evaluation. The evaluation managers need to be vigilant for undue signs of distortion aimed at personal gain.