Jointly published by Akadémiai Kiadó, Budapest
and Springer, Dordrecht

Scientometrics,
Vol. 62, No. 2 (2005) 269–284

The measurement of scientific activity: Research directions in linking philosophy of science and metrics of science and technology outputs

ELIEZER GEISLER

Stuart Graduate School of Business, Illinois Institute of Technology, Chicago (USA)

The application of the measurement of scientific and technical activities has been a lengthy process of the appropriate metrics and the assignment of the standards and benchmarks for their usage. Although some studies have addressed issues of the management of science and technology and their relation to scientometrics and infometrics, there is nevertheless a need to consider the linkages between the conceptual background of scientific generation and progress – and the measurement of its process and outcomes.

This paper first reviews the three main approaches to the generation and progress of human knowledge in general and scientific activity in particular. These approaches are reviewed in terms of the demands they would make on the measurement of scientific process and outputs. The paper then examines the currently used categories of metrics, and arrives at several conclusions.

The paper provides an analysis of these conclusions and their implications to the generation and utilization of metrics of science and its outcomes. The review of the conceptual or philosophical foundations for the measurement of science offers an in-depth examination, resulting in the correlation of these foundations with the metrics we now use to measure science and its outcomes. The paper suggests research directions for a much needed link between theories of science and knowledge, and the application of metrics used to measure them. Finally, the paper offers several hypotheses and proposes potential empirical studies.

Introduction

The application of the measurement of scientific and technical activities and outcomes has been a lengthy and challenging process of selection of the appropriate metrics and the assignment of the standards and benchmarks for their usage (ARROW, 1980; GEISLER, 2000; GEISLER, 2002; POLLOCK & CRUZ, 1999). Much of this research was based on the conceptual framework that emerged from given disciplinary areas. Economic Theory produced concepts and metrics that explored the items of concern to the discipline (ARROW, 1980; BOZEMAN & MELKERS, 1993; GRILICHES, 1998; MANSFIELD, 1972). Outputs from research and development (R&D) and science and technology (S&T) were linked to productivity of organizations, sales of new products,

Received August 22, 2004

Address for correspondence:
ELIEZER GEISLER
Stuart Graduate School of Business, Illinois Institute of Technology
565 W. Adams Street, Chicago, IL 60661, USA
E-mail: geisler@stuart.iit.edu

0138–9130/2005/US $ 20.00

profitability, and similar economic variables (FREEMAN, 1987; GRILICHES, 1998). In this paper, R&D and S&T are used interchangeably. Although they are distinct and different concepts and activities, researchers tend to use similar metrics in their measurement. Also, the arguments in this paper apply to both the R&D and S&T activities. Other disciplinary areas, such as organization theory, sociology, and information science, have also generated studies of the link between S&T activities and outputs and such variables as performance of individuals and groups in technology organizations.

Currently, the state of the art in the measurement of scientific and technical outcomes is such that there is a lack of a conceptual scheme or an underlying theory. In the absence of such a framework, the result is a disparate array of indicators and measures. Each set reflects the influence of the underlying discipline, so preference for indicators and measures is guided by the variables of concern in the discipline. Organization theorists measure those variables that are pertinent to the link between S&T and organization design, whereas sociologists link S&T outcomes to social phenomena. Information theorists address the relationship between S&T and knowledge and information creation.

This state of affairs has produced distinct lines of inquiry, in many instances conflicting and unable to converge (GEISLER, 2000). The sets of indicators that are lumped together under the rubric of "science indicators" are a mere aggregation of distinct groups of metrics. MCGINNIS (1978) provided a coherent and illustrative criticism of the federal project of science indicators, and the research based on these indicators. He argued that they are a poor model of science, partly because they omit "the entire cognitive and social structures of science" (p. 20).

This state of knowledge that ignores the cognitive structure of science is as prevalent today as it was 25 years ago. Researchers dip into a pool of existing metrics to measure the outcomes from S&T or R&D, with regard only to the variables they wish to study in the impacted organizations.

In parallel to the studies of the outcomes from S&T and R&D activities, there has been a considerable progress in the philosophy of science. This field of inquiry is concerned with the structure and progress of science and scientific activity. Scholars in this area explored such questions as: what constitutes scientific activity? What are the criteria for acceptance of research as scientific? What is the role of scientific programs in the evolution of science?

The basic questions of how science is structured and how it evolves and grows have been addressed by three approaches. One was Karl Popper's "Acritical rationalism" by which conjectures about scientific phenomena are constantly challenged empirically (POPPER, 1962; POPPER, 1995). The second includes the work by Lakatos, Carnap, and Kuhn. These scholars argued for a more programmatic, even social approach to science, and viewed it as a complex activity. The third approach, exemplified by the work of Feyerabend, argued for a less rigorous and more humanistic view of scientific progress (FEYERABEND, 2000).

Although in this body of research there exist appropriate versions of a conceptual framework for the nature of scientific outcomes, and in general the outcomes from R&D/S&T; there has not been – to date – a meaningful convergence of the two bodies of knowledge. The measurement of scientific activity and approaches to the structure and progress of science have remained on separate tracks (CAMPBELL & MUIR, 1990; GEISLER, 1995; KOSTOFF, 1997).

The purpose of the research proposed here is to fill this gap and to attempt a convergence of the two bodies of knowledge. The research is designed to examine the key concepts in the philosophy of science (on the structure and progress of science) within the three different approaches – and to relate them to the body of measures and metrics used to evaluate the outcomes from R&D and S&T.

This is an exploratory research project. It attempts to discover the possible connections between two bodies of knowledge that should have – in the opinion of the proposer – a measure of convergence. This exploratory research is aimed at discovering and identifying those conceptual frameworks in the philosophy of science under study that may serve as a foundation for the generation of metrics of R&D and S&T (GEISLER, 2000; GEISLER, 2002). As it now stands, metrics are applied in a somewhat opportunistic manner. They are tailored to the variables that they are supposed to measure, and these are engendered by the disciplinary boundaries of the phenomenon under investigation. For example, the outcomes of S&T are measured by economists in terms of the patents thus generated and the impact of these on economic activity of their organization (WERNER & SOUDER, 1997). Information scientists measure the outcomes of S&T in bibliometric terms, and the knowledge they produce (KNORR-CETINA, 1999; PRICE, 1976).

Recently, CHEN (2003) has suggested that the measurement of science and mapping of its frontiers has greatly evolved. He argued that "to apply science to science itself, we need to understand the nature of scientific activities, the philosophy and the sociology of science" (p. 3). Part of this argument is the aim of this proposed research: to provide a link between concepts in the philosophy of science (the structure of scientific activities and their progress) and the measurement of scientific outcomes.

Background research

Philosophy of science

The structure and progress of science have been philosophically explained by three distinct approaches. The first is "critical rationalism" or a logical-empirical approach

(POPPER, 1962). Scientific activities are viewed as having a logical structure. Conjectures about scientific phenomena are challenged empirically, with the object of refuting them. This is also a minimalist approach, in which propositions (and theories) are assessed individually. Science progresses through the constant flow of conjectures and their refutations (CAMPBELL & MUIR, 1990).

The second approach is characterized by the work of LAKATOS (1994), CARNAP (1995) and KUHN (1996). These scholars argued that a more pragmatic, social and historical approach needs to be considered. Programs of science, rather than individual theories or propositions, need to be addressed as the unit of research into the structure and progress of science. KUHN (1996) argued that paradigms of science are the normal structure of scientific activities. Changes in paradigm mark revolutionary changes that characterize scientific growth. Others in this school considered the interactions within scientific programs among scientists and their social framework as explanation for the progress of science (CARNAP, 1995a).

The third approach holds that science progresses through an unorganized model of human creativity and serendipity (FEYERABEND, 2000). Logical structures are considered a framework imposed by researchers (outsiders) rather than an integral part of the scientific activity.

Some philosophers of science in the aforementioned approaches attempted incursions into potential measurements of scientific outcomes. KUHN (1996) suggested that perhaps citation analysis would be helpful in explaining how paradigms change. Such attempts remained within the framework of bibliometrics, grounded in the direct outcomes from scientific activity, such as publications and their citations (GARFIELD et al., 1978). There are no systematic linkages, beyond bibliometrics, with other categories of metrics of S&T (GEISLER, 2000).

Similarly, scholars studying metrics of S&T tended to consider their downstream impacts, rather than their conceptual and theoretical foundations (ARROW, 1980; HAUSER & KATZ, 1998; MANSFIELD, 1972). Information scientists and scholars engaged in scientometric studies have examined co-citation maps as potential explanations of theoretical and conceptual transformations in science (CHEN, 2003). THAGARD (1992) tried to explain the emergence and perseverance of the continental drift theory by conceptual maps driven by co-citations. But, these attempts still lacked an underlying conceptual framework that supports the relationships proposed by these authors.

Metrics of S&T

The second and parallel literature is concerned with metrics of science and technology (GEISLER, 1995a; GEISLER, 2002). In this literature there are seven categories or sets of metrics used to assess the outcomes from S&T: (1) economic and financial; (2) commercial and business; (3) bibliometrics; (4) patents; (5) peer-review; (6) organizational, strategic, and managerial; and (7) stages of outcomes (GEISLER, 2000). Each set of metrics measures the outcomes from S&T from a different perspective and within a different time frame along the continuum of innovation. Bibliometrics measure the outputs that are proximal to the scientific activity, that is, the client outcomes that science generates. Conversely, commercial and business metrics measure the longer-term outcomes from S&T, as they impinge upon the business enterprise, downstream the continuum of innovation (GEISLER, 2001a).

These categories of metrics of S&T are normally applied opportunistically, depending on the perspective of the evaluator and the portion of the innovation continuum that is the object of evaluation (KOSTOFF, 1997). When a given category of metrics lends itself to measurement and quantification, and when such metrics offer an adequate assessment of S&T outputs, these metrics will be selected by the evaluators (KOSTOFF & GEISLER, 1999; PRICE, 1976).

The existing taxonomies of outcomes from S&T are at best crude. They relate to the various formats in which outcomes from S&T impact other entities – those who supported the original scientific and technological activity, and those who are the downstream recipients, users, and impactees of the S&T outcomes.

The choice of the set of metrics is therefore driven by those entities which are *external* to the S&T-performing unit, and whose impacts from S&T we wish to evaluate (GRILICHES, 1998; HAUSER & KATZ, 1998). There is not an underlying theoretical framework that explains the use of a given category of metrics. There is a lack of a framework that would link these categories and taxonomies of metrics to the theories of scientific progress and knowledge creation (HAWKING, 2002; KNORR-CETINA, 1999).

Conceptual framework

A priori versus a posteriori

The study described in this paper is designed to help replace the existing *a posteriori* reliance on criteria for selection of metrics of S&T outcomes, with *a priori* criteria based on a conceptual framework, as foundation for measurement. In the current state of the art, metrics are chosen based on the ultimate variables they are meant to measure. The choice among the pools of economic metrics, scientometrics (bibliographic indicators) or organizational/managerial metrics is *a posteriori* – reflecting those variables that the outcomes from S&T will be measuring and whose impacts will be the key to the selection (e.g., how patents affect economic growth, or how bibliometric measures, such as publications and citations, affect the prestige of the impacted organization). The results from this study would contribute to a different – *a priori* – approach. Metrics will be chosen based on a conceptual background anchored in concepts and constructs from philosophy of science – about how scientific activity is structured and how it progresses.

This study therefore is based on the possibilities that would emerge from the convergence of the two bodies of knowledge, philosophy of science, and the metrics of S&T. There are two key possibilities. The first is the introduction of theoretical foundations as the background for measuring the outcomes from S&T. Currently there is a lack of a theory, in an a priori manner, that can serve as a foundation for measurement efforts. The second possibility is the enhanced ability to construct macro-indicators with a theoretical background (GEISLER, 1995a; HAUSER & KATZ, 1998; KOSTOFF & GEISLER, 1999). When indicators (or measures) of S&T outcomes are selected based on *a posteriori* impact, there is a lack of a theoretical and conceptual foundation that would support the assembly of such indicators into a macro-indicator–across categories of measures. For example, one can construct a macro indicator of S&T outcomes as the aggregation of economic, bibliometric, and organizational measures (BOZEMAN & MELKERS, 1993; GEISLER, 1994).

This study is focused on eliciting concepts and constructs in the philosophy of science, while respecting the different approaches in this body of literature. The aim of this study is to explore the three main approaches to identify the relevant constructs and to link them to the current state of the art in the measurement of S&T outcomes.

The results from this study will have considerable impacts on the academic context of S&T measurement, and on its applications in government and industry. The results from this study would contribute to the academic literature by creating the model of linkages between the two aforementioned bodies of knowledge. These *linkages* or *bridges* between metrics of S&T and the possible philosophical foundations will open various avenues for further research. In addition, they will also provide a theoretical and conceptual background for S&T measurement activities conducted by government and industrial firms. The effort that produces the *Science Indicators* reports will be able to incorporate the philosophical foundations for its creation and the selection of indicators. Companies will do likewise in the design of their systems of S&T metrics.

Within each category of metrics, hence within each discipline's boundaries, reliance on a theoretical framework will strengthen the legitimacy of the choice of indicators and their validity in measuring the outcomes from S&T (GARFIELD et al., 1978; GEISLER, 1995). Each approach would yield selected constructs. An attempt will be made to suggest underlying constructs that cut across the different approaches (ELSTER, 1979; SJOBERG, 1968).

The conceptual framework

The conceptual framework of this study is shown in Figure 1. The concepts developed by philosophers of science will be considered in the three categories or models that emerge from this literature. While distinguishing between concepts that describe discovery, justification, and outcomes, the focus will be on the models of "critical rationalism," as in Popper's philosophy, the programmatic model (or approach) and the humanistic approach (FEYERABEND, 2000). The questions will be asked: Do these different concepts – each in its own conceptual womb – have implications for and can then serve as theoretical foundations for categories of metrics used to assess the progress of S&T? The key hypothesis of this research is that this question can be answered in the affirmative.

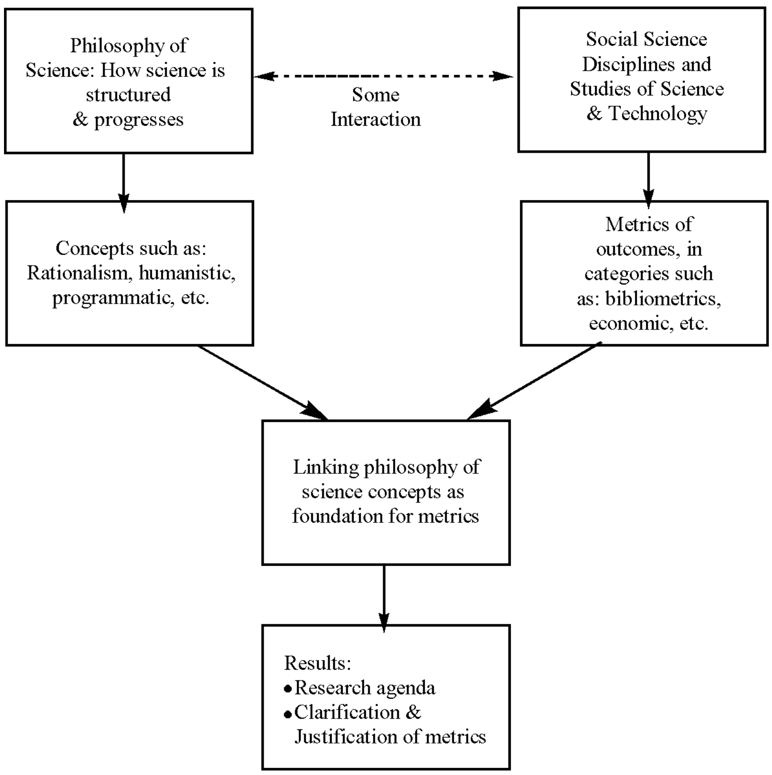

Figure 1. Conceptual framework of the convergence of two bodies of knowledge

Some authors, particularly those interested in methodological issues, have considered extensions of some concepts in the philosophy of science to elucidate their models of knowledge creation, its progress, and its impacts (BOUVIER, 2002; BUNGE, 2001; CAMPBELL, 1996; COOK & CAMPBELL, 1979). Much of this research focused on *process*, by exploring stages of progression and outputs (GEISLER, 2001; GEISLER, 2002), historical events and their linkages (HAWKING, 2002), and the rules by which these systems of knowledge creation tend to operate (BLACHOWICZ, 1998; BUNGE, 2000). Yet, little has been done to merge the two bodies of knowledge with regards to the link between these concepts and the outcomes from science and technology (HEISENBERG, 1958; LAKATOS, 1999).

The aim of the conceptual framework of this study is to attempt to establish the linkage between the different concepts and metrics of outcomes in the general and the more specific modes. In general, are there links between these concepts and the measurement of outcomes from science, and by extension, technology? In particular, are there links between these concepts and specific categories of metrics, so that these categories can thus be theoretically founded on such concepts? This study hypothesized that both instances are possible.

Convergence of philosophy of science and S&T metrics

In this paper the conceptual frameworks suggested by philosophers of science are linked to S&T metrics. Such convergence is done by examining individual approaches to philosophy of science. The method will thus consist of a set of three scenarios in which concepts extracted from the philosophy of science literature will be associated in conjunction with metrics of S&T outcomes. The links identified and their strengths will be assessed by the author and his interpretation of the convergence (SJOBERG & NETT, 1968; THAGARD, 1992; GEISLER, 2000).

"Critical rationalism"

This approach argued that scientific activities have an inherent logical structure, and that conjectures about scientific phenomena are challenged empirically, with the intent of refuting them. Science thus progresses via the flow of conjectures and their immediate refutations of these theories and propositions (POPPER, 1962; 1995).

The immediacy of refutation, before the theory has had an opportunity to evolve, or the scientists who formulated it have had the time and effort to work with it , has been the subject of criticisms on part of those who pertain to the second approach (LAKATOS, 1994; CARNAP, 1995). But, this notion of immediacy also supports the measurement of scientific activity with immediate or proximal metrics. This convergence is shown in Figure 2.

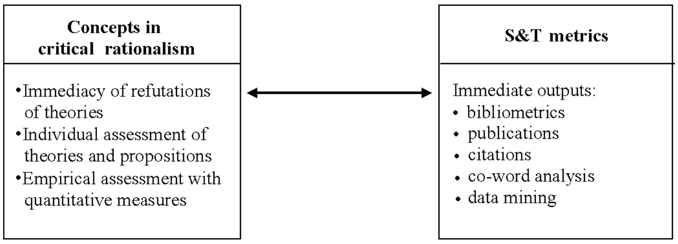

Figure 2. Convergence of critical rationalism and S&T metrics

The convergence of the concepts of "critical rationalism" with immediate or proximal measures of scientific activity reflects the ability to assess individual theories and propositions. Bibliometrics offer a good measure of a proximal adequacy of scientific effort. A theory or even a scientist who produces below a certain level of quantity and quality (via a measure of citations, for example) will not be an acceptable asset in the scientific evaluation yardstick. This is similar to the refutation of individual theories by putting them to the test of empirical acceptability (GEISLER, 1995; CAMPBELL, 1996; CHEN, 2003).

Programmatic and social progress of science

This approach is characterized by the work of LAKATOS (1994, 1999), CARNAP (1995, 1995a) and KUHN (1996). The progress of science is achieved through the intricate processes of social interaction of the scientists involved in scientific activities. They act as a social group, influencing each other. The unit of analysis for the evaluation of scientific outcomes is not the individual theory, but the *program* of research, usually involving several scientists, often as an interdisciplinary group. This approach allows the evaluation to assume a longer-term view of scientific outcomes, and to consider their application and implications downstream on social and economic entities.

The social aspect of this approach offers an explanation to the revolutions in science, in which paradigms are replaced and new directions in research are thus established. The emphasis in this approach is on the process by which scientists interact with each other within their research program, with others in the community of science, and with stakeholders in the larger sphere of science, for example: funding entities, regulators, and impactees (GRILICHES, 1998; MANSFIELD, 1972; THAGARD, 1992).

As it includes the human side and the social (or group) side of the scientific effort, this approach corresponds to a wide range of metrics, in various categories. The convergence is shown in Figure 3.

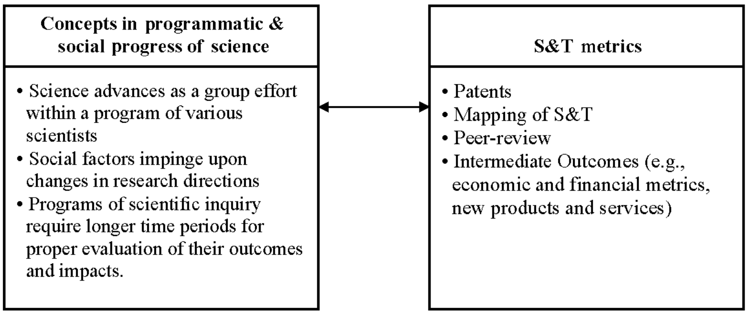

Figure 3. Convergence of programmatic and social approach to science, and S&T metrics

Viewed as a program, rather than the individual hypothesis or theory, scientific endeavor is now the province of a group of researchers working toward the advancement of science. Such a group may be physically connected within the same organization, or be composed of researchers in various locations, working on similar problems (ARROW, 1980; GEISLER, 1999).

The metrics that would attend to this perspective are peer review, mapping of S&T, and economic/financial metrics. Peer review is the metric that best encompasses the social interactions and social impacts of scientists upon one another. By means of the peer-review process researchers are able to promote or to block discoveries and to influence the 'paradigms' of scientific inquiry (KUHN, 1996; LAKATOS, 1994).

The longer time-frame necessary for a program of science to be effectively evaluated provides the basis for the assessment of S&T by economic and financial metrics. These are measures of intermediate outcomes, where S&T impinges upon economic organizations by means of new products and processes, patents, and their contribution through sales and profits (FREEMAN, 1987; GEISLER, 2001a; GRILICHES, 1998).

Humanistic perspective

The term "humanistic" is presented in a generic format, with the meaning of a less rigorous view of scientific activity. Science and scientists who engage in it are evaluated by the contributions they produced for humanity and society. Science is viewed as a vocation, a form of artistic and humanistic expression of innovativeness and creativity – rather than a rigorous and analytical structuring of the scientific endeavor (BOUVIER, 2002; FEYERABEND, 2000; GEISLER, 2001).

This perspective suggests that science is an activity that evolves without a structural framework, based on individual and group creative impacts. The outcomes from this effort are therefore distributed outputs that may form a cohesive field of study through the eyes of the external evaluator, and society benefits from this effort by its emergence, not necessarily via a pre-determined mode of exploitation of research results (BLACHOWICZ, 1998; BUNGE, 2001; GEISLER, 1995; HEISENBERG, 1958). Figure 4 depicts the link between concepts of this perspective and S&T metrics.

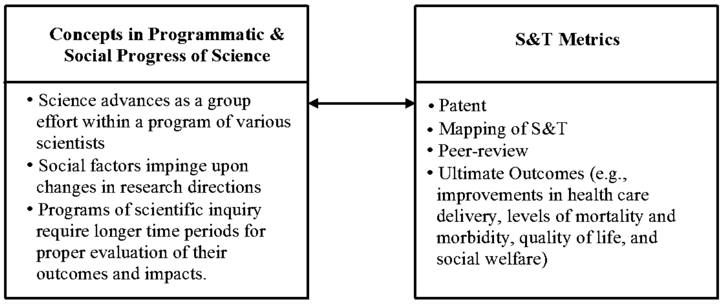

Figure 4. Convergence of the humanistic perspective and S&T metrics

The humanistic perspective offers a basis for linking S&T with metrics such as ultimate benefits for society, in the areas of health, education, and quality of life. Although not specifically advocated by philosophers of science who hold this perspective, the humanistic view of the scientific endeavor invariably leads to a strong link between the outcomes from science and social goals. As shown in Figure 4 above, scientific outcomes contribute to societal improvements.

Research agenda

The linkages between the three approaches to philosophy of science and S&T metrics that are listed above provide an initial consideration of such relationships. This preliminary effort engenders a possible research agenda, designed to explore, in more detail, the said linkages. The following research questions may be advanced:

(1) How can the convergence between approaches to philosophy of science and S&T metrics be studied empirically?
(2) Is it possible to form macro-indicators or macro-metrics of S&T outcomes that would correspond to specific approaches in the philosophy of science?
(3) What is the strength of the convergences and how strong is the reliance of S&T metrics on philosophy of science (as a theoretical foundation for these metrics)?
(4) Are the key differences among the approaches to philosophy of science also reflected in the different sets of metrics of S&T outcomes?
(5) How will the approaches to philosophy of science provide a theoretical basis for S&T metrics?

Hypotheses and potential empirical studies

For each of the three instances of convergence of a stream in philosophy of science and S&T metrics (Figures 2-4) this paper advances a hypothesis and a potential empirical study. These sets of hypotheses and empirical investigations lay the ground for future studies that would employ case studies or samples of performers of S&T activities for which such convergence is possible.

Critical rationalism

In this area of convergence, the following hypothesis may illustrate the possibility of employing empirical methods.

H_1: "The scientific performance of individuals in their organizations will be assessed by the immediate outputs from these individuals' effort, so that there is a positive relation between such outputs and organizational success."

This hypothesis contends that scientists and their organizations will be evaluated by the sole use of proximal or immediate metrics (such as bibliometrics). This implies assessment at the expense of downstream impacts and benefits that may emanate from their work (GEISLER, 2004). The immediacy of refutation of theories also applies to the evaluation of individual and organizational S&T outputs. As long as there is a credible and acceptable metric for evaluation (such as bibliometrics) the evaluation is considered complete, valid, and terminal. The implications are far-reaching, because promising effort and performers may be negatively evaluated in this fashion, and their work and future contributions discarded (FEYERABEND, 2000; FREEMAN, 1987; GEISLER, 2002).

The critical rationalism method for assessment and refutation of theories thus presents a justification for the evaluation of S&T with a minimalist and proximal set of metrics. This method disregards the temporal potential of S&T outputs to generate impacts and benefits beyond the short-term results. Thus, as in the case of theories assessed by critical rationalism, the issue becomes the validity and reliability of the methods and metrics used to evaluate S&T performance. Bibliometrics, for example, will be put under added scrutiny due to the finality of the function they perform in the evaluation procedure (GEISLER, 2001a; GRILICHES, 1998).

A possible empirical testing of H_1 would be to examine a sample of scientists and to assess their performance by the use of immediate outputs and downstream outputs (GEISLER, 2000; 2002a). The preponderance of outputs beyond such immediate outputs as bibliometrics may indicate that organizational performance and success depend on *both* immediate and downstream outputs (GEISLER, 2002; MANSFIELD, 1972; KOSTOFF, 1997).

Programmatic and social progress of science

H_2: "Social factors and the effects of programs and groups of scientists help determine the successful contributions of scientific research to sponsoring organizations, to the economy, and to society."

As shown in Figure 3, the metrics used to evaluate scientific outputs would be primarily downstream metrics (intermediate outputs), such as S&T mapping and economic metrics – combined with peer-review and count of patents. These metrics measure the transformation of results from scientific activity, beyond bibliometrics as measures of proximal outputs. Thus, peer review provides a social dimension of how scientists assess each other in a format that contends with biases and programmatic influences (DEWETT & DENISI, 2004; SWYGART-HOBAUGH, 2004; THELWALL, 2004; VAN RAAN, 2004).

Empirical testing of H_2 would consist of a study of teams of researchers within and between distinct programs of study (KOSTOFF, 1997; KOSTOFF & GEISLER, 1999). The outputs of these teams will be measured by intermediate outcomes (such as new products and services, and economic measures of success), and by downstream ultimate outputs. The latter will measure contributions to the economy and society. This study would assess the hypothesized relationship between social factors and the effects of group and research program dynamics – and the downstream contributions and impacts from this research (CHEN & HICKS, 2004; EGGHE & ROUSSEAU, 2004; GEISLER, 2004; LAKATOS, 1994).

Humanistic perspective

This perspective may be studied by the following hypothesis:

H_3: "Individuals and research teams contribute to the goals of society so that the more distributed research outputs among practicing scientists, the more their effort will contribute to societal goals and welfare."

This hypothesis may be tested empirically by examining the impacts of research in a selected area of scientific inquiry, and tracing them back to the scientists and program who produced these results. Specifically, this study would select a topical area in a scientific discipline and assess its impacts in the social arena (KOSTOFF & GEISLER, 1999; GEISLER, 1994). Such a topical area would be selected in a manner that it would allow for the identification of the contributions of individual scientists in different programs, even across strict disciplinary boundaries (BUNGE, 2000; CAMPBELL & MUIR, 1990; CARNAP, 1995a; CHEN, 2003).

This empirical study will potentially yield some findings as to the relationship between degree of distribution of scientific effort among scientists and research programs – and ultimate social benefits, such as contribution to improved quality of life, reduced mortality and morbidity, and economic welfare.

Conclusions

There is a long and arduous way from showing a convergence between philosophy of science and S&T metrics and having concepts in philosophy of science serve as theoretical foundations for S&T metrics. In this paper the linkages between them have been proposed. There now remains the task of establishing such convergence as a mode of justifying the use of S&T metrics (HAUSER & KATZ, 1998; HAWKING, 2002; POLLOCK, 1999).

The basic question is whether we can employ approaches to philosophy of science as a conceptual framework by which one can analyze the application of selected metrics of S&T outcomes. This will assist in answering such questions as why certain metrics *should* be used, and how does the assembly of metrics into macro-indicators relate to a specific perspective or approach in the philosophy of science (KNORR-CETINA, 1999). At the very least, the intellectual contribution in this effort will be the establishment of some sense of conceptual and analytical order into the selection and usage of S&T metrics.

The hypotheses and proposed empirical studies outlined in this paper offer an initial step in making the connection between streams in philosophy of science and S&T metrics. As the hypotheses show, the differences among the streams of philosophy of science may be empirically depicted in the stage of S&T metrics: from immediate to intermediate to ultimate outputs. Each stream of philosophy of science converges, and can be described, by a set of S&T metrics at a given stage of the temporal and conceptual continuum from proximal to the ultimate outputs of science.

Such a characteristic of the streams of philosophy of science may suggest that in addition to conceptual (philosophical) differences among the streams, there are also differences in the temporal and conceptual sets of output metrics that measure their impacts. Thus, the selection and usage of a given set of S&T metrics is related to and supported by a distinct perspective in the philosophy of science.

*

An earlier version was prepared for the 9th International Conference on Scientometrics & Infometrics, Beijing, August 25-29, 2003.

References

ARROW, K. (1980), A Proposed conceptual framework for indicators of R&D inputs, outputs, and industrial innovation. Paper presented at the *Proceedings of the OECD Science and Technology Indicators Conference,* Paris.

BLACHOWICZ, J. (1998), *Of Two Minds: The Nature of Inquiry*. Albany, NY: State University of New York Press.

BOUVIER, A. (2002), An epistemological plea for methodological individualism and rational choice theory in cognitive rhetoric. *Philosophy of the Social Sciences,* 32 (1) : 51–70.

BOZEMAN, B., J. MELKERS (1993), (Eds), *Evaluating R&D Impacts: Methods and Practice*. Boston: Kluwer Academic Publishers.

BUNGE, M. (2000), Systemism: The alternative to individualism and holism. *Journal of Socio-Economics,* 29 (2) : 147–157.

BUNGE, M. (2001), Systems and emergence, rationality, and imprecision, free-wheeling and evidence, science and ideology. *Philosophy of the Social Sciences,* 31 (3) : 404–423.

CAMPBELL, D. T. (1996), Unresolved issues in measurement validity: An autobiographical overview. *Psychological Assessment,* 8 (4) : 363–368.

CAMPBELL, D. T., D. MUIR (1990), Methodology and epistemology for social sciences. *Sociological Inquiry,* 60 (3) : 439–442.

CARNAP, R. (1995), *An Introduction to the Philosophy of Science*. New York: Dover Publications.

CARNAP, R. 1995a, *The Unity of Science*. Bristol, UK: Thoemmes Press.

CHEN, C. (2003), *Mapping Scientific Frontiers*. London: Springer-Verlag.

CHEN, C., D. HICKS (2004), Tracing knowledge diffusion, *Scientometrics*, 59 (2) : 199–211.

COOK, T., D. T. CAMPBELL (1979), *Quasi-Experimentation: Design and Analysis Issues for Field Settings*. Chicago: Rand McNally.

DEWETT, T., A. DENISI (2004), Exploring scholarly reputation: It's more than just productivity, *Scientometrics*, 60 (2) : 249–272.

EGGHE, L., R. ROUSSEAU (2004), How to measure own-group preference: A Novel approach to a sociometric problem, *Scientometrics*, 59 (2) : 233–252.

ELSTER, J. (1979), *Ulysses and the Sirens: Studies in Rationality and Irrationality*. Cambridge, UK: Cambridge University Press.

FEYERABEND, P. (2000), *Against Method*. New York: Verso.

FREEMAN, C. (1987), (Ed.), *Output Measurement in Science and Technology*. Amsterdam: North-Holland.

GARFIELD, E., M. MALIN, H. SMALL (1978), Citation data as science indicators, In: Y. ELKANA (Ed.), *Towards a Metric of Science*, New York: John Wiley & Sons.

GEISLER, E. (1994), Key output indicators in performance evaluation of research and development organizations. *Technological Forecasting and Social Change,* 47 (1) : 189–203.

GEISLER, E. (1995), Measuring the unquantifiable: Issues in the use of indicators in unstructured phenomena. *International Journal of Operations and Quantitative Management,* 1 (2) : 145–161.

GEISLER, E. (1995a), An integrated cost-performance model of research and development evaluation. *Omega*, 23 (3) : 281–294.

GEISLER, E. (1999), *Methodology, Theory, and Knowledge in the Managerial and Organizational Sciences : Actions and Consequences*. Westport, CT: Greenwood Publishing Group.

GEISLER, E. (2000), *The Metrics of Science and Technology*. Westport, CT: Greenwood Publishing Group.

GEISLER, E. (2001), Good-bye Dodo bird (Raphus Cucullatus): Why social knowledge is cumulative, expansive, and nonevolutionary. *Journal of Management Inquiry,* 10 (1).

GEISLER, E. (2001a), *Creating Value with Science and Technology,* Westport, CT: Greenwood Publishing Group.

GEISLER, E. (2002), The metrics of technology evaluation: Where we stand and where we should go from here. *International Journal of Technology Management,* 24 (4) : 341–374.

GEISLER, E. (2004), Commercialization of federal technology to industry: An empirical investigation of inter-sector cooperation, paper under review.

GRILICHES, Z. (1998), *R&D and Productivity: The Econometric Evidence*. Chicago: University of Chicago Press.

HAUSER, J., G. KATZ (1998), Metrics: You are what you measure! *European Management Journal,* 16 (5) : 517–528.

HAWKING, S. (2002), *On the Shoulders of Giants*. Philadelphia: Running Press.

HEISENBERG, W. K. (1958), *Physics and Philosophy*. New York: Harper Brothers.

KNORR-CETINA, K. (1999), *Epistemic Cultures: How the Sciences Make Knowledge*. Cambridge, MA: Harvard University Press.

KOSTOFF, R. (1997), *Handbook of Research Impact Assessment*. Vol. DTIC-A296021. Washington, DC: U.S. Government Printing Office.

KOSTOFF, R., E. GEISLER (1999), Strategic management and implementation of textual data mining in government organizations. *Technology Analysis & Strategic Management,* 11 (4) : 493–525.

KUHN, T. (1996), *The Structure of Scientific Revolutions*. Chicago: University of Chicago Press.

LAKATOS, I. (1994) *The Methodology of Scientific Research Programs*. New York: Cambridge University Press.

LAKATOS, I. (1999), *Proofs and Refutations: The Logic of Mathematical Discovery*. New York: Cambridge University Press.

MANSFIELD, E. (1972), R&D's contribution to the economic growth of the nation. *Research Management,* 15 (3) : 30-46.

MCGINNIS, R. (1978), Science indicators – 1976: A critique. *Society for Social Studies of Science,* 3 (4) : 14–29.

POLLOCK, J., J. CRUZ (1999), *Contemporary Theories of Knowledge*. Lanham, MD: Rowman & Littlefield Publishers.

POPPER, K. (1962), *Conjectures and Refutations*. New York: Basic Books.

POPPER, K. (1995), *Objective Knowledge: An Evolutionary Approach*. Oxford: Clarendon Press, Revised Edition.

PRICE, D. (1976), A general policy of bibliometric and other cumulative advantage processes. *Journal of the American Society for Information Science,* 27 (3) : 292–306.

SJOBERG, G., R. NETT (1968), *A Methodology for Social Research*. New York: Harper & Row.

SWYGART-HOBAUGH, A. (2004), A citation analysis of the quantitative-qualitative methods debate's reflection in sociology research: Implications for library collection development, *Library Collections, Acquisitions and Technical Services*, 28 (2) : 180–195.

THAGARD, P. (1992), *Conceptual Revolution*. Princeton, NJ: Princeton University Press.

THELWALL, M. (2004), Weak benchmarking indicators for formative and semi-evaluative assessment of research, *Research Evaluation*, 13 (1) : 63–68.

VAN RAAN, A. (2004), Sleeping Beauties in science, *Scientometrics*, 59 (3) : 467–472.

WERNER, B., W. SOUDER (1997), Measuring R&D performance: State of the art. *Research-Technology Management,* 40 (2) : 34–42.