Review of General Psychology
2006, Vol. 10, No. 2, 98–112

1089-2680/06/$12.00 DOI: 10.1037/1089-2680.10.2.98

Scientific Status of Disciplines, Individuals, and Ideas: Empirical Analyses of the Potential Impact of Theory

Dean Keith Simonton
University of California, Davis

The place of theory in scientific research can be subjected to empirical investigation. This possibility is illustrated by examining three issues. First, what determines a scientific discipline's placement in a hypothesized hierarchy of the sciences? This was addressed in an analysis of the characteristics that distinguish various disciplines, including attributes bearing an explicit connection to the role of theory. Second, what individual research programs are most likely to have a long-term impact on a scientific discipline? This was examined by looking at how thematic organization and theoretical orientation influence a scientist's disciplinary visibility. Third, what are the features of scientific publications that render some more successful in terms of long-term influence? This question was addressed by examining how theoretical content determines the impact of journal articles.

Keywords: theory, science, disciplines, research, impact

Science is a unique blend of rather distinctive intellectual activities. Although science emerged out of natural philosophy, it departs from philosophy in placing major emphasis on empirical data. Ultimately, the truth or falsity of a scientific proposition is decided not by the logical coherence of some syllogistic line of reasoning but rather by direct observation or experiment. Yet it is clear that true science is something more than mere fact gathering. Science is not like history, for example, where detailed particulars—the names, dates, and places—have intrinsic importance. Instead, in the sciences the observational and experimental data have significance largely with respect to abstract propositions. These abstractions, whether they be called hypotheses or predictions, transcend the temporal and spatial boundaries of any given scientist, instrument, or laboratory. That is, scientific statements claim a high degree of universality. History may not repeat itself, rendering every fact unique, whereas science cannot exist without repetition, or rather replication. If a finding cannot be replicated, it cannot exist as a valid scientific result. Hence, hypotheses and predictions are almost always framed in such a way as to render them independent of a particular person, time, or place.

Moreover, science appears to favor a specific kind of replicable and abstract datum: The observational or experimental result should also be of theoretical value or interest. The fact should acquire significance within the context of a general theory, model, or system. Otherwise empirical research would represent nothing more than "random fact gathering." In the absence of some theoretical framework, there is no way of separating critical findings from trivial results. Given the infinite complexity of the real world, completely theory-free investigations would produce massive and chaotic inventories of unconnected data, facts that would have hardly more meaning than the names and numbers in a telephone directory. It is for this reason that the great names of science are almost always those scientists who were supreme theoreticians. Copernicus, Newton, Darwin, and Einstein are prime examples. And even in those instances where a great theorist might have also made significant empirical contributions, those contributions are seldom viewed as having equal importance. Newton's *Principia* is rated above his *Optics*, just as Darwin's *Origin of Species* is valued more highly than his far more voluminous work on barnacles. To be sure, theories, models, and systems must fit the facts.

Correspondence concerning this article should be addressed to Dean Keith Simonton, Department of Psychology, One Shields Avenue, University of California, Davis, California 95616-8686. E-mail: dksimonton@ucdavis.edu

Again, that criterion is what separates science from natural philosophy. Yet even then it can be argued that a theory that has been disconfirmed often has higher status than a fact that is later proven to be wrong. Ptolemy's geocentric model of planetary motion has a better reputation than his woefully outdated geographic contributions. The former was eventually superceded, the latter ignored.

But all this is pure speculation. What is the relative importance of fact and theory in successful science? Philosophers of science have offered sundry sorts of theoretical conjectures on the subject, but is it possible to address this question in a scientific manner? I answer in the affirmative. Besides being feasible, some research has already shed some light on this issue. This research addresses three questions. First, what determines a scientific discipline's placement in a hypothesized hierarchy of the sciences? Second, what individual research programs are most likely to have a long-term impact on a scientific discipline? Third, what are the features of scientific publications that render some more successful in terms of long-term influence? This literature is reviewed below.

Disciplines: High-Prestige Sciences

Ever since the time of Comte (1839–1842/1855) scholars have speculated about a hypothesized hierarchy of the sciences. Presumably some disciplines are more scientific than others, and the diverse disciplines can be ranked according the degree to which they are "hard" or "soft." Unfortunately, these rankings often suffer from several problems (Simonton, 2002). Some are highly subjective and qualitative, such as Comte's classic effort. Others apply only a single criterion that probably has insufficient reliability to provide a reliable evaluation (e.g., Hedges, 1987; Rosenthal, 1990). Even those attempts that rely on multiple criteria may fail to perform the statistical analyses necessary to determine their convergence on a single consensus (e.g., Cole, 1983). However, one recent empirical investigation has removed these problems by applying appropriate analyses to multiple criteria (Simonton, 2004). Furthermore, the results of this inquiry, though not specifically directed at evaluating the role of theory, have definite implications for any such evaluation.

Measures

The investigation began with previously published measures of the characteristics of several scientific disciplines (e.g., Cole, 1983; McDowell, 1982; Roeckelein, 1997; Schachter, Christenfeld, Ravina, & Bilous, 1991; Smith, Best, Stubbs, Johnston, & Archibald, 2000; Suls & Fletcher, 1983). For purpose of analysis, these measures were divided into two categories: primary and secondary.

Primary measures. These assessed four core disciplines that represent a range of hard and soft fields. In particular, the measures all assessed the disciplines of physics, chemistry, psychology, and sociology. The measures could evaluate more than these four disciplines, but they had to assess at least these four. In any case, the following seven measures qualified as primary criteria:

1. *Theories-to-laws ratio*—Roeckelein (1997, Table 2, p. 137) assessed the number of theories and the number of laws mentioned in introductory textbooks in physics, chemistry, psychology, anthropology, and sociology. These counts were then used to compute the ratio of theories to laws, the higher the ratio the more "soft" is the discipline. That is, exact sciences have many more laws in proportion to theories (see also Roeckelein, 1996).

2. *Consultation rate*—The next criterion was a consultation measure based on Festinger's social comparison theory (Suls & Fletcher, 1983, Table 1, p. 578). According to this theory, when people are uncertain about their beliefs or performance, they are more likely to engage in social comparison with similar others. The specific measure was the number of colleagues recognized in the acknowledgment section adjusted for the number of authors. In other words, the measure is independent of the number of collaborators. As this number increases, the apparent uncertainty about the quality of one's work also increases. This score was available for physics, chemistry, psychology, and sociology.

3. *Obsolescence rate*—Based on the relative frequency of citations to older publications, McDowell (1982) determined the rate at which knowledge becomes obsolete for the disciplines of physics, chemistry, biology, sociology, psychology, history, and English. The specific measure used here was his calculation of the expected publication cost of interrupting a career for just 1 year (McDowell, 1982, Table 2, p. 757). For example, if the career is interrupted for a single year (e.g., administrative work, parental or health leave), the output of physicists will be cut by about 17% whereas the productivity of psychologists would be cut by about 10% (because physicists would have much more "catching up on the literature" to do before they can resuscitate their careers).

4. *Graph prominence*—Cleveland (1984) assessed the extent that graphs appear in articles published in the professional journals, demonstrating that graphs are more extensively used in the "hard" disciplines (see also Best, Smith, & Stubbs, 2001; Smith, Best, Stubbs, Archibald, & Roberson-Nay, 2002). The specific disciplines were physics, chemistry, biology, medicine, psychology, economics, and sociology. Although Cleveland (1984) did not aggregate the findings for the disciplines, this aggregation was carried out in Smith et al. (2000).

5. *Early impact rate*—In Table 2 Cole (1983) provided the "proportion of scientists under 35 whose work received more than the mean number of citations for their field" (p. 118). Those fields that incorporate most quickly the work of young scientists are assumed to rank higher in the hierarchy because such disciplines have a stronger consensus about what can be regarded a significant contribution to the field. The disciplines covered were mathematics, physics, chemistry, geology, psychology, and sociology.

6. *Peer evaluation consensus*—Cole's (1983) Table 3 provided data indicating the "consensus on evaluating scientists by field" (p. 120), where 60 scientists per field were rated by colleagues in the same discipline. The consensus was gauged by the mean standard deviation of the ratings, the lower the standard deviation the higher the consensus. The disciplines in this case were physics, chemistry, biochemistry, psychology, and sociology.

7. *Citation concentration*—The "concentration of citations to research articles" was presented in Table 5 of Cole (1983, p. 122). The citations were to journals in mathematics, physics, chemistry, biochemistry, geology, psychology, and sociology. If the citations are all concentrated in a single article, the disciplinary consensus must be very high, scientists concurring on what contributions deserve the status of "citation classics." In contrast, if the citations are more evenly distributed across articles, then the consensus must be minimal. In the case of completely even distribution, in fact, the citations received by articles would not differ from chance expectation.

Because the above seven variables were measured on rather different scales, the raw scores were standardized to z scores ($M = 0$, $SD = 1$). In addition, those variables that were reverse indicators—namely the theories-to-laws ratio, the consultation rate, and the peer evaluation consensus—were inverted by reversing the sign of the standardized scores.

Secondary measures. The measures in the second set all have one thing in common: However many disciplines to which they are applied, they have a missing value for at least one of the four disciplines evaluated by the primary measures. These secondary measures are useful for validating the results obtained from the primary measures. There were five indicators in this group:

1. *Lecture disfluency*—Schachter et al. (1991) determined the rate of filled pauses ("uh," "er," and "um") during classroom lectures for undergraduate courses in mathematics, chemistry, biology, psychology, economics, sociology, political science, philosophy, art history, and En-

glish. As the number of pause words per minute increases, the degree of speech disfluency also increases, which presumably reflects the degree to which a discipline is less formal, structured, and factual. This interpretation is justified by the fact that the same set of lecturers did not differ in disfluency when speaking on a common subject (viz., teaching). Hence, it is not a matter of the more inarticulate scientists being attracted to the less rigorous disciplines.

2. *Citation immediacy*—Cole (1983, Table 8, p. 126) calculated the extent to which the references in published articles were confined to recent work. In other words, the calculation gauges whether the citations emphasize contemporary research is emphasized over classic studies. Scores on this immediacy factor were available for physics, chemistry, biochemistry, geology, and psychology. These immediacy scores reflect a discipline's scientific rigor insofar as earlier research becomes more quickly assimilated into the core body of accepted findings that no longer require citation (Cole, 1983; Kuhn, 1970; McDowell, 1982). For instance, no modern physical scientist needs to give a citation whenever he or she uses the law of energy conservation in a derivation.

3. *Anticipation frequency*—Hagstrom (1974, Table 1, p. 3) reported the results of a survey of 1,718 scientists who were asked to report whether their work had been anticipated by other scientists. The percentage of scientists who had this experience at least once during their career course was gauged for mathematics, physics (combining theoretical and experimental), chemistry, and biology (combining experimental and other). The greater the frequency of anticipation, the higher the consensus became on what are deemed the important and unimportant problems in a discipline.

4. *Age at Nobel Prize*—Stephan and Leven (1993, Table 1, p. 395) provided the median age at which scientists received Nobel prizes in the fields of chemistry, physics, and medicine (from 1901–1992). Using the information provided at the official Nobel Prize site (http://www.nobel.se) the same statistic was obtained for the recipients of the economics prize (from 1969 to 2001). The logic behind including this indicator is the same as the early impact rate measure among the primary predictors. The more codified a discipline is, the sooner it can recognize when a scientist has made an exceptional contribution to the field.

5. *Rated disciplinary hardness*—Smith et al. (2000) had psychologists rate disciplines on the degree to which they could be considered "hard" versus "soft." The respondents used a 10-point Likert scale, with 10 indicating the highest degree of hardness. Seven disciplines were so rated, namely, physics, chemistry, biology, medicine, psychology, economics, and sociology. Smith et al. (2000) showed that this subjective assessment correlated .97 with Cleveland's (1984) measure of graph use. In addition, the investigators showed that this hardness assessment correlated .94 with an independent measure of paradigm development in various disciplines (Ashar & Shapiro, 1990). This measure, although subjective, was included to determine whether the objective assessments concur with more intuitive attitudes about the relative status of different scientific disciplines.

As before, the preceding measures were all standardized to a mean of 0 and a standard deviation of 1. Moreover, lecture disfluency was inverted by multiplying by −1. This was necessary because disfluency is hypothesized to be a negative indicator of a discipline's scientific status. It should be noted that, unlike the case of the primary indicators, there was sometimes little overlap in the disciplines covered by these five indicators. In fact, some measures had as little as two disciplines in common, a figure too small to compute meaningful correlations. Even so, these measures will prove useful in validating the primary indicators.

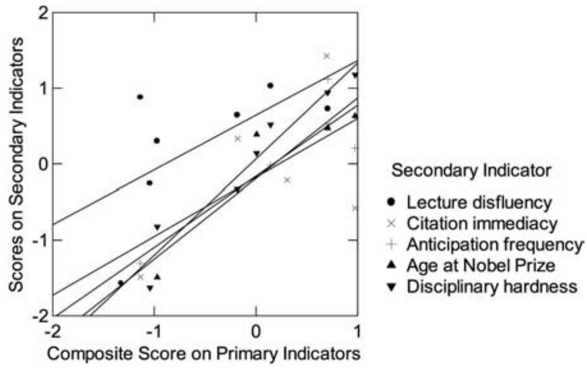

Figure 1. Scatterplot showing the relation between a discipline's score on the composite measure and indicators of a discipline's lecture disfluency, citation immediacy, anticipation frequency, age at Nobel Prize, and rated disciplinary hardness (from Simonton, 2004). The composite measure is defined by the discipline's theories-to-laws ratio, consultation rate, obsolescence rate, graph prominence, early impact rate, peer evaluation consensus, and citation concentration. Regression lines of least-squares best fit are also shown.

Analyses

The analyses began by generating a composite measure from the primary indicators using the four core disciplines of physics, chemistry, psychology, and sociology to calibrate an analytical baseline. The first step was to calculate the correlations among the seven measures for just these four sciences, the resulting correlations ranging between .63 and .998. These correlations were then subjected to a principal components analysis.[1] Only one component had an eigenvalue exceeding unity, and that lone component accounted for 86% of the total variance. Moreover, the loadings on the first component were uniformly high, ranging from .86 to .99. The specific loadings were as follows: theories-to-laws ratio .99, consultation rate .99, graph prominence .96, peer evaluation consensus .93, early impact rate .88, citation concentration .87, and obsolescence rate .86. As a consequence, the standardized scores across all seven measures were averaged to produce a linear composite. The internal-consistency reliability (Cronbach's alpha) for this composite was .96.

The next step was to extend this linear composite to all disciplines that contained at least one nonmissing value on the seven primary indicators. This was accomplished by simply averaging the standardized scores across all indicators with nonmissing values for a given discipline. This means that a discipline's score on the linear composite may represent anywhere between one and seven scores. Of course, the expected measurement error will be greater for those disciplines that have more missing values. The ratings based on a single component criterion would be the least reliable. In any case, the resulting composite measure was restandardized to a zero mean and unit standard deviation.

To validate the resulting composite indicator, the scores on the secondary measures were plotted as a function of scores on the composite primary measure. The outcome is shown in Figure 1, which also gives the lines of best least-squares fit. If no association existed, these lines would be horizontal, whereas lines inclined by about 45° would represent perfect positive relations. Clearly, a strong association exists between the composite of the primary indicators and each of the separate secondary indicators. In particular, higher scores on the composite are

[1] Because the number of variables exceeds the number of cases, it is impossible to extract more than three principal components (i.e., the number of nonzero positive eigenvalues cannot exceed $N - 1$). Nonetheless, the results still amply surpass what must be expected under a null (random) model in which the eigenvalues for the first three components would all be equal. If there were two underlying dimensions rather than one, the first two components would have roughly equal eigenvalues. Contrary to either of these scenarios, the first eigenvalue is over 8 times the size of the second and 21 times the size of the third.

associated with lower lecture disfluency, higher concentration of citations on more recent literature, more frequent experiences of anticipation, the greater youthfulness of Nobel Prize recipients, and higher rated disciplinary hardness. The corresponding Pearson product–moment correlation coefficients range between .60 and .97, with a median of .88. To obtain a better idea of the correspondence between primary and secondary indicators, the five secondary measures were also collapsed into a single composite in the same manner as the primary composite. That is, a new score is created for each discipline by averaging across the indicators for which there are nonmissing scores. This secondary composite correlates .87 ($n = 11$, $p = .0004$) with the primary composite, thereby confirming statistically what is so apparent graphically in Figure 1. The 12 primary and secondary indicators reflect a coherent latent variable on which disciplines can be reliably differentiated.

Finally, because five disciplines have nonmissing values on at least three of the primary indicators, it is possible to provide fairly reliable rankings for this subset of the 13 studied. Figure 2 shows the outcome. Physics, chemistry, biology, psychology, and sociology are arrayed according to standardized composite score and rank. The composite scores for physics, chemistry, psychology, and biology are based on all seven primary indicators, and thus have an internal-consistency reliability of .96. Because the standard error of measurement is equal to the square root of one minus the reliability coefficient, the error for these four disciplines is only 0.2, or just one fifth of a standard deviation. The composite score for biology, in contrast, was based on only three indicators, with a reliability of .89. Even so, the standard error of measurement is still reasonably

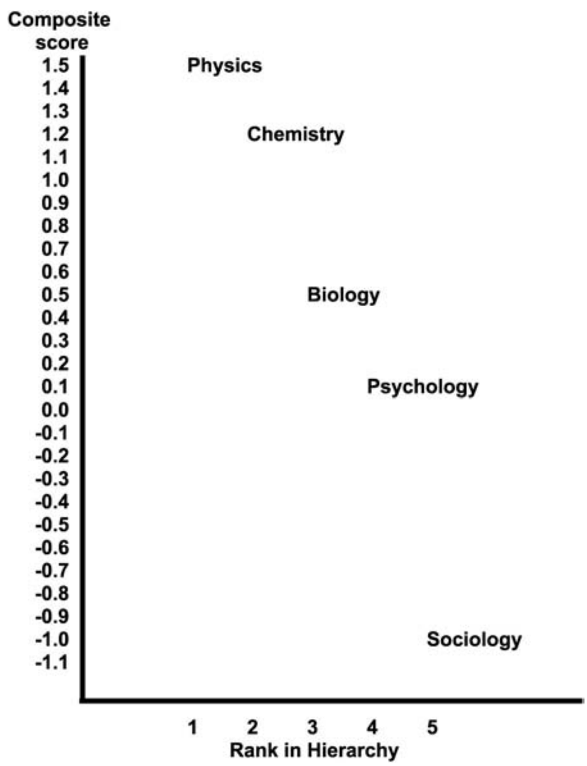

Figure 2. The disciplines of physics, chemistry, biology, psychology, and sociology placed in a hierarchy of the sciences (from Simonton, 2004). The horizontal axis indicates the rank and the vertical axis indicates the composite score on the seven primary indicators (i.e., the discipline's theories-to-laws ratio, consultation rate, obsolescence rate, graph prominence, early impact rate, peer evaluation consensus, and citation concentration).

small, namely 0.3, or about a third of a standard deviation. Consequently, the ordinal placement of these five disciplines is reasonably secure.

Interpretation

Judging from Figure 2, it is apparent that scientific disciplines can be ordered in close conformity to expectation. In particular, the supposedly "hard" natural sciences score higher than the presumably "soft" social sciences.[2] Moreover, if the primary and secondary measures are subjected to additional scrutiny, it is clear that the disciplines are at least partly differentiated according to the degree that a given science is governed by a paradigm. According to Kuhn (1970), a science is paradigmatic insofar as

> some accepted examples of actual scientific practice—examples which include law, theory, application, and instrumentation together—provide models from which spring particular coherent traditions of scientific research. These are the traditions which the historian describes under such rubrics as "Ptolemaic astronomy" (or "Copernican"), "Aristotelian dynamics" (or "Newtonian"), "corpuscular optics" (or "wave optics"), and so on. The study of paradigms, including many that are far more specialized than those named illustratively above, is what mainly prepares the student for membership in the particular scientific community with which he will later practice. . . . Men whose research is based on shared paradigms are committed to the same rules and standards for scientific practice. That commitment and the apparent consensus it produces are prerequisites for normal science, that is, for the genesis and continuation of a particular research tradition. (pp. 10–11)

Because of the potent consensus in paradigmatic disciplines, they should display lower consultation rates, earlier impact rates, higher agreement in peer evaluation, greater citation concentration and immediacy, higher anticipation frequency, faster obsolescence rates, an earlier age at winning a Nobel Prize, and even lower lecture disfluency. More importantly, it should be clear that strong theories have a crucial function in defining this consensus. Those theories define a discipline's central nomenclature and concepts, and they determine what facts are important and what not. Furthermore, theories provide the foundation for generating the predictions and hypotheses that guide empirical research in a highly specific direction. Therefore, the ordering of the five disciplines in Figure 2 can also be interpreted as a ranking based on the prominence of theory.

Admittedly, one finding might seem inconsistent with this conclusion: The lower theories-to-laws ratio of the more paradigmatic disciplines. Yet there is no inconsistency. Kuhn (1970) argued that preparadigmatic disciplines suffered from a profusion of theories, whereas truly paradigmatic disciplines had relatively few theories on which there was considerable consensus. Moreover, one reason why those theories were so strongly held is that they were intimately associated with many well-established laws. Hence, highly paradigmatic sciences would necessarily have a smaller ratio of theories to laws. By comparison, in the soft sciences, such as psychology and sociology, theories can proliferate all too easily without any grounding in widely accepted laws.

Individuals: High-Impact Research Programs

In the preceding section I presented empirical evidence that the most scientific of the sciences tend to be most theoretical. Or more precisely, the hard disciplines possess strong theories that provide a consensus on what are the key concepts and questions that underlie scientific research. Does the same connection hold at the individual level? Within a given discipline, no matter how paradigmatic, are the greater scientists those who are also more theoretically driven? The answer to this question comes from empirical research on the characteristics of high-impact research programs. These attributes concern both thematic organization and theoretical orientation.

Thematic Organization

According to Kuhn (1970), the researchers in preparadigmatic disciplines are engaged in random fact gathering. Because no firm theoretical position separates wheat from chaff, all facts

[2] Although the composite scores are based on fewer than three criteria, it is worthwhile to point out that the humanities are placed far below sociology. For example, scholarship in English is farther removed from sociology than sociology is from psychology. Thus, it is apparent that the criteria by which the sciences are here judged also distinguish the sciences from the nonsciences.

become equally important. Accordingly, findings gather helter-skelter, without rhyme or reason. In contrast, in highly paradigmatic disciplines scientists are engaged in "puzzle-solving" research that closely follows theoretical dictates, and thus the collective research effort is more strongly coordinated, and the results more cohesive and cumulative.

There is some empirical support for the conclusion that high-impact research programs operate in much the same manner. The support comes from a study of the factors that affect the long-term disciplinary impact of 69 highly eminent but deceased American psychologists (Simonton, 1992). Among the potential determinants was the thematic structure of each psychologist's research program. This was gauged via a computerized content analysis of the titles of each psychologist's collective publications (as given in Watson, 1974). This measure was the type-token ratio, an index of the ratio of distinct words to the total number of different words. This content analytical measure was found to correlate with alternative indicators of the psychologist's impact. In particular, the type-token ratio was negatively correlated with three alternative citation indicators of posthumous impact (for validity of these measures, see Rushton, 1984). The indicators were the total number of citations ($r = -.31$, $p < .05$), the number of publications cited at least once ($r = -.34$, $p < .01$), and the number of citations to the single most cited work ($r = -.22$, $p = .07$, two-tailed tests; all three indicators log-transformed). In addition, the type-token ratio was also negatively correlated with a measure of posthumous reputation ($r = -.30$, $p < .05$, after partialing out year of birth). The latter measure was a composite of three separate indicators of a psychologist's long-term influence (using Annin, Boring, & Watson, 1968; Zusne & Dailey, 1982; coefficient alpha for composite was .89).

How do these negative correlations endorse the conjectured pattern for high-impact research programs? A high type-token ratio means that the proportion of different words is very large. Since words in the titles of publications indicate topics, concepts, and methods, a high ratio signifies that the psychologist's career was devoted to investigating a great diversity of issues using a variety of techniques. In contrast, if the titles representing a psychologist's life work display a low type-token ratio, so that the number of unique words is proportionately smaller, then the individual's research was more concentrated on a well-defined set of topics, concepts, and methods. The obvious conclusion is that the high-impact psychologists—those whose work is most frequently cited many years later and who enjoy the most conspicuous posthumous reputations—tend to have had research programs that were much more coherent and focused.

Naturally, because the discipline of psychology falls on the lower end of the hierarchy presented in Figure 2, one might ask whether results for these 69 psychologists, no matter how eminent, apply in equal force to scientists in the more theory-driven disciplines, such as physics, chemistry, and biology. Feist (1997) provided an affirmative answer. To be specific, he showed that a similar set of relationships held for 99 contemporary physicists, chemists, and biologists affiliated with major U.S. research universities. After calculating the type-token ratio using the titles listed in their curriculum vitae, he found that this indicator was negatively related to the quantity of research, the total number of citations received, election to the National Academic of Sciences, and an indicator of global eminence (consisting of peer ratings of creativity and historical significance, professional visibility, and the prestige of the highest honor received). Consequently, it apparently does not matter whether scientists hail from a paradigmatic or preparadigmatic discipline. The more paradigmatic research programs expect a higher impact of the resulting work.

These computerized content analyses are corroborated by other studies showing that highly successful scientists tend to exhibit an exceptional degree of continuity in their research programs (Crane, 1965; Garvey & Tomita, 1972). Nonetheless, it must be emphasized that this continuity does not mean that the scientist is conducting the same investigation over and over again with only minor variations. On the contrary, influential researchers tend to engage in several projects simultaneously, each dealing with a different problem or issue (Hargens, 1978; Root-Bernstein, Bernstein, & Garnier, 1993; Simon, 1974; Taylor, Locke, Lee, & Gist, 1984). However, these diverse investigations are interconnected to constitute what Gruber (1989) called a "network of enterprises." Great scientists neither focus on a single narrow topic nor flip randomly around from topic to topic without rhyme or reason. Instead, the various subjects that constitute a highly successful research program are interconnected with each other, sometimes in subtle ways. Permeating almost everything is a central theme or theoretical preoccupation that creates a high degree of cohesion. In many ways the situation is analogous to what is seen in highly paradigmatic scientific disciplines. In the early 19th century physics was dominated by Newtonian theory, yet that dominance did not mean that all physicists conducted research on the same topics using identical methods. On the contrary, Newtonian physics involved a network of enterprise at the disciplinary level, each puzzle-solving scientist attempting to extend and elaborate the basis paradigm to various phenomena. Yet despite the tremendous range of investigations inspired by Newtonian mechanics, everyone maintained a common set of ideas that helped decide what facts were important, and what not, who was doing good work, and who mundane, and so forth.

Theoretical Orientation

The inferences of the previous section are vulnerable to criticism. The criticism is that the inferences are just that, inferential. The inference assumed that the thematic organization of research programs reflected the cohesive influence of theory. Although this assumption was founded on an analogy with the characteristics of paradigmatic sciences, the analogy is only an analogy, not equivalence. As a consequence, it is necessary to examine the impact of theory in a more direct fashion. Such a direct examination was carried out in another empirical study of 54 eminent psychologists (Simonton, 2000). This investigation was based on an earlier inquiry into the theoretical stances taken by 54 distinguished psychologists (Coan, 1968, 1979; see also Coan & Zagona, 1962). Using the judgments of 232 experts, these psychologists were ultimately assessed on the following six bipolar dimensions:

1. *Objectivistic versus Subjectivistic*—Emphasis on observable behavior versus emphasis on subjective experience (e.g., Watson, Pavlov, Skinner, and Hull vs. Jung, Brentano, Adler, Piaget, Fechner, and Janet).

2. *Quantitative versus Qualitative*—Emphasis on mathematics, statistics, and precision versus emphasis on qualitative attributes and processes (e.g., Estes, Thurstone, Spearman, Binet, and Ebbinghaus vs. Freud, Charcot, Wertheimer, Sullivan, and Köhler).

3. *Elementaristic versus Holistic*—Emphasis on molecular analysis versus emphasis on molar analysis (e.g., Spence, Titchener, Estes, Hull, Wundt, Pavlov, and Skinner vs. Goldstein, Koffka, G. Allport, Lewin, and Rogers).

4. *Impersonal versus Personal*—Emphasis on the nomothetic, deterministic, abstract, and tightly controlled versus emphasis on the idiographic, emotional, and the unconscious (e.g., Hull, Skinner, Titchener, and G. E. Müller vs. Rorschach, Adler, Jung, Janet, G. Allport, and Charcot).[3]

5. *Static versus Dynamic*—Emphasis on the normative and stable versus emphasis on motivation, emotion, and the self (e.g., Wundt, Mach, Fechner, Spearman, and Külpe vs. McDougall, Mowrer, Freud, and James).

6. *Exogenist versus Endogenist*—Emphasis on environmental determinants and social influences versus emphasis on biological determinants and heredity (e.g., Skinner, Angell, Hull, Rogers, and Watson vs. Galton, Freud, Hall, McDougall, and Cannon).

The correlations among these six factors are sufficiently high as to suggest the existence of higher-order factors (i.e., the *r*s range from .04 to .67, with a mean of .36 and a median of .32).

[3] Although this dimension was first styled the "transpersonal versus personal" (Coan, 1968) and later "apersonal versus personal" (Coan, 1979), the term "impersonal" appears far more appropriate than either transpersonal or apersonal, and so it was substituted here throughout this article. It is not only standard English, but highly descriptive of the factor besides.

To examine the possibilities, the factor intercorrelations were subjected to a principal components analysis (Simonton, 2000). Only the first two components were retained (using according to the eigenvalue >1 criterion). These were then subjected to an oblimin rotation, which thereby allows for an oblique factor solution ($\gamma = 0.0$). The loadings for the unrotated and rotated solutions are shown in Table 1.

There are two different ways these results can be interpreted. On the one hand, the first principal component accounts for nearly half of the total variance, with no loading below .49. This may be considered a general factor that pits elementaristic, objectivistic, quantitative, exogenist, impersonal, and static psychologists against their holistic, subjectivistic, qualitative, personal, endogenist, and dynamic colleagues. On the other hand, the oblimin rotation yields a two-factor solution with a fairly clear interpretation. The first factor includes the objectivistic, quantitative, and elementaristic dimensions, whereas the second includes the static, impersonal, and exogenist dimensions. The first group stresses the preferred analytical approach, whereas the second emphasizes underlying processes. However, in line with the unrotated solution, the two factors have a reasonably high correlation—almost as high as the lowest salient factor loading.

There are several ways of dealing with these results, but the most revealing in the context of the present article concentrates on the general factor in the unrotated solution (Simonton, 2000). Each psychologist's score on this factor was calculated by defining a composite measure that was simply the sum of the scores on the separate dimensions (and standardized into *z* scores). The internal-consistency reliability coefficient (alpha) for the resulting composite measure was .85, which indicates that the general factor shows a respectable amount of cohesiveness. The 54 eminent psychologists can be reliably placed along a dimension that contrasts objectivistic, quantitative, elementaristic, impersonal, static, and exogenist theorists with their subjectivistic, qualitative, holistic, personal, dynamic, and endogenist counterparts.

Table 1
Factor Loadings for Principal Components and Oblimin Rotated Factors

Dimension	Unrotated I	Unrotated II	Rotated I	Rotated II
Objectivistic	.79	−.39	.90	−.07
Quantitative	.78	−.33	.85	−.02
Elementaristic	.83	−.21	.80	.13
Impersonal	.64	.46	.14	.73
Static	.49	.70	−.16	.90
Exogenist	.53	.19	.26	.41
% Total variance	47.45	17.28	38.73	26.00

Note. The factor intercorrelation for the oblimin solution is .39. Taken from Simonton (2000).

So how do these scores relate to a psychologist's long-term impact? To respond to this query, each psychologist was again assessed on the number of citations they received in the *Social Sciences Citation Index* (log-transformed; see Simonton, 2000). Curiously, however, this index of total citations did not show a statistically significant correlation with the general factor scores. Before it can be concluded that the two variables are unrelated, it is first necessary to consider two complications. First, because the 54 psychologists' birth years span over a century (from Fechner to Estes), it is advisable to control for any cohort effects. Second, it is conceivable that the relationship between the two variables is curvilinear rather than linear, a possibility suggested by inspecting the scatterplot. Accordingly, a multiple regression analysis was conducted introducing (a) year of birth as a control and (b) both linear and quadratic functions of the general factor to assess nonlinearity.

The results were striking: Long-term impact, as assessed by total citations, was a U-shaped function of a psychologist's position on the general factor ($b = 0.24$, $\beta = 0.26$, $t = 2.05$, $p < .05$). In other words, those of the 54 with the most citations over the long term are those who are either extremely objectivistic, quantitative, elementaristic, impersonal, static, and exogenist or extremely subjectivistic, qualitative, holistic, personal, dynamic, and endogenist. Although superimposed over this curvilinear component is a negative linear component, the latter does not quite satisfy conventional levels of statistical significance ($b = -0.24$, $\beta = -0.22$, $t = 1.77$, $p = .083$). The upshot is a U curve with a slight leaning toward a backward-J function. Figure 3 displays the scatterplot with the curve superimposed (Simonton, 2000).

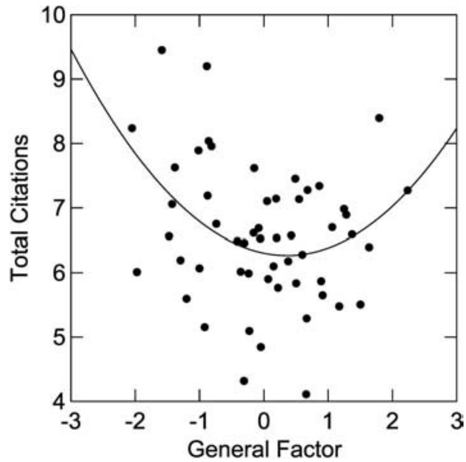

Figure 3. Scatterplot of the relation between the general factor and total citations for 54 eminent psychologists (from Simonton, 2000). Also shown is the best-fitting quadratic function defining the curvilinear backward-J curve describing the association between the two variables.

An empirical finding such as that shown in Figure 3 can have a multitude of interpretations (Simonton, 2000). For instance, it could be argued that long-term impact depends on a psychologist adopting extremist positions on the issues that divide our discipline, a phenomenon that has parallels in philosophy (Simonton, 1976). Yet another possible account ties in better with the questions addressed in this article. Perhaps psychology is not a single coherent discipline but rather represents two separate sciences somewhat artificially amalgamated. Reflecting in part Snow's (1960) "two cultures," psychology is both a natural science with close connections with biology (as witnessed in Figure 2) and a human science with close connections with the humanities (see also Kimble, 1984). Natural-science research in psychology is guided by theories that are elementaristic, objectivistic, quantitative, exogenist, impersonal, and static, whereas human-science research in the field is informed by theories that are holistic, subjectivistic, qualitative, personal, endogenist, and dynamic. Theories in each of these psychologies are internally coherent and consistent, and thus provide a solid basis for structuring individual research programs. To provide concrete illustrations, the high-impact psychologists with a human-science orientation are Freud, Jung, Adler, James, Allport, and Rogers, whereas the high-impact psychologists with a natural-science orientation include Skinner, Harlow, Thurstone, and Estes. In contrast, psychologists who conduct their research in a manner that tries to cut across these two alternative frameworks, who try to mix and match the components of the theoretical orientations, will be less likely to have long-term impact. These psychologists find themselves situated at the bottom of the J-curve. Members of this group are J. R. Angell, G. E. Müller, and J. M. Cattell, the also-rans of eminent psychologists.

Products: High-Citation Journal Articles

I began this discussion with scientific disciplines as the unit of analysis and then turned to individual scientists as the analytical unit. It is possible to shrink the unit still more. After all, ultimately a scientist's impact is founded on the influence of his or her publications, especially the articles they publish in professional journals. Moreover, these articles vary immensely regarding impact. At one extreme are those rare papers that are cited so many times by scientific colleagues that they become "citation classics" that receive hundreds, even thousands of citations (Garfield, 1987; Price, 1965). At the other extreme are articles that receive no citations at all in the professional literature, a fate shared by all too many scientific publications. To give a notion of the dispersion, of 783,339 papers published in scientific journals in 1981, 81% were cited 10 times or less, and 47% were not cited at all between 1981 and June 1997 (Redner, 1998). So what are the reasons for this incredible range in citation impact?

Potential Answers to This Question Come From 3 Sources

1. One possible clue to an answer comes from the criteria that are applied in evaluating manuscripts that have been submitted for publication in peer-reviewed journals. For instance, a survey asked 66 editors of major journals to rank order 15 potential criteria for evaluating whether submitted manuscripts should be accepted or rejected (Wolff, 1970; see also Wolff, 1973). The criterion of "theoretical model" ranked sixth, the criteria of

"contribution to knowledge," "research design," "objectivity in reporting results," "statistical analyses," and "writing style and readability" having greater importance. Although theory emerged in the top half of the list—exceeding such criteria as "manuscript length," "author's status and reputation," "punctuation," and "institutional affiliation"—it appears less crucial than purely empirical criteria.

2. A more direct approach is to ask scientists to specify what they consider to constitute the attributes of high-impact publications. For example, in an investigation psychologists were asked to identify and rate the characteristics of highly influential articles (Sternberg & Gordeeva, 1996). A factor analysis then revealed the existence of six dimensions: "Quality of presentation," "theoretical significance," "practical significance," "substantive interest," and "value for future research." Although a criterion of theoretical content again emerged, it remained only one of a half-dozen standards for evaluating publications.

3. The most direct approach is to have scientists rate actual journal articles according to various criteria of quality. For instance, one study had researchers evaluate publications that they themselves had cited in their own work (Shadish, 1989). Besides an overall assessment of quality, the articles were assessed on over two dozen attributes. Although a large number of criteria were found to correlate with the quality evaluation, very few of these had anything to do with the theoretical merits of the article. The item that came closest was an assessment "Used a method or a theoretical perspective that you think is currently unusual or especially innovative," which correlated .39 with overall quality ($p < .05$). Yet this statement clearly confounds methodological and theoretical attributes. Even worse, this criterion, however confounded, did not correlate with an index of citation impact. On the contrary, the correlation was slightly even if nonsignificantly negative ($r = -.10$, $p > .05$). Accordingly, it would seem that theory plays even less of a role in an article's long-term impact than might be inferred from the previous two investigations.

Are we to conclude that theory has a very small part to play in the long-term impact of single publications? I think no. Such a conclusion would be premature for two reasons. First, all of the studies just discussed were based on psychological research. It is possible that theory might have a stronger influence in those disciplines that enjoy stronger theories, such as the highly paradigmatic sciences of physics and chemistry. Second, it is conceivable that even in the latter disciplines theory has minimal relevance for the paradoxical reason that theory has such a crucial place in the conduct of research. According to Kuhn (1970), the scientists operating within highly paradigmatic domains will be engaged in puzzle-solving within the context of the prevailing theory. If true, then there will be very little variation in the theoretical rationale of separate journal articles. Indeed, it could be the case that the theoretical aspects of research exhibit more variation in the less paradigmatic disciplines precisely because it is possible to carry out research that has no foundation whatsoever in strong theory.

The foregoing discussion focused on journal articles as the primary means of scientific communication. Yet such articles are not the only vehicle for the publication of scientific ideas, nor are they necessarily the most influential. For instance, a study of 69 eminent American psychologists found evidence that books are more high-impact than articles (Simonton, 1992). For each psychologist the proportion of his or her book to total output was calculated. This measure was then correlated with the number of cited publications, the total number of citations, and the number of citations to his or her single most cited work (using a five-year accumulation for 1971–1975). In all three cases the correlations were statistically and substantively significant (i.e., about 10% of the variance was explained; r's range from .29 to .33). The superior impact of the more ambitious publications was also shown by looking at the psychologist's most frequently cited work. Although books only accounted for 17% of all the publications credited to these 69 psychologists, books represented 45% of those works that received the most citations.

Corroborative results are found in another study using a different methodology (Heyduk & Fenigstein, 1984). The investigators sent letters to eminent psychologists asking them to identify those “texts or articles. . .which have significantly influenced your work and thought, both past and present, in your major area of psychology” (p. 556). As many as 10 works could be so identified by each survey respondent. Yet extremely few articles were mentioned. And in every case but one, when a scientific paper was deemed influential, a book or monograph by the same author proved even more so. Only one author out of the 39 most influential psychologists had more impact through an article rather than a book. That means that fewer than 3% of these eminent contributors staked their fame on an article rather than a book. Furthermore, 92% of the works that influenced eminent psychologists were books or monographs, leaving only 8% to be credited to articles.

How is this disparity relevant to the issue at hand? In those sciences that are highly codified in mathematical symbols, such as physics and chemistry, it is possible to make major theoretical advances within the confines of the journal article. In contrast, theoretical contributions in sciences that have yet to receive such mathematical codification will be less concise and concentrated. As a consequence, a book-length treatment may provide the only practical means for theory development. Hence, the higher impact of books over articles may reflect the greater theoretical content of books. Naturally, this inference is mere conjecture. And certainly many books are no more than compendiums of empirical findings. Even so, additional research may discover that a primary function of books in disciplines like psychology and sociology is to promote theory rather than report fact.

Conclusion

I have just scrutinized the relation between theory and impact using three distinct levels of analysis. At the level of disciplines I showed that the status of a science can be ranked to a large extent according to the prominence of theory. At the level of individuals, it was seen that a scientist’s long-term impact could be associated with the extent to which his or her research is theory driven. At the level of products the connection between theory and impact was more tenuous. Nevertheless, theoretical content is used as one criterion for evaluating journal articles, and the higher impact of books over articles might reflect differences in theoretical content, at least in disciplines that have not yet undergone mathematical codification.

It is obvious that much more research needs to be done. Furthermore, additional inquiries must be theoretical, not just empirical. It seems ironic that the research on this question is not itself theory-driven. Instead, the investigations were instances of unadulterated empiricism. We currently lack an explicit theory that can provide strong guidance to empirical inquiries. The closest any theory comes to providing this desideratum is Kuhn’s (1970) ideas about paradigms, ideas that have provided some inspiration for work regarding the relative status of disciplines (Cole, 1983). Yet Kuhn’s conceptions of science are not universally accepted among historians and philosophers of science (Gholson & Barker, 1985; Lakatos, 1978; Laudan, 1977), nor is there any explicit psychological foundation for his theories. What is needed is a cognitive psychology of science that can specify the place that theory has in scientific problem solving. There has been abundant research on the role of lower-level analogies and hypotheses in scientific thinking (e.g., Gentner & Jeziorski, 1989; Mynatt, Doherty, & Tweney, 1978), but much less on higher-order cognitive structures and systems (cf., Thagard, 1992). In other words, a complete psychology of science will probably have to develop a comprehensive theory of scientific theories.

Future research must attend to a closely related issue: What constitutes a good scientific theory? Although philosophers of science have considered this question at great length, their approach tends to be analytical rather than empirical. There is no reason to believe that their conclusions and prescriptions will be any more valid than those of philosophers in other specialty areas whose ideas have withered in the face of hard experimental fact (e.g., the fate of philosophical esthetics when confronted by experimental esthetics; see Berlyne, 1971, 1974). Meehl (1992) has outlined some of the methods by which to conduct an “empirical, history-based philosophy of science.” In simple terms, the various characteristics of theories can be objectively quantified and then introduced as potential predictors of any theory's likely long-term impact. Once the predictors are identified, they can provide guidance regarding the best procedures in theory construction. It is possible that a sold appreciation of what constitutes a good theory will also have implications for understanding the psychological processes underlying scientific advance. These processes may entail a complex mixture of cognitive, developmental, differential, and even social psychological mechanisms. Hence, by understanding the role of theory in science, we may eventually acquire a better appreciation of the function of psychology in the scientific enterprise.

References

Annin, E. L., Boring, E. G., & Watson, R. I. (1968). Important psychologists, 1600–1967. *Journal of the History of the Behavioral Sciences, 4,* 303–315.

Ashar, H., & Shapiro, J. Z. (1990). Are retrenchment decisions rational? The role of information in times of budgetary stress. *Journal of Higher Education, 61,* 123–141.

Berlyne, D. E. (1971). *Aesthetics and psychobiology*. New York: Appleton-Century-Crofts.

Berlyne, D. E. (Ed.). (1974). *Studies in the new experimental aesthetics.* Washington, DC: Hemisphere.

Best, L. A., Smith, L. D., & Stubbs, D. A. (2001). Graph use in psychology and other sciences. *Behavioral Processes, 54,* 155–165.

Cleveland, W. S. (1984). Graphs in scientific publications. *American Statistician, 38,* 261–269.

Coan, R. W. (1968). Dimensions of psychological theory. *American Psychologist, 23,* 715–722.

Coan, R. W. (1979). *Psychologists: Personal and theoretical pathways*. New York: Irvington Publishers.

Coan, R. W., & Zagona, S. V. (1962). Contemporary ratings of psychological theorists. *Psychological Record, 12,* 315–322.

Cole, S. (1983). The hierarchy of the sciences? *American Journal of Sociology, 89,* 111–139.

Comte, A. (1855). *The positive philosophy of Auguste Comte* (H. Martineau, Trans.). New York: Blanchard. (Original work published 1839–1842)

Crane, D. (1965). Scientists at major and minor universities: A study of productivity and recognition. *American Sociological Review, 30,* 699–714.

Feist, G. J. (1997). Quantity, quality, and depth of research as influences on scientific eminence: Is quantity most important? *Creativity Research Journal, 10,* 325–335.

Garfield, E. (1987). Mapping the world of science: Is citation analysis a legitimate evaluation tool? In D. N. Jackson & J. P. Rushton (Eds.), *Scientific excellence: Origins and assessment* (pp. 98–128). Beverly Hills, CA: Sage.

Garvey, W. D., & Tomita, K. (1972). Continuity of productivity by scientists in the years 1968–1971. *Science Studies, 2,* 379–383.

Gentner, D., & Jeziorski, M. (1989). Historical shifts in the use of analogy in science. In B. Gholson, W. R. Shadish, Jr., R. A. Neimeyer, & A. C. Houts (Eds.), *The psychology of science: Contributions to metascience* (pp. 296–325). Cambridge: Cambridge University Press.

Gholson, B., & Barker, P. (1985). Kuhn, Lakatos, and Laudan: Applications to the history of physics and psychology. *American Psychologist, 40,* 755–769.

Gruber, H. E. (1989). Networks of enterprise in creative scientific work. In B. Gholson, W. R. Shadish, Jr., R. A. Neimeyer, & A. C. Houts (Eds.), *The psychology of science: Contributions to metascience* (pp. 246–265). Cambridge: Cambridge University Press.

Hagstrom, W. O. (1974). Competition in science. *American Sociological Review, 39,* 1–18.

Hargens, L. L. (1978). Relations between work habits, research technologies, and eminence in science. *Sociology of Work and Occupations, 5,* 97–112.

Hedges, L. V. (1987). How hard is hard science, how soft is soft science? *American Psychologist, 42,* 443–455.

Heyduk, R. G., & Fenigstein, A. (1984). Influential works and authors in psychology: A survey of eminent psychologists. *American Psychologist, 39,* 556–559.

Kimble, G. A. (1984). Psychology's two cultures. *American Psychologist, 39,* 833–839.

Kuhn, T. S. (1970). *The structure of scientific revolutions* (2nd ed.). Chicago: University of Chicago Press.

Lakatos, I. (1978). *The methodology of scientific research programs.* Cambridge, England: Cambridge University Press.

Laudan, L. (1977). *Progress and its problems.* Berkeley: University of California Press.

McDowell, J. M. (1982). Obsolescence of knowledge and career publication profiles: Some evidence of differences among fields in costs of interrupted careers. *American Economic Review, 72,* 752–768.

Meehl, P. E. (1992). Cliometric metatheory: The actuarial approach to empirical, history-based philosophy of science. *Psychological Reports, 71,* 339–467.

Mynatt, C. R., Doherty, M. E., & Tweney, R. D. (1978). Consequences of confirmation and disconfirmation in a simulated research environment.

Quarterly Journal of Experimental Psychology, 30, 395–406.

Nobel Prize (n.d.). Accessed from http://www.nobel.se.

Price, D. (1965). Networks of scientific papers. *Science, 149,* 510–515.

Redner, S. (1998). How popular is your paper? An empirical study of the citation distribution. *European Physical Journal B, 4,* 131–134.

Roeckelein, J. E. (1996). Citation of laws and theories in textbooks across 112 years of psychology. *Psychological Reports, 79,* 979–998.

Roeckelein, J. E. (1997). Psychology among the sciences: Comparisons of numbers of theories and laws cited in textbooks. *Psychological Reports, 80,* 131–141.

Root-Bernstein, R. S., Bernstein, M., & Garnier, H. (1993). Identification of scientists making long-term, high-impact contributions, with notes on their methods of working. *Creativity Research Journal, 6,* 329–343.

Rosenthal, R. (1990). How are we doing in soft psychology? *American Psychologist, 45,* 775–777.

Rushton, J. P. (1984). Evaluating research eminence in psychology: The construct validity of citation counts. *Bulletin of the British Psychological Society, 37,* 33–36.

Schachter, S., Christenfeld, N., Ravina, B., & Bilous, F. (1991). Speech disfluency and the structure of knowledge. *Journal of Personality and Social Psychology, 60,* 362–367.

Shadish, W. R., Jr. (1989). The perception and evaluation of quality in science. In B. Gholson, W. R. Shadish, Jr., R. A. Neimeyer, & A. C. Houts (Eds.), *The psychology of science: Contributions to metascience* (pp. 383–426). Cambridge: Cambridge University Press.

Simon, R. J. (1974). The work habits of eminent scientists. *Sociology of Work and Occupations, 1,* 327–335.

Simonton, D. K. (1976). Philosophical eminence, beliefs, and zeitgeist: An individual-generational analysis. *Journal of Personality and Social Psychology, 34,* 630–640.

Simonton, D. K. (1992). Leaders of American psychology, 1879–1967: Career development, creative output, and professional achievement. *Journal of Personality and Social Psychology, 62,* 5–17.

Simonton, D. K. (2000). Methodological and theoretical orientation and the long-term disciplinary impact of 54 eminent psychologists. *Review of General Psychology, 4,* 1–13.

Simonton, D. K. (2002). *Great psychologists and their times: Scientific insights into psychology's history.* Washington, DC: APA Books.

Simonton, D. K. (2004). Psychology's status as a scientific discipline: Its empirical placement within an implicit hierarchy of the sciences. *Review of General Psychology, 8,* 59–67.

Smith, L. D., Best, L. A., Stubbs, D. A., Archibald, A. B., & Roberson-Nay, R. (2002). Constructing knowledge: The role of graphs and tables in hard and soft psychology. *American Psychologist, 57,* 749–761.

Smith, L. D., Best, L. A., Stubbs, D. A., Johnston, J., & Archibald, A. B. (2000). Scientific graphs and the hierarchy of the sciences. *Social Studies of Science, 30,* 73–94.

Snow, C. P. (1960). *The two cultures and the scientific revolution.* New York: Cambridge University Press.

Stephan, P. E., & Levin, S. G. (1993). Age and the Nobel Prize revisited. *Scientometrics, 28,* 387–399.

Sternberg, R. J., & Gordeeva, T. (1996). The anatomy of impact: What makes an article influential? *Psychological Science, 7,* 69–75.

Suls, J., & Fletcher, B. (1983). Social comparison in the social and physical sciences: An archival study. *Journal of Personality and Social Psychology, 44,* 575–580.

Taylor, M. S., Locke, E. A., Lee, C., & Gist, M. E. (1984). Type A behavior and faculty research productivity: What are the mechanisms? *Organizational Behavior and Human Performance, 34,* 402–418.

Thagard, P. (1992). *Conceptual revolutions.* Princeton, NJ: Princeton University Press.

Watson, R. I., Sr. (1974). *Eminent contributors to psychology* (Vol. 1). New York: Springer.

Wolff, W. M. (1970). A study of criteria for journal manuscripts. *American Psychologist, 25,* 636–639.

Wolff, W. M. (1973). Publication problems in psychology and an explicit evaluation schema for manuscripts. *American Psychologists, 28,* 257–639.

Zusne, L., & Dailey, D. P. (1982). History of psychology texts as measuring instruments of eminence in psychology. *Revista de Historia de la Psicología, 3,* 7–42.

Received January 31, 2006
Accepted February 8, 2006 ■