Beware the Undiscovered Genius

Wendy M. Williams and Stephen J. Ceci

Department of Human Development

Cornell University

There is nothing like serving on a faculty grievance committee to give one a robust appreciation for the shortcomings of the faculty performance review system. At our university, a professor lodged a complaint stating that he had been denied well-deserved pay raises for more than five years. Evidence showed that his department chair had steadfastly refused his demands and ignored his pleas for fair treatment. Why?

The tenured professor making the complaint had not published anything for nearly a decade, choosing instead to immerse himself in the task of translating and interpreting a single stanza from a work of classical literature. Because the chair expected more, he bestowed below-average pay increases. The professor was outraged; he declared himself an undiscovered genius whose magnum opus would only be ready for publication at the end of his career, just a few more decades in the making. A great artist, he explained, cannot be judged fairly by the bureaucrats and bean-counters in his department. A life’s work, he claimed, is best appreciated after-the-fact, historically, as a legacy to the field.

What a sticky wicket for a grievance committee. After all, nobody wants to stifle creative genius in the arts and sciences. It is also true that brilliant works often are years in the making. But how can an academic department function within such a framework? Do we reward everyone,
assuming that those who do not produce finished products and who are inevitably “still at it” actually deserve as much status and material compensation as those who are productive year in and year out? What is better: work that is published in peer-reviewed journals, and critiqued (although some may consider it flawed) or work that escapes judgment because it is never complete? Who works harder, who accomplishes more, and how can we tell?

Evaluating performance is a central issue throughout all fields of research in the academy, the crux of hiring, tenure, promotion, and pay decisions. Done well, the institution benefits by having a system that truly maximizes meaningful performance in domains that matter. Done poorly, the best people flee because they are undercompensated, and promoted later then they should be (or not at all), while the less able soak up resources. The variability in how we evaluate performance in the academy is staggering. Some administrators stress publishing one type of material or another in peer-reviewed journals or books; others emphasize large federal or private grants; some reward teaching large courses and earning high ratings; others look for national service and awards and fellowship status in prestigious professional organizations. Too much depends on a haphazard, unreliable, and inadequately monitored system, particularly when it comes to evaluating scholarship.

We absolutely can and must do better. Academics are a systematic lot who know how to research an issue. The question of how best to evaluate research has not escaped scrutiny; it has generated considerable data. By examining these data we can build a better system to evaluate scholarly success in academe. Two guiding principles developed and tested through empirical research on human performance are first, that the best predictor of future performance is past
performance, and second, that high impact is associated with high productivity (1,2). On both theoretical and empirical grounds, the single best predictor of success in science is productivity, starting even before the Ph.D. is awarded. (3)

We can all point to examples of people who bucked these trends: the recluse in the corner office who shocked the college by publishing her first book at sixty, hailed by the scientific community as a masterpiece; the mid-career scholar who suddenly struck gold in his lab after a couple of decades spent collecting data instead of publishing it. These folks generate heated mailroom conversation and a tall dose of envy among their colleagues. But they are clearly exceptions, and there is no defensible reason for using their statistically aberrant career trajectories as the basis for crafting fair evaluation policies.

How do we handle this seemingly contradictory evidence about predicting future impact and success? Fortunately, there is a solution, one that will not make the undiscovered geniuses happy. People who produce do so for a variety of reasons--the desire to achieve, personal quest for fulfillment, financial or power aspirations--but at the core, an individual’s level of productivity results from what psychologist Dean Simonton has termed “initial creative potential.” This is an individual’s creative spark and energy level--a manifestation of intellectual drive and the quest to understand--as expressed through creative output. Noting that Edison held 1,093 patents and Picasso executed more than 20,000 paintings, drawings, and pieces of sculpture, Simonton studied creative contributions of diverse scholars from diverse fields in the arts and sciences. He responded to the pervasive beliefs that creative careers are almost infinitely varied and that there seem to be no secure regularities describing how creative
productivity varies across individuals or fluctuates over the lifespan.

Simonton demonstrated convincingly that these ideas are groundless: “the accumulated body of evidence shows that certain consistent relationships underlie what superficially appears to be a prohibitively complex phenomenon.” He provided multiple sources of evidence (analysis of citations, awards from professional associations, and peer recognition, to name just three) and showed that high impact results from high productivity (4). Prolific scholars may publish some duds, but just as Babe Ruth held a record for strike-outs as well as home-runs, they also publish more gems.

What does it mean in practical terms to operationalize the scientific finding that the best predictor of future performance is past performance? It means that a historically unproductive professor who is given extra departmental resources (such as research assistants, forgiveness of teaching assignments, departmental leaves, seed grants, additional lab space, or new equipment, in addition to healthy annual pay increases) will usually continue to be unproductive. Promise after promise will be made, but--as Annie sang--“tomorrow is always a day away.” Conversely, a formerly productive professor who suffers a personal or health-related setback or family tragedy will, after a short time, be productive once again. Bad hiring decisions cannot be made good simply by dumping more resources into underperforming scientists to galvanize their productivity. Administrators might wish it were so, but the data do not bear it out. Dead wood generally remains dead; productive researchers tend to flourish, grow, and bear fruit.

The high-productivity researcher has certain characteristics. On average, each publication is
cited more as is the person’s work as a whole; the individual receives more career awards; the published work is better recognized with awards and honors—all indicators of peer recognition. Administrators must not be seduced by claims of undiscovered genius, rationalizations on the basis of health and life-course disruptions, or claims of insufficient time and diminishing resources. The underlying relationship between productivity and impact, revealed through solid empiricism (2), will bear itself out. There will be exceptions, but there is no scientifically adequate way to predict them. Our universities and research centers are weakened if we fall victim to these romantic ideas about genius and intellectual productivity.

Why might administrators and others resist this message? One possibility is the awkwardness involved in confronting bad personnel decisions. Someone was hired; he or she is not performing. Rather than accept blame for a bad decision, it is simpler to invest more resources in the hopes of a turnaround. Another reason is that many administrators are no longer able to do research because they are consumed by managing large budgets, programs, and facilities. Some may fear the label of bean-counter, a derisive term flung about by researchers who tend often to belong to the less productive camp. There is the criticism that relying on quantitative measures of productivity is tantamount to abandoning the responsibility to form reasoned opinions about the scholarly quality of research. Not all prolific work is good work; Simonton notes that John Gray, a contemporary of Charles Darwin, published close to 900 works, none of which was particularly influential. Again, such aberrant careers are not a good basis for crafting policies, as the vast majority of productive researchers have high impact, and this relationship is found across most fields of scholarship and for myriad methods of operationalizing the dependent variables.(5) In the physical sciences, for example, quantity and quality are highly correlated
(with r’s between .7 and .9), a finding documented by many researchers since the seminal analysis by S. Cole and J. R. Cole over thirty years ago(6). Even when the dependent variable was operationalized not as total citations to all articles, but simply as number of citations to the top three articles, the correlation still reached an impressive .72 in a sample of physicists (6).

Then there is the criticism that one is stifling creative expression with formulaic expectations and equations: how can we possibly quantify genius? This message carries with it an assumption that it is wrong to create short-term meritocracies and differential rewards based on productivity. This meritocracy based on productivity is a concept that appears to be resisted by many administrators who instead prefer to parse out merit pay equally— in other words, irrespective of actual merit. Those queuing up to lodge a complaint are silenced by a memo stating that everyone got the same. For the administrator this is the course of least resistance. The more vocal adversaries of any administrator are seemingly those who are less productive and spend more time complaining than publishing. In the short run, it is faster and easier to give in to the squeaky wheels, but extremely costly in the long run.

The issue of a faculty member’s total contribution to a university goes well beyond the evaluation of just research. Many ways exist in which to make valuable contributions. Rational distributions of resources can reflect effective teaching and service, solid administration and leadership, and numerous other factors. We speak here to the issue of evaluating research, a key and central component of life in the academy and independent research centers, and one which generates considerable confusion and enmity. The message from empirical research is clear, if only we would heed it: Beware the undiscovered genius.
References and Notes


5. It has been demonstrated empirically that the ratio of “hits” (major, highly cited articles) to total productivity remains stable over the course of scientists’ careers (see note 4), resulting in a relationship between quality of publications and overall quantity that is highly linear (adding quadratic and higher-order terms to the regression equation does not account for greater variance), positive, and stochastic in the sense that sizeable variance remains unexplained (see note 2). This relationship is developmentally invariant over a scientist’s lifetime; when the number of high quality articles is regressed on total productivity, the slope of the regression does not typically change during a scientist’s different career epochs, even though there may be good years and bad years over brief periods. (see note 2)