On the surface, it seems obvious—and, indeed, it is commonly assumed—that the best predictors of the scientific impact of a project are the novelty, methodology, feasibility, and importance of the proposed studies (as assessed by peer review groups). Accordingly, funding decisions by the NIH, AHA, and other bodies are based primarily on what the applicants say in their proposals. What the applicants have actually done in the past (i.e., their track record of productivity or lack thereof) is generally seen as a secondary factor; reviewers are reluctant to emphasize it lest they may be accused of being biased in favor or against an investigator. Although the evaluation of a grant proposal is supposed to include the qualifications of the investigators and, in the case of competitive renewals, the progress made in the previous funding period, these factors are not paramount, and past productivity remains peripheral in the overall assessment. Reviewers commonly assume that the applicants will do what they say they will do and that, if they do it, they will publish the results of their work. Projects, we are told, must be evaluated on their own merit, not on the basis of what the applicants did (or did not do) in the past. What matters most, we hear, is the future.

The future, however, is determined to a very large extent by the past. Furthermore, as the Italian proverb quoted at the beginning of this article reminds us, saying something and doing it are two very different things, which are separated by a vast chasm. Another, related concept is that encapsulated by the Italian proverb “Tra il dire e il fare c’è di mezzo il mare” (“Between saying and doing there is the sea”).

Ab actu ad posse valet illatio
(Latin proverb, loosely translated as “From the past it is possible to infer the future”)

Facta non verba
(Latin motto, “Deeds, not words”)

On the surface, it seems obvious—and, indeed, it is commonly assumed—that the best predictors of the scientific impact of a project are the novelty, methodology, feasibility, and importance of the proposed studies (as assessed by peer review groups). Accordingly, funding decisions by the NIH, AHA, and other bodies are based primarily on what the applicants say in their proposals. What the applicants have actually done in the past (i.e., their track record of productivity or lack thereof) is generally seen as a secondary factor; reviewers are reluctant to emphasize it lest they may be accused of being biased in favor or against an investigator. Although the evaluation of a grant proposal is supposed to include the qualifications of the investigators and, in the case of competitive renewals, the progress made in the previous funding period, these factors are not paramount, and past productivity remains peripheral in the overall assessment. Reviewers commonly assume that the applicants will do what they say they will do and that, if they do it, they will publish the results of their work. Projects, we are told, must be evaluated on their own merit, not on the basis of what the applicants did (or did not do) in the past. What matters most, we hear, is the future.

The future, however, is determined to a very large extent by the past. Furthermore, as the Italian proverb quoted at the beginning of this article reminds us, saying something and doing it are two very different things, which are separated by a vast chasm. Another, related concept is that encapsulated by the popular American adage “Actions speak louder than words” (I would say, “much louder than words”): that is, what applicants have actually done carries more weight than what they say they will do. (A precursor to the American adage is the Latin “Facta non verba”.) Do you want to know what someone will do tomorrow? Your best guess is to look at what he/she is doing today or has done yesterday. One does not need bibliometric analyses, mathematical formulas, or computer models to grasp this simple concept; the correlation of past behavior with future behavior is rooted in human nature. This was well known to the ancient Romans, who coined the proverb “Ab actu ad posse valet illatio” (“From the past it is possible to infer the future”). Yes, people can change, but major changes happen rarely and when they do, it is usually a gradual process. I think it is the increasing appreciation of these facts that is fueling a reassessment of how grant recipients should be selected to attain optimal productivity. For more than a half century, funding decisions have been underlain by the tenets summarized in the previous paragraph, that is, grants have been awarded primarily on the basis of what the applicants say; recently, however, the validity of this approach has been questioned, leading the NIH to seek objective evidence of effectiveness and to consider alternative funding models. This is a welcome development.

Problems With the Current System
For many years, I have advocated a change in the system used to allocate research grants. I believe the current method is inadequate, for several reasons (Table 1). Instead of working in the laboratory to generate new knowledge, investigators must spend a significant part of their time (at least 40% according to some estimates) writing long and burdensome grant applications in which they are expected to provide a detailed description of what they plan to do in the next 5 years; actually, taking into account the lag from submission to award initiation, this means 2 to 6 years after the application is submitted—an interval long enough to render some studies obsolete (vide infra). In addition, since the grant scoring process typically favors applications with extensive preliminary data, it is not uncommon for the proposed studies to have been performed, at least in part, by the time the application is submitted. (It has been said many times that “the best grants are those that have already been done”.) The Center for Scientific Review and its reviewers must therefore spend a significant amount of time and money to evaluate studies that may have already been done, or may never be done, or may be done in a manner that is materially different from what was proposed and reviewed. Although the purpose of a grant is to perform a specific set of investigations, the investigator is often not held accountable for their completion and can renew the grant even if those studies have not been done. Clearly, there are significant problems with this system, as explained below (Table 1). First, a good scientist does not know exactly what he/she will be doing in 2 to 6 years because many things are likely to
change during this time. If someone knows exactly what studies he/she will be doing 5 years from now, chances are that his/her research is not particularly trail-blazing and does not move the field forward. Science changes very rapidly, at an ever faster pace. Ideas and techniques that were timely a year ago may be outdated today. (This is particularly true for basic research; clinical and population research change more slowly.) So, the reviewers’ expectation that investigators provide solid preliminary data to justify experiments that will be done 5 years later (which sometimes means the proposed studies have, for the most part, been done when the application is submitted) is not exactly a recipe for creative or innovative work. Asking basic investigators to provide a detailed description (including specific experimental protocols, methods, number of animals, etc) of studies that they will do 2, 4, or 6 years into the future is, at best, an academic exercise that consumes energy and time but bears little resemblance to reality (ie, to the studies that the investigators will actually do). Yet, this partially imaginary account of the applicant’s work is used by reviewers and granting agencies as the basis for funding decisions.

Another major problem with the current system is that there is a poor correlation between the quality of an application and the productivity resulting from the grant. A great application does not necessarily translate into a successful project. Deftly-crafted applications can enthral reviewers and get funded; however, the set of skills required to prepare a good application is not the same that is required to carry out the work and publish it. Some investigators are gifted writers and even visionaries, but poor executors. They may prepare beautiful applications that have all of the requisites for excellence (viz, the proposed studies are innovative, mechanistic, feasible, and potentially of high impact), but then they may fail to deliver, for a variety of reasons; for example, they may lack the energy to carry out the proposed studies, they may encounter technical difficulties that they are unable to overcome, they may not be able to lead the research team, they may lose focus, etc. In other cases, they may do the work described in the application but then lack the motivation to write it and publish it, which is equally unsatisfactory because, for all practical purposes, what is not published does not exist. In all of these situations, despite a well-written proposal, the final outcome is the same: the awardee fails to accomplish the purpose of the grant, which is to transduce dollars into new knowledge useful to the scientific community.

(Parenthetically, it is very important to define “productivity.” This term, as used herein, does not mean number of articles published; it means number of publications that significantly advance human knowledge. One article that is innovative, important, and comprehensive may advance the research field more than several less outstanding articles and thus may denote greater productivity than several lesser articles.)

The success of a grant is measured by its ability to promote high-quality publications, that is, publications that advance the field. So, the key question that funding agencies face is: what is the best criterion to identify applications that are most likely to generate high-quality publications? Is it the quality of the content (novelty of the idea, importance of the problem, soundness of the methods, quality of the approach)? These are all very important aspects, but, as explained above, the proposed studies may never be completed or may never be published. Is it the quality of the presentation (elegance of the prose, organization of the content, clarity of the exposition, cogency, and eloquence of the arguments)? These are all laudable features, but they do not necessarily lead to generation of new, important knowledge: a well-crafted grant application can meet all of these standards, and yet result in little or no new significant knowledge.

So, if a great application does not necessarily predict a great outcome, what does?

Rationale for Funding People Rather Than Projects

Here, it is important to distinguish between midcareer/senior investigators, who have had ample time to make scientific contributions, and early-career investigators, who have not had such an opportunity. For midcareer or senior investigators (ie, individuals who have been faculty members for at least 10 years or so), I believe that the recent record of productivity is a more reliable predictor of outcome than the quality of the application. The reason is that, in science, the best predictor of future performance is past performance. Does a reviewer want to know how productive an applicant is likely to be in the next 5 years? The reviewer should look, first and foremost, at what the applicant has done in the past 5 years. This criterion is, of course, not infallible, and all of us can think of exceptions; nevertheless, in the aggregate, I believe it is better than anything else we have to predict the future performance of an investigator.

A grant application is a promise. Is there any better criterion to gauge the credibility of a promise than to examine whether the person who makes it delivered on his/her previous promises? A useful analogy could be that of a homeowner who is looking for a contractor to do work on his/her house. Many contractors may promise outstanding results and present excellent, well-articulated, and eloquent plans; however, the most important question the homeowner asks is “How has this contractor performed in previous jobs?” A similar approach should be used by granting agencies before they invest money in a research proposal.

In a manner analogous to the law of inertia that governs nature, productive people tend to remain productive; conversely, unproductive people tend to remain unproductive. (The definition of “productivity” is given earlier in this Perspective.) A 10-year track record should be sufficient to sort these two categories with reasonable accuracy. If an investigator has not been very productive for the past 10 years, the chances that he/she will suddenly become highly productive are very slim. By the same token, someone who has been highly productive for 10 years is likely to continue to be so for the next 5 years. (Again, there are exceptions, but they do not invalidate these general patterns.) From an epistemological perspective,
the concept that recent productivity trumps the quality of the application in predicting whether an investigator will generate new significant knowledge is not just a theoretically plausible construct predicated on deductive reasoning, but also an inductive formulation supported by emerging empirical evidence; for example, recent work by Lauer and colleagues demonstrates that the prior publication productivity of NHLBI grantees (measured by the number of articles and citations of the applicant in the 5 years that preceded the start of a grant) predicts grant-specific citation impact, whereas the percentile ranking of the application does not.

Possible Approaches to Funding People

If past publication productivity is the best predictor we have, how can we use it to select grant recipients among applicants who fall into the midcareer/senior category? One can envision different scenarios depending on how much weight one wishes to place on past productivity. A radical approach would be to base funding decisions solely on it; in this case, applicants would be asked to submit a list of their publications in the past, say, 5 years, highlighting their significance, together with a brief outline of the studies that they intend to perform in the next 5 years. Award of a new 5-year grant would be based largely on the amount of significant new knowledge that the applicant has generated in the previous 5 years. (In my opinion, awarding grants for periods longer than 5 years would not be wise.) As mentioned earlier in this essay, “productivity” should not be measured by the number of articles but by the number of high-quality, high-impact articles that significantly advance the field. This is, of course, a subjective judgment that the reviewers of the application will have to make. There exist objective parameters of impact, such as number of article downloads from the website, number of citations (corrected for the specific field of research), and other bibliometric measures, that can aid in the process, but they are no substitute for the judgment of expert reviewers regarding whether the applicant’s work has advanced his/her field; that is, whether, directly or indirectly, it has produced paradigm shifts, illuminated the pathogenesis or pathophysiology of disease, or laid the foundation for new therapeutic or diagnostic modalities. It is important to stress that assessment of the productivity of an investigator should be based only on the quality and impact of his/her publications; other accomplishments (awards, honorary titles, administrative positions, etc) should not be taken into account, as they do not directly measure generation of significant new knowledge.

A less radical approach would be to use a composite score in which the weight assigned to past productivity is the dominant, but not the only, factor. In this scenario (which is a variation of the composite score proposed by Eugene Braunwald in the News & Views article in this issue), applicants would still be required to describe the specific studies that they propose (possibly in abbreviated form and with less detail than in the present system). A composite score would be developed that takes into account both the application and the track record of the applicant, with the latter constituting more than 50% of the total score. For example, 60% could be accounted for by the track record and the remaining 40% by the novelty, feasibility, methodology, and impact of the proposal. This second, more conservative model may be appropriately used as an initial experiment to determine whether it improves the correlation between peer review assessment of the proposal/proposer and subsequent productivity. If it does, the weight assigned to the applicant’s track record could be gradually increased in subsequent iterations of this formula.

Which model should be used? Should funding decisions be based solely on past productivity? Should the merits of the proposal be interpreted solely based on what is already known? Should the potential risk of failure be accounted for? Should the potential benefits of a larger investment be considered? Should the proposal be evaluated based on the probability of success? Despite these differences, the key factors are the same: productivity, innovation, and impact. In making funding decisions, reviewers need to consider these factors and to balance them with the need to avoid premature and deleterious decisions. As Lauer proposes, randomized trials comparing new systems with the current system, and between themselves, would enable meaningful conclusions and evidence-based approaches to allocation of grant funding. Although this would obviously require time, it is the most sensible strategy. The stakes are high that any change is bound to be met with fierce opposition. Objective evidence is necessary to overcome resistance and achieve the necessary buy-in from the scientific community.

It may also be sensible to use different formulas for different types of applications. For example, when it comes to clinical and population studies, I believe that the specific research protocols that are being proposed need to be described in the application and taken into account in making funding decisions, as is done currently. Because of the safety and ethical considerations associated with human studies, reviewers need to know exactly what the applicants propose to do. In addition, as noted above, methods and techniques change less rapidly in clinical/population research than in basic research, and so the problem of proposing methods that become quickly outdated is less severe in the former.

Regardless of the exact formulas used to evaluate applications, it is important that funding be awarded for 5 years, so that investigators have adequate time to accomplish their goals without the need to deliver results in 2 to 3 years. Currently, the average duration of NHLBI grants is ≈4 years; since most investigators wish to avoid a funding hiatus, they must apply for a competitive renewal 2 to 3 years after the inception of the grant—a timeline that precludes long-term plans and stifles innovation and experimentation.

Early-Career Investigators

It is important to stress that the new funding criteria described above should not be applied to early-career investigators (ie, investigators who have been faculty members for less than 10 years). It is obvious that for them, the track record of publications cannot be used as the sole or main criterion for awarding grants, because they have not had sufficient time to establish a track record. The productivity of students and trainees is difficult to interpret because it is heavily dependent on mentors and environment; it is only after these fledgling investigators “leave the nest” and start their own program that they reveal their true colors, and this usually requires several years of faculty appointment. A system that funds individuals rather than projects would be unfair to beginner scientists because it would pit them against more senior scientists who would have an obvious advantage; indeed, this may be the reason...
why some institutions that fund people, such as the Welcome Trust, have reported a progressive “graying” of their grantees. The traditional peer-review system based on the quality of the application is probably the best for early-career investigators. Their applications should be evaluated and funded separately from those of more senior colleagues. In fact, it may be wise to stratify applicants into different groups depending on their career stage (early career, midcareer, and senior), so that investigators compete with other investigators of similar seniority.

Can Productivity Be Measured?
If past productivity becomes the major criterion for awarding grants, what exact measure of productivity would be most useful in informing funding decisions in a manner that is objective and reliable? A variety of metrics has been proposed to predict a researcher’s future performance, but no agreement has been reached as to their usefulness and validity. For example, it has been reported that the future h-index of an investigator is moderately predictable on the basis of the current h-index, number of articles, publications in prestigious journals, and diversity of the journals in which the articles are published; however, the h-index would not be appropriate to inform funding decisions because this metric has several flaws, including its intrinsic autocorrelation and the fact that its predictive power is heavily dependent on career age and is least accurate for young researchers. The annual number of citations at the time of prediction has been claimed to be the best predictor of future citations. The number of publications and citations in the previous 5 years has been found to correlate with the citation impact of a grant. Other variables or formulas will undoubtedly be proposed in the future. Defining the exact metric(s) to be used to predict objectively an applicant’s postaward performance is outside the scope of this Perspective; the important concept here is that, as documented by Kaltman et al, past publication performance can be used to garner objective, unbiased insights into an applicant’s potential for future publication success.

Whether citations are the optimal measure of grant success is also debatable. “Niche” research areas generate fewer citations, even when the work is highly innovative and impactful within those areas. It is well known that basic research tends to be cited less than clinical research, which is why it is unfair to compare impact factors of basic and clinical journals. The journal in which an article happens to be published (a decision usually made only by three to four people: the editor and the reviewers) also affects its citations. Nevertheless, there are no other widely accepted measures of grant impact, that is, of how much someone’s work has advanced knowledge and medicine and has influenced other investigators. In general, citations do measure the interest of the community in published work and offer the advantage of being a relatively objective metric. If they are corrected for the specific research field, citations can be a useful and objective (although not perfect) measure of postaward performance.

As already mentioned, however, it is important to bear in mind that bibliometric measures of productivity can be used as supporting evidence but cannot and should not replace the primary measure of productivity: peer-review evaluation of the scientific impact of an applicant’s work on the field, as defined above.

Potential Problems With Funding People
No peer-review system is perfect, and there are weaknesses with any method used to evaluate applicants or applications. A common criticism of a system that funds people rather than projects is that it may favor senior, more established investigators. I do not believe this concern to be founded. First, as mentioned above, early-career investigators should be evaluated separately from more senior investigators; the funds for these two categories should also be separate. Second, “the rich” would not get richer unless they continue to perform at a high level of productivity; if senior investigators do not remain productive, they would lose their grants. Admittedly, there is a danger that Study Sections may evaluate the productivity of an investigator by the number of publications rather than by their quality and impact; appropriate measures should be taken to avoid this undesirable outcome. High-quality reviewers, selected on the basis of their scientific stature and credentials, are indispensable for the new system (or any system) to succeed. Accomplished scientists are optimally poised to recognize the potential of other accomplished scientists.

Previous Studies and Opinions
Taken together, the considerations expounded in this essay support the conclusion that it is perilous to rely on the characteristics of the application as the principal criterion to predict postaward performance. Others have also voiced concern that the excellence of a grant proposal is a poor predictor of subsequent success. Importantly, as mentioned above, empirical evidence exists to support the thesis espoused in this article. In a recent study published in Circulation Research, Danthi et al have found essentially no association between the quality of R01 grant applications to the NHLBI (measured by percentile ranking) and their scientific impact (measured by the number of publications and citations), providing objective evidence that peer-review assessment of an application does not correlate with postaward productivity. In contrast, this group found a significant association between the past productivity of the investigator and the scientific impact of the grant.

The Fondation Leducq program offers investigators greater freedom than the typical NIH grant; its Transatlantic Networks of Excellence are oriented toward long-term goals rather than short-term or predefined results. Investigators are given considerable latitude and independence in pursuing their objectives, including the ability to change the research studies, to restructure the investigator network, and to reallocate funds. Arguably the most relevant example of successful implementation of a system that funds people rather than projects is the Howard Hughes Medical Institute (HHMI). Azoulay et al compared HHMI investigators with a cadre of NIH-funded investigators with similar professional stature and accomplishments (as determined by awards, leadership positions, recognitions, etc) and concluded that the performance of HHMI investigators was superior in terms of impact, creativity, and productivity. The authors stated: “HHMI investigators produce high-impact articles at a much higher rate than a control group of similarly accomplished NIH-funded scientists. Moreover, the direction of their research changes in ways that suggest the program induces them to explore novel lines of inquiry.”

The
greater productivity and creativity of HHMI investigators was ascribed to their longer review cycles and greater freedom to experiment. It makes sense that not tying investigators to performing specific studies may promote more creative work.

Conclusions
Granting agencies should move toward implementing a new system that funds people rather than projects, but only for midcareer and senior investigators, not for early-career investigators. Changes should be made cautiously and, preferably, on the basis of randomized trials comparing different funding systems. It is important that funding decisions be based solely on scientific productivity, not on name recognition, personality, stature in the field, administrative achievements, or other extraneous criteria. For midcareer or senior investigators, this new system might offer many advantages (Table 2). It might predict more accurately whether the grant will generate new knowledge that advances the field. It might promote creativity by lengthening the funding period from ≈4 to 5 years and by liberating investigators from the shackles of doing experiments that were conceived several years ago and may be outdated. It might save considerable time and effort to investigators who, therefore, will be able to focus on their research rather than on writing grant applications. It might save considerable time to grant reviewers, because evaluating applications will be easier and faster. It might also save taxpayers’ dollars by reducing the total labor necessary to review grant applications.

The time has come for a paradigm shift in the evaluation of research grants and in the selection of grant recipients. The epistemic analysis discussed in this essay provides a rational foundation for a new approach. Regardless of the specific formulas used, what applicants do should matter more than what they promise. Facta non verba.

Acknowledgments
I wish to thank Aruni Bhatnagar, A.J. Marian, Michael S. Lauer, and Steven P. Jones for helpful discussion.

References

Key Words: funding ■ National Institutes of Health (US) ■ research support
Roberto Bolli

Actions Speak Much Louder Than Words: For Midcareer and Senior Investigators, the Track Record of Productivity Should Be Paramount in Selecting Grant Recipients

Roberto Bolli

*Circ Res.* 2014;115:962-966; originally published online November 4, 2014; doi: 10.1161/CIRCRESAHA.114.305556

*Circulation Research* is published by the American Heart Association, 7272 Greenville Avenue, Dallas, TX 75231
Copyright © 2014 American Heart Association, Inc. All rights reserved.
Print ISSN: 0009-7330. Online ISSN: 1524-4571

The online version of this article, along with updated information and services, is located on the World Wide Web at:
http://circres.ahajournals.org/content/115/12/962

Permissions: Requests for permissions to reproduce figures, tables, or portions of articles originally published in *Circulation Research* can be obtained via RightsLink, a service of the Copyright Clearance Center, not the Editorial Office. Once the online version of the published article for which permission is being requested is located, click Request Permissions in the middle column of the Web page under Services. Further information about this process is available in the Permissions and Rights Question and Answer document.

Reprints: Information about reprints can be found online at:
http://www.lww.com/reprints

Subscriptions: Information about subscribing to *Circulation Research* is online at:
http://circres.ahajournals.org//subscriptions/