

The Evaluation of Research Merit versus the Evaluation of Funding of Research

Michael Scriven

The Evaluation Center—Western Michigan University

The evaluation of research and researchers is an example of a fairly basic kind of evaluation. It normally involves either a ranking or a grading (aka., rating) of research projects or personnel for merit, worth, or significance, and these are tasks that we know a good deal about doing.

But the evaluation of research funding is another kind of animal altogether. It aims for an apportionment or allocation decision, which is either something essentially different from evaluation or, with a stretch, a highly complex kind of evaluation decision. It is certainly a decision that depends on more than one kind of basic evaluation, but it depends on them in a way that has never been reduced to a formula or computer program.

To a substantial degree, the major efforts that are being made by a number of countries to allocate governmental research funding to researchers and research projects in a rational way have been confounded by a failure to make this distinction. They tend to suggest, and be attacked as if this suggestion is correct,

that they are engaged in the evaluation of research, when in fact they are, quite rightly, engaged in working out the best way to allocate research funds.

It's easy to see the difference in simple cases. Suppose we are reviewing a number of requests for research funding, a task not unlike the task of deciding how to fund a number of researchers for the immediate future. Suppose that one of the proposals, Project A, is clearly the best of the bunch; but that it would use up virtually all of the available funds. It is not at all obvious that it should be the one funded. In other words, the correct allocation decision is not simply to fund by merit. Well, it might be argued, that's clear enough; but it just means that we ought to fund by cost-effectiveness, not just effectiveness. Suppose, then, that the probable payoffs from the top project are by far the largest in terms of return on investment as well as in terms of its merit. It still does not follow that the correct decision is to fund it alone. The strategy of putting all one's eggs in one basket, or maximax (on one definition of that term) is not demonstrably optimal. It risks the loss of all possibility of payoff if Project A fails; and it risks the loss of all personnel involved in the other proposals if they, lacking other support, turn to other fields during the time before A fails. These are big risks in the real world, and plain caution, or conservatism, or minimax, counsel other choices.

What is the correct decision in such a case? It depends on other variables, including variables that describe the context of the decision and notably refer to the levels of needs the number of other funding sources, with some consideration of the political climate of the country. So it will not be the same for two different countries, or employers, or agencies. But the first lesson to be learnt is not to think or talk as if what you are doing, if your job is to find the right solution here, is primarily or mainly to evaluate the research or researchers in the country. You will

certainly be doing that; but not only or principally that, and you should avoid thinking that because it will divert your attention from the real task (another relevant variable concerns the divisibility of A: does it have parts that could be funded now, with completion in the next funding period?)

Four concluding notes. First, the evaluation of research, although it's a fairly basic task, is not simple to do well and involves at least one major trap into which many have fallen—the trap of failing to distinguish concurrent or retrospective (aka. ex post facto) evaluation from predictive (aka. ex ante evaluation), two very different activities (see note 4). Second, the exact relation of resources, risks, needs, and political preferences to solve an allocation problem may not have been, and may never be reducible to a precise quantitative formula, but it can be explicated rather thoroughly, to the point where a training process that produces high inter-judge reliability and a reasonable level of validity is obtained. So there is light at the end of the tunnel—as long as you don't try to get to it by breaking through the walls too soon. Do that and you are likely to find the tunnel at that point is under water, and you will drown. Third, the logical model for allocation or apportionment is the one often used in the evaluation of investment portfolios and the comparison is quite enlightening. Besides reminding one that there is a rational way to go about it—or at least a difference between reasonable and highly unreasonable ways to go about it—it illustrates very well how needs and risks come in.

Forth, in funding future research by a researcher, or research group, while past track record is a major indicator, there are four contra-indicators that make its use inappropriate. Look only at the past eighteen months production. (i) If it shows a sharp downslope, either burnout or (ii) personal life trauma is likely and unless ruled out by further evidence, prior productivity becomes irrelevant. (iii) Examine

Michael Scriven

the content of the recent research; if it has shifted fields, the same applies. (iv) If the journals used have shifted sharply, the same applies. In all these cases, while the presumption of continued good, relevant research is refuted, it can be reinstated by strong refutatory evidence.

With thanks to Chris Coryn for significant improvements.