Carnegie Mellon University
Research Showcase @ CMU

Department of Social and Decision Sciences | Dietrich College of Humanities and Social Sciences

2000

Scientific Management of Science?

Baruch Fischhoff
Carnegie Mellon University, baruch@cmu.edu

Follow this and additional works at: http://repository.cmu.edu/sds

This Article is brought to you for free and open access by the Dietrich College of Humanities and Social Sciences at Research Showcase @ CMU. It has been accepted for inclusion in Department of Social and Decision Sciences by an authorized administrator of Research Showcase @ CMU. For more information, please contact research-showcase@andrew.cmu.edu.

Policy Sciences **33**: 73–87, 2000.
© 2000 *Kluwer Academic Publishers. Printed in the Netherlands.*

Scientific management of science?

BARUCH FISCHHOFF
Department of Social and Decision Sciences, Department of Engineering and Public Policy, Carnegie Mellon University, Pittsburgh, PA 15213-3890, U.S.A.
E-mail: baruch@cmu.edu

Abstract. A framework is advanced for allocating research resources, based on the value of the information that proposed projects are expected to produce. As an organizing device, the framework uses integrated assessments, showing the relationships between the predictors of outcomes arising in important decisions (e.g., interest rates, mortality rates, crop yields, crime probabilities). Proposed projects are evaluated in terms of their ability to sharpen estimates either of those variables or of the relationships among them. This approach is intended to allow diverse forms of science to show their potential impacts – while encouraging them to integrate their work. Where suitable estimates are available, the expected value of the information from alternative studies can be computed and compared. However, even at a qualitative level, formal analyses can improve the efficiency and integration of research programs.

Scientific management of science?

Science requires resources. At the very least, it needs the time and energies of skilled individuals. Often, it requires money as well – in order to secure the facilities, release time, travel expenses, and other things that money can bring.

As a result, allocating resources is a task that, in one way or another, occupies all scientists. As individuals, they must make choices about which projects to pursue and which to abandon, when to take on additional students and when to concentrate on current ones, and whether to spend a week in the lab or at a workshop. If they choose well, scientists increase their chances of living rewarding lives, producing useful results, and leaving a legacy of admiring students and colleagues – not to mention receiving tenure, grants, patents, and other tangible benefits.

When individual scientists choose poorly, they are the primary losers. Although funders might lament not having invested their resources more wisely, the opportunity costs of single projects are seldom that large. The stakes increase when investments are made in areas, rather than in individual projects. Sciences (and scientists) could do quite well for themselves, even as sponsors receive a poor return on substantial investments. This essay offers a framework for managing the allocation of research resources, drawing upon the tools of decision theory (Clemen, 1991; Grayston, 1960; Green and Swets, 1966; Keeney, Hammond and Raiffa, 1998; Morgan and Henrion, 1990; Raiffa, 1968) and experiences with priority setting (National Research Council, 1995a, b; 1998a, b).

Selection criteria

Science today has multiple sponsors, including society as a whole (acting through government agencies), interest groups (acting through research foundations), or firms (acting through in-house and extramural research programs). The projects that each funds should pass two criteria: (a) the sponsors consider the topics worth understanding and (b) the relevant sciences can make enough progress to justify the investment.

The former criterion is a matter of the funders' tastes, needs, ethics, etc. These values could be widely shared (e.g., curing a major disease) or fairly specialized (e.g., determining the roots of a small language group). They might even be controversial. For example, the social sciences sometimes run afoul of people who resent empirical challenges to their intuitive (or ideological) claims about human nature. A pluralistic society may need multiple funding sources, in order to satisfy its citizens' diverse needs.

The latter criterion requires candid input from scientists, regarding what they can produce, if provided resources. If scientists promise more than they can deliver, then they may divert resources from more worthy causes. If scientists understate their capabilities, resources may go to less worthy causes (including, perhaps, sciences less candid about their limitations).

Analytical options and hesitations

If these two sets of assessments exist, then there are formal methods for their combination (Grayson, 1960; Platt, 1964). These procedures calculate the expected value of the information from each proposed research project or program, valued in terms of what funders want to learn. They are, however, rarely used. Rather, scientists and research managers render educated judgments about which investments make the most sense to them (National Research Council, 1995a, b; Tengs, 1998).

On the surface at least, there are good reasons to avoid formalization. Analysis requires quantification of both criteria. Neither task is for the faint of heart. Each could distort funding if done poorly. A numerical approach to values could unduly favor research focused on readily quantified outcomes (e.g., economic benefits, deaths averted), and create an unhealthy demand for quantification procedures. Those concerned about hard-to-measure outcomes (e.g., endangered species with no commercial uses, peace of mind, understanding of ancient civilizations) may see an unsavory choice between adopting poor measures and being ignored altogether.

Estimating what research is likely to reveal about these valued outcomes faces its own problems. Some scientists dislike assigning probabilities to anything but relative frequencies of previously observed events (and not the outcomes of possible experiments). Some scientists dislike predicting the yield of studies whose very design depends on other, yet-to-be conducted studies. When scientists are willing to make predictions, careful elicitation is needed to capture their beliefs (Morgan and Henrion, 1990), which may still prove to be poor guides to the future. Studies have sometimes found expert forecasts to be overconfident or in wild disagreement. That seems especially likely in situations lacking the conditions needed for learning how to assess uncertainty: prompt, unambiguous feedback that rewards candor (Fischhoff, 1977; Kahneman, Slovic and Tversky, 1982; Kammen and Hassendzahl, 1999; Morgan and Keith, 1995; Murphy and Brown, 1984; Yates, 1990). A lifetime (or professional career) of making value judgments and assessing uncertainties provides no guarantee of being able to express oneself in the language of analysis (Fischhoff, 1980; 1989).

Alternatives to analysis

These risks of analysis must then be compared to those of the alternative. In the traditional, informal process (a) funders specify their general priorities, (b) scientists propose projects promising relevant progress, (c) funders support the seeming 'best buys,' and (d) researchers do the best they can with what they get, while positioning themselves to make even better promises for the next funding cycle.

Scientists can shape this process by influencing either what the funders value or what they expect to get from research. Like other citizens, scientists can lobby for particular topics. They might point to the severity of a health problem (e.g., asthma) or to the beauty of an intellectual pursuit (e.g., measuring the dark matter in the universe). If funders accept these tastes, then these sciences will benefit. Or, scientists can influence expectations by showing how their work or discipline can achieve funders' goals.

Science's many successes show that these informal processes have some validity. However, it is easy to imagine their weaknesses. Scientists can distort the goal-setting process by assuming unwarranted moral authority, asserting their priorities as those of society's. Scientists can distort the assessment of research opportunities, by making unrealistic promises. They may be deluding themselves, overly enthused about what they hope to find, or they may be acting strategically in order to get disproportionate attention.

Biased claims need not distort resource allocation – if all sciences have equal ability to make their case. The process will be suboptimal, however, if some disciplines hog the limelight, have particularly inflated self-assessments, criticize colleagues with particular harshness, or find the politics of funding particularly distasteful.

Once out of kilter, an allocation system may be hard to fix. Both funders and recipients have strong incentives to defend past expenditures. Moreover, funding can create a 'treatment effect' that is difficult to detect or undo (Einhorn, 1982): receiving resources increases scientists' chances of producing interesting results, recruiting talented students, and discovering serendipitous findings. As a result, funded scientists will tend to be more productive, prominent, etc., than their unfunded colleagues. One knows that those other scientists would have done something had they had the resources. However, such possibilities are inevitably less tangible than actual research results, whose existence helps to define what science and society value.

Planned and unplanned analysis

In addition to their near-term performance, resource-allocation processes must be judged on their ability to learn. Science is a great experiment. As such, it needs to have explicit expectations in order to get the prompt, unambiguous feedback needed to evaluate and improve its practices. Without an analytical approach, fifty years from now, science will still be guided by the best guesses of those invited to the critical allocation meetings. Those guesses are expert judgments, but judgments nonetheless. The hypotheses underlying them need to be clearly stated and evaluated, if the management of science is to follow the canons of science.

Arguably, 'trusting us' has served society well. However, that claim remains a matter of judgment. To the extent that allocation choices are made in a black box, it is harder to defend choices and demonstrate performance when funders must justify their portfolios to their constituencies (be they voters, trustees, etc.) (NRC, 1998b). When those justifications fail, the price just might be a slower increase in funding. However, it could also be the chaos created by mechanical reliance on imperfect productivity measures (e.g., number of publications, patents, citations, graduate students, experiments, matching industry contracts) capable of capturing only some parts of some sciences. Another is political meddling (e.g., legislative earmarking of funds), more capable of representing what people want from science than what they can get.

Easing into analysis

Decision theory, economics, operations research, and management science all offer formal methods for analyzing the informational yield of possible actions. Yet, it is hard to find a whisper of them in any allocation process or performance review. Hubris and defensiveness may account for some of this reluctance. However, there can also be more legitimate reasons: scientists are busy, with little time to mess with new formalisms, especially ones that might stand between them and their intuitions. Even scientists may prefer to make choices that feel right to choices that emerge from an unfamiliar algorithm, whatever its pedigree. Because analytical intuitions accrue slowly, an abrupt procedural change may, indeed, be imprudent.

Moreover, even for the analytically fluent, productive change could not come very fast. Allocation decisions regarding research resources are fundamentally complex. A formal allocation procedure with requisite variety must also be complex. It cannot be developed overnight.

Thus, neither method nor users are ready for one another. One way out of this stalemate is to cultivate them in tandem. That is, begin with a version of the formal approach that is simple enough to make sense to practicing scientists, while still being compatible with more complex analyses. Let scientists try it out, adapt it to their needs, and challenge the analysts to improve it. Over time, scientists and analysts may learn one another's language and concerns.

One way to connect models and intuitions is through examples, instantiating common situations. Over time, these stories can become templates for characterizing real-world problems (e.g., 'That looks like an inclined plane,' 'You're committing the base-rate fallacy,' 'Those incentives create an agency problem.') Such pattern-matching would be a natural stepping stone to learning how to add the unique features of specific applications and, eventually, solving novel problems from first principles.

To this end, the next section proposes a general analytical approach for setting research priorities, followed by two examples. These examples attempt to illustrate the approach's logic in two domains, showing how it could accommodate a variety of research projects. The conclusion considers the institutional steps for implementing such an approach.

The value of scientific information

A general method should allow any potentially relevant science to articulate the case for investing in it, with that effort advancing the pursuit of knowledge, without costing too much. It should help scientists to connect what interests them with what matters to funders, creating a clearer trail from the lab to the world.

Decision analysis provides the most general approach to valuing the yield of information-gathering exercises. It is known, appropriately enough, as *value-of-information* analysis (Clemen, 1991; Morgan and Henrion, 1990; Raiffa, 1968). It asks how the expected value of a decision is expected to change if one waits for the results of a study. The message of those results might be specific – as when a test clarifies a medical diagnosis and the corresponding treatment. Or, it might be general – as when research produces a diagnostic test that could improve many surgery decisions. The same logic applies whatever the focal impact (e.g., dollars, peace of mind, QALYs).

The logic of the analysis is straightforward in principle. First, identify each link between the research results and the valued outcomes. Then, assess the expected impacts of the research on reducing the uncertainty in those links. The most valuable projects are those that contribute the most to understanding (and perhaps affecting) those focal outcomes. Such thinking is implicit in scientists' thinking; why else would they choose their projects? However, rendering it explicit is a nontrivial chore.

Some basic researchers may find it a disagreeable chore, insofar as such consequentialist thinking offends their personal commitment to knowledge for knowledge's sake. Moreover, they must work harder than more applied researchers, to construct the inferential chains from research to practice. The intrinsic motivation for making this effort is understanding better the context for their research and its potentially useful connections. The extrinsic motivation is obtaining their rightful share of research resources generated by practical concerns. Analysis allows them to show how basic research results could affect many decisions, thereby having a large aggregate impact, even if the expected contribution to each decision is small. Problem-focused research programs should feel an obligation to show just what difference their results could make. Otherwise, they are vulnerable to the charge of exploiting a sponsor's concerns in order to satisfy their personal curiosity.

As mentioned, one can accept the idea of analysis in principle, but reject it in practice, if it seems unfair or unintuitive. The following section attempts to address these concerns with a concrete example, showing how different sciences can demonstrate their relevance in value-of-information terms.

Information extraction – a prospecting metaphor

Oil prospectors are forever looking for additional sources. With limited equipment, personnel, and capital, prospectors need to decide where to drill. If they drill in the center of known fields, then they are likely to hit oil, but unlikely to learn much new. As a result, the expected information-to-cost ratio should be low (but not zero, as long as surprises are possible). Drilling near the periphery of known fields has a lower probability of finding oil, but higher informational value for what is found. Moreover, even dry holes can be informative (e.g., bounding the field, sharpening reserve estimates, clarifying geological structures). The practical value of that improved understanding is reflected in the improved efficiency of subsequent exploration decisions (Grayson, 1960).

Those calculations require quantitative estimates of the parameters in those decisions. Some are readily available, such as current day rates for rigs and crews. Others can be estimated from historical patterns, such as the days per exploratory drill, costs per subsequent development drill, and the market value of new oil from such formations. Still other estimates require expert judgment, such as the chances of finding specific features at various depths, given the partially known geology.

Formal analyses would calculate the expected value of the information from a proposed test drill for the set of contingent decisions. If that value exceeds the expected costs, then it passes a *cost-benefit* test, measured in money or another metric. It is then worth comparing to other possible test drills in order to identify the most *cost-effective options* – providing the greatest expected information yield, relative to their cost. Those calculations will reflect commercial considerations as well as scientific and engineering ones. For example, if a firm's lease is about to expire, exploratory drilling will only have value if it can inform the bidding for lease renewal.

In addition to elaborating their knowledge of existing fields, oil prospectors sink exploratory holes in relatively unfamiliar areas, hoping for pleasant surprises. They might see a good probability of a modest deposit, a small probability of a large one, or just a major area of ignorance. As site-specific knowledge decreases, basic geology becomes more important. Hence, there can be practical value to research that sharpens estimates of the probability of finding commercially viable deposits in a fundamental rock type.

The value of data-gathering actions (drilling, seismic work, etc.) depends on prospectors' ability to interpret what they get. As a result, there is value to activities that enhance that ability. Indeed, some such activities are routinely figured into the costs of drilling (e.g., hiring consultants, conducting lab work on core samples, installing software for visualizing underground structures). Those who invest commercially in creating methods for signal extraction perform, in effect, generalized value-of-information analyses. They gamble that someone will pay for their visualization software or their method for training workers to use it. Of course, projects that are attractive in principle may fail to fulfill their promise because of poor marketing or superior competing products. Projects may also succeed beyond their realistic dreams if unanticipated uses emerge.

As research becomes more fundamental, computing its value requires more thought. There are more links between research and any given application, as well as more applications to be analyzed. For example, improved science education could improve the pool of young petroleum engineers. Their greater productivity would have economic value for prospecting firms. Those who provide that education should get credit for it, as well as for its contribution to other industries. So should those who created the curriculum and those who performed the fundamental science underlying it. Such research might provide a better return than more direct investments, especially ones that are useless without adequate personnel. Similarly dispersed value might come from solving non-work problems (e.g., substance abuse, child care), so that workers can focus on their jobs and extract information more efficiently.

Just framing these question properly requires an intellectual effort, not to mention providing the needed parameter estimates. Nonetheless, with sufficient thought, a value-of-information perspective can formalize these potential impacts, helping scientists to articulate – and evaluate – claims that they often make informally. With further modeling efforts, one might even evaluate allocation procedures (e.g., individual-investigator versus center grants, intramural versus extramural research). Such studies are a long way from the oil patch. However, an analytical approach could clarify how their impacts trickle down to signal extraction in prospecting, as well as which intermediate activities might improve that flow (Galvin, 1998; Jones, Fischhoff and Lach, 1998, in press).

Integrated assessment

The success of such comprehensive analyses depends on being able to identify, and then quantify, these relationships. To do so can be a challenging task, requiring the integration of results from diverse fields. Basic research can require characterizing multiple links in multiple contexts. For example, improved estimates of the inflation rate can aid many economic decisions; such improvements are one benefit of research into consumers' beliefs about future economic conditions. Anticipated immune system robustness plays a role in many medical decisions; predictions about it might be improved by a better understanding of basic biology, the effects of socioeconomic status and age on health, etc. Thus, estimating the value of research related to inflation and immune competence may involve understanding these 'upstream' and 'downstream' connections.

The emerging discipline for such analyses is often called *integrated assessment*. Among the domains where it has been pursued most vigorously are complex environmental problems, such as acid rain and climate change (Dowlatabadi and Morgan, 1993, in press; Schneider, 1997). These models provide a common platform for the contributing disciplines to pool their knowledge, identify gaps, and establish research priorities. These are, typically, *reduced-form* models. They attempt to capture the key relations among the outputs of the relevant sciences, capitalizing on the detailed research without being drowned by it. Such models could be represented in various ways. A common formalism is *influence diagrams* (Clemen, 1991; Howard, 1989). These are directed graphs, in which each node represents a variable. An arrow connects two nodes, if the predicted value of the variable at its head depends on the estimated value of the variable at its tail.

Figure 1 presents such a model, integrating a research project on the risk posed by cryptosporidium in domestic water supplies (Casman et al., 1999). Problems typically arise when run-off carries these parasites from land grazed by infected livestock. When ingested, they can cause intestinal problems, which are discomfiting for healthy people, potentially lethal for immunocompromised ones. Estimating these risks requires inputs from microbiology (dose-response relationships), civil engineering (filtration and testing), ecology (land use), communications (attention to 'boil water' warnings), and psychology (perceived risk), among other disciplines. The computational version of this model specifies values for the variable(s) at each node and for the dependencies among them. In some cases, these estimates are extracted from empirical studies; in others, they are but expert judgments. In some cases, the relevant research has no other uses (e.g., the effectiveness of filtration systems designed for this specific parasite); its total value could be calculated from models like this, adapted to the conditions of different water supplies. In other cases, the research needed for this model could sharpen other models as well (e.g., determinants of the credibility and clarity of warnings).

One advantage of the influence-diagram representation for integrated assess-

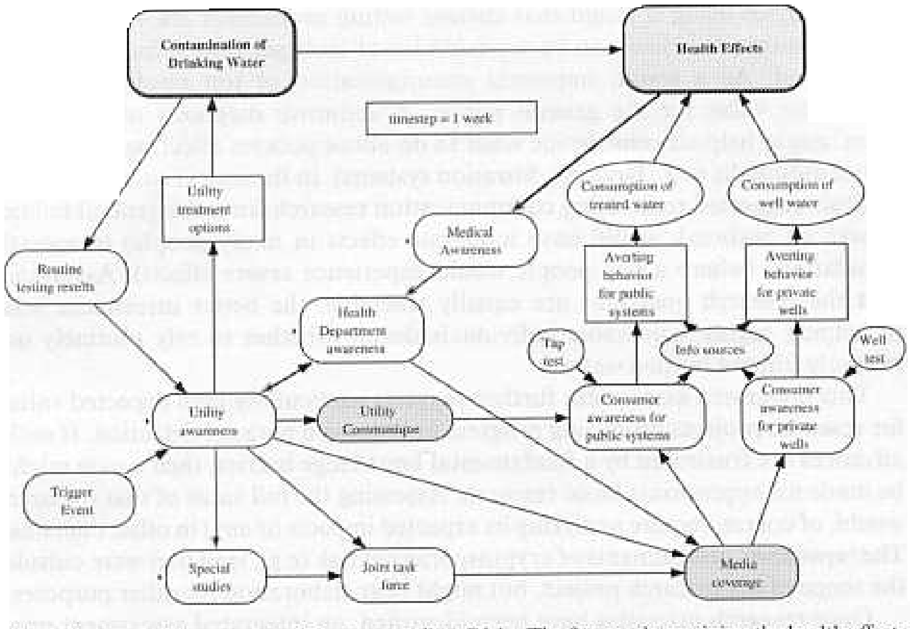

Fig. 1. Integrated Assessment of *Cryptosporidium* Risks. The factors determining the health effects (disease, death) arising from a 'trigger event' (e.g., flooding of an area grazed by infected cattle). These include the responses of individuals (e.g., averting behaviors, such as using bottled water) and institutions (e.g., communiqués from water utilities), as well as the efficacy of engineering practices (e.g., routine testing).
Source: Casman et al. (1999).

ments is its ability to accommodate diverse forms of evidence. A second is that its graphic form facilitates the cognitive task of figuring out how the pieces fit together. Scientists can make the case for the value of their research by pointing to the influence diagram and saying, 'Here is where I have something to add.' Having done so, they can proceed to formalize their claims, knowing that they have a place at the table. Even if no numbers are run, translating such claims into model terms forces clarity about the variables (e.g., what do we mean by 'inflation rate,' 'immune system competence,' or 'warning'?) and their relationships (e.g., avoiding recursive loops, with A predicting B predicting A, perhaps through some intermediate variables). Inevitably, formalization requires at least rough quantitative estimates – otherwise, anything that might conceivably be relevant could find its way into the model (and the research agenda). These order-of-magnitude judgments provide a transition to more quantitative analyses.

Research management

Figure 1 was created as the integrating core of a project focused on reducing cryptosporidium risks by better communication with water consumers. However, analyses using it found that current testing procedures are so ineffective that an outbreak is likely to have passed (or at least peaked) before its source is detected. As a result, improved communication of test results has little immediate value for the general public. A definitive diagnosis of 'what hit them' might help citizens decide what to do about policies affecting the risk of future outbreaks (e.g., land use, filtration systems). In the short run, though, the analysis suggested redirecting communication research from the general public (where an outbreak would have moderate effects in many people) to special populations (where a few people would experience severe effects). Assuming that the research questions are equally tractable, the better investment was in helping highly vulnerable individuals decide whether to rely routinely on properly filtered bottled water.

This integrated assessment further suggests particularly high expected value for research projects promising progress in improved parasite detection. If such advances are frustrated by a fundamental knowledge barrier, then a case might be made for appropriate basic research. Assessing the full value of that research would, of course, require analyzing its expected impacts (if any) in other domains. The 'upstream' determinants of cryptosporidium risk (e.g., land use) were outside the scope of the research project, but might bear elaboration for other purposes.

Once research priorities have been identified, an integrated assessment provides a basis for analyzing a research enterprise's ability to address them. At a minimum, one can ask whether anyone in the scientific community is supported to understand each element in the influence diagram (Jones et al., 1998, in press). More ambitiously, one can ask whether the investments are commensurate with the opportunities. As mentioned, relevant research could attempt to sharpen estimates of either the variables at the nodes or the dependencies between them. Variable estimates can be improved both by measuring them and by developing better measurement methods (thereby increasing the yield of each measurement). Dependency estimates can be improved both by measuring them and by refining the theories linking them (thereby increasing understanding of what has been measured). The efficiency of research might be improved by infrastructure investments, such as recruiting, training, archiving, and software. Some of the latter activities might also improve the yield of research in multiple areas. Those contributions would, then, have to be traced in the integrated assessments coordinating those other areas.

If they hope to maximize (and demonstrate) the usefulness of their programs, research managers need, somehow, to accomplish the work of integrated assessment. That is, they need to identify their domain, determine which of its nodes and links are most important, and ensure appropriate investments. If a program serves a single client (e.g., oil prospectors), then its integrated assessment may, initially at least, look much like that of its client. However, as the analysis progresses, it may identify important opportunities lying outside that client's conventional investments and, perhaps, the research program's boundaries. For example, specialists in prospecting might lack expertise in training, just as training specialists might not understand the needs of that industry.

When there is a mismatch between research priorities and institutional capacities, a management response is needed. One such response is transferring funds to programs that can address the research need, with appropriate joint supervision. Another is adding 'absorptive capacity,' in the sense of staff qualified to solicit and evaluate bridging proposals. More ambitious actions are needed when domains share a need, such as infrastructure that supports several sciences. Integrated assessment can clarify the multiple impacts of such investments. Such analyses are steps toward answering the recurrent strategic management question of whether funding levels are appropriate.

Conclusion

The current system for funding science is largely reactive. It changes when some research managers or decision makers sense a need and manage to make a case for it. Although this system has served us well, it is inherently vulnerable: some research groups may be disproportionately effective advocates, while others lack efficient mechanisms for getting organized or are too busy (or pessimistic) to try. Some practitioners and scientists may not recognize their mutual relevance, leaving opportunities unaddressed. The system as a whole may have undue inertia, favoring existing programs over proposed ones (if only because its procedures for evaluating proposals best suit familiar forms of research) (House Committee on Science, 1998; National Research Council, 1995a; 1998b; Tengs, 1998).

By facilitating value-of-information analyses, integrated assessments offer partial solutions to these problems. Research portfolios will be clearer if their elements are characterized in formal terms, showing the uncertainties that each project is meant to reduce. By accommodating diverse research activities, such analyses should create a relatively level playing field. Moreover, they should encourage researchers to enter it by helping them to tell their story – about how their work relates to practical concerns and to other research. The exercise might even suggest ways to reshape their research in order to have a better story to tell.[1]

Anything that gets researchers talking in better-coordinated, more practical terms should help research managers explain their programs to funders. Exposing the logic of a program to public scrutiny might encourage charges of funding imbalances. However, it also provides an orderly way to address them. Differential funding rates can be appropriate in a value-of-information sense. A topic is justifiably neglected if reducing the associated uncertainty would make little difference or if there are few cost-effective ideas for doing so (Davies, 1996). Integrated assessment provides a structure for making (and evaluating) such arguments, as well as for more quantitative analyses, should they be warranted.

Integrated assessments will, naturally, look somewhat different for problem-oriented and discipline-oriented research programs. The former will typically work backward from an objective (e.g., more oil, shortened hospital stays) toward more fundamental topics. The latter will work forward from core interests and competencies toward practical implications of possible research results.

A research program's structure should be more accessible to researchers when laid out in integrated-assessment terms compared to the topic list of a conventional program announcement. They might find that their work is not as pertinent as they had imagined, when they thought about being 'useful' in general terms. They might also find ways to make their research more relevant, without violating its overall mission.

Looking across the integrated assessments for different problems may reveal unproductive duplication, unrealized complementarity, repeatedly neglected opportunities, and overemphasized topics. For example, the information yield from clinical trials declines when participants drop out and when they stay in, but fail to follow the protocol (i.e., take their medication, avoid other drugs, keep experience logs). Seeing this problem recur in multiple contexts might prompt a research initiative on such behavior. Such basic knowledge might be so valuable that the sponsors of clinical trials might voluntarily 'tithe' for its creation. If they needed further convincing, one could 'run the numbers' to calculate the return on that investment.

The US NSF's current Human Capital Initiative reflects a claim by social scientists to have a portfolio of projects that could significantly advance the national economy, by improving its workers' productivity. Generally speaking, these projects aim to increase mastery of the intellectual skills involved in extracting information. Formalizing this claim requires practitioners to be explicit about their needs and researchers to be explicit about their capabilities. Completing the analysis might prompt research on impacts (e.g., how effective are training programs or improved computer displays?), on bridges (e.g., how can cognitive research be made more useful to industrial trainers?) and on the assessment process (e.g., do researchers over- or underestimate the yield from their work?).

A work plan

People don't, won't, and probably shouldn't change how they think and work overnight. Given the importance of maintaining cognitive control over one's affairs, changing too fast can be dangerous. The purpose of this article has been to make the case that some change is needed – and that some slow progress is possible, using integrated assessment to implement a value-of-information perspective. Given the complexity of the issues, developing methods and cultivating understanding in the scientific community would have to proceed in tandem. The first steps in that direction might include.

Worked examples

Because people master principles through examples, well-developed case studies are a priority. Ideally, they would address problems of sufficient complexity to test the method and include roles for multiple sciences. Oil prospecting is one (obvious) candidate. Global change research is another. Much of its funding has been generated by promises of helping to solve practical problems. It should benefit from analyses clarifying the rationale for its budget allocations and demonstrating the connections between research and reality (National Research Council, 1999). Its researchers are pioneers in the development of integrated assessment, including research on how to assess the scientific uncertainty central to quantitative analyses.

FAQs

The very mention of an unfamiliar procedure will evoke preconceptions that shape people's thinking about it. If mistaken, these beliefs can interfere with mastering its details and even with willingness to try to do so. The most frequently asked questions might be collected and addressed directly in explanatory material (after being tested for comprehensibility). Likely questions include: (a) Doesn't the approach place too high a premium on economic outcomes? (b) How can you place a value on research that creates delight in the beauty of the universe? (c) How do you value graduate training for students who do and do not stay in the basic science for which they were trained? (d) How can you compare clinical trials with basic bioscience?[2]

Basic research

Adopting a scientific approach to the management of science may require an investment in that science. Some will be spent in the conduct of analyses per se. Some will be spent in helping individual sciences develop relevant models. Done correctly, this effort should not feel like an odious exercise in mechanistic accountability, threatening to contort science for measurement sake. Rather, it should provide information that scientists themselves would like to know (e.g., how well can we estimate the yield from our experiments? How do we compare the value of center and single-investigator grants?). Thus, there should places for both *basic applied* research (addressing fundamental questions arising from applications) and *applied basic* research (making use of what we know already) (Baddeley, 1979).

Initial procedures

Finally, a plan is needed for gradual introduction and testing. There should be time to learn – and to complain when the approach is not (yet) working. Scientists willing to try an analytical approach should not be penalized for exposing themselves to such scrutiny. The process itself should be evaluated critically, held to the standards that it attempts to implement. Its introduction could begin by shadowing conventional allocation practices, supplanting them over time as the approach takes shape and achieves credibility. Integrated assessments are scalable in the sense that simple qualitative models can evolve into more complex and quantitative ones. Properly executed (and explained), the approach should be seen as facilitating intellectual activities that scientists are attempting to accomplish anyway.

Acknowledgements

Preparation of this manuscript was supported by the Environmental Protection Agency, National Science Foundation, and National Institute for Allergies and Infectious Disease. This is gratefully acknowledged, as are the thoughtful comments of Barbara Torrey, David Vislosky, and two anonymous reviewers. The views expressed are those of the author.

Notes

1. Fischhoff, Downs and Bruine de Bruin (1998) offer such an encompassing perspective for the domain of infectious disease, focused on those transmitted through sexual contact.
2. Brief answers (in order): (a) One can evaluate contributions in terms of any outcome that is given explicit expression. (b) If delight is not an end in itself (hence entitled to resources dedicated to purely noninstrumental science), it can still be evaluated as a contributor to recruiting talented, committed individuals to science, increasing the productivity of their disciplines. (c) Training improves one's ability to extract information from relevant situations. (d) If the research is justified on practical grounds, then the potential broad impacts of basic bioscience need to be sketched in terms comparable to those of focused clinical trials.

References

Baddeley, A. (1979). 'Applied cognitive and cognitive applied research,' in L. G. Nilsson, ed., *Perspectives on Memory Research*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Casman, E., B. Fischhoff, C. Palmgren, M. Small and F. Wu (1999). *Integrated Assessment of* Cryptosporidium *risks*. Manuscript under editorial review.

Clemen, R. (1991). *Making Hard Decisions: An Introduction to Decision Analysis*. Boston, MA: PWS-Kent.

Davies, C., ed. (1996). *Comparing Environmental Risks*. Washington, DC: Resources for the Future.

Dowlatabadi, H. and M. G. Morgan (1993). 'Integrated assessment of climate change,' *Science* 259 (5103): 1813.

Dowlatabadi, H. and M. G. Morgan (in press). *Integrated Assessment*. New York: Cambridge University Press.

Einhorn, H. J. (1982). 'Learning from experience and suboptimal rules in decision making,' in D. Kahneman, P. Slovic and A. Tversky, eds., *Judgment under Uncertainty: Heuristics and Biases*. New York: Cambridge University Press, pp. 268–283.

Fischhoff, B. (1977). 'Cost-benefit analysis and the art of motorcycle maintenance,' *Policy Sciences* 8: 177–202.

Fischhoff, B. (1980). 'Clinical decision analysis,' *Operations Research* 28: 28–43.

Fischhoff, B. (1989). 'Eliciting knowledge for analytical representation,' *IEEE Transactions on Systems, Man and Cybernetics* 13: 448–461.

Fischhoff, B., J. Downs and W. Bruine de Bruin (1998). 'Adolescent vulnerability: A framework for behavioral interventions,' *Applied and Preventive Psychology* 7: 77–94.

Galvin, R. (1998). Editorial. *Science* 280: 803.

Grayson, J. R. (1960). *Decisions under Uncertainty*. Cambridge, MA: Harvard School of Business Administration.

Green, D. M. and J. A. Swets (1966). *Signal Detection Theory and Psychophysics*. New York: Wiley.

House Committee on Science (1998, September 24). *Unlocking Our Future: Toward a New Science Policy*. Washington, DC: Author.

Howard, R. A. (1989). 'Knowledge maps,' *Management Science* 35: 903–922.

Jones, S., B. Fischhoff and D. Lach (1998). 'An integrated impact assessment for the effects of climate change on the Pacific Northwest salmon fishery. *Impact Assessment and Project Appraisal* 16, 227–237.

Jones, S., B. Fischhoff and D. Lach (in press). 'Evaluating the usefulness of climate-change research for policy decisions,' *Climate Change*.

Kahneman, D., P. Slovic and A. Tversky, eds. (1982). *Judgment under Uncertainty: Heuristics and Biases*. New York: Cambridge University Press.

Kammen, D. and D. Hassendzahl (1999). *Shall We Risk It*? Princeton, NJ: Princeton University Press.

Keeney, R. L., J. S. Hammond and H. Raiffa (1998). *Smart Choices: A Practical Guide to Making Better Decisions*. Cambridge, MA: Harvard Business School.

Morgan, M. G. and M. Henrion (1990). *Uncertainty*. New York: Cambridge University Press.

Morgan, M. G. and D. W. Keith (1995). 'Subjective judgments by climate experts,' *Environmental Science and Technology*, 29: 468A–476A.

Murphy, A. H. and B. G. Brown (1984). 'A comparative evaluation of objective and subjective weather forecasts in the United States.' *Journal of Forecasting* 3, 369–393.

National Research Council (1995a). *Allocating Federal Funds for Science and Technology*. Washington, DC: National Academy Press.

National Research Council (1995b). *AIDS and Behavior*. Washington, DC: National Academy Press.

National Research Council (1998a). *Investing in Research Infrastructure for the Behavioral and Social Sciences*. Washington, DC: National Academy Press.

National Research Council (1998b). *Scientific Opportunity and Public Input: Priority Setting for NIH*. Washington, DC: National Academy Press.

National Research Council (1999). *Making Climate Forecasts Matter*. Washington, DC: National Academy Press.

Platt, J. R. (1964). 'Strong inference,' *Science*, 146: 347–353.

Raiffa, H. (1968). *Decision Analysis: Introductory Lectures on Choices under Uncertainty*. Reading, MA: Addison-Wesley.

Schneider, S. H. (1997). 'Integrated assessment modeling of global climate change: Transparent rational tool for policy making or opaque screen hiding value-laden assumptions?' *Environmental Modeling and Assessment* 2: 229–249.

Tengs, T. O. (1998). *Planning for Serendipity: A New Strategy to Prosper from Health Research*. Washington, DC: Progressive Policy Institute.

Yates, J. F. (1990). *Judgment and Decision Making*. New York: Wiley.