## Letters

### Scientific Integrity

The Policy Forum "Government and quality in science" by Bernard D. Davis (10 Nov., p. 736) expresses concern that the newly established offices in the Public Health Service (The Office of Scientific Integrity and the Office of Scientific Integrity Review) will "become involved in increasingly detailed management of the practice of science." In Davis' view, such a concern arises from the stated role of these offices to promote high standards of scientific conduct, which he interprets to mean that the offices will be dictating on matters of scientific judgment or quality, rather than limiting their activities to scientific misconduct.

The promotion of responsible scientific conduct is a responsibility shared by the scientific community at large, grantee and applicant institutions, professional and academic associations, and all Public Health Service (PHS) components supporting research. It is entirely appropriate that the PHS offices will play a catalytic role in fostering the development of standards for research conduct. A successful collaboration between the federal and the scientific-academic communities in developing such standards is the best protection against regulatory or legislative remedies.

There are no immediate plans to implement the Institute of Medicine proposal for requiring institutions receiving PHS research grants to have policies and procedures to encourage responsible research practices (*1*). It should be noted however that the recently issued "Final Rule" (*2*) requiring institutions to have policies and procedures for inquiring into and investigating scientific misconduct concludes with a statement that institutions "shall foster a research environment that discourages misconduct. . . ." It is the response of institutions that will demonstrate whether there is a need for more formalized requirements for prevention and education activities.

The existing peer-review process is the forum for judgments about the quality of research. However, it is important for the Office of Science Integrity and the Office of Scientific Integrity Review, in collaboration with the scientific community, to do a better job of spelling out what is unacceptable scientific behavior. The limitation of the definition of scientific misconduct to only falsification and plagiarism, as proposed by Davis, would miss a range of unacceptable behaviors that have already been judged by scientific investigative panels to constitute misconduct. Standards for the responsible conduct of science should include the clearest possible statements of what is unacceptable behavior, which requires a further elaboration, not limitation, of the definition of scientific misconduct. A refined definition of scientific misconduct would not "casually fold in questions of quality or of error," as Davis fears, but would in fact serve to more clearly separate differences in scientific judgment or honest error from misconduct.

The establishment of the Office of Scientific Integrity Review within the Office of the Assistant Secretary for Health provides a vital PHS-wide oversight role for scientific integrity activities and indicates the importance placed by the Department of Health and Human Services on dealing with scientific misconduct. We fully intend to continue working with the scientific and institutional communities in discussing such im-

portant issues as where responsibilities are properly vested for promoting the responsible conduct of research and the proper form of standards and guidelines to foster integrity in science.

JAMES O. MASON
*Assistant Secretary for Health* and
*Acting Surgeon General,*
*Department of Health and Human Services,*
*Public Health Service, Washington, DC 20201*
LYLE W. BIVENS
*Acting Director,*
*Office of Scientific Integrity Review,*
*Office of Assistant Secretary for Health,*
*Department of Health and Human Services,*
*Rockville, MD 20852*

REFERENCES

1. *The Responsible Conduct of Research in the Health Sciences* (National Academy Press, Washington, DC, 1989), p. 25.
2. "Responsibility of awardee and applicant institutions for dealing with and reporting possible misconduct in science," *Fed. Reg.* **54** (No. 151), 32446 (8 August 1989).

*Response*: I am delighted by the comments of Mason and Bivens reassuring us that the new federal offices concerned with scientific integrity do not plan at this time to require formal institutional efforts to encourage responsible research practices. Nevertheless, the authors do not accept the proposal, in my article, that the government should draw a sharp line between fraud (that is, fabrication and falsification of data) and other undesirable practices.

When the rather open-ended word "misconduct" began to replace "fraud" and "plagiarism," President Howard Schachman of the Federation of American Societies for Experimental Biology, and its Public Affairs Committee (of which I was a member), opposed the shift. We obviously lost the battle, and my Policy Forum did not aim at trying to renew it. But the letter from Mason and Bivens illustrates the problem that the shift created: the government is still seeking the precise definition that the law needs. The search is difficult, because the term merges into questions of judgment and quality. For example, the present definition includes "practices that deviate seriously from those that are commonly accepted within the scientific community"—a definition that seemed to our committee far too loose.

The comments by Mason and Bivens further understate the danger of excessive governmental involvement by describing the role of the new offices as promoting "high standards of scientific conduct"—a phrase that would seem to contrast proper conduct with misconduct. But the charter for these offices assigns them a rather different responsibility: promoting "high standards of laboratory and clinical investigations in science through a prevention and education program." This phrase, which clearly gets beyond misconduct into the area of quality, was the main cause of my concern, and it still is. While it is gratifying that the current officials in charge evidently have no intention to delve into this area, a later official might feel obligated to follow the letter of the law. This charge to the offices therefore deserves reevaluation.

We are dealing here with a gray area—and the lighter the shade of gray, the more difficult it is for the government, however laudable its intention to serve as a catalyst, to avoid imposing a rigidity that would do more harm than good. I certainly agree with Mason and Bivens that government as well as scientists and their organizations share responsibility for promoting responsible conduct; but it does not follow that all these groups share the whole range of responsibilities implied by this broad term. Mason and

Bivens recognize that the scientific community has a role in determining where responsibilities are properly vested. We will no doubt need continual discussion to ensure that the proper lines are maintained. Because of the intensely personal nature of scientific research, and because students learn standards from the behavior of their preceptors and colleagues, just as children do from their parents and their other contacts, the discussion will always face an ancient and general problem: where should the law end and personal morality begin in setting standards of conduct?

In justifying their position, Mason and Bivens note that "scientific investigative panels" have judged misconduct to include a range of unacceptable behaviors beyond falsification and plagiarism. To my knowledge, the most prominent support for this view (and hence the main focus of my article) was the report of the Institute of Medicine (IOM). I would therefore emphasize that the many researchers with whom I have discussed the matter uniformly disagree with the IOM recommendations. Even though such issues as carelessness, bad judgment, and improper distribution of credit are perpetual problems in science, few scientists seem to believe it would be helpful for government to try to prevent them.

BERNARD D. DAVIS
*Bacterial Physiology Unit,*
*Harvard Medical School, Boston, MA 02115*

## Clinical and Actuarial Judgment

In their article, "Clinical versus actuarial judgment" (31 Mar., p. 1668), Robyn M. Dawes *et al.* address an important issue. But it is an issue that now extends well beyond psychiatric and clinical prediction. True, Meehl's landmark book (*1*) limited itself to clinical psychology, as does much of the article by Dawes *et al.*, but the question of whether to use the "head" (clinical intuition) versus the "formula" (actuarial or mechanical information combination)—to borrow Meehl's apt terms (*2*)—is equally relevant for medicine (*3*), engineering (*4*), auditing (*5*), management (*6*), polygraphy (*7*), and, as Newell and Simon (*8*) clearly show, for most decisions and choices made in ill-structured problem domains. Moreover, the dilemma they pose of using either the head *or* the formula is no longer the main focus of contemporary decision research. Rather, the focus has long ago shifted to evaluating the use of *both* modes of information combination in tandem.

This trend of combining judgmental with formal modes of information processing probably started in psychology with the suggestions of Edwards (*9*) and of Sawyer (*10*) that experts contribute to predictions by providing intuitions about appropriate judgmental quantities that are best aggregated mechanically. Such mechanical procedures are in constant need of judgmental monitoring (*11*). Dawes (*12*) has made important contributions to this literature, as have many others (*13*), including Meehl (*14*), whose observation, in this regard, is worth repeating (*14*, pp. 372–373).

> 95% of the ordinary decisions made by working practitioners . . . [in mental health settings] . . . are not comparable in richness and subtlety to that of a good psychoanalytic hour . . . [but] . . . when you check out at the supermarket, you don't eyeball the heap of purchases and say to the clerk, "well it looks to me as if it's about $17.00 worth; what do you think?" The clerk adds it up.

It seems, then, that Dawes as well as Meehl advocates the less divisive (than the title suggests) strategy of using the head *and* the formula, depending on whether the decision problem lends itself more readily to intuitive judgment or to mechanical combination. Faust, too, does not appear to have given up entirely on clinical intuition. Otherwise, why would he have provided a set of cognitive correctives in a recent article on human jugement (*15*)? The correctives were designed to help "clinicians to better serve their clients" (*15*, pp. 426–428).

These polemics aside, it is essential to note that the idea that began with the mechanical aggregation of judgmental inputs has been followed up by contemporary decision analysis, a technology that facilitates decisions that will outperform either a purely clinical or a purely actuarial mode. Decision analysis, a variant of Bayesian thinking, is a formal technique that incorporates Bayes' theorem, but adds three essential components (*16, 17*). Stated here as questions, these are (i) In my judgment, can this decision problem be decomposed into simpler segments? (ii) What are the consequences of alternative actions of the decision? and (iii) What are the uncertainties in the environment relevant to the actions and their consequences? By focusing on the resolution of these questions by means of a technique that uses both the head and the formula, contemporary decision analysis, which has been applied in a large variety of domains (*3, 16, 17*), avoids favoring either extreme of the clinical-actuarial dichotomy. It does so by blending formal logic with intuitive insight (*18*). This blend, it has been argued (*17*), yields better results than the use alone of either the head *or* the formula.

BENJAMIN KLEINMUNTZ
*Department of Psychology,*
*University of Illinois, Chicago, IL 60680*

REFERENCES AND NOTES

1. P. E. Meehl, *Clinical Versus Statistical Prediction* (Univ. of Minnesota Press, Minneapolis, MN, 1954).
2. ———, *J. Couns. Psychol.* **4**, 268 (1954).
3. A. S. Elstein *et al.*, *Am. J. Med.* **80**, 246 (1986).
4. K. R. Hammond *et al.*, *IEEE Trans.* SMC-17 (1987).
5. R. H. Ashton *et al.*, in *Research Opportunities in Auditing*, L. R. Abde-khalik and I. Solomon, Eds. (American Accounting Association, Auditing Section, Sarasota, FL, 1989), pp. 95–132.
6. D. N. Kleinmuntz, *Manage. Sci.* **31**, 680 (1987).
7. B. Kleinmuntz, *Nature* **308**, 349 (1984).
8. A. Newell and H. A. Simon, *Human Problem Solving* (Prentice-Hall, Englewood Cliffs, NJ, 1972); H. A. Simon, *Ann. Rev. Psychol.* **30**, 363 (1979).
9. W. Edwards, *Hum. Fact.* **4**, 59 (1962).
10. J. Sawyer, *Psychol. Bull.* **66**, 178 (1966).
11. L. Savage, *The Foundations of Statistics* (Wiley, New York, 1954); W. Edwards, in *Computer Diagnosis and Diagnostic Methods*, J. A. Jacquez, Ed. (Thomas, Springfield, IL, 1972); pp. 139–151.
12. R. M. Dawes, *Am. Psychol.* **26**, 180 (1971).
13. H. J. Einhorn, *Organ Behav. Hum. Perf.* **7**, 86 (1972); R. M. Hogarth, *Psychol. Bull.* **90**, 197 (1981); B. Kleinmuntz, *Science* **139**, 416 (1963); C. Camerer, *Organ Behav. Hum. Perf.* **27**, 411 (1981).
14. P. E. Meehl, *J. Pers. Assess.* **50**, 370 (1986).
15. D. Faust, *Prof. Psychol. Res. Pract.* **17**, 420 (1986).
16. R. A. Howard, in *Proceedings of the Fourth International Conference on Operations Research*, S. B. Hertz and J. Melese, Eds. (Wiley, New York, 1966), pp. 555–571; *Manage. Sci.* **34**, 679 (1988); R. L. Keeney, *Oper. Res.* **30**, 803 (1982); H. Raiffa, *Decision Analysis: Introductory Lectures on Choices Under Uncertainty* (Addision-Wesley, Reading, MA, 1968).
17. D. von Winterfeldt and W. Edwards, *Decision Analysis and Behavorial Research* (Cambridge Univ. Press, New York, 1986).
18. D. N. Kleinmuntz, in *Insights in Decision Making: A Tribute to Hillel J. Einhorn*, R. M. Hogarth, Ed. (Univ. of Chicago Press, Chicago, IL, in press).
19. Preparation of this comment was partially supported by National Library of Medicine grant No. 1RO1 LM04583-03.

*Response*: Kleinmuntz correctly states that the clinical-actuarial issue extends beyond the domain we covered; he has himself made distinguished contributions to this literature. His comments could, however, create erroneous impressions about research outcomes and our views in the domain on which we focused—the diagnosis and prediction of human conditions and behavior.

To restate the problem, if one assumes the option of using the clinical, actuarial, or clinical-actuarial approach (to which Kleinmuntz refers, respectively, as the head, the formula, or the two in combination), which judgment strategy leads to the most accurate diagnoses or predictions of human conditions and behavior? The literature shows, overwhelmingly, that the accuracy of the actuarial method equals or exceeds that of the clinical method. The limited research comparing the actuarial and clinical-actuarial approaches also favors the former strategy. Generalization or lack of generalization to other problem realms does not change the evidence in the domain of human outcomes. This large and consistent body of scientific evidence is so important precisely because the intuition that the research would or should turn out otherwise is so compelling.