Letters

Scientific Evidence

Although I agree with Bert Black's premise in his article "Evolving legal standards for the admissibility of scientific evidence" (25 Mar., p. 1508), I feel that his definition of the problem is incomplete. It causes his analysis to be skewed and ultimately leads him to a flawed conclusion. Specifically, Black postulates that the reason our recent judicial decisions in the toxic tort field are "wholly out of keeping with accepted scientific knowledge" is because "[t]he law looks to science for answers to factual questions that lie beyond the understanding and knowledge of *non*scientists" (emphasis added). While such a statement, standing alone, is accurate, Black does not consider the more important fact that the scientific "answers" to most toxic tort problems lie beyond the understanding and knowledge of scientists as well as nonscientists. More simply, the nonscientists' primary problem is not in their miscomprehension of established science, but in their failure to understand the limitations of what science can prove. Alvin Weinberg recognized this fundamental problem more than 17 years ago (*1*).

> The point missed . . . is that the seemingly simple question 'What is the effect on human health of very low levels of physical insult?' can be stated in scientific terms; it can, so to speak, be asked of science, yet it cannot be answered by science. I have . . . proposed the name *trans-scientific* for such questions that seemingly are part of science yet in fact transcend science.

Thus, when judges do rule rigidly as Black advocates, they exclude much of the relevant research and chip away many of the larger pieces in the overall scientific puzzle, leaving the jury with virtually no evidence to consider. While Black appears to appreciate that the courts' "refusal to judge an expert's opinions according to the criteria of [that expert's] profession can lead to results that clearly conflict with accepted scientific knowledge," he nevertheless commends judges who take science into their own hands. Specifically, Black cites the Agent Orange case as exemplifying correct judicial declarations of what scientific knowledge is. He applauds New York Federal Judge Jack Weinstein's exclusion of all toxicity studies done on animals since "[t]here was no evidence that the plaintiffs had been exposed to the high animal study concentrations, and in any event the differences between humans and other species meant the studies were more likely to mislead than to help the jury [and thus their prejudicial effect outweighed their relevancy]."

As our experiences with asbestos, the Dalkon Shield, diethylstilbesterol, and the swine flu vaccination have demonstrated, the court's insistence on quantitative, definitive, scientific answers for complex causation questions not only neglects the overwhelming experimental evidence, but necessitates that a statistically significant number of persons be exposed to the hazard, and die or be injured as a result of their exposure. A successful epidemiological study must then be published documenting these deaths and injuries. Rather than encourage a nonscientist judge to exclude scientific experimentation simply because the experiments cannot provide definitive answers, then, the standards for admitting evidence should be loosened in accordance with the limits and practices of science. The jury should be the ultimate arbiter, considering and weighing all evidence relevant to the causation question, including inherently less definitive trans-scientific research.

WENDY E. WAGNER
Lands and Natural Resources Division,
Department of Justice,
110 Pennsylvania Avenue, NW,
Washington, DC 20530

REFERENCES

1. A. M. Weinberg, *Science* **174**, 546 (1971).

Black argues that the judicial system frequently misuses science. He believes judges often have not examined the underlying methodology to verify that expert testimony truly represented expert consensus. His solution is to have judges "delve into the reasoning behind an expert's conclusions and require that this reasoning reflect accepted scientific practice." Black's objective is for "verdicts consistent with scientific reality."

While agreeing with Black's basic concern, I believe his solution is likely to exacerbate the problem. Judges are not trained in science, and the cross-examination advocacy system is a poor process with which to establish scientific conclusions. Black illustrates the inherent difficulties a scientist faces in the judicial system by his example of the meaning of a 90% chance. As judges conclude they must delve into science, they become activists in determining the meaning of the science (*1*). This can be seen in an article (*2*) by a judge of the U.S. Court of Appeals for the District of Columbia Circuit (where many appeals of federal agency decisions are taken).

> [T]he judicial responsibility is . . . to ensure that the agency's decision-making is thorough and within the bounds of reason. The agency's decisional record must disclose . . . the agency's precise reasons for resolving conflicts in the evidence. This includes the basis for selecting one scientific point of view rather than another.

A better solution might be for professional societies to establish a "friend of the court" system. The AAAS, the American Physical Society, the Institute of Electrical and Electronics Engineers, and so forth could set up panels that would use consensus agreement to address issues raised by courts. Then when a technical issue arose in a court, a representative of the "friend of the court" panel from the pertinent society could provide the consensus technical judgment. The witness and the associated panel would be paid for by the court (*3*). This approach would lay the responsibility for presenting science credibly directly on the scientists, where it belongs.

JOHN F. AHEARNE
Resources for the Future,
1616 P Street, NW,
Washington, DC 20036

REFERENCES AND NOTES

1. J. F. Ahearne, *Prog. Nucl. Energy* 7, 77 (1981).
2. D. L. Bazelon, *Science* **208**, 661 (1980).
3. Individuals would not be precluded from appearing as witnesses for either side, but should be labeled "advocacy" rather than "expert" witnesses.

Response: In my article, I argued that courts should more actively examine the way in which scientists reach the conclusions about which they testify. The comments of Wagner and Ahearne essentially bracket my position. Wagner would abandon scientific standards, while Ahearne would have scientific professional societies become more involved in the judicial process.

Ahearne's suggestion relates to the crucial question of where courts should turn for advice when faced with disputed scientific testimony, and I do not think that it is necessarily inconsistent with my article. His discussion, however, focuses more on disputes about regulations than on tort litigation. Because the existing Federal Rules of Evidence empower a court to appoint its own expert (*1*), judges could to a large extent already take the Ahearne approach in appropriate tort litigation cases.

In contrast to Ahearne, Wagner advances an argument that is logically inconsistent and fundamentally wrong. She wants juries to answer questions admittedly beyond the understanding of either scientists or nonscientists. On the issue of causation in a toxic tort case, the basic question is whether or not the exposure for which the defendant is responsible more likely than not caused the plaintiff's disease or injury (*2*, pp. 764–766); and if the plaintiff cannot introduce admissible evidence sufficient to answer this question, he or she should lose as a matter of law. Juries should not be asked to do the impossible. It defies logic to argue, as Wagner does, that lay jurors can somehow answer the unanswerable.

Indeed, one is hard pressed to imagine what would constitute "trans-scientific" evidence and what a jury would do with it. I can think of three scenarios. First, an expert witness might testify, on the basis of unproved and rejected theories like clinical ecology (*3*, pp. 689–691), that the exposure at issue did in fact more likely than not cause the disease at issue. Because such testimony is patently out of keeping with accepted scientific practice, it could only mislead and confuse a jury, and it should not be admitted. Second, a witness might present data or research results, along with speculative conclusions about ill-defined and completely unquantified potential risks associated with the exposure; but if the expert cannot reach a more definitive conclusion, nonexpert jurors can hardly be expected to do better. Finally, an expert might validly conclude that there is some chance, but less than 50%, that the exposure caused the disease. The jury would then have to reinterpret this testimony to find that it somehow satisfies the legal "more likely than not" requirement; but again, if the expert cannot reach this conclusion, there is no rational reason to think that a jury could.

The examples cited by Wagner bring no clarity to her argument. They demonstrate no need for relaxing the rules of evidence. Plaintiffs regularly win asbestos, Dalkon Shield, and diethylstilbesterol cases, and regularly won swine flu vaccine cases, using science, not trans-science. Wagner laments the fact that people were injured before the problems with these products were discovered, but her solution would convert virtually any vague "trans-scientific" speculation into the basis for a lawsuit. Her solution would also radically change tort law by making companies pay for countless illnesses and injuries that would have occurred even if the companies had never made the accused products.

Wagner would solve a largely nonexistent problem with a totally unwarranted reshaping of the law. She may be impatient with the limitations of science, but impatience does not justify the relaxation of both scientific and legal standards. Trans-science is not science, and it has no place in the law.

BERT BLACK
Venable, Baetjer and Howard,
1800 Mercantile Bank & Trust Building,
2 Hopkins Plaza, Baltimore, MD 21201-2978

REFERENCES AND NOTES

1. Rule 706 of the Federal Rules of Evidence allows a court to appoint an expert witness of its own selection and to have the parties pay the cost.

2. B. Black and D. E. Lilienfeld, *Fordham Law Rev*. **52**, 732 (1984).
3. B. Black, *ibid*. **56**, 595 (1988).

Mathematics Achievement

Colin Norman's article "Math education: A mixed picture" (News & Comment, 22 July, p. 408) gives a misleading impression of both the source of the deficits in mathematics achievement of black and Hispanic students and the educational policies required to correct the deficits. When grade equivalent units are used to report scores on achievement tests, one's expectation is that groups with lower means in an early grade will fall farther and farther behind with progression up the grades. This expectation derives, however, from the nature of the grade equivalent units of measurement for which variances in racially homogeneous populations increase markedly with grade (and age) for all kinds of academic achievement. Variances of measures of physique also increase during the grade school years and reflect primarily normal physical growth, not nutritional differences.

To determine whether a minority group is falling farther behind majority whites during the school years, it is more defensible to use a standard score scale. This scale reveals whether the relative position of an individual or subgroup in the total sample changes as grade or age increases. In standard score units academic or intellectual deficits that appear early, including those appearing in preschool, tend to remain constant. In a particular instance, if standard score means do draw apart, one can be reasonably certain that there is a real "falling behind" phenomenon.

In spite of an initial disclaimer to the contrary, Norman also discusses correlations of students' mathematics achievement with attitudes of parents, teachers, counselors, and the students themselves as if those correlations represented causal relations. Describing the correlations of parental attitudes and encouragement with children's performance in mathematics as the former having "the strongest influence" or the correlation of the student's own liking for the subject with performance as the former being an "important factor" should be avoided. Educational changes based on inferences concerning causation from mere correlations are likely to be fruitless.

LLOYD G. HUMPHREYS
Department of Psychology,
University of Illinois at
Urbana-Champaign,
Champaign, IL 61820

Norman reports the finding of the National Science Foundation that "the disparity [between boys and girls on standardized mathematics tests in high school may be] unique to the SAT [Scholastic Aptitude Test, Mathematical] itself." Our data indicate that this is not correct. Although girls tend to achieve at least as well as boys on in-school tests, even in most of the subjects where they are getting better grades than boys they tend to average less on nationally standardized tests. Some mean difference results favoring males, in standard deviation units, for large numbers of cases are: Preliminary Scholastic Aptitude Test, Mathematical, 0.37; American College Testing Program Mathematics, 0.34; College Board Achievement Test, Mathematics Level I, 0.39, Level II, 0.38, and Physics, 0.59; Advanced Placement Program, Calculus Level AB, 0.20, Calculus Level BC, 0.18, and Computer Science, 0.50; quantitative score of Medical College Admissions Test, 0.37; quantitative score of Graduate Management Admissions Test, 0.43; Graduate Record Examination, Quantitative, 0.67, advanced test in mathematics, 0.71, and advanced test in political science, 0.76 (*1*).

Whereas most of these mean differences are not huge, there can be strong effects when applicants are selected partly because of their high test scores. For example, the ratio of males to females taking the computer science test is about 5.7 to 1, and even among the examinees twice as large a percentage of males as of females score 4 or 5 on the 5-point scale. For the European History test of the College Board Achievement Tests (effect size 0.63 in 1985) the examinee ratio, favoring males, for scores of 700 or more was 5.2.

We now know a great deal about the existence and magnitude of differences between males and females on at least 86 nationally standardized cognitive tests, but little about why they occur and how to lessen them. Current data provide promising leads for research about the whys and hows.

JULIAN C. STANLEY
Study of Mathematically Precocious Youth,
Johns Hopkins University,
Baltimore, MD 21218

REFERENCES

1. J. C. Stanley, "Sex differences on cognitive tests: A summary," paper presented at the annual meeting of the American Educational Research Association, Washington, DC, 24 April 1987; C. P. Benbow and J. C. Stanley, *Am. Educ. Res. J.* **19**, 598 (1982); *Science* **222**, 1029 (1983).

Erratum: In Leslie Roberts' article "New targets for human gene therapy" (19 Aug., p. 906), the collaborators from Tufts–New England Medical Center were unintentionally omitted. They are David E. Johnston and Douglas M. Jefferson.