

- The Plain Dealer* (Cleveland) editorial: "A Sad Case of Scientific Hubris" (27 March 1991)
- New York Times* (among others, after the OSI Draft Report): "Crucial Research Data in Report Biologist Signed Are Held Fake" (21 March 1991)
- "Hero in Exposing Science Hoax Paid Dearly" (22 March 1991)
- "Upholding a Challenge to the Status Quo" (24 March 1991)
- "A Scientific Watergate" (Editorial, 26 March 1991)
- "How Charges of Lab Fraud Grew Into a Cause Célèbre" (26 March 1991)
- "Can the Scales of Justice be Calibrated for Scientific Fraud?" (31 March 1991)
- "Rockefeller U. Is Hurt as President Feels Pain of Science-Fraud Case" (1 April 1991)
- TIME, "Science Under Siege: Crisis in the Labs" (26 August 1991)
- U.S. News and World Report: "The Best of America; Setting Wrongs Right—Congress's Most Feared Democrat" (26 August—2 September 1991)
- New Republic*: "THE SCIENCE MOB—Fraud, Complacency, and Secrecy in the Scientific Establishment", by Philip J. Hiltz (18 May 1992)
- Newsweek*: "A Plague of Science Fraud" by Robert Bell (15 June 1992)

On the other hand, throughout the Baltimore affair, the *Wall Street Journal* was systematically against Dingell and pro-Baltimore. See for instance:

- "Science in Stripes" ("REVIEW AND OUTLOOK"), 18 April 1988 (p. 24). The gist of the article is: "There is no free lunch. Becoming so thoroughly dependent on federal funding leads unavoidably to conflicts with the funders, who are politicians and who behave like politicians. Up to now the research community has been willing to try to paper over fundamental problems created through the funding process by accepting more oversight, more regulations, more review boards. The result over time is likely to be less significant science."
- "*Latest Chapter in the Fine Science Of the Smear*" by Paul Gigot, 5 May 1989
- "Fraud vs. error: The Dingelling of Science", by Bernard Davis, 8 May 1989
- "The Science Police" ("REVIEW AND OUTLOOK"), 15 May 1989
- "Politics and Science" (Editorial), 29 March 1991
- "Dingell Gets Baltimore" (Editorial), 4 December 1991

## On Being a Whistleblower: The Needleman Case

Claire B. Ernhart

*Case Western Reserve University*

Sandra Scarr  
*University of Virginia*

David F. Geneson  
*Hunton & Williams*

We believe that members of the scientific community have a primary obligation to promote integrity in research and that this obligation includes a duty to report observations that suggest misconduct to agencies that are empowered to examine and evaluate such evidence. Consonant with this responsibility, we became whistleblowers in the case of Herbert Needleman. His 1979 study (Needleman et al., 1979), on the effects of low-level lead exposure on children, is widely cited and highly influential in the formulation of public policy on lead. The opportunity we had to examine subject selection and data analyses from this study was prematurely halted by efforts to prevent disclosure of our observations. Nevertheless, what we saw left us with serious concerns. We hope that the events here summarized will contribute to revisions of process by which allegations of scientific misconduct are handled and that such revisions will result in less damage to scientists who speak out.

Key words: misconduct, misrepresentation, due process, whistleblower

The increasing incidence of reports of scientific misconduct is troublesome to the public as well as to the scientific community. Even a few instances of misconduct can erode public trust in science. If we are to sustain the respect of the public when there are allegations of misdeeds, procedures for evaluating evidence must be credible (Teich & Frankel, 1992). The widely publicized

Requests for reprints should be sent to Claire B. Ernhart, Department of Psychiatry, MetroHealth Medical Center, 2500 MetroHealth Drive, Cleveland, OH 44109-1998.

difficulties arising from investigations into the cases of Robert Gallo (the AIDS virus) and Therza Imanishi-Kari (the David Baltimore case) highlight the importance of the problem (Hoke, 1992).

Policies of the National Institutes of Health (NIH<sup>1</sup>) and the National Science Foundation for the resolution of misconduct charges are presently being questioned on several fronts. It has been recommended that full legal hearings be available for researchers who request them (American Psychological Association, 1992; Hamilton, 1992). Recommendations also include a change in the definition of misconduct, which is discussed later. The proposals do not discuss problems that arise in the quasi-judicial investigations conducted at the university level. This portion of the process also merits attention so that the rights and welfare of those involved can be protected. We are particularly concerned about the protection of those who file legitimate complaints about possible misconduct (i.e., the whistleblowers). The position of professors who serve on inquiry and investigation panels is also anomalous. They serve in their roles as academicians but face many of the problems of judicial proceedings without having the authority to compel production of testimony and evidence.

The authority to conduct reviews of misconduct is now the subject of a class action suit by a plaintiff who is challenging the right of NIH and his university to conduct an investigation of his work (*Needleman v. Healy*, 1992). We, who are the whistleblowers in the Needleman case, herein describe our experiences and observations. We hope that this report and others in this issue will be useful in the development of revised policy.

#### THE INVESTIGATOR

Herbert Needleman is well known for his work on the effects of lead exposure on children. He is an active crusader for the position that lead exposure at low levels is harmful. He has received considerable attention in the media and has published numerous articles in professional journals and chapters in edited books. In all forums, he has consistently emphasized material that supports his stance and has downplayed and disparaged information that is not consistent with his position. He does not publish data from his own findings that do not support his position (Needleman, 1981).

Needleman has testified before Congress; consulted for the Environmental Protection Agency (EPA), the Centers for Disease Control, and the Agency for

<sup>1</sup>The Office of Scientific Integrity has been moved from NIH to the Public Health Service and remained the Office of Research Integrity. Policies regarding scientific misconduct are under review.

Toxic Substances and Disease Registry; and is a major author of government policy documents. He has testified for plaintiffs in many tort cases in which children have claimed damages as a result of lead exposure. He has obtained grant support from NIH and EPA to conduct further research on this topic. In *Science*, he was called the "Get-the-Lead-Out Guru" (Palca, 1991).

Needleman's activities have been important in setting federal policies that will be very costly. Given the strong likelihood that these policies will result in a diversion of funds from programs that can be helpful to children, the consequences of the publication of misleading reports in the scientific literature and in the media are far-reaching.

#### BACKGROUND FOR OUR REVIEW

The major item that catapulted Needleman into prominence was the article (Needleman et al., 1979) at the center of our concern. This article was disturbing to us and to others because an exceptionally large number of subjects was eliminated from analysis and because some aspects of the procedure were not well described. Limitations of the methods used were not recognized or discussed in the article. Some problems were noted by one of us (Ernhart, Landa, & Schell, 1981) and attributed to methodological naiveté. Collegial efforts were made at professional meetings and through letters in journals to explore these problems. Needleman's responses were not informative, and nagging concerns remained.

In 1982 and 1983, the EPA conducted workshops to assist in preparing a revision of the air quality criteria for lead. Because the Needleman work was questioned at the workshops and because he attacked Ernhart's research, the project officer chose to have the two studies evaluated to determine if they were to be given credence in the review.

In the spring of 1983, an expert panel on pediatric neurobehavioral evaluations was appointed to review these studies, among others. The panel included Scarr and others with relevant methodological expertise who were not engaged in lead research. Ernhart submitted all of her data and conducted reanalyses requested by the panel. Cooperation by Needleman was limited. The panel did not have sufficient time and was not given the opportunity to explore all of the problems in his work. Nevertheless, the panel believed that the difficulties in the study went well beyond the usual complications resulting from conducting research on human subjects whose exposure and behaviors were beyond the investigator's control. The panel concluded:

In summary, at this time, based on questionable Pb [lead] exposure categorization and subject exclusion methods, problems with missing data, and concerns

regarding the statistical analyses employed and selected for reporting, the Committee concludes that the study results, as reported in the Needleman, et al. (1979) paper, neither support nor refute the hypothesis that low or moderate levels of Pb exposure lead to cognitive or other behavioral impairments in children. (United States Environmental Protection Agency, 1983, p. 38)

The panel was disbanded. In April 1984, EPA's Clean Air Scientific Advisory Committee (CASAC) reviewed Needleman's response to the panel's report and some additional analyses he conducted. Recently, Needleman (1991a) asserted that one of EPA's statisticians, Hugh Pitcher, was given the data at that time, reanalyzed them, determined that the findings were creditable, and gave the results to CASAC. Pitcher denies this (University of Pittsburgh, 1992). His report to CASAC was limited to comments on analytic issues.

In spite of the strength of the conclusion rendered by the expert panel, CASAC decided that the study was not so seriously flawed as to exclude its findings from consideration in policymaking decisions. The impression we had, as we observed that CASAC meeting, was that few members knew the details of the situation and that the decision had been preplanned. In early January 1992, we filed a request under the Freedom of Information Act to obtain and review documents related to the CASAC decision. As of September 1992, this material had not been provided. From another source, we learned that, in August 1983, while the EPA panel was investigating his work, Needleman was given a grant by EPA. This grant yielded one obscure publication (Atkinson, Crocker, & Needleman, 1987), which consists, ironically, of an other inappropriately conducted reanalysis of the same dataset.

#### THE SUPERFUND CASE

In 1990, Scarr and Ernhart were retained independently by attorneys for different defendants in a Superfund case. Needleman was retained by the plaintiff (i.e., the Department of Justice acting for EPA). We each expressed concerns to the respective attorneys regarding the integrity of Needleman's work. Because he would be relying on his own data, the attorneys used the discovery process to request that we have access to these data. The court approved the request.

Our September 1990 visit to Needleman's office to review computer outputs of analyses had a number of outlandish elements. The attorney who accompanied us was forced to wait in a public hallway because he represented a lead-using industry. Although he was present, Needleman did not speak to us; communications were relayed through the attorney for the Department of Justice who apparently was present for this purpose. We were not given access

to raw data or to a code book for the computer printouts. Requests for explanations for cryptic marks on the printouts were refused. We were not allowed to photocopy; instead we handcopied pages of data to substantiate our observations. We saw no materials related to earlier analyses known to have been made (spring 1977, December 1977, February 1978), and we did not see any data or analyses for subsequent reports, including the 11-year follow-up study.

On the second day, we were asked to rescind our rights of freedom of speech by agreeing that we would not discuss our findings other than in the case being tried. We refused, took our notes, and left. Our report (Ernhart & Scarr, 1990), completed in October 1990, was not distributed pending settlement of the case.

Shortly before the Superfund case was settled in November 1990, the Department of Justice, at Needleman's behest, sought again to have us return or destroy our notes and report. There is no precedent for such an order. We retained counsel, and in January 1991, the case was argued in the Federal Court for the District of Utah, Central Division. In his April 1991 decision in our favor, Chief Judge Bruce S. Jenkins (*United States v. Sharon Steel*, 1991, p. 7) wrote:

Finally there are public policy reasons regarding the suppression of information and knowledge, particularly knowledge of a scientific nature. The pursuit of scientific knowledge is, in theory, an open process. There is something inherently distasteful and unseemly in secreting either the fruits or seeds of scientific endeavors. This is especially true here, where there appears to be little professional or economic justification for the order sought by the plaintiff. . . . The only harm or prejudice asserted by Dr. Needleman appears to be the risk of academic criticism and potential misuse of his data by others. Exactly how this misuse will occur is not clear. This is insufficient justification for the relief sought by plaintiff. Plaintiff's motion is therefore DENIED.

#### THE REPORT

In the report (Ernhart & Scarr, 1990, pp. 58-59), we described the problems that we had been able to discern as serious. Our summary, here condensed, included the following points:

1. There was improper control of confounding. The final analytic models did not include covariates, notably child's age, that should have been included. Age was negatively related to the test scores, including the normed IQ scores, and to lead level. Initial analyses with age entered

yielded few significant results. Analyses without child's age followed, yielding significant models, some of which were published. Data in widely reproduced graphs were not adjusted for confounding.

2. There was exclusion of the data of 40% to 50% of the 270 children who were tested. Exclusions were related to the likelihood of finding significant effects. One group of 16, 18, or 19 (reports differ) children who were excluded as "lead poisoned" were not impaired in intelligence ( $M$  IQ = 100,  $SD$  = 16).
3. No consideration was given to the risk of Type I error, given the very large number of variables analyzed. Results of analyses that did not support the hypothesis were not reported.

Early analyses in the dated computer output included data for the full sample of the 270 children who had been tested. The initial analyses of covariance, which included age and other covariates, showed little in the way of statistical significance. These were followed by others that moved toward the results that were published. The progression of analyses contributed to a sense of deliberateness in a shift to misleading results.

#### INQUIRY AND INVESTIGATION

Our report was not thorough or complete. We had only a day and a half with only a portion of the records, and our handcopied notes were subject to error. Nevertheless, we felt that what we had been able to document warranted further investigation.

After the federal court decision, we sent the report to NIH—OSI on May 7, 1991. The staff at the Office of Scientific Integrity (OSI) reviewed the material and in October 1991 (Watkins, 1992) requested that a formal inquiry be conducted by the Integrity Office at the University of Pittsburgh, where Needleman is a professor.

We became aware in December 1991 that the inquiry panel had completed its report when reporters contacted us for interviews. Although Needleman made the report (or a portion thereof) available to reporters and his associates, confidentiality requirements prevented the university and OSI from providing us with copies. Reporters told us that the inquiry panel had recommended a full investigation of Needleman's research, as they had been unable to rule out misconduct. Six months later, the inquiry panel report became public through the NIH reply to Needleman's suit against NIH and the University of Pittsburgh.

#### FINDINGS: THE INQUIRY REPORT

The Needleman Inquiry Panel (1991), a distinguished group of Needleman's colleagues, examined published reports, computer printouts, and data files. It's conclusions, submitted in December 1991, were:

In terms of the basic data management, the panel was impressed with the accuracy and completeness of the records reviewed with one exception—the loss of the main data tape, apparently without any provision for a back-up.

In the subject selection and classification area, the panel was concerned that the application of the selection and classification procedure criteria described in the 1979 report (and repeated in the 1990 report) did not appear to correspond to what in fact occurred. Indeed a third of the subjects included in the analyzed sample did not seem to meet the criteria, while many of those apparently eligible were neither included nor could be shown to have a reason for exclusion. Unfortunately, Dr. Needleman was unable on three occasions (in his "response to OSI inquiry" and in two interviews 11/22/91 and 12/3/91) to specify the exact selection procedure used. These concerns are further compounded by the subsequent exclusion of subjects for a variety of reasons (i.e., the reduction of the sample size from 270 to 158), an issue much discussed in previous critiques of Dr. Needleman's study. However, it is apparent that the inappropriate selection of subjects doesn't, as far as can be ascertained, bias the study in favor of Dr. Needleman's hypothesis, suggesting an absence of scientific rigor rather than the presence of fraudulent science. *Based on the subject selection and classification problems alone it is doubtful whether the 1979/1990 report represents a fair and accurate ascertainment of the relationship between IQ and dentine lead levels in Chelsea and Somerville 1st and 2nd graders.*

Finally, the data analysis appears to be almost mechanically conducted with insufficient thought given to the inclusion of appropriate variables. *The final covariate models chosen for presentation do not seem to represent the inclusion of the most important covariates, but do include co-variables with little or no contribution.* The basis for choice of the final models are neither obvious nor clearly explained. It is thus difficult to ascribe such choices to deficient analysis or misrepresentation, but the panel felt the modeling to be questionable and that age, sex, father's education and the examiner effects were neither sufficiently explored in the analysis nor sufficiently presented and discussed in the reports.

In view of the above findings and additional concerns raised by other reviewers of Dr. Needleman's work, *the panel recommends that a full investigation be held to determine if the apparent inappropriate selection of cases and incomplete presentations of results constitutes research misconduct.* (Needleman Inquiry Panel, 1991, pp. ii–iii; italics added)

The University of Pittsburgh promptly appointed an investigation board, which began work in early 1992. We did not participate in the reviews by OSI

or the inquiry panel at the university. Our only involvement in the Pittsburgh investigation of Needleman's possible misconduct consisted of appearing at the unprecedented open hearing that Needleman demanded during the investigation phase.

#### THE HEARING

##### Preliminary Activity

When the inquiry panel recommended a full and formal investigation, the university appointed a board of five professors, not including those who had served on the inquiry panel, to conduct the investigation. The university's policy requires that a formal hearing be held during the course of an investigation. We were notified by the University Research Integrity Officer that a hearing would be held on April 13-14, 1992, but we were not asked to be present. There seemed to be no need for our participation because our findings were in the report we filed with NIH-OSI. Our report had been superseded by the more thorough inquiry report.

The university's policy also specified that the hearing be closed for the protection of the respondent (in this case Needleman) and the witnesses. Needleman objected vociferously to a closed hearing. He solicited support from the faculty and others and included his complaints in his suit against NIH and the university. The chancellor of the university yielded and ruled that the hearing could be open.

The idea of an open hearing appeals to Americans' sense of justice.<sup>2</sup> This quasi-judicial hearing at Pittsburgh, however, was not to be conducted with due process or with any clear precedents for procedure. Our interests as witnesses were not considered. Because we were familiar with Needleman's notorious behavior in other forums, we were concerned that having the hearing open would result in an unproductive and acrimonious performance planned to intimidate witnesses and members of the board in the presence of the media. We were right.

We were contacted by the university less than a week before the hearing and asked to be present on the first morning. Needleman demanded "to confront his accusers," we were told, and it was not clear to the chairperson of the investigation board that the hearing could proceed without us. It was only as we interacted with the chairperson in making arrangements that we learned the identity of the board members. We also learned then that the board was, in fact, working diligently on the investigation.

<sup>2</sup>The significance of this idea in American thought is reflected in the Sixth Amendment to the Constitution, which ensures the right to a speedy and open trial.

Because this was not a judicial procedure, we asked the university's attorney to seek legal protection for us so that we could be candid in our testimony. (Because they were employees of the university, members of the board were personally protected by the university against the risk of litigation.) Exploration of our request for legal protection by the respective attorneys consumed several days. Ultimately, the university refused to indemnify us against litigation. Three days before the meeting, we decided that as a matter of principle we had an obligation to respond to questions about our findings. We decided to attend, accompanied by our attorney, even without legal protection. Such protection was a serious concern because Needleman had already sued us once and had recently filed his suit against NIH and the university.

##### Conduct of the Hearing

As a result of confidentiality policies, we went to the open hearing with no knowledge of the findings of the inquiry panel. After introductory statements and a few questions from the board, Needleman was allowed to question us. Few questions had any connection with his research or with our review of his research. Topics included details of studies we had published in the 1970s, other clients of our attorney's firm, a lawsuit in which Ernhart had testified 4 years ago, and details of a talk that Scarr had given. Some questions were insultingly simple (e.g., he quizzed us about the computation of IQ and what is meant by confounding), and he then made sarcastic comments about the responses.

Several times we asked the chairperson if questioning could be limited to relevant material. We were told that Needleman could ask whatever he wished but that we were not obliged to answer. We balked at contributing to a process that made us targets of insults and that became more and more abusive. We refused to answer a number of questions and carefully limited our responses because of the risk of litigation. In this meeting, as elsewhere, Needleman diverted attention from problems in his research by saying that we were acting for the lead industry. His attorney, who under NIH and University of Pittsburgh regulations was not permitted to participate, was in fact allowed to make extensive comments, which included references to us as "the industry witnesses."

Three of the six other witnesses provided testimonials about Needleman; their connection to the study was very limited. One co-author was questioned in detail by the board about shifting criteria for classifying subjects. An EPA witness testified that he had conducted additional statistical analyses in the fall of 1991; he said the analyses indicated that Needleman's original conclusion was valid, even if his analyses were not optimal. It was not clear whether he had been given the same data as those used for the 1979 *New England Journal of Medicine* article (Needleman et al., 1979). Another statistician commented

on analyses done by the inquiry panel. Because neither the inquiry report nor written copies of additional analyses were provided to us, the value of these analyses was not readily apparent at the time.

Most co-investigators and staff who were actually involved in the conduct of the study were not at the hearing. The most important missing witness was Charles Gunnoe, the psychologist who had participated closely in the data collection and who had conducted most of the data analyses. In the final session, members of the board questioned Needleman. Once again, penetrating questions about the selection, exclusion, and classification of subjects indicated that the board found this to be a matter of considerable concern.

The open hearing may have pleased Needleman and entertained the audience, but the information produced was quite limited.<sup>3</sup> We found that open misconduct hearings, conducted by a university without due process or legal rules of procedure, can be untidy, acrimonious, threatening to unprotected witnesses, and not very productive. We came away feeling that we had been badly used, not by the board but by a proceeding that did not include rules to prevent irrelevant and insulting questioning and that did not protect witnesses from future litigation. The sense that the hearing was not very enlightening was shared by members of the press (Palca, 1992; Taylor, 1992).

The window that the hearing provided indicated that the study was receiving an investigation that was more thorough than was possible either by the EPA expert panel or by us in our brief visit to Pittsburgh. The burden of the knowledge of the problems in this work lay with people who were competent and who were able to learn more about it than we did. Our involvement became history. Or so we thought.

#### FINDINGS: THE INVESTIGATION REPORT

We now have the Needleman Hearing Board Final Report, but, at the request of the university, it remains confidential while Needleman completes his appeals. We have also been kept apprised of the progress of the final report and the appeals as they move through the university hierarchy. The report and recommendations have been accepted by the Dean of the Medical School. They have now been transmitted, with another set of appeals, to the Provost of the university. The Provost has been trying to appoint a panel of the faculty to review the appeals but has yielded to Needleman's complaints about the composition of various suggested panels of his fellow faculty members.

<sup>3</sup>In accordance with the University of Pittsburgh policy, we have been given the report and recommendations of the hearing board. Ironically, one line of inquiry that the board pursued had not been anticipated in our report, and the members of the board were unaware that we had important documentation that might have helped in the resolution of the specific issue.

In the meantime, Needleman has released a statement saying that he was found not guilty of misconduct. This decision followed from the definition of *scientific misconduct* the board was required to use. He has not, to our knowledge, circulated the final report to the public.

#### A LEGAL PERSPECTIVE<sup>4</sup>

From a legal standpoint, inquiry/investigation procedures for misconduct investigations have not been adequately defined by NIH or, in this particular instance, by the University of Pittsburgh. They are, as might be expected, neither fish nor fowl, as evidenced by their susceptibility to manipulation by Needleman who, as a result of their inadequacy, was able to run afoul of good common sense and propriety during the investigative proceedings. Based on this experience, it is imperative that henceforth such procedures be comprehensively and accurately defined, both substantively and procedurally, to ensure that all involved are aware of the requirements and know what to expect in such proceedings.

#### Delegation of Authority

Under the existent policies prepared by NIH for OSI (now Office of Research Integrity [ORI]), a misconduct inquiry/investigation can be delegated to the university that is the present location of the scientist under scrutiny. The first legal question such a mandate raises is whether a decision to request an investigation at a university, other than that where the questioned conduct occurred, can adequately deal with the examination of conduct and activities engaged in at another university and thereby restricts the investigators' ability to obtain competent evidence over a broad range of inquiry. Investigatory panels have no capability to seek and obtain evidence formally from the other school.<sup>5</sup> Concomitant is the issue of whether the potential political value of a nonoffensive outcome—effectively a not guilty finding—at the present school can affect the adequacy or results of the investigation on behalf of OSI (ORI). It is hoped that this would not be the case, but venue and politics have been known to affect the outcome of proceedings.

<sup>4</sup>Although most of this article reflects the experiences of Scarr and Ernhart, this section was written primarily by Geneson.

<sup>5</sup>Follow-up study data on the 1979 sample (Needleman, Schell, Bellinger, Leviton, & Alfred, 1990) are located at Harvard University, as are data from a prospective study for which Needleman is co-principal investigator.



### Inadequate Authority

Assuming the best, the process by which a university will proceed remains fraught with questions and pitfalls. What should be—if there were adequate procedural safeguards and authority—essentially a private inquiry, done in confidence to protect both the subject of the investigation and the witnesses to the wrongdoing, can, as in the Needleman situation, deteriorate into an unusual and inappropriate public hearing. A potentially useful nonpublic investigative hearing in this instance was turned into guerrilla theater that did not reflect the panel's actual investigative efforts, which occurred both before and after the public "event." Moreover, the creation of a public proceeding generates the need for the panel—acting in its role as a special master<sup>6</sup>—to assume instead the role of judge, a responsibility for which it is neither charged nor equipped.

In the Pittsburgh open hearing, despite the chairperson's best effort, Needleman was abusive of and inappropriate in his conduct toward witnesses Scarr and Ernhart. Of course, because of the nonformal nature of the proceedings, the chairperson had no specific authority to control and no apparent experience in controlling this quasi-judicial proceeding. What should have been effectively a grand jury<sup>7</sup> inquiry conducted in private, without external influences, with its results presented to a final arbiter—the Dean—for review and response, deteriorated, at least externally, when it was forced to interact with the public and media.<sup>8</sup> Such is the consequence of inadequate procedural safeguards.

### Litigation Versus Investigation

Another aspect is the interaction of the investigative process with the "regular" legal system. First, the question must be addressed as to whether NIH can, in fact, "downstream" its investigative responsibility and any legal authority that its right of investigation might carry. There is no identifiable predicate suggesting that it can legally delegate this responsibility and concomitant authority to another entity—no regulation, not even a policy statement. Without formal

<sup>6</sup>A special master is a person or, in this case, a group whose unique background and/or knowledge makes them useful as a designated fact finder to whom a judge or other legal authority can delegate the responsibility and procedural authority to investigate and report on specific matters.

<sup>7</sup>A grand jury is a deliberative body whose work is done in secret and whose responsibility is typically to investigate and determine whether there is adequate evidence in support of bringing formal charges for trial.

<sup>8</sup>In a tribute to their professionalism and skills, the chairperson and board carried on in the investigation very effectively and did as thorough a job as was possible given the procedural limitations.

authority to investigate, a university's panel and witnesses remain susceptible to lawsuits predicated on any number of legal theories, without the protection of qualified legal immunities<sup>9</sup> that go with such proceedings when authorized under law. In this case, Needleman, in fact, sued everyone in sight except, remarkably, the complaining witnesses. Further, the ability to sue over the existence, much less the conduct, of the investigation simultaneously with its occurrence created the possibility of legal blackmail, which effectively would have been unavailable if the investigation had used its own independent legal authority.

### No Power to Compel

Next, the question must be considered as to whether the NIH, a delegated university, or the accused should be granted the power, through some legal vehicle, to compel witnesses' appearances for sworn testimony and production of relevant records. In the Needleman case, only moral suasion was available to the panel, effectively precluding any substantive inquiry into the underpinnings of the Needleman study's population, testing methods, blinding, and so on irrespective of whether this was part of the investigative panel's mandate. Moreover, testimony not under oath is intrinsically suspect, because dissembling carries no penalty, particularly in this type of investigation in which self-protection and protection of career and reputation in the scientific community are the order of the day.

From a procedural standpoint, despite intellectual skill and expertise, university professors do not instantly become judges without portfolio just because a panel is created to investigate and they are appointed as investigators. Without the development of comprehensive procedures to handle the investigative process and proceedings, due process and legal procedures are abandoned, not necessarily by design but by unfamiliarity.<sup>10</sup> To obviate this

<sup>9</sup>Qualified immunities are legal protections that inure to an individual or entity as a result of the legal proceedings in which they take part or as a function of the job they hold. Examples are the qualified immunity of witnesses in court proceedings to be effectively immune from suit for their truthful testimony, the broad immunity from suit with which a prosecutor is cloaked for his or her actions in the scope of his or her duties, and the limited immunity an employer enjoys in being able to evaluate his or her employee candidly.

<sup>10</sup>The Administrative Procedures Act (the "APA"), Title 5 United States Code, Sections 701 et seq., enumerates in detail the processes and procedures for judicial review available to a person aggrieved by the decisions or actions of an agency of the United States. It is not reasonable to suggest that the government can, at its whim or discretion, empower universities with judicial authority under the APA. Alternatively, if the school's investigative process were to be in tandem with the government agency, or if legislation were enacted to create a legal "right of action" for universities to seek judicial enforcement of investigative procedures done at the government's direction, the procedures available in other matters under the APA might be an effective model for misconduct investigations.

problem, it is necessary that the panel assume more effectively the role of special master and be provided with adequate guidance in the processes and procedures of an investigation of this type. In this regard, the university's scientific integrity officer and his or her legal counsel become important in developing and tailoring the process and procedures to follow the actual tenets of the law, it is hoped, under the direction and/or guidance of a trained prosecutor/investigator—either from NIH or retained by a university.

#### Appeals

Finally, the procedures in terms of appellate review must be more accurately defined. The ability to appeal a panel's findings indefinitely under a university's policies is a patent circumvention of the investigation's process and effect. If scientific misconduct or any variation thereof—a lesser included offense—is found by the investigative body, the conduct must necessarily be disclosed, if only to undermine the perpetrator's effectiveness in continuing to attempt to influence improperly the scientific community. Indefinite litigation over a panel's findings serves no purpose except to dilute the value of the investigative process and its results.

#### Summary

The authority of the hearing board was circumscribed, and it was necessarily handicapped by an inadequate legal foundation for effectuating the procedures and processes to fulfill its mission. It is thus not surprising that the Needleman investigation was incomplete. Without the authority to compel document production, witness appearances, and sworn testimony of what actually occurred during the crucial period of Needleman's creation/evolution of the data set, the board was obliged to accept the data set as valid. Given the problems with other information and representations both within the study and given to the board to "support" the study, such concessions may not be warranted. This limitation of the investigation is important, because the credibility of further analyses depends on the integrity of the data collection procedure.

Albeit, the board did evaluate comprehensively the accuracy or inaccuracy of some of the key representations about the study's methodology made by Needleman and others. In large measure, the board's effectiveness was very much a result of the hard work and skill of its members and their staff, which helped to make up for the failings of the inadequately defined and inadequately empowered system under which they were required to work. With this said, serious considerations should be given to legally defining, formalizing, and improving the scientific misconduct investigative system and its processes. It is foolish to assume that any given investigative panel would display the judiciousness shown in this instance.

Furthermore, the board did not and could not fulfill the broader mandate of fully investigating Needleman's research from the inception of the 1979 study (Needleman et al., 1979) to the present.<sup>11</sup> The investigation was limited to the 1979 study data provided by Needleman and to the major publication we had questioned following the 1990 Ernhart-Scarr visit to his office. There was no investigation of the 1990 follow-up study, which used the same subject pool as the 1979 study, or of the ongoing Boston prospective study for which Needleman is co-investigator. There was also no investigation into unexplained recent alterations of a widely reproduced graph of data from the 1979 study (Needleman, Leviton, & Bellinger, 1982). Other publications of questionable data also failed to receive scrutiny. Undoubtedly, with adequate legal empowerment, the investigation would and could have fulfilled the more comprehensive mandate.

### DID WHAT WE OBSERVED RISE TO THE THRESHOLD OF MISCONDUCT?

#### What Is Misconduct?

Several times during the hearing, members of the board expressed concern about the differentiation of misconduct and poor science. The definition of *misconduct* is a thorny issue (Hamilton, 1992). Tort law (Goodstein, 1992) requires the proof of five elements in cases of civil fraud. Two of these, reasonable belief by a plaintiff and resulting damage, are not generally critical for research fraud. The other elements are (a) the representation was false, (b) the defendant knew that the representation was false, and (c) that the defendant intended to induce belief in the representation. The requirement of proof of *intent* to misrepresent (i.e., that the misrepresentation is deliberate) makes this a stringent definition.

In addition to plagiarism, fabrication and falsification of data, the present NIH-NSF definition (Department of Health and Human Services, 1989) includes "other practices that seriously deviate from those that are commonly accepted within the scientific community for proposing, conducting, or reporting research," which does not specify intent. This definition is also included in the University of Pittsburgh Research Integrity Policy. This definition is broad and could sweep into it deviations based on a reasonable difference of opinion about methods and/or creative scientific innovation.

A proposed revision (American Psychological Association, 1992, p. 6) nar-

<sup>11</sup>The Research Integrity Policy of the University of Pittsburgh states, "Consideration shall be given by the hearing board to the review of all research with which the respondent is involved."



rows the definition of *research fraud* to avoid this problem and to require demonstration of intent: "evidence of plagiarism, the fabrication or intentional falsification of data, research procedures or data analysis; or other *deliberate misrepresentations in proposing, conducting, reporting or reviewing research*" (italics added). The shift from *deviation*, which may reflect a difference of opinion about research strategy, to *misrepresentation* should be helpful in determining scientific misconduct in specific cases.

#### Fraud as Misrepresentation

The only evidence we have seen of possible fabrication or embellishment of data in Needleman's work is an unexplained revision, with new data points added, of a previously published graph. Needleman has also failed to report the results of other research when the results were not favorable. Underreporting may also be considered scientific misconduct (Chalmers, 1990) because it misrepresents research results. There were certainly clear deviations from usual and customary practice in the conduct of the 1979 study. Do the deviations in this work reflect an honest difference of opinion about methods? In the introduction to the 1979 article, Needleman et al. acknowledged principles of methodology that were quite appropriate. Nonconcordance between these principles and the published report was noted shortly after its publication (Ernhart et al., 1981). Needleman's deviations were not creative—at least not in a positive sense. His deviations from ordinary scientific practice produced reports that misrepresented the research results.

#### Intellectual Misconduct

*Science* recently discussed the effect of intellectual passion on the behavior of scientists (Marshall, 1992, p. 621). Scientists interviewed acknowledged that intellectual conflicts of interest are pervasive.

The key difference among scientists, they said, is not between those who have conflicts and those who do not, but in how the potential conflicts are handled—whether the researcher has the detachment required to be the severest critic of his or her own work—in the absence of that self-reflection, an advocate becomes so deeply invested that it's almost impossible to let go, even in the face of contrary evidence.

Extending this line of thought, Goodstein (1992) suggested that the perpetrators of fraud think they know the answer and that they simply take shortcuts—help things along, so to speak. Following usual and customary procedures may be seen as slowing things down, particularly if the results are not as expected.

A scientist may rationalize an unwillingness to be self-critical because of a belief that his or her conclusions represent underlying "truth." This kind of prejudgment is suggested in a statement (Needleman, 1983, p. 243) explaining the age difference between the high- and low-lead groups: "I wanted my technicians to get some experience with *normal children*, so I arranged, without their knowledge, for the first group of children to be low lead children." An absolutist and unquestioning posture on the effects of lead is seen in most of Needleman's public statements. Needleman's views stand in marked contrast to the more balanced approach taken by others (Fulton et al., 1987; Smith, 1989; Winneke, Brockhaus, Ewers, Krämer, & Neuf, 1990).

#### LIFE AS WHISTLEBLOWERS

We sent our report (Ernhart & Scarr, 1990) to NIH-OSI for investigation because we believe that we have an affirmative obligation as scientists and as members of the academic community to speak out against improper and deliberately misleading practices in the conduct and reporting of research.

In contrast to some other whistleblowers, we are established scientists and tenured professors. We believed our risk of truly serious reprisals to be much lower than that experienced by whistleblowers who are less established in the sciences. Nevertheless, this situation has been costly in anxiety, time, and money. These costs are not as severe as the loss of research positions that some whistleblowers have faced, but they are real.

We have been subject to a number of egregious actions overtly due to our role in this matter. Several letters, which falsely asserted that Ernhart was in conflict of interest by her actions in the matter, were written by the officers of the International Society of Environmental Epidemiology, of which she is a member. The society's letterhead was used to distribute these defaming letters to government officials, to members of Congress, to administrators at the University of Pittsburgh, and to members of the society.<sup>12</sup> Descriptions of other letters, some for publication in newspaper letters columns, are omitted from this report. These letters, along with a rude and sexist parody in a newspaper and an anonymously sent cartoon, are too bizarre to describe.

During the inquiry period, Scarr was invited to submit a manuscript on low-level lead exposure to *The Harvard Mental Health Letter*. The manuscript was rejected, however, as being "too political." Other manuscripts have been given unusually severe reviews. A very serious concern is the possibility of biased reviews of grant applications.

<sup>12</sup>This letter also asserted that epidemiological statisticians, and not psychologists, should evaluate Needleman's research. Needleman's analyses, conducted by a psychologist, consisted primarily of *t* tests, correlations, and analyses of covariance.

Other adverse events appear to be due to prejudice related to taking on the role of whistleblower. Suspicions about our motives arose continually in interviews with the press and even from colleagues who wondered why we would ever become involved in such a mess.

We have been the recipients of considerable ugliness, in the hearing and elsewhere, from Needleman. He repeatedly charges in all possible forums that we are employees of the lead industry. We should not have to make a statement of denial. This is slander. But we are not the only ones who have suffered his insults. In his response to the inquiry panel, Needleman (1991b, p. 2) tried to discredit the panel members, all competent professors and researchers, by charging that their report "is superficial, incomplete, tendentious, shallow in its scholarship, naive and incorrect in its application of statistical principles, and wrong in its conclusions."

In his lawsuit (*Needleman v. Healy*, 1992), filed while the investigation was underway, he charged that two members of the investigation board had "an appearance of impropriety," one because he "served in a high level position in a professional organization" with Scarr and the other because he reviewed evidence for a tobacco company regarding an experimental cigarette (p. 16). He further argued that the members of the board, who are renowned experts in research design, statistics, child development, neurology, environmental science, and sociology, were not qualified to evaluate his work. Despite such exaggerations, Needleman has managed to convince some otherwise responsible scientists, as well as reporters, of his allegations. Some reporters have ignored or distorted responses to their questions and have suggested that we acted out of some kind of personal vendetta.

Our families have been supportive, but they have not been untouched. One of Scarr's daughters was told by a professor that her mother was "slime." This kind of attack hurts. We also worry that our families will be hurt by the mounting costs.

One of the few positive features of this situation is that the publicity has resulted in the receipt of letters, phone calls, and after-lecture comments from pediatricians, psychologists, public health officials, and others who say that they were concerned about problems they saw in the Needleman work but had not known that anyone else had made similar observations. The irony of this is that, for the most part, these are not people who are doing research on the topic. With a few exceptions, other lead-effects researchers who have discussed in private the problems we have seen have been very reluctant to speak up. One indicated candidly that he did not want to jeopardize his research support. We sympathize with him.

We have each had the support of loyal friends, people who know us and know that the ugly claims are not so. Furthermore, we have each other for occasional "handholding." We were able to retain the services of David Geneson, a highly competent attorney who has been very supportive. We have needed these supports.

## CONCLUSION

The inquiry report indicated that our suspicions of possible misconduct were quite reasonable. The members of the panel, who had more access than we had to the materials, shared our skepticism about the methods for the selection and classification of subjects and the choice of covariate models. Even though our observations were supported, we doubt that the reprisals have ended.

Much is being written at present about misconduct in research. There are those (Abelson, 1992) who believe that traditional methods such as peer review and an increased emphasis on the mentoring of junior scientists may be sufficient to decrease misconduct. The problems that have surfaced in this and other cases involving well-known researchers would probably not have been averted by increased emphasis on these remedies. Reviewers of grant applications and of manuscripts submitted to journals can evaluate only the information they are given by the investigator. Deliberate misrepresentation is difficult to detect when one cannot examine the sources of the information.

In most cases of misconduct, the evidence of wrongdoing would not come to light without reports made by whistleblowers. In view of the risks to whistleblowers, it seems likely that a considerable amount of misconduct is never reported. The costs and tribulations to the whistleblower can be severe (Committee on Government Operations, 1990), and there are few rewards other than the sense that one has done what is right. Teich and Frankel (1992, pp. 21-23) stated that reporting misconduct contributes to science and should be viewed in this light rather than as a betrayal of a colleague. Deceptive work harms all scientists and, as in the present case, can have important societal repercussions. We hope that this article and others in this issue will contribute to a change in attitude toward those who do not look away when they observe apparent violations of ethical procedures in research.

Teich and Frankel further recommended that the allegations of scientists who speak out be taken seriously and explored effectively. Our report was taken seriously by OSI and by the ORI at the University of Pittsburgh, which conducted a valuable investigation, despite the lack of clear legal procedures. Much as we dislike regulation, judicial procedures, and bureaucracy, however, the difficulties involved in resolving some cases, including this one, point to a need for carefully structured policies and procedures that enable thorough investigation and protect the rights of all involved.

## REFERENCES

- Abelson, P. H. (1992). Integrity of the research process. *Science*, 256, 1257.  
 American Psychological Association. (1992). Advisory committee proposes changes in PHS response to scientific misconduct, fraud charges. *Science Agenda*, 5(3), 6.  
 Atkinson, S. E., Crocker, T. D., & Needleman, H. L. (1987). The importance of specification

- uncertainty and intolerance to measurement error in a study of the impact of dentine lead on children's IQs. *International Journal of Environmental Studies*, 29, 127-138.
- Chalmers, I. (1990). Underreporting research is scientific misconduct. *Journal of the American Medical Association*, 263, 1405-1408.
- Committee on Government Operations. (1990). *Are scientific misconduct and conflicts of interest hazardous to our health?* (101st Congress, 2nd session, 19th report). Washington, DC: U.S. Government Printing Office.
- Department of Health and Human Services. (1989). *Responsibilities of PHS awardee and applicant institutions for dealing with and reporting possible misconduct in science: Final rule Federal Register 54 (August 8): 32446-32451.*
- Ernhart, C. B., Landa, B., & Schell, N. B. (1981). Subclinical levels of lead and developmental deficit—A multivariate follow-up reassessment. *Pediatrics*, 67, 911-919.
- Ernhart, C. B., & Scarr, S. W. (1990, October). *Report on the research of Dr. Herbert Needleman based on samples reported in the 1979 article in the New England Journal of Medicine.* (Available from C. B. Ernhart, Department of Psychiatry, MetroHealth Medical Center, Cleveland, OH 44109 or from S. W. Scarr, Department of Psychology, University of Virginia, Charlottesville, VA 22903)
- Fulton, M., Thompson, G., Hunter, R., Raab, G., Laxen, D., & Hepburn, W. (1987). Influence of blood lead on the ability and attainment of children in Edinburgh. *Lancet*, i, 1221-1226.
- Goodstein, D. (1992, March 2). What do we mean when we use the term 'science' fraud? *The Scientist*, pp. 11, 13.
- Hamilton, D. P. (1992). Reorganization of OSI now a reality. *Science*, 256, 1383.
- Hoke, F. (1992, June 22). NAS scientific misconduct report draws harsh criticism. *The Scientist*, pp. 3, 10.
- Marshall, E. (1992). When does intellectual passion become conflict of interest? *Science*, 257, 620-621.
- Needleman, H. L. (1981). Studies in children exposed to low levels of lead (Report to the Environmental Protection Agency, Contract 68-02-2217, EPA Document 600/1-81-066).
- Needleman, H. L. (1983). Low level lead exposure and neuropsychological performance. In M. Rutter & R. R. Jones (Eds.), *Lead versus health* (pp. 229-248). Chichester, England: Wiley.
- Needleman, H. L. (1991a). Lead control. *Science*, 254, 500-501.
- Needleman, H. L. (1991b). [Response to the Needleman Inquiry Panel Report, addressed to George M. Bernier, Jr., MD, Dean of the School of Medicine, University of Pittsburgh, Pittsburgh, PA. Appended to *Federal Defendants' Motion to Dismiss or in the Alternative for Summary Judgement in Needleman v. Healy* No. 92-0749, W.D. Pa.] (Available on request from the authors of this article)
- Needleman, H. L., Gunnoe, C., Leviton, A., Reed, R., Peresic, H., Maher, C., & Barrett, P. (1979). Deficits in psychological and classroom performance in children with elevated dentine lead levels. *New England Journal of Medicine*, 300, 689-695.
- Needleman v. Healy No. 92-0749, W.D. Pa. Filed March 25, 1992.
- Needleman, H. L., Leviton, A., & Bellinger, D. (1982). Lead associated intellectual deficit. *New England Journal of Medicine*, 306, 367.
- Needleman, H. L., Schell, A., Bellinger, D., Leviton, A., & Alfred, E. N. (1990). The long-term effects of exposure to low doses of lead in childhood. *New England Journal of Medicine*, 322, 83-88.
- Needleman Inquiry Panel. (1991). *Needleman inquiry final report*. Pittsburgh, PA: University of Pittsburgh.
- Palca, J. (1991). Get-the-lead-out guru challenged. *Science*, 253, 842-844.
- Palca, J. (1992). Lead researcher confronts accusers in public hearing. *Science*, 256, 437-438.
- Smith, M. (1989). The effects of low level lead exposure on children. In M. A. Smith, L. D. Grant, & A. I. Sors (Eds.), *Lead exposure and child development: An international assessment* (pp. 3-47). Dordrecht, The Netherlands: Kluwer.
- Taylor, R. (1992). Unprecedented open misconduct hearing held at Pittsburgh. *The Journal of NIH Research*, 4(5), 44-46.
- Teich, A. H., & Frankel, M. S. (1992). *Good science and responsible scientists*. Washington, DC: American Association for the Advancement of Science.
- United States Environmental Protection Agency. (1983). *Independent peer review of selected studies concerning neurobehavioral effects of lead exposures in nominally asymptomatic children: Official report of findings and recommendations of an interdisciplinary expert review committee* (EPA-600/8-83-028A). Research Triangle Park, NC: Author.
- United States v. Sharon Steel. (1991). Memorandum Opinion and Order, No. 86-C-924, District Court, Utah.
- University of Pittsburgh. (1992). [Transcript]. *Public hearing of scientific misconduct allegations against Herbert Needleman, M.D.* Pittsburgh, PA: University of Pittsburgh School of Medicine.
- Watkins, C. A. (1992). Declaration in re: Needleman v. Healy, No 92-749, W.D., Pa.
- Winneke, G., Brockhaus, A., Ewers, U., Krämer, U., & Neuf, M. (1990). Results from the European multicenter study on lead neurotoxicity in children: Implications for risk assessment. *Neurotoxicology and Teratology*, 12, 553-559.