

PEDIATRICS®

OFFICIAL JOURNAL OF THE AMERICAN ACADEMY OF PEDIATRICS

Salem Comes to the National Institutes of Health: Notes From Inside the Crucible of Scientific Integrity

Herbert L. Needleman
Pediatrics 1992;90:977-981

The online version of this article, along with updated information and services, is located on the World Wide Web at:

<http://www.pediatrics.org>

PEDIATRICS is the official journal of the American Academy of Pediatrics. A monthly publication, it has been published continuously since 1948. PEDIATRICS is owned, published, and trademarked by the American Academy of Pediatrics, 141 Northwest Point Boulevard, Elk Grove Village, Illinois, 60007. Copyright © 1992 by the American Academy of Pediatrics. All rights reserved. Print ISSN: 0031-4005. Online ISSN: 1098-4275.

American Academy of Pediatrics

DEDICATED TO THE HEALTH OF ALL CHILDREN™



Salem Comes to the National Institutes of Health: Notes From Inside the Crucible of Scientific Integrity

Herbert L. Needleman, MD

ABBREVIATIONS. NIH, National Institutes of Health; EPA, Environmental Protection Agency; WISC-R, Wechsler Intelligence Scale for Children-Revised.

Many readers of *Pediatrics* may have only a dim idea of the combative arena in which environmental research is conducted. Probably, very few have had the experience of being investigated for scientific misconduct. My aim in reviewing these two topics is to provide a preventive road map to others and to reveal some inadequacies and inequities in the investigative process. It is necessary, to accomplish this, to be direct and specific. Tact is sacrificed here for the sake of clear instruction.

In 1972 I published 700 words in *Nature* reporting that Philadelphia inner-city children had higher dentine lead levels than suburban children.¹ The paper suggested that the tooth might be a useful marker to estimate body lead burden after exposure had ended. I did not know then that I was taking the first step toward being investigated for scientific misconduct by my university and the National Institutes of Health (NIH) Office of Scientific Integrity.

The Environmental Protection Agency (EPA) asked me to present the 1972 tooth lead paper in Amsterdam at an international meeting on lead. I was unprepared by my past attendance at pediatric meetings for what I encountered there. This was no scholarly debate on the toxicology and epidemiology of lead; this was war. The speakers did not behave like academics hoping to embellish their reputations by parading the results of their last 6 months in the lab. These stakes were much higher.

Arrayed against each other were a small and defensive group of environmentalists and health scientists on one side, and on the other the representatives of the gasoline companies, including such formidable entities as El DuPont, Associated Octel, Dutch Shell, and Ethyl Corporation of America. Any paper suggesting that lead was toxic at lower doses immediately faced a vocal and well-prepared troop that rose in concert to attack the speaker. My 10-minute talk was not spared; giving it marked the beginning of my post-graduate education.

This encounter pushed me, on returning to the United States, to look into the history of lead research.

I found that my experience was not new. Two Australians, A. J. Turner and J. L. Gibson, who first described childhood lead poisoning in Brisbane in 1892, were derogated by industry and by a segment of the medical community. When Randolph Byers, one of the earliest pediatric neurologists, first suggested in 1943 that some school dysfunction might be due to undiagnosed lead toxicity, he was threatened with a million dollar lawsuit by Lead Industries Association.² Clair Patterson, the geochemist credited with dating the age of the earth, was publicly vilified as a crank by the industry and had his career threatened when he suggested that civilization had raised everyone's body lead burdens to 1000 times that of our ancient ancestors (personal communication, 1992). All of the early research in lead toxicity was funded by the industry, who had a tight grip on what the public was permitted to know.³

Reading these records vividly brought back an experience I had when I was in medical school. One summer I worked as a laborer at the Deepwater, NJ, DuPont plant, where tetraethyl lead had been synthesized years before. Workers were forbidden to carry matches, and when the smoking whistle blew at 10 AM and 2 PM, we poured out of our buildings by the hundreds to collect at wooden smoking shacks in open areas. There we lined up at two glowing cigar lighters imbedded in the shack wall. While I smoked two cigarettes back-to-back in the 15-minute break, I inspected my coworkers.

Off to the side sat a few older men, obviously slow and clumsy, staring silently into middle space. When they did speak, they seemed remote and out of touch. A veteran worker told me that they were from "The House of Butterflies." They had been poisoned while making tetraethyl lead. Years later, I would read in the *American Journal of Public Health* that during the early stages of tetraethyl lead production at Deepwater,⁴ there had been an outbreak of poisoning among the work force. More than 300 men had been affected, often with full-blown psychotic symptoms; at least 4 had died. Affected workers were frequently seen brushing hallucinated insects off their bodies, hence the name. Production was temporarily stopped by the Public Health Service, but this ban was lifted after a superficial investigation. These damaged men were some of the survivors.

Years later, having satisfied myself that the tooth was a valid marker of past exposure, with Alan Leviton and Bob Reed, I studied a sample of children who were asymptomatic for lead, classifying them by dentine lead levels. The data showed that after con-

From the University of Pittsburgh School of Medicine, Pittsburgh, PA.
Received for publication Aug 12, 1992; accepted Aug 12, 1992.
Reprint requests to (H.L.N.) Suite 305 Iroquois Bldg, 3600 Forbes Ave, Pittsburgh, PA 15213.
PEDIATRICS (ISSN 0031 4005). Copyright © 1992 by the American Academy of Pediatrics.

trolling for a number of covariates, children with elevated lead in their teeth scored lower on tests of psychometric IQ, speech and language function, and on measures of attention.⁵ The study seemed to respond to a number of research difficulties that had until then vexed the field, and as a result it received considerable attention. The lead industry, in the form of the International Lead Zinc Research Organization, was uncharacteristically silent for about 6 months. Then they began to call for copies of my original data. I declined. I had seen what had happened to good data when massaged and distorted by industry technicians, and while I was happy to share my data with any bona fide scientist—and did—I was not willing to include the lead industry.

In 1982, the EPA began to rewrite the Air Lead Standard. I was asked to participate. Also invited was Dr Claire Ernhart, a psychologist who had published a paper in 1974⁶ that reported that lead was associated with lower IQ in a group of Long Island black preschoolers. In 1981, she published a paper (in this journal) which criticized my study and said that when followed into the first grade, the lead effect she had previously reported was no longer significant.⁷ Close examination of the paper showed that school-age blood lead levels were in fact significantly related to IQ. Ernhart dismissed this finding as due to chance, and stated that: "If there are, in fact, behavioral and intellectual sequelae of low levels of lead burden. . . these effects are minimal." Shortly after that paper she became a grantee of the International Lead Zinc Research Organization and began to speak against controlling lead in the environment. When there was a move to put lead back in gasoline, Ernhart appeared in testimony for Lead Industry Associates, asserting that there was no valid health reason to ban its use.⁸

The industry began to raise public questions about the integrity of my studies. In 1983, EPA's Clean Air Scientific Advisory Committee thoroughly reviewed industry-generated charges that my work was flawed. They concluded:

A pioneering general population study was reported by Needleman et al (1979). . . Significant effects ($p < .05$) were reported for full scale WISC-R [Wechsler Intelligence Scale for Children-Revised] scores, WISC-R verbal IQ scores, for 9 of 11 classroom behavioral scale items, and several experimental measures of perceptual motor function. . . .

Reanalyses carried out in response to the Committee's recommendations have been reported by Needleman (1984), Needleman et al (1985) and US EPA's Office of Policy Analysis (1984) as confirming the published findings on significant associations between elevated dentine lead levels and decrements in IQ. . . .⁹

I thought that this official statement had finally and permanently sealed the argument. I could have not conceived that these same charges would be resuscitated 7 years later.

In 1990, an attorney from the Department of Justice asked me to participate in what he described to me as a landmark suit brought under the Superfund Act against three lead polluters in Midvale, UT. Among the witnesses for the defense were Dr Ernhart and Dr Sandra Scarr. Scarr had been a member of the government committee that had reviewed my work for EPA. She now appeared in a different role, this time

on behalf of the lead industry, reviving the same charges that had been settled in 1986. They came to my lab for 2 days to examine my raw data in preparation for the trial.

Before going to trial, the case was settled. Sixty-three million dollars was awarded to the federal government to clean up the mine site. After the case was settled, I found out that Scarr and Ernhart had written a lengthy document accusing me of unscientific behavior. They maintained that their conclusions grew out of their examination of my printouts. This document was forwarded to NIH's office of scientific misconduct by David Genesson, an attorney for the Washington, DC, law firm of Hunton and Williams. It was also given to defense lawyers in a number of lead damage cases. I had encountered the name of Hunton and Williams before. This firm had represented Ethyl Corporation of America and El DuPont, contesting the regulation of lead additives in federal court and before the EPA and the Federal Trade Commission. In reading the Scarr/Ernhart document, I found numerous allegations and hints of unscientific behavior.

As I perceived them, their major criticisms of my work were (1) that I did not properly control for confounding; (2) that I selected cases in a biased fashion; and (3) that multiple tests were done, and this could lead to positive associations on the basis of chance.

These kinds of issues are generally considered methodological disagreements and are fought out in the pages of journals; I could not understand why they were defined by my critics as scientific misconduct. Similar criticisms were raised before the EPA in 1982 and dismissed. These facts notwithstanding, in October of 1991, I was notified by the Dean of my medical school that an inquiry into charges of misconduct was being done at the instruction of NIH's Office of Scientific Integrity.

When the proceedings began, I was confident that the printouts would be examined, that I would explain how I analyzed the data, and that like the EPA, the university would rapidly put matters right. I thought this would end this matter quickly and permanently. But the university's behavior seemed odd and troubling. They chose to ignore a number of rather obvious facts that I repeatedly brought to their attention: that the charges were initially raised by two individuals who had been supported by the lead industry; that they had been raised before and dismissed by the Clean Air Science Advisory Committee of the EPA; that my work had been replicated more than 12 times since its publication; and that I had shared my data with other scientists in the past.

Instead, the preliminary Inquiry Panel issued a strange report. The Panel stated that it "found no evidence of fraud, falsification or plagiarism," but inexplicably added that it "is not able at this time to exclude the possibility rule of scientific misconduct in terms of misrepresentation." The report argued that the models I chose were selected to optimize a lead effect, and that I may have selected cases in a biased fashion. The report presented no evidence in support of this assertion, only conjecture.

I rebutted their charges in a letter to the Dean and showed that the charge of misrepresentation was based on false evidence. The Dean declined to review my letter. Instead, he turned it over to the Panel for comment. They also did not respond to any of the facts that I raised in the letter. Instead, they stated that the material I supplied in rebuttal of the report of the Inquiry Panel was "not directly relevant." They recommended a full investigation.

During the time the investigation was being arranged, I requested of the Dean that the Hearing Board he appointed include experts of international standing in the fields of behavioral toxicology and epidemiology. This was denied. I was told that there was no need for this expertise in the two disciplines that my work spanned. I requested that the hearings be open to the university community and the press. Again, this was denied. I asked that two members of the Hearing Board be replaced for possible conflict of interest. One, Dr Robert McCall, was a developmental psychologist whose appointments on many professional committees overlapped with Dr Scarr, and who frequently cited her work in support of his. The second, Dr Herbert Rosenkranz, had been Director of the Environmental Sciences Center at Case Western Reserve University, where Dr Ernhart was a faculty member. This request was also denied.

I began to feel uneasy and increasingly certain that if the case were reviewed *in camera*, I would be found guilty of something. I went before the Faculty Assembly of the university and requested their support in my demand for open hearings. The faculty emphatically supported me. The Assembly passed a unanimous resolution asking the university to open the hearings. At the Faculty Senate, a representative of the administration argued against open hearings, because, he said, it was necessary to "protect the process." The "need to protect the process" was a phrase I was to hear repeated many more times. I argued that the process did not have a nervous system; that it was people who required protection; and that the given reason that hearings were closed was to protect the reputation of the accused. I was in this instance the accused, and I wanted the hearings to be open. The Senate unanimously voted for open hearings.

Pressure began to build on the administration, and I began to receive letters of support from colleagues around the country. Six eminent health scientists, Frank Oski, Arthur Upton, Samuel Epstein, Philip Landrigan, David Bellinger, and Bernard Weiss sponsored a petition to the Chancellor demanding open hearings. It listed almost 400 scientists' signatures. I filed a complaint in federal court asking for open hearings. Reluctantly, for the first time in its history, the university agreed to open hearings.

My accusers, who until then had been quite public and emphatic in their allegations, and who had said that they would willingly come to Pittsburgh to be questioned by me, reversed their field. They were now reluctant to attend. After lengthy negotiations with the administration, they agreed to attend the hearings as witnesses.

The hearing room was filled with scientists, faculty, and members of the local and national press. My

accusers became surprisingly reticent. Dr Scarr, in a lecture at the Massachusetts Mental Health Center,¹⁰ said: "What we have done is to report. . . Dr Needleman to the Office of Scientific Integrity at the NIH, because we feel there are significant deviations from normal scientific practice here and we feel that the data has been massaged, to put it mildly. . . ." Now, in an open hearing, she revised her complaint to say that she merely "had suspicions" that I had consciously manipulated the data to present a false case.

Both witnesses were accompanied by their attorney, Mr David Genesson of Hunton and Williams of Washington, DC. When I asked Dr Ernhart who was paying her legal bills, she refused to answer. She stated that she did not know that Hunton and Williams had represented El DuPont and Ethyl Corporation of America before the Food and Drug Administration and Federal Trade Commission. In the newspaper the next day, it was reported that there was a "trust fund" established to cover my accusers' legal expenses, but that Scarr and Ernhart did not know who had contributed to it.

During my examination of my accusers, it became clear that a different standard, perhaps an ad hoc standard, was being applied to my work as contrasted to theirs. One of the charges raised by my accusers was that I did not control for age in evaluating the effect of lead on IQ. I pointed out in my cross-examination that the WISC-R IQ was age-adjusted.

DR NEEDLEMAN: Isn't the Wechsler age adjusted?

DR ERNHART: The norming of the Wechsler is age adjusted. . . norming alone is not sufficient to handle age variation. . .

DR NEEDLEMAN: So it would be better to enter age into the model?

DR ERNHART: Yes. . .

DR NEEDLEMAN: In your 1981 paper did you put age into the model?

DR ERNHART: My study is irrelevant to the issues here today.¹¹ [Ernhart had not controlled for age.]

Since Ernhart had raised these criticisms of my work in 1981, and examined my printouts in 1990, I asked her whether it was not true that she had concluded that my study misrepresented the data before she had ever examined my data. Her answer was intriguing.

DR ERNHART: On advice of counsel, I'm not answering that question.

Another claim was that I excluded subjects on the basis of head injury or history of exposure or being non-English speaking *after* I knew their IQ scores, in order to maximize the effect of lead. In the hearing I showed her a piece of computer code from my printout that headed every data analysis. Translated, it said: "Select if lead level equal high or low, and head injury equal 'no,' and plumbism equal 'no' and English is the first and only language in the home." This proved conclusively that the subjects were excluded on criteria that were identified before the study was begun, and that the exclusion was executed by computer without any human judgment. Because Dr Ernhart had spent 2 days with my printouts as part of the Midvale suit, I asked whether she had seen this piece of code.

DR NEEDLEMAN: . . . does this look familiar to you at all?

DR ERNHART: I don't recall having seen this.

This piece of computer code appeared 24 times in the printouts I furnished them. It is difficult to see how it could have been overlooked by anyone looking for problems in case selection.

There was a general retreat by both witness in the degree of certainty with which they indicted me. Scarr, who had been direct and accusatory in a lecture at Harvard, was much less sure about whether I committed scientific misconduct in the public hearing. I asked her about it directly:

DR NEEDLEMAN: Are you certain that you are right when you say I selected the cases consciously knowing the outcome in relation to lead?

DR SCARR: I know you had the opportunity to do that. I don't know what you did.

At the conclusion of the cross-examination, Dr William Cooley, Chairman of the Hearing Board, who had frequently advised my accusers that they were not required to answer my questions, addressed himself to Dr Scarr:

I believe that, if I may ask a clarifying question, it is my impression that you have gone on record here today as essentially indicating that you had ample basis for being suspicious of the scientific work that's under consideration here, but have no specific charges of misconduct.

DR SCARR: Yes, that's correct.

The 2-day hearings were widely reported in the lay press^{12,13} and in *Science*¹⁴ and the *Journal of the National Institutes of Health*. Two months later, on May 20, 1992, the Hearing Board unanimously found no evidence of scientific misconduct.

What is there to be learned from this story? I believe that the spectrum of those behaviors labeled as misconduct in scientific enterprises is disturbingly common and that both the public and the scientific enterprise needs to be protected from inferior or dishonest studies that open the door to procedures or pharmaceuticals of dubious efficacy or that distort our understanding of the way that nature works. I believe that because of the intensely competitive business that science has become, the ethos in which young scientists are socialized and the actual work is conducted has fundamentally changed, and not for the better. Young scientists are regularly exposed to the gap between the professed idealistic standards of practice and the actual, often cynical, conduct of grant getting, data collecting, interpreting, and publishing. There needs to be better policing of our profession.

But the entire tangled process of identifying putative cases of scientific misconduct, and of fairly judging them, is open to abuse at a number of points. If my case illuminates anything, it shows that the federal investigative process can be rather easily exploited by commercial interests to cloud the consensus about a toxicant's dangers, can slow the regulatory pace, can damage an investigator's credibility, and can keep him tied up almost to the exclusion of any scientific output for long stretches of time, while defending himself.

Some way must be found to screen out frivolous or harassing charges of misconduct and shield investigators from this form of tribulation. Once an inquiry

or investigation has begun, it should operate under formal principles of due process. The option as to whether the investigation is open should lie with the accused. If an open hearing is requested, it should be freely granted. One should not be required to fight for this long-honored right. Certainly there is stigma and embarrassment attached to this charge; these are trivial compared with the risks that attend closeted star-chamber proceedings. One can live with embarrassment.

The charges should be given in specific written form to the accused party. They should take the shape of single valued propositions that can be disproven. Vague charges of guilt are out of place in a free society. The accused should have an attorney of his or her choice furnished by the university. The rules of evidence and the burden of proof should be clearly defined. Full and unhindered cross-examination of the accusers should be allowed. Each authority, whether university, hospital, or research institute, should have an ombudsman group with official, not advisory status. At my university, there is a standing committee on academic freedom which serves this role, but it has little official standing. A majority of the members of any investigative panel should be constituted from experts outside the university. Full disclosure to avoid conflicts should be required. These should be chosen in the same fashion as a jury, with challenges for cause allowed.

What can a young investigator do to avoid this unpleasantness? First, be honest. I do not intend this to be facetious. Begin by avoiding work that you believe is clouded by proprietary interests. Avoid contract work to fill our your salary or the department's budget. I say this recognizing that this is a difficult imperative, particularly for young investigators in difficult funding times, but much of this work can carry pressure, even if unstated, to find a certain effect. Recognize the pressure that accompanies the need to produce a publishable study or a given effect. Evaluate what the cost to you might be. In choosing a mentor, select one whose value system places honest science over publishable results.

Discuss with your associates steps to take to minimize bias, conscious or unconscious. Consult a good biometrician or epidemiologist about these questions early in the planning of the project. Record these discussions in a bound book. Remember that years later you may be asked to defend your choices of methods. Keep your data in two secure places, and document the means taken to find, classify and scale subjects and any changes in protocol. In a recent paper, Freedland and Carney¹⁵ polled a group of highly regarded investigators and found that a majority had trouble recalling the methods used to classify patients. Keep minutes of staff meetings, and document discussion of problems. Consult with experts in the difficult methodological areas. Ask them for written comments. Be skeptical of your conclusions. Write up and submit negative studies for publication. Be modest in your claims.

Finally, work to reform the system at every level. Discuss these issues in research conferences, at institutional review board meetings, and at meetings of

scientific societies. Do not avoid difficult areas of investigation. Take risks. If scientists exclusively choose the safe routes, avoid controversial research problems, and play only minor variations of someone else's themes, they voluntarily turn themselves into technicians. Our craft will indeed be in peril. Find and nurture good colleagues who will insist on the best from you, tell you when you are wrong, and stand with you in a difficult time. They are truly treasures, and their friendship will endure and sustain you past all confusion and pain. This article is a deeply felt thank you note to the many valued men and women who did precisely that for me.

REFERENCES

1. Needleman HL, Tuncay O, Shapiro IM. Lead levels in deciduous teeth of urban and suburban American children. *Nature*. 1972;235:111-112
2. Byers K. *Memoirs*. 1986
3. Graebner W. Hegemony through science: information engineering and lead toxicology 1925-1965. In: Rosner D, Markowitz G, eds. *Dying for Work: Workers' Safety and Health in Twentieth-Century America*. Bloomington, IN: Indiana University Press; 1987
4. Rosner D, Markowitz G. A 'Gift of God'? The public health controversy over leaded gasoline during the 1920s. *Am J Public Health*. 1985;75:344-352
5. Needleman HL, Gunnoe C, Leviton A, et al. Deficits in psychological and classroom performance of children with elevated dentine lead levels. *N Engl J Med*. 1979;300:689-695
6. Perino J, Ernhart SB. The relation of subclinical lead exposure to cognitive and sensorimotor impairment in black preschoolers. *J Learn Disabil*. 1974;7:616-620
7. Ernhart CB, Landa B, Schell NB. Subclinical levels of lead and developmental deficit: a multivariate follow-up reassessment. *Pediatrics*. 1981;67:911-919
8. Ernhart CB. Testimony before the Environmental Protection Agency. April 15, 1982
9. Environmental Protection Agency. *Review of the National Ambient Air Quality Standards for Lead: Assessment of Scientific and Technical Information*. Office of Air Quality Planning and Standards; March 1989
10. Scarr S. Transcript of Grand Rounds, Massachusetts Mental Health Center. October 30, 1991

11. Transcript of Scientific Integrity Hearings, University of Pittsburgh, PA. April 13-14, 1992
12. Hilt P. Hearing is held on scientist's lead-poison data. *New York Times*. April 15, 1992
13. Putka G. Professor's data on lead levels cleared by panel. *Wall Street Journal*. May 27, 1992
14. Palca J. Lead researcher confronts accuser in public hearing. *Science*. 1992;256:437-438
15. Freedland KE, Carney RM. Data management and accountability in behavioral and biomedical research. *Am Psychol*. 1992;47:640-645

Editor's Note

I asked Dr Needleman to write up his experience with the court system and the National Institutes of Health Office of Scientific Integrity. I tried to follow this case in the press, but I didn't find this very satisfactory. If you're searching for truth you rarely find it in newspapers. Now that I've read Dr Needleman's story I have a clearer idea of his ordeal, but I am confused. Dr Needleman believes he has been found not guilty. The government (Environmental Protection Agency) and other scientists also believe this, but others may not (see page 978, the preliminary report of the Inquiry Panel).

How long must this go on? Has Dr Needleman been victimized over a difference of opinion about the quality of his science?

Editors are exposed daily to conflicting opinions. It has never occurred to me to take such matters to court to be settled! Conflicting opinions are common and very important in science. Truth doesn't emerge easily. Many studies are often needed before one side convinces the other that they are right. Scientific debates can't be settled in courts!

I expect that we will hear the opinions and viewpoints of others about this in our Letters to the Editor column in the next issue of *Pediatrics*.

J. F. L., MD

WRITING AND THINKING

As anyone knows who has ever sat down to write, writing is thinking. The thought not only precedes the word, it follows it too: we do not know what we mean to say until, after many trials and errors, we have found the words. The purpose of writing well is thinking well.

Raimi J. By any other name. *Columbia Magazine*. 1991; (Spring):36-37.

Submitted by Student

Salem Comes to the National Institutes of Health: Notes From Inside the Crucible of Scientific Integrity

Herbert L. Needleman

Pediatrics 1992;90:977-981

Updated Information & Services	including high-resolution figures, can be found at: http://www.pediatrics.org
Citations	This article has been cited by 3 HighWire-hosted articles: http://www.pediatrics.org#otherarticles
Permissions & Licensing	Information about reproducing this article in parts (figures, tables) or in its entirety can be found online at: http://www.pediatrics.org/misc/Permissions.shtml
Reprints	Information about ordering reprints can be found online: http://www.pediatrics.org/misc/reprints.shtml

American Academy of Pediatrics

DEDICATED TO THE HEALTH OF ALL CHILDREN™

