COMMENTARY

Science, Public Policy, and a Critic's Dilemma

Sandra Scarr, Ph.D.*

Prevention of environmental harm to children is a worthy objective, widely shared by those with different views on, and participation in, controversies about low-level lead research. In the September 1992 issue of *The PSR Quarterly*, Needleman [1] reiterated his misrepresentations of his own and others' research results, failed to note the many experts who disagree with his conclusions, and confused public policy issues. Paul Mushak [2], in the same issue, lauded Needleman's 1979 research, although he was a signatory to the unanimous 1983 EPA Expert Committee report that dismissed Needleman's 1979 paper for faulty methods and improper conclusions. Needleman's colleagues at the University of Pittsburgh have found him guilty of deliberate misrepresentations [3] in his publications about low-level lead effects. Science can inform public policy only if reliable results are accurately presented in methodologically sound reports. Premature claims of scientific concensus are not acceptable, even if they influence health policies in directions one advocates.

EVOLUTION OF LEAD POISONING

Lead poisoning is no longer a diagnosis; it is a phrase that has been redefined into a multitiered condition. In 1975, recommendations were made for regular monitoring of asymptomatic children with blood lead levels between 30 µg/dl and 80 µg/dl, and environmental, and dietary measures were recommended. In 1985, the level of concern in children was lowered to 25 µg/dl, and in 1991 the CDC proposed that the safe blood lead level be lowered to 10 µg/dl. Children with lead levels above 10 µg/dl began to be called lead poisoned by regulatory agencies and their experts.

Blood lead surveys from 1935 to 1987 in the United States show that average blood lead levels dropped considerably over this period. Although the surveys were taken in different parts of the country, using different methods of assaying lead levels, the results are remarkably consistent in demonstrating a secular decline in blood lead levels from an average of about 30 µg/dl of blood in 1935 to about 7 µg/dl in 1987, a fourfold decrease. The current national health survey of children is expected to show even lower national lead levels in the 1990s.

The redefinition of lead poisoning relied on expert advice from Needleman, Mushak, some other U.S. investigators, and environmental activists, whose recent conclusions are that no level of lead exposure is safe. By relying on these consultants, the Centers for Disease Control (CDC), the Environmental Protection Agency (EPA), and other regulatory agencies have erroneously concluded that studies of low levels of lead exposure yield consistent, coherent findings of adverse effects on neurobehavioral development [4,5]. The agencies also relied on a Needleman study whose reports are not credible [3,6-8].

Needleman's claim that "the lead industry" raised issues about his lack of scientific integrity through the EPA Expert Committee in 1983 [6] and again through the Office of Scientific Integrity of the National Institutes of Health in 1991 is false. Despite Needleman's oft repeated charges, I have never received research support from any entity connected in any way with lead. I have never studied lead

1051-2438/1993/0301-0027$03.00/0

*SS is Commonwealth Professor of Psychology, University of Virginia, Charlottesville, Virginia. Address correspondence to Sandra Scarr, Ph.D., Department of Psychology, Gilmer Hall, University of Virginia, Carlottesville, VA 22903-2477

effects on children, and I do not speak for the lead industry. Because of, and subsequent to, my 1983 involvement in the EPA review of Needleman's data and of low-level lead research, I was asked in 1990 to testify about low-lead effects on children's neurobehavioral development. My appraisal of the literature agrees far more with those of Fulton [9], Smith [10], and Winneke [11] than with Needleman. As do several other reviewers of the research evidence, I have doubts that firm conclusions can be drawn from low-level lead research, and doubts about the validilty of Needleman's evidence and the soundness of his policy recommendations.

The Expert Committee never publicly questioned Needleman's integrity and limited its comments to flaws in the research, identifying five categories of error in Needleman's methodology—in calculation, subject selection, statistical analyses, missing data, and lack of adequate control for confounding variables. In the September 1992 issue of *The PSR Quarterly*, Needleman claims these errors were fixed in his reanalyses and in those done in 1991 at EPA. One can readily see that only two types of error could be fixed (calculations and analyses). Errors of subject omission, missing data, and omissions of adequate controls for confounding could not be addressed after data collection was complete. The Committee concluded with the following statement:

> In summary, at this time, based on questionable 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 reported in the Needleman et al. (1979) paper, neither support nor refute the hypothesis that low or moderate levels of Pb exposures lead to cognitive or other behavioral impairments in children [6]

No members of this Expert Committee, including me, had any connection with lead research or the lead industry. Despite his current praise of Needleman's research, Paul Mushak, a member of the EPA Expert Committee, signed the unanimous Committee report.

Needleman's oft published Teacher Ratings' graph [1] is again presented with no controls for any confounding variables, although the text does not make this important point clear. This graph lacks any control for socioeconomic status, parental IQ, education, family income, etc. As will be noted, most cross-sectional studies report some association between lead level and neurobehavioral outcomes *before* important covariates have been controlled. Adequate controls, as noted by Fulton [9], Smith [10], and Winneke [11], typically reduce lead effects to nil or to very small effects that are often not statistically reliable.

The probabilities Needleman presented [1] for graduation from high school are based on samples, including seven children with reported plumbism, of 0 to 7 cases. From this he claims a sevenfold increase in high school drop-out as a result of low-level lead exposure. Given the inadequate controls for confounding in this study, a more modest conclusion seems warranted.

PROBLEMS OF NONEXPERIMENTAL RESEARCH

The major problem in all field research, where the investigator cannot exercise experimental control over the application of the "treatment" (in this case lead exposure), is that the treatment is not randomly assigned. This problem is not unique to lead research; it affects all kind of real-world investigations from the effects of child care arrangements and schools to studies of work environments. People are not randomly assigned to these treatments; they carry into the treatment their own characteristics. Lead exposure covaries with respect to many other parental and child characteristics and characteristics of their environments that independently affect the same, adverse child outcomes that investigators hypothesize are due to lead exposure.

Higher lead exposure, hypothesized to decrease IQ and other measures of neurobehavioral functioning, is correlated with lower parental occupational status, educational achievements, income, IQs, nutrition, quality of parent-child interaction, stimulation in the home, and so forth, all of which have been shown to have negative effects on children's development, in the absence of lead exposure. In other words, the child's heredity and environment are correlated with lead exposure such that it is extremely difficult, if not theoretically impossible, to disentangle the effects of the child's broader biology and ecology from lead exposure. For example, lower IQ mothers tend to be poorer housekeepers, so that their children are more likely to be exposed to more lead [12]; they also have, on average, lower IQ children, live in poorer circumstances in general, and provide a less stimulating environment for the child.

Only if there were powerful and consistent results across correlational studies of diverse populations should one attribute causality to differences in lead exposure, apart from the other, well-studied heredi-

tary and environmental effects on children's behavioral development. Two kinds of studies need to be considered: Cross-sectional and prospective, longitudinal designs. Cross-sectional studies collect data at one point in time, whereas longitudinal studies measure the same persons on two or more occasions. Prospective, longitudinal studies may permit stronger inferences about developmental consequences of lead exposure and are therefore preferable to cross-sectional research designs.

CROSS-SECTIONAL RESEARCH: OTHER INVESTIGATORS' CONCLUSIONS

The consensus about low-level lead effects to which Mushak referred [2] consists only of U.S. governmental agencies for which he and Needleman consult and for whose reports they are authors. Needleman continues to misrepresent the low-level lead research literature. His selective appraisals and conclusions differ significantly from those of other leading investigators, particularly in Europe. Three notable examples are Drs. Mary Fulton, Marjorie Smith, and Gerhard Winneke. A major review of the world literature ended with:

> Studies of the effects of lead on children are, for necessary ethical reasons, nonexperimental in design. Such studies, particularly when they involve multivariate social data, pose problems in drawing causative inferences, and lead studies are no exception. Despite the wealth of data, and including data from experimental studies with animals, it is still not possible to conclude with any certainty that lead at low levels is affecting performance or behaviour of children. It is clear, however, that any differences in a measure such as IQ, which may be attributed to lead are likely to be small, accounting for about 1% or 2% of the variance [10].

Others' summaries of issues and conclusions agree with Smith, not Needleman or Mushak. Most cross-sectional studies find an association between higher levels of lead exposure in the absence of any controls for other variables. Once confounded variables, such as family socioeconomic status, quality of home environment, and parental IQ, are entered, most find either no association between higher lead levels and poorer behavioral development, or at most a very small effect, less than 1% the variance in children's scores.

In the large Edinburgh study, with a sample of 501, Fulton [9] could reliably detect very small effects of variation in blood lead levels, after careful control of confounding. She was especially sensitive to Needleman's issue of over-control. Her conclusions were modest and scientifically appropriate:

> Although the coefficients of blood-lead are significant, the size of the effect on the scores is small compared with the effect of other factors. For the BASC (British Ability Scale Combined) score, only 0.9% of a total of 45.5% variance explained by the covariates in the optimal regression model can be attributed to the effects of lead [9].

The most important cross-sectional study in the past few years is the *The World Health Organization Study*. The neuropsychologic status and blood lead levels of 1,879 children in eight European countries were sampled by Winneke and colleagues [11]. Blood lead levels ranged from less than 5 µg/dl to more than 60 µg/dl. Neurobehavioral and WISC results were very small and inconsistent across countries.

> [The Figure] depicts WISC results (mean of four subtests) from all the individual studies. Except for group 1[1], no systematic decline of WISC standard scores with increasing blood-lead concentrations can be inferred from [the figure]. No significant associations with blood lead were found for less standardized measures of visual-motor integration, namely the Trail-Making test, and for ratings of general behavior by parents and teachers. The strongest and most consistent effects were observed in established clinical tests of visual-motor integration, namely both GFT versions, as well as in serial choice reaction performance, namely the Vienna Reaction Device. The degree of association was sig-

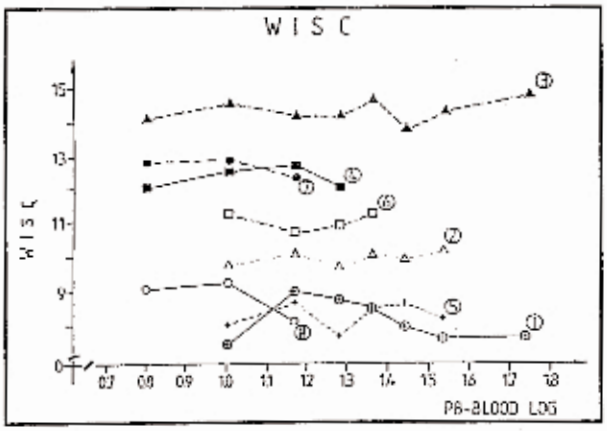

FIGURE. Dose-response information for individual studies with WISC results (mean of four subtests) as the behavioral endpoint. The abscissa is log PbB (µg/dl) [11].

[1] Data from the one sample that showed results were published separately (Hatzakis A, Kokkevi A, Maravelias S, et al., 1989, see ref. 11); the other seven studies with no results have not been published.

nificant or highly significant for these variables, although the variance explained by PbB (r^2) never exceeded 0.8%. Psychometric intelligence, as assessed by means of the WISC was also affected by lead exposure, although the effects were inconsistent across studies and the overall degree of association was only of borderline significance [11].

Variation in lead levels from <5 µg/dl to >60 µg/dl accounted for less than 1% of variation in IQ or any other neurobehavioral measures, including teacher and parent ratings of behavior problems. Even with very large samples and poor control of confounding variable, WHO study results do not support Needleman's or Mushak's conclusions. The cross-sectional literature suffers from inconsistent results and, as described earlier, problems of confounding the many family determinants of behavioral development with lead exposure.

THE PROSPECTIVE STUDIES

Five U.S. and Australian prospective studies were begun in the early 1980s to address the childhood effects of low levels of lead exposure. The studies, from which dozens of reports have been published, are known by their locations and major investigators: Boston (Bellinger and Needleman), Cincinnati (Dietrich), Cleveland (Ernhart), Port Pirie (McMichael), and Sydney (Cooney and McBride). The five studies are similar in design; all collected blood lead levels in the prenatal or natal period or both and at intervals during the infant and preschool years. Several have also reported dentine lead measures. Similar measures of cognitive development were used. Potentially confounding variables, such as parental IQ, socioeconomic status, and quality of the caretaking environment, were used as covariates in analyses. Not all collected parental IQ data, however. It was expected that these sizeable studies would yield consistent effects. Instead, the studies have published selected and very mixed results. Overall, they have obtained only chance levels of findings (approximately 5% of results "significant" at the 5% level of confidence), and the studies do not replicate each others' results.

A recent report [13] from the Port Pirie study of WISC-R scores at age 7 reports that no antenatal blood leads or blood lead values after the age of 4 were related to IQ results, but that cumulative lead exposure measured from birth to 15 months, birth to 2, 3, and 4 years were associated with decrements in

Table: Summary of Lead Effects in Five Prospective Studies[1]

Pb Sample (N studies)	Infant Mental Development[2] 6	12	24 Months	Preschool Mental Development 4-5 Years McCarthy Scales[3](3) GCI	V	P-P	Q	M	MO	Kaufman Battery/WPPSI[3](2) MPC	SEQ	SIM	NV	ACH
Maternal(4)	2	1	1	0	0	0	0	0	0	0	0	0	0	0
Cord(5)	1	1	1	0	0	0	0	0	0	0	0	0	0	0
Delivery(1)	0	0	1	0	0	0	0	0	0	0	0	0	0	0
First 6 mos.(5)	1	0	0	1	0	0	0	0	0	1	1	1	1	1
First Year(4)[4]	–	0	0	0	0	0	0	0	0	0	0	0	0	0
Second Year(5)[4]	–	–	0	1	0	1	0	1	0	0	0	0	0	0
Third Year(5)	–	–	–	2	0	2	0	0	0	0	0	0	0	0
Fourth Year(4)	–	–	–	0	0	0	0	0	0	0	0	0	0	0
Fifth Year(4)	–	–	–	0	0	1	0	0	0	0	0	0	0	0
All[5] current and prior PbB	–	–	–	1	0	1	0	0	0	0	0	0	0	0

[1] The studies are known by their locations: Boston, Cincinnati, Cleveland, Port Pirie, and Sydney. Entries are the number of the five studies that reported a statistically reliable result at alpha > .05.

[2] All studies used the Bayley Infant Test of Mental Development. One used the KID Test and obtained positive results at 6 months for maternal blood lead and results contrary to the hypothesis for infant 6-months blood lead.

[3] It should be remembered that the subtest scales are all positively correlated with each other about .40 and correlated with the total CGI about .70, so that results are not independent. The same is true for the Kaufman Battery and the WPPSI.

[4] One study sampled twice; the results are combined here.

[5] Other attempts were made to integrate blood levels across 6 to 57 months,12 to 57 months, 18 to 57 months, and 24 to 57 months, with no positive results.

verbal and full-scale IQ scores. Of 10 WISC-R subscales, the only statistically significant scale differences were found for on Performance scale, Block Design, and one Verbal scale, Information. The effects were larger for girls than for boys. It should be noted that the vast majority of children in the Port Pirie cohort are considered "lead poisoned" under the new CDC guidelines; the 2-year blood lead values for the four PbB quartile groups in this study average 13.0 µg/dl, 18.6 µg/dl, 24.2 µg/dl, and 33.5 µg/dl. The authors conclude that "the deleterious effects of lead are not large, and that only a small fraction of the overall variation in IQ can be attributed to lead exposure" [14]. Nonetheless, as Mushak noted, there may be possible social consequences for a population if such levels of blood lead do have even small effects on individuals.

To provide an overview of the five prospective studies, Ernhart [15] reviewed all reports available for each study. All lead measures and behavioral outcome measures collected longitudinally from antenatal to age 6, and reported to early 1992, were organized in a matrix for each study. A summary of results reported as statistically reliable at the $p = .05$ level appear as a proportion of total possible results in the Table.

First, it is apparent that only 4.7% (25 of 530) of possible results obtained by the five studies reached statistical significance at the 5% level, entirely chance findings. Published reports often ignore the many measures and analyses that raise the probabilities of finding "significant" results by chance. Positive results that Needleman reports [1] for the Boston and Port Pirie studies must be evaluated in the larger context of all the data collected and, presumably, analyzed. Second, the statistically significant results that are reported are inconsistent across the five studies. In no case do more than two of the five studies report the same results; most predicted effects of lead are not obtained; and published reports consistently highlight the few positive results to the exclusion of null results. In a recent evaluation of methodology conducted by scientists at the CDC and the Center for Environmental Health and Injury Control, the prospective studies judged highest in quality had fewer positive results than those lower in overall quality [14].

META-ANALYSIS

It has been argued that all of the reported results from cross-sectional and prospective studies can reasonably be combined by meta-analysis, which sums effects across studies. Winneke, however, noted that the eight WHO studies' results do not fit Needleman's claims from his meta-analysis [11]. The major problem with the meta-analysis conducted by Needleman and Gatsonis [16] is that they lumped the divers ages at which blood and tooth leads were collected and combined across developmental measures, as though they showed a consistent picture.

CDC scientists [14] reported recently that meta-analysis could not be applied to 35 reports from the five prospective studies, summarized above. Meta-analysis could not be applied because of inconsistencies in the studies methods and inconsistencies in their results. "Consequently, definitive conclusions regarding the effects of low-level body burdens of lead could not be determined from the longitudinal data" [14]. These are the same conclusions that I reached in 1983 and again in 1990, after examining the cross-sectional and longitudinal research.

SCIENCE AND PUBLIC POLICY

Science requires dispassionate scrutiny [17]. Advocacy, even for the perceived public interest, is an uncomfortable companion to science, because it may pit social change against research findings. Public policies should be informed by good science, fairly reported and evaluated. Consideration of low-level lead effects in a dispassionate perspective could lead one to discount individual effects, as Mushak did, and to focus on populations. If the "real" effect of low- to moderate-level lead on IQ and other behavioral measures is from 0.0% to 1.0% of the variance, then the cost of lead abatement can be rationally compared to the benefit of improved performance in a population. The costs and benefits of lead abatement can be compared to other public programs, such as prenatal maternal care, child inoculations, and Head Start.

One *can* be a scientist-advocate without doing more harm than good to science or the public interest. Perhaps the hardest part is to persist in presenting the science as I see it, which seems to set me against my own environmental concerns. Like readers of this journal, I want to reduce environmental hazards to children, but my evaluation of the research does not lead me to feel alarmed over current levels of lead in the population of the United States. ❧

REFERENCES

1. Needleman HL. Childhood lead poisoning: man-made and eradicable. PSR Quarterly 1992;2:130-134.

2. Mushak P. The landmark Needleman study of childhood lead poisoning: scientific and social aftermath. PSR Quarterly 1992;2:165-170.
3. Herbert L. Needleman vs. Bernadine Healy et al. Amended complaint for declaratory and injunctive relief and damages. In the United States District Court for the Western District of Pennsylvania, Civil Action No. 92-0749, October 27, 1992.
4. Centers for Disease Control. Preventing lead poisoning in young children: a statement by the Centers for Disease Control. Atlanta, GA: US Department of Health and Human Services, 1991.
5. Agency for Toxic Substances and Disease Registry. The nature and extent of lead poisoning in children in the United States: a report to Congress. Atlanta, GA: US Department of Health and Human Services, 1988.
6. Expert Committee on Pediatric Neurobehavioral Evaluations. 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. Washington, DC: Environmental Protection Agency, 1983.
7. Inquiry Panel Final Report, University of Pittsburgh, December 18, 1991, available from the Office of Research Integrity, University of Pittsburgh, Pittsburgh, PA 15213.
8. Ernhart CB, Scarr S. Report on the Research of Dr. Herbert Needleman Based on Samples Reported in the 1979 Article in the N Engl J Med, October 1990 (available from authors).
9. Fulton M, Thompson G, Hunter R, Raab G, Laxen D, Hepburn W. Influence of blood lead on the ability and attainment of children in Edinburgh. Lancet 1987;8544:1221-1226.
10. Smith M. The effects of low-level lead exposure on children. In: Smith MA, Grant LD, Sors AI, eds. Lead exposure and child development: An international assessment. Boston: Kluwer Academic Publishers, 1989;3-47.
11. Winneke G, Brockhaus A, Ewers U, Kramer U, Neuf M. Results for the European multicenter study on lead neurotoxicity in children: implications for risk assessment. Neurotoxicol and Teratol 1990;12:553-559.
12. Hawk BA, Schroeder S, Robinson, et al. Relation of lead and social factors to IQ of low-SES children: a partial replication. American Journal of Mental Deficiency 1986;91:178-183.
13. Baghurst PA, McMichael AJ, Wigg NR, et al. Environmental exposure to lead and children's intelligence at the age of seven years. N Engl J Med 1992;327:1279-1284.
14. Thacker SB, Hoffman DA, Smith J, Steinberg K, Zack M. Effect of low-level body burdens of lead on the mental development of children: limits of meta-analysis in a review of longitudinal data. Arch Environ Health 1992; 47:336-346.
15. Ernhart CB. A critical review of low-level prenatal lead exposure in the human: 2. Effects on the developing child. Reproductive Toxicology 1992;6:21-40.
16. Needleman HL, Gatsonis G. Low-level lead exposure and the IQ of children: a meta-analysis of modern studies. JAMA 1990;263:673-678.
17. Marshall E. When does intellectual passion become conflict of interest? Science 1992;257:620-621.