ECAO-CD-81-2
IIA.K.2

EPA-600/8-83-028A

UNITED STATES ENVIRONMENTAL PROTECTION AGENCY
Environmental Criteria and Assessment Office (MD-52)
Research Triangle Park, North Carolina 27711

DATE: November 30, 1983

SUBJECT: Corrigenda for the First External Review Draft of 1983 Revised EPA Criteria Document, Air Quality Criteria for Lead

FROM: Lester D. Grant, Director *Lester D. Grant*
ECAO/RTP, U.S. EPA (MD-52)

TO: Recipients of the subject first external review draft of the 1983 revised EPA Lead Criteria Document

Copies of the first External Review Draft for the EPA Document Air Quality Criteria for Lead were recently made available for a ninety-day public comment period (October 15, 1983 - January 15, 1983), as announced in the Federal Register.

The External Review Draft of the Lead Document circulated to you and other recipients did not contain Appendix 12-C, a detailed report of the Expert Committee on Pediatric Neurobehavioral Evaluations (which was convened by ECAO/RTP to provide in-depth evaluations of studies of associations between low-level lead exposure and neuropsychologic deficits in children reported by Drs. Ernhart, Needleman, and their respective colleagues). Copies of Appendix 12-C were withheld pending reconvening of the "Neurobehavioral Evaluations" Committee in order to consider comments by Drs. Ernhart and Needleman on a preliminary draft of the Committee's report and in order for the Committee to take into account newly available published and unpublished information pertaining to the subject studies in carrying out final revision of their report. A copy of the Committee's final report is enclosed for insertion as Appendix 12-C following other Chapter 12 materials in the External Review Draft of the Lead Document recently provided to you.

Another committee was convened by ECAO/RTP in late September to evaluate certain German studies (by Drs. Kirchgessner and Reichlmayr-Lais) reporting evidence interpreted by those investigators as being indicative of beneficial effects of lead at very low exposure levels. The "Essentiality" Committee has recently completed their report evaluating the subject studies, and the report is enclosed herein. That report constitutes Appendix 12-A and is to replace the existing critique of the subject studies, which appeared as Appendix 12-A in the recently circulated External Review Draft of the Lead Document.

In addition to the above appendices, a corrigenda is also enclosed which lists corrections/notations for text and tables in the circulated four volumes of the subject draft Lead Criteria Document. The corrections noted are restricted to those thought to be crucial for accuracy or understanding of the information presented and to indicate changes apropos to the insertion of the above new appendices. Minor typographical errors or editorial changes are generally not included. There are no changes indicated to be made for Chapters 2-6 or 8 for the document.

We apologize for the unfortunate delay in our being able to circulate the above Appendix materials (for both Appendix 12-A and 12-C) to you. We also recognize the importance of their contents in terms of their crucial utility in helping to resolve certain key issues of much relevance for the development of criteria and standards for lead. In view of our delay in circulating these important materials, we are currently processing a Federal Register notice announcing a one-month extension of the Public Comment Period for that draft document to February 15, 1984. We do not anticipate any problems in having that extension and its announcement approved in time for publication in the Federal Register during the first or second week of December, 1983. We hope that this information will assist in your planning of work efforts connected with preparation of public comments.

First External Review Draft for EPA
Lead Criteria Document (1983): Corrigenda

Chapter 1

Page	Line	Correction/Notation
1-35	Table 1-4 footnote.	References cited in Table 1-4, but not provided in Section 1.14 (reference list for Chapter 1), are available in the cited Nriagu (1978b) paper.
1-36	6-7 up*	Delete "..., due complex" from end of sentence.
1-37	9 up	Replace "...mg/g..." with "...μg/g..."
1-42	13 up	Change to read: "...lead contribute differently to each of these dietary groups (Figure 1-1)."
1-48	2 up	Change to read: "...1971 Annual Report of the United Kingdom..."
1-50	1	Replace "...Association..." with "...Organization..."
1-52	23-24	Change to read: "...about 100 μg of lead is consumed daily by each American. For all Americans, this amounts to only 8 tons/year, or 0.001-0.01 percent of the total environmental contamination."
1-57	7 up	Change "...(Nriagu, 1978)" to read: "...(Nriagu, 1978a)"
1-65	1 up	Change "...(Nriagu, 1978)" to read: "...(Nriagu, 1978a)"
1-66	3,7,14 up	Change "mg" to "μg"
1-69	20	Change "primary" to "non-circumpulpal"
1-98	18	Change "1000 μg/dl" to "1000 μg/g"
1-116	2 up 1 up	Change "human" to "pediatric" Change "human studies" to "studies of children"
1-141	Table 1-20	Downward arrow should be inserted below "Vitamin D metabolism interference" entry, indicating that the vitamin D effect occurs down to 10-15 μg/dl blood lead.
1-52	12 up	"maxim safe level" should read "maximum safe level"

*Number of lines up from bottom of page (other entries are for number of lines from top of page).

Lead Document Corrigenda (continued)

Chapter 7

Page	Line	Correction/Notation
7-63	4	Change mm^3 to m^3
7-64	21	Change μ/m^3 to $\mu g/m^3$
7-69	1-2 up	Change last sentence of page to read: "For all Americans, this amounts to only 8 tons/year or 0.001-0.01 percent of the total environmental contamination."

Chapter 9

Page	Line	Correction/Notation
9-12	8	Change "secondary (circumpulpal)" to "circumpulpal"

Chapter 10

Page	Line	Correction/Notation
10-5	13 up	Change "often" to "after"
10-15	12	Line should read: "difficult to measure, and reliable values have become available only recently (see Chapter 9)."
10-20	19	Add "Shapiro et al., 1978" after "Winneke et al. 1981"
10-42	3-7 up	Sentence should read: "In a related study (Grant et al., 1980), rats were exposed to lead _in utero_, through weaning, and up to 9 months of age at the dosing range used in the Kimmel et al. study (0.5 to 250 ppm in the dams' drinking water until weaning of pups; then the same levels in the weanlings' drinking water). These animals showed a blood lead range of 5 to 67 µg/dl."
10-45	23	Line should read: "remains the one readily accessible measure that can demonstrate in a relative way the relationship of various effects to increase in exposure."

Chapter 11

Page	Line	Correction/Notation
11-104	17	Change "statistical relationship" to "significant relationship"
11-110	2 and 6 up	Change "per mg/g" to "per 1000 µg/g"

Chapter 11 (continued)

Page	Line	Correction/Notation
11-142	Table 11-59	Numerical table entries are in μg/dl units.
11D-23	Appendix D3	Insert revised Appendix D3 (attached single sheet) in place of comparable page in Appendix 11-D.

Chapter 12

Page	Line	Correction/Notation
12-1	19-23	Replace last two sentences of second paragraph with: "An evaluation of these studies by an expert committee convened by EPA in September, 1983, is contained in Appendix 12-A. The committee's report notes methodological problems with the studies, which preclude acceptance of the reported findings as demonstrating the essentiality of lead. These studies are, therefore, not considered further in the present document.
12-48	11 11-12 up	Change "(10 to 20)" to "(10 of 20)." Change "Reports of low blood levels..." to read "Reports of effects at low blood levels..."
12-60	4	Change "...history of pica, as well..." to read "...history of pica for paint and plaster, as well as..."
12-62	18-on	Replace the entire last paragraph of page 12-62 with: "The Perino and Ernhart (1974) and Ernhart et al. (1981) studies were evaluated by an expert committee convened by EPA in March, 1983. The committee's report (see Appendix 12-C) notes methodological problems which preclude acceptance of the analyses and findings published by Perino and Ernhart (1974) and Ernhart et al. (1981). The committee's report, further, recommends that the Ernhart data set be reanalyzed, including longitudinal analyses of data for subjects evaluated in both the Perino and Ernhart (1974) and Ernhart et al. (1981) studies. Pending resolution of methodological problems with the Ernhart data set and/or publication of adequate reanalyses, the subject studies are not considered further in this document."
12-64	5	Change "primary" to "non-circumpulpal"
12-65	7-on	After the first sentence ending with "...(see Appendix 12-C)," replace the rest of the paragraph with the following: "The committee's report notes methodological problems which preclude acceptance of the published analyses and findings reported either (1) by Needleman et al. (1979) or (2) in subsequent papers by Needleman

Chapter 12 (continued)

Page	Line	Correction/Notation
12-65	7-on	and coworkers concerning additional analyses of the same data set. The committee's report also recommends that the Needleman data set be reanalyzed. Pending resolution of methodological problems with the Needleman data set and/or publication of adequate re-analyses, the subject studies are not considered further in this document."
12-67	16 up	Change "-0.06" to "+0.06"
12-69	1 up	Add the following sentence: "The Landrigan et al. (1975) and McNeil and Ptasnick (1975) studies are, therefore, not considered further in this document."
12-83	Table 12-3	For Overmann (1977), Pb concentration should read: "5, 15, or 45 mg/kg." For Winneke et al. (1977), Pb concentration should read: "745 mg/kg (diet)."
12-84	Table 12-3	For Dietz et al. (1978), Pb concentration should read: "0.025%." For Cory-Schlecta and Thompson (1979), Pb concentration should read: "(1)0.0025, (2)0.015, or (3)0.05%." For Cory-Schelacta et al. (1981), Pb concentration should read: "(1)0.005 or (2)0.015%."
12-87	Table 12-3	For Milar et al. (1981), Pb concentration should read: "mg/kg b.w. (gavage)." For Nation et al. (1982), Pb concentration should read: "mg/kg b.w." For Winneke et al. (1982), Pb concentration should read: "0.08, 0.025, or 0.075%."
12-88	Table 12-3	Under Abbreviations, add: "b.w. body weight"
12-90	Table 12-4	For Rice and Willes (1979), Pb concentration should read: "μg/kg b.w." Under abbreviations, add: "b.w. body weight"
12-91	5 9	Change "or to "of" Change "change level" to "chance level"
12-101	Table 12-5	Under exposure protocol, the 3rd entry should end with "PND 20" rather than "PNDO"
12-115	3 5	Change "human" to "pediatric" Change "human studies" to "studies of children"
12-125	5	The 4th sentence of paragraph should read: "Proteinuria occurred in two patients."
12-216	1	Change "as much as" to "inasmuch as"

Chapter 12 (continued)

Page	Line	Correction/Notation
12-229	7	Change "human" to "pediatric"
	8	Change "human studies" to "studies of children"
12A-1	Appendix 12-A	Replace Appendix 12-A (pp. 12A-1 to 12A-7) with enclosed copy of "Essentiality" Committee's report labeled as Appendix 12-A.
12C-1	Appendix 12-C	Insert enclosed copy of "Neurobehavioral Evaluations" Committee's report labeled as Appendix 12-C.

Chapter 13

Page	Line	Correction/Notation
13-2	4 up	Delete "which" from line
13-22	5	Change "...about 2, may represent..." to "...about 2 μg/dl per 1000 μg/g may represent..."
13-24	Table 13-6	Entries for PbB values are in μg/dl units
13-25	Table 13-7	Entries for PbB values are in μg/dl units
13-26	Table 13-8	Entries for PbB values are in μg/dl units
13-32	Table 13-10	Under first column, "10" should be "10 μg/dl." Under last column, a downward arrow should be added immediately below the "Vitamin D metabolism interference" entry, to indicate that the vitamin D effect occurs down to 10-15 μg/dl blood lead.
13-44	12	In line 7 of the second conclusion, change "...lead contribution can be..." to read "...lead contribution to human blood lead levels can be..."
	1-2 up	Change "maxim safe level" to "maximum safe level"

Appendix D3

List of Attendees at March 10-11 and March 30-31, 1983
meeting of
NHANES II TIME TREND ANALYSIS REVIEW GROUP

Committee Members

Joan Rosenblatt (Chairman)
National Bureau of Standards

J. Richard Landis
University of Michigan

Roderick Little
Bureau of the Census

Richard Royall
John Hopkins University

Harry Smith, Jr.
Mt. Sinai School of Medicine

Invited Discussants

Joel Schwartz
U.S. EPA

J. Lee Annest
NCHS

Jean Roberts*
NCHS

James Pirkle
CDC

Vernon Houk†
CDC

Ben Forte
Ethyl Corporation

Chuck Pfieffer
DuPont

Ron Snee
DuPont

Asa Janney
ICF

EPA Staff

David Weil (Meeting Co-ordinator)
U.S. EPA

Dennis Kotchmar*
U.S. EPA

Vic Hasselblad
U.S. EPA

Allen Marcus
U.S. EPA

Observers

Earl Bryant*
NCHS

Trena Ezzote*
NCHS

Mary Kovar*
NCHS

Bob Casady*
NCHS

Robert Murphy
NCHS

Jack Pierrard*
DuPont

Kathryn Mahaffey*
FDA

*attended March 10-11 meeting only.
†attended March 30-31 meeting only.

☆U.S. GOVERNMENT PRINTING OFFICE: 1983—659-017/7239

EPA-600/8-83-028A

INDEPENDENT PEER REVIEW OF SELECTED STUDIES BY DRS. KIRCHGESSNER AND REICHLMAYR-LAIS CONCERNING THE POSSIBLE NUTRITIONAL ESSENTIALITY OF LEAD:

Official Report of Findings and Recommendations of an Interdisciplinary Expert Review Committee

Presented by

Expert Committee on Trace Metal Essentiality

to

Dr. Lester D. Grant, Director
Environmental Criteria and Assessment Office
United States Environmental Protection Agency
Research Triangle Park, North Carolina

November, 1983

The materials contained in this report were generated as the result of critical evaluations and deliberations by the members (listed below) of the Expert Committee on Trace Metal Essentiality. All members concur with and endorse the findings and recommendations contained in the present report as representing the collective sense of the Committee.

Dr. F. William Sunderman, Jr. (Chairman)
Professor, Departments of Laboratory Medicine and Pharmacology
University of Connecticut School of Medicine
Farmington, CT 06232

Dr. M. R. Spivey Fox
Chief, Nutrient Interaction Section
Division of Nutrition
U.S. Food and Drug Administration
Washington, DC 20204

Dr. Kathryn Mahaffey
Chief, Priorities and Research Analysis
National Institute of Occupational Safety and Health
Cincinnati, OH 45226

Dr. Forrest Nielsen
Research Chemist
Human Nutrition Research Center
U.S. Department of Agriculture
Box 7166 University Station
Grand Forks, ND 58202

Dr. Orville Levander
Research Chemist
Beltsville Human Nutrition Research Center
U.S. Department of Agriculture
Beltsville, MD 20705

Dr. Walter Mertz
Director, Beltsville Human Nutrition Research Center
U.S. Department of Agriculture
Beltsville, MD 20705

Dr. Ekhard Ziegler
Professor, Department of Pediatrics
University of Iowa Hospital
Iowa City, IA 52242

TABLE OF CONTENTS

PREFACE

The Expert Committee on Trace Metal Essentiality was appointed in August, 1983 by the Environmental Criteria and Assessment Office (ECAO) of EPA, to evaluate the studies of Drs. M. Kirchgessner and A. M. Reichlmayr-Lais on the possible nutritional essentiality of lead. The Committee was provided with all relevant papers by the authors, the critiques of these papers prepared by Dr. Paul Mushak in his capacity as a consulting author of the revised Air Quality Criteria Document for Lead, and all correspondence between ECAO and Drs. Kirchgessner and Reichlmayr-Lais. Attachment 1 contains a complete list of the materials reviewed by the Committee.

The Committee convened on September 29, 1983 at the ECAO facilities in Research Triangle Park, NC. Present at the meeting were all but one member of the Committee (K.M.), Drs. Anna Reichlmayr-Lais and E. Grassmann (substituting for Professor Kirchgessner), Dr. Mushak, EPA staff, and observers from various interest groups. A complete list of attendees may be found in Attachment 2.

Following a presentation by Dr. Reichlmayr-Lais, in which she reviewed her published data as well as experiments in progress, all meeting attendees were given an opportunity to address the Committee. The Committee then pursued specific lines of questioning to its satisfaction and retired to executive session to draft its final report.

The Committee was charged with critically evaluating the studies of Kirchgessner and Reichlmayr-Lais and determining whether or not they supported the concept of a nutritional essentiality of lead. Their findings and recommendations are contained in this consensus report; views expressed by the members of the Committee in this report are their own and are not necessarily those of the institutions with which they are affiliated.

INTRODUCTION

The Committee commends Drs. Kirchgessner and Reichlmayr-Lais for their pioneering work, which is at the frontier of current research on trace metal nutrition, and wishes to express their appreciation to them for their cooperation in the Committee's efforts to assess their findings.

CRITICAL COMMENTS

The Reichlmayr-Lais and Kirchgessner data that were available for review were derived from two experiments. Based upon the published and oral descriptions of the experiments, members of the Committee expressed reservations about specific facets of experimental design, execution, and documentation, including the following:

(1) No selenium or chromium was added to the basal diet, nor were the concentrations of selenium or chromium measured in the diet to indicate nutritional adequacy of those essential elements. (In discussion, Dr. Reichlmayr-Lais indicated that Se, Cr, and other essential elements are being added to the diets in two experiments that are currently in progress.)

(2) The sole source of fat in the basal diet was coconut oil, which might render the rats deficient in essential fatty acids. (In discussion, Dr. Reichlmayr-Lais indicated that linoleic acid is being added to the diets in the experiments that are in progress.)

(3) The possibility exists that chelant residues (EDTA and APDC) may persist in the basal diet despite the extensive extraction procedures that were employed. Documentation of the EDTA or APDC concentrations was lacking and the Committee considered that HPLC or radiotracer experiments would be advisable in order to address these concerns.

(4) As the basal diet was prepared, iron supplements were added in an aqueous mixture with several other inorganic ingredients (e.g., KI, $CuSO_4$). Under the conditions of drying at 50°C, oxidation-reduction reactions could occur that might affect the bioavailability of iron. The Committee considered that this potential problem should be addressed in future experiments.

(5) The method of blood collection by decapitation and draining into a test tube via a funnel raised concerns owing to the potential for contamination by other body fluids (e.g., gastric fluid, spinal fluid, lymph).

(6) The results of lead analyses of blood and tissues of the experimental animals have not been reported. (Dr. Reichlmayr-Lais indicated that such analyses are being attempted in current experiments.)

(7) The possibility that lead supplementation of the basal diet might affect its palatability was not addressed in the experiments, either by pilot trials or by measurements of food intake.

(8) Lead supplementation of the basal diet was performed only at a single (relatively high) concentration of 1 ppm. Further experiments at graded levels of Pb supplementation are desirable in order to establish a dose-effect relationship.

(9) The statistical methods that were used to analyze the data in the two experiments were not described in sufficient detail; the application of multiple t-tests may be a cause for concern, and the various reports contain inconsistencies in numbers of experimental animals per group. These matters might advantageously be clarified in a consolidated report of each experiment. (Dr. Reichlmayr-Lais indicated that such a consolidated report is in press.)

RECOMMENDATIONS AND CONCLUSIONS

In view of the concerns that are listed above, the Committee reached the following conclusions and recommendations:

1. The Kirchgessner and Reichlmayr-Lais data furnish evidence that is consistent with and, in some opinions, indicative of a nutritional essentiality of lead for rats.

2. The evidence is not sufficient to establish nutritional essentiality of lead for rats.

3. To address the basic issue of nutritional essentiality of lead, additional evidence needs to be obtained under different conditions in the laboratory of Kirchgessner-Reichlmayr-Lais, as well as by independent investigators; additional species should also be examined.

The Committee emphasizes the difference that apparently exists between lead concentrations that are of concern from a toxicologic viewpoint and those that might possibly be of nutritional concern. Hence the Committee does not perceive any practical incompatibility between (a) efforts to reduce Pb in the human environment to concentrations that are unassociated with toxic effects and (b) efforts to define the potential nutritional essentiality of lead. The Committee recognizes that current public health concerns for humans are those of lead toxicity.

ATTACHMENT 1

The following materials were considered by the Committee in their deliberations:

1. Reichlmayr-Lais, A. M. and Kirchgessner, M. (1981) Zur essentialitat von blei fur das tierische Wachstum. [Why lead is essential for animal growth.] Z. Tierphysiol. Tierernaehr. Futtermittelkd. 46:1-8.

2. Reichlmayr-Lais, A. M. and Kirchgessner, M. (1981) Depletions studien zur essentialitat von blei an wachsenden ratten. [Depletion studies on the essential nature of lead in growing rats.] Arch. Tierernaehr., 31:731-737.

3. Reichlmayr-Lais, A. M. and Kirchgessner, M. (1981) Eisenkupfer- und zinkgehalte in neugeborenen sowie in leber und milz wachsender ratten bei alimentarem blei-mangel. [Iron-, copper- and zinc contents in newborns as well as in the liver and spleen of growing rats in the case of alimentary lead deficiency.] Z. Tierphysiol. Tierernaehr. Futtermittelkd. 46:8-14.

4. Kirchgessner, M. and Reichlmayr-Lais, A. M. (1980) Lead deficiency and its effects on growth and metabolism. Presented at TEMA-4 Meeting; May; Perth, Australia.

5. Reichlmayr-Lais, A. M. and Kirchgessner, M. (1981) Activities-veranderungen verschiedener enzyme im alimentaren blei-mangel. [Activity changes of different enzymes in alimentary lead deficiency.] Z. Tierphysiol. Tierernaehr. Futtermittelk 46:145-150.

6. Kirchgessner, M. and Reichlmayr-Lais, A. M. (1981) Changes of iron concentration and iron-binding capacity in serum resulting from alimentary lead deficiency. Biol. Trace Elem. Res. 3:279-285.

7. Kirchgessner, M. and Reichlmayr-Lais, A. M. (1981) Retention, absorbierbarkeit und intermeditare neifugbarkeit von eisen bei alimentarem bleimangel. [Retention, absorbability and intermediate availability of iron in the case of alimentary lead deficiency.] Int. J. Vitam. Nutr. Res. 51:421-424.

8. Reichlmayr-Lais, A. M. and Kirchgessner, M. (1981) Katalase- und coeruloplasmin -acktivitat im blut bzw. serum von ratten in blei-mangel. [Catalase and coeruloplasmin activity in blood and serum of rats with lead deficiency.] Zentralbl. Veterinaermed. Reihe A 28:410-414.

9. Kirchgessner, M. and Reichlmayr-Lais, A. M. (1982) Konzentrationen verscheidener stoffwechsel-metaboliten im experimentellen bleimangel. [Concentration of different metabolites resulting from experimental lead deficiency.] Ann. Nutr. Metab. 26:50-55.

10. Reichlmayr-Lais, A. M. and Kirchgessner, M. (1981) Hematologische veranderungen bei alimentarem blei mangel. [Hematological changes in the case of alimentary lead deficiency.] Ann. Nutr. Metab. 25:281-288.

11. Schwarz, K. (1973) New essential trace elements (Sn, V, F, Si): progress report and outlook. Proceedings International Conference Trace Element Metabolism in Animals (TEMA) II. Madison, Wisconsin. Edited by W. G. Hoekstra, J. W. Suttie, H. E. Gantner, and W. Mertz, University Park Press, Baltimore, MD.

12. Pallauf, J., and Kirchgessner, M. (1971) Herstellung der gereinigten halbsynthetischen diat. [Production of the purified semi-synthetic diet.] Z. Tierphysiol. Tierernaehr. Futtermittelkd. 23:128-139

13. Schnegg, A. (1975) Dissertation, T. U. Munchen. Excerpt on diet from Mr. Schnegg's dissertation, provided by Professor Kirchgessner.

14. Kirchgessner, M. and Schwarz, W. A. (1976) Zum einfluss von zinkmangel und unter-schiedlichen zinkzulagen auf resorption und retention des zinks bei milchkühen. [Concerning the influence of zinc deficiency and different zinc additions on resorption and retention of zinc in milk cows.] Arch. Tierernaehrung. 26:3-16.

15. Mertz, Walter (1981) The essential trace elements. Science (Washington, D.C.) 213:1332-1338.

16. Mushak, P. (1982) [Appendix 11-A, draft Air Quality Criteria Document for Lead]. August 9. Assessment of studies reporting the potential essentiality of lead. Available for inspection at U.S. Environmental Protection Agency, Environmental Criteria and Assessment Office, Research Triangle Park, N.C.

17. Kirchgessner, M. and Reichlmayr-Lais, A. M. (1982) [Rebuttal to Appendix 11-A]. September 2. Available for inspection at U.S. Environmental Protection Agency, Environmental Criteria and Assessment Office, Research Triangle Park, N.C.

18. Weil, D. (1982) [Letter to M. Kirchgessner]. October 14. Available for inspection at U.S. Environmental Protection Agency, Environmental Criteria and Assessment Office, Research Triangle Park, N.C.

19. Kirchgessner, M. (1982) [Reply to D. Weil]. October 26. Available for inspection at U.S. Environmental Protection Agency, Environmental Criteria and Assessment Office, Research Triangle Park, N.C.

20. Mushak, P. (1983) [Appendix 12-A, draft Air Quality Criteria Document for Lead]. January 5. Assessment of studies reporting data regarding the potential essentiality of lead. Available for inspection at U.S. Environmental Protection Agency, Environmental Criteria and Assessment Office, Research Triangle Park, N.C.

21. Grant, L. D. (1983) [Letter to M. Kirchgessner]. February 15. Available for inspection at U.S. Environmental Protection Agency, Environmental Criteria and Assessment Office, Research Triangle Park, N.C.

22. Kirchgessner, M. (1983) [Reply to L. D. Grant]. March 28. Available for inspection at U.S. Environmental Protection Agency, Environmental Criteria and Assessment Office, Research Triangle Park, NC.

23. Grassmann, E., Kirchgessner, M. and Hampel, G. (1970) Zur kupferdepletion bei ratten und kuten mit athylendiamin-tetraazetat und adenin. [Copper depletion in rats and chicks as produced by ethylenediamine-tetraacetate and adenine.] Arch. Tierernaehr. 20:537-544.

24. Grassmann, E. (1976). Zur verwertung verschiedener eisenverbindungen bei der ratte. [The utilization of various iron compounds in the rat.] Zentralbl. Veterinaermed. Reihe A 23:292-306.

25. Schnegg, A. and Kirchgessner, M. (1977) Zur differentialdiagnose von Fe- und Ni-mangel durch bestimmung einiger enzymabtivitaten. [Differential diagnosis of Fe and Ni deficiencies by determining some enzyme activities.] Zentralbl. Veterinaermed. Reihe A 24:242-247.

26. Schnegg, A. and Kirchgesser, M. J. (1977) Aktivitatsanderungen von enzymen der leber und xiere im nickel-bzw. eisin-mangel. [Changes in liver and kidney enzyme activities during nickel or iron deficiency.] Z. Tierphysiol. Tierernaehr. Futtermittelkd. 38:300-205.

27. Schnegg, A. and Kirchgessner, M. (1977) Konzentrationsanderungen einiger substrate in serum und leber bei Ni-bzw. Fe-mangel. [Concentration changes in some serum and liver substrates with Ni and Fe deficiency.] Z. Tierphysiol. Tierernaehr. Futtermittelkd. 39:247-251.

28. Schnegg, A. and Kirchgessner, M. (1977) Alkalische und saure phosphatase-ackivitat in leber und serum bei Ni-bzw. Fe-mangel. [Alkaline and acid phosphatase activity in the liver and serum with Ni versus Fe deficiency.] Int. Z. Vitam. Ernaehrungsforsch. 47:274-276.

29. Nielsen, F. (1983) [Letter to M. Davis]. May 19. Available for inspection at U.S. Environmental Protection Agency, Environmental Criteria and Assessment Office, Research Triangle Park, NC.

30. Mushak, P. (1983) [Appendix 12-A, draft Air Quality Criteria Document for Lead]. July 1. Assessment of studies reporting the potential essentiality of lead. Available for inspection at U.S. Environmental Protection Agency, Environmental Criteria and Assessment Office, Research Triangle Park, NC.

ATTACHMENT 2

List of attendees at September 29, 1983 meeting of the Expert Committee on Trace Metal Essentiality:

PANEL MEMBERS

Dr. F. W. Sunderman, Jr. (Chairman)
University of Connecticut, School of Medicine

Dr. Forrest Nielsen
USDA

Dr. M. R. Spivey Fox
FDA

Dr. Orville Levander
USDA

Dr. Walter Mertz
USDA

Dr. Ekhard Ziegler
University of Iowa

Dr. Kathryn Mahaffey*
NIOSH

INVITED DISCUSSANTS

Dr. Paul Mushak
University of North Carolina

Dr. Anna M. Reichlmayr-Lais
Technical University of Munich
Federal Republic of Germany

Dr. E. Grassmann (substituting for Dr. M. Kirchgessner*)
Technical University of Munich
Federal Republic of Germany

EPA STAFF

Dr. David Weil (Meeting Coordinator)
EPA/ECAO

Mr. Jeff Cohen
EPA/OAQPS

Dr. J. Michael Davis
EPA/ECAO

Dr. Robert Elias
EPA/ECAO

Dr. Lester Grant
EPA/ECAO

PUBLIC OBSERVERS

Dr. Gary Ter Haar
Ethyl Corporation

Dr. Elizabeth Lightfoot
Ethyl Corporation

Dr. Jerry Cole
ILZRO

Dr. Magnus Piscator
Karolinska Institute

*not present at meeting

☆U.S. GOVERNMENT PRINTING OFFICE: 1983—659-017/7241

ECAO-CD-81-2
IIA.K.2

EPA-600/8-83-028A

APPENDIX 12-C

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

Presented by

Expert Committee on Pediatric
Neurobehavioral Evaluations

To:

Dr. Lester D. Grant, Director
Environmental Criteria and Assessment Office
United States Environmental Protection Agency
Research Triangle Park, North Carolina

November 14, 1983

The materials contained in this report were generated as a result of critical evaluations and deliberations concerning the subject studies in the course of review of them by members of the Expert Committee on Pediatric Neurobehavioral Evaluations. The members of the Committee (listed below) unanimously concur with and endorse the findings and recommendations contained in the present report as representing the collective sense of the Committee.

Expert Committee on Pediatric
Neurobehavioral Evaluations

Dr. Lyle Jones,
Alumni Distinguished Professor
Dept. of Psychology and
Director, L. L. Thurstone
Psychometric Laboratory
University of North Carolina
Chapel Hill, NC 27514

Dr. Lloyd Humphreys, Professor
Dept. of Psychology and
Educational Psychology
University of Illinois
Champaign, IL 61820

Dr. Paul Mushak, Associate Professor
Dept. of Pathology and Co-Director,
Environmental Toxicology Research Program
University of North Carolina
Chapel Hill, NC 27514

Dr. Richard Weinberg, Professor
Dept.of Educational Psychology
and Co-Director, Center for
Early Education and Development
University of Minnesota
Minneapolis, MN 55455

Dr. Larry Kupper, Professor
Dept. of Biostatistics
School of Public Health
University of North Carolina
Chapel Hill, NC 27514

Dr. Sandra Scarr, Commonwealth
Professor, Dept. of Psychology
University of Virginia
Charlottesville, VA 22901

PREFACE

As part of the periodic (5-year) review and revision of criteria for the National Ambient Air Quality Standards (NAAQS) for lead established in 1978, the EPA Environmental Criteria and Assessment Office (ECAO/RTP) initiated in 1982 an intensive, critical evaluation of pertinent scientific information concerning health effects associated with lead (Pb) exposure. Of considerable importance in that regard are certain published (and related unpublished) studies from several different research groups, which provide data that have been interpreted as demonstrating significant associations between neuropsychologic deficits (e.g., impaired cognitive development) or other neurobehavioral effects (e.g., poorer classroom behavior) and lead exposures in otherwise apparently asymptomatic children. The findings and interpretation of such studies have become a matter of great controversy, especially among those research scientists directly involved in the conduct and reporting of the subject studies.

In an effort to resolve major points of controversy concerning some of the most important and controversial of the subject studies, an interdisciplinary Expert Committee on Pediatric Neurobehavioral Evaluations was convened by Dr. Lester D. Grant (Director of ECAO/RTP) starting in March, 1983, to provide independent peer review of selected studies and to make recommendations concerning how particular study results should be most appropriately interpreted or, possibly, reanalyzed before final interpretation. The Committee comprised internationally recognized experts in the areas of: child development, psychometric techniques, biostatistics, lead exposure measurement techniques, and overall aspects of lead pharmacokinetics and toxicology. The present report contains a series of critiques of interrelated sets of selected studies conducted during the 1970s and early 1980s.

The Committee focused on answering the following four general questions in reviewing each of the sets of studies:

(1) Were the studies appropriately designed and conducted (including data collection and statistical analyses) so as to allow for scientifically sound testing of the main hypotheses posed regarding possible associations between lead exposure and neurobehavioral effects (e.g., poorer classroom behavior, IQ deficits, etc.) in children?

(2) To what extent do the particular data, statistical analyses, and results obtained support the conclusions stated in the published papers (or other related materials regarding each study), and what caveats or limitations should most appropriately be stated as applying to such conclusions?

(3) Are there other conclusions that might be appropriately drawn (given the particular design, data collection, and statistical analyses employed in each study) and/or are there other appropriate approaches to the analysis of the data collected that would be expected to yield further meaningful and important information concerning the hypothesis that low-level lead exposure leads to neurobehavioral deficits in children?

(4) To what extent do the published studies allow for meaningful conclusions to be drawn regarding quantitative exposure-effect or dose-response relationships between any observed neurobehavioral effects and specific levels of lead exposure (as defined by either dentine or blood lead concentrations as indices of exposure)?

In the course of deliberating on general issues such as those posed above, the Committee considered more specific questions or points as appropriate for each of the studies reviewed. Many of the specific questions posed were presented in letters from Dr. Grant to the investigators (see attachment to this report, for a listing of letters). At initial meetings of the Committee in March, 1983, these and other questions were discussed with the senior investigators responsible for the conduct of particular studies, and some additional, unpublished information was provided by the investigators to the Committee to assist in accomplishing as complete an evaluation of each study as possible at the time of review. A preliminary draft of the Committee's report was provided to Drs. Ernhart and Needleman in September, 1983. The Committee reconvened in October, 1983, at which time written comments submitted by Drs. Ernhart and Needleman were considered by the Committee in making revisions in the report.

The Committee members thank the investigators for taking time to meet with us, for their assistance in providing and discussing information beyond that included in the published reports of their studies, and for calling to our attention certain factual errors in the preliminary draft of our report. The Committee hopes that the ensuing critiques of specific studies both (1) help to resolve legitimate controversy regarding the most appropriate interpretation(s) of the subject study results and (2) provide constructive criticisms and recommendations that are of value in carrying out reanalysis of certain subject data sets which hold promise for providing more definitive outcomes than those thus far reported for the studies in the published literature.

TABLE OF CONTENTS

SECTION | PAGE

SUMMARY

The Expert Committee on Pediatric Neurobehavioral Evaluations reviewed two independent sets of studies by: (1) Dr. Claire Ernhart and colleagues and (2) Dr. Herbert Needleman and colleagues. The studies evaluated possible associations between low-level lead (Pb) exposures and neuropsychological deficits in children who were otherwise apparently asymptomatic.

The Perino and Ernhart (1974) study evaluated relationships between blood Pb levels in a sample of 80 inner city black children (aged 3-5 yr) and IQ scores determined by the McCarthy Scales of Cognitive Abilities. Small but significant associations between lead exposure and lower IQ scores were reported, based on regression analyses. The Committee found the blood Pb measures were of acceptable reliability, as were also the psychometric measures for children. However, errors now have been discovered in the data analyzed for that report. In addition, confounding variables may not have been adequately measured, and the statistical analyses did not deal adequately with confounding variables. The Committee concludes, therefore, that the study results, as published by Perino and Ernhart (1974), neither confirm nor refute the hypothesis that low-level Pb exposure in children leads to neuropsychologic deficits.

Ernhart et al. (1981), in a follow-up study, reassessed blood Pb levels and neuropsychologic function in a subset of the same children 5 years later. The McCarthy Scales were again used, along with school reading tests and teacher ratings of classroom behavior. Small but statistically significant negative correlations were found between school-age blood Pb levels and scores on some McCarthy subscales, controlling for certain confounders. No significant associations remained if results were deleted for one "outlier" with markedly elevated dentine Pb beyond other values for the higher Pb group. The Committee found the psychometric measures to be acceptable, but the blood Pb sampling method raised questions about the reliability of the reported blood Pb levels. In addition, the statistical analyses did not adequately control for confounding factors. The Committee concludes, therefore, that the Ernhart et al. (1981) results neither confirm nor refute the hypothesis that low-level Pb exposure in children is partially responsible for neuropsychologic deficits. The Committee recommends that longitudinal analyses be carried out, using data from both the Perino and Ernhart (1974) and Ernhart et al. (1981) follow-up studies.

The Committee also reviewed a doctoral dissertation prepared by J. Yamins (1976) under Dr. Ernhart's direction. The Yamins study attempted to replicate certain aspects of the findings reported by Perino and Ernhart (1974), but used different psychometric measures and a different population of children. A major problem was the method of blood sampling, i.e., collection onto filter paper, which requires correction for hematocrit. Hematocrit levels

apparently are not available for data reanalysis. Although Yamins reported small but significant effects of lead exposure on some indices of cognitive functioning (taking age into account), the Committee found it difficult to place much confidence in such findings because of the failure to control adequately for confounding variables (besides age).

Results from an epidemiological study conducted by Needleman and colleagues were reported or discussed in: Needleman et al. (1979), Burchfiel et al. (1980), Needleman (1981), Needleman (1982), Needleman et al. (1982), Bellinger and Needleman (1983), and Needleman (1983). The main set of analyses was presented by Needleman et al. (1979). The study entailed neuropsychologic evaluations for more than 2000 first- and second-grade (mainly white) students. Lead exposure was indexed by dentine Pb in deciduous teeth. The classroom behavior of each child submitting a tooth was rated by the child's teacher. Some children, falling within the highest and lowest deciles for dentine Pb measured in one or more of their teeth, underwent more in-depth neuropsychologic evaluations, including use of an individual standardized measure of intellectual abilities (the WISC-R) to estimate IQ levels and tests of academic achievement, auditory and language processing, visual-motor reflexes, attentional performance, and motor coordination. Needleman et al. (1979) reported a relationship between first tooth dentine Pb values and percentages of students receiving poor classroom behavior ratings, which he has interpreted (Needleman, 1983) as "a strong dose-response relationship." Children in the high-Pb group (top 10% of dentine Pb levels) were also reported to have statistically significantly lower IQ scores (especially verbal IQ) than the low-Pb group (lowest 10% of dentine Pb values), taking into account five covariates in an analysis of covariance. The high-Pb children were also reported to do more poorly on certain other neurobehavioral tasks.

The Committee concludes that the relationship between dentine Pb levels and teachers' ratings of classroom behavior cannot be safely attributed to the effects of Pb, due to: (1) reservations regarding the adequacy of classification of subjects into Pb exposure categories using only the first dentine Pb value obtained for each child and (2) failure to control adequately for effects of confounding variables. The Committee also concludes that the reported results concerning the effects of lead on IQ and other behavioral neuropsychologic abilities measured for the low-Pb and high-Pb groups must be questioned, due to: (1) errors made in calculations of certain parental IQ scores entered as a control variable in analyses of covariance; (2) failure to take age and father's education into account adequately in the analyses of covariance; (3) the failure to employ a reliable strategy for the control of confounding variables; (4) concerns regarding missing data for subjects included in the analyses; and (5) questions about possible bias due to exclusion of large numbers of provisionally eligible subjects from statistical analyses. The Committee concludes, therefore, that the study results, as published by Needleman et al. (1979), neither confirm nor refute the hypothesis that low-level Pb exposure in children leads to neuropsychologic deficits.

The publications by Needleman (1982), Needleman et al. (1982), Bellinger and Needleman (1983), and Needleman (1983) describe further analyses of the same data set reported by Needleman et al. (1979). Burchfiel et al. (1980) reported analyses of certain psychometric data together with additional data on electrophysiological (EEG) measures for a subset of the high-Pb and low-Pb children from the Needleman et al. (1979) study. The above reservations regarding the basic analyses reported by Needleman et al. (1979) apply also to the analyses reported by Burchfiel et al. (1980), Needleman (1982), Needleman et al. (1982), Bellinger and Needleman (1983), and Needleman (1983). Similar reservations apply to analyses of another data set (Needleman, 1981). The Committee recommends that the entire Needleman data set be reanalyzed, correcting for errors in data calculation and entry, using better Pb exposure classification, and appropriately adjusting for confounding factors.

In addition to evaluating the studies of Ernhart and Needleman, the Committee reviewed available reports (some published and others as yet unpublished) of other studies from the United States and Europe. Although an exhaustive, in-depth evaluation of the world literature on low-level Pb exposure was beyond the current charge to the Committee, we note that new studies reported in the spring and summer of 1983, with only a few exceptions, failed to find significant association between low-level Pb exposure and neuropsychologic deficits, once control variables were taken into account.

From its review of the recent research literature covered in this report, the Committee concludes that: (1) in the absence of control for other variables, a negative association between Pb exposure and neuropsychologic functioning has been established; (2) the extent of this negative association is reduced or eliminated when confounding factors are appropriately controlled; and (3) the Committee knows of no studies that, to date, have validly established (after proper control for confounding variables) a relationship between low-level Pb exposure and neuropsychologic deficits in children.

INTRODUCTION

In approaching its task, the Committee was faced first with establishing criteria for research studies, the results of which may be accepted as evidence pertinent to determining the influence of Pb exposures on cognitive functioning in apparently asymptomatic children. Because children live and mature in a complex socio-cultural milieu that affects them in many diverse ways, isolation of a definitive cause, e.g., lead exposure, for neuropsychological problems in children is extremely difficult. Under these circumstances, what kind of research design is necessary or adequate to produce pertinent evidence?

The problem of determining the effects of Pb on cognitive functioning is viewed as an instance of a general class of dosage-response problems. Alternative research designs with which to approach such problems include:

(i) randomized clinical trials;
(ii) cross-sectional observational study of individuals from groups known to vary in exposure (dosage);
(iii) longitudinal study of the same individuals over time;
(iv) a time series of observations on different sets of individuals who are members of groups known to differ in exposure (dosage).

A. Alternative Research Designs

1. Randomized Clinical Trials

There is no question that randomized clinical trials, properly conducted, provide evidence that is highly relevant to the research question. Neither is there any question that the experimental administration of Pb to human subjects is unethical, and not to be considered. This highly effective research design, then, simply cannot be adopted to address the question of the effect of Pb on human cognitive functioning.

2. Cross-Sectional Designs

The bulk of published work assessing Pb effects on human cognitive functioning has entailed the cross-sectional study of a sample of children. A serious complicating feature of the design results from typical empirical findings of association between low or moderate levels of lead exposure, on the one hand, and such background variables as parental IQ, parental education, quality of home environment, family size, etc., all of which are known to be correlated with children's cognitive performance. Under what conditions, then, might this design yield valid conclusions about the effect of Pb on cognition? Three possibilities appear to exist, as follows:

(a) Were low or moderate levels of Pb consistently found to be negatively correlated with cognitive performance, while all potential confounding background variables were negligibly correlated with cognitive performance, then a valid conclusion would be that Pb is responsible for the cognitive deficits. However, the premise generally appears to be false: published studies on Pb, consistent with research literature in child psychology, report sizable correlations between cognitive performance and a host of background variables.

(b) Were low or moderate levels of Pb consistently found to be negligibly correlated with cognitive functioning, regardless of the pattern of association between cognitive function and confounded background variables, then it would be fairly safe to conclude that the differences in Pb levels are not importantly related to cognitive performance. Again, however, the premise generally appears to be false: most published studies on Pb report significant correlations between Pb and cognitive test scores unadjusted for other key confounders.

(c) Consider a study designed so as to provide a factor analysis of interrelations among variables. It might be found that cognitive performance is represented on one factor along with noncognitive variables that are not appreciably associated with cognitive performance in the absence of Pb. In the presence of Pb, however, such noncognitive variables might be hypothesized to be associated with cognitive function. Lead would be the primary defining variable on such a factor. Noncognitive variables that would be appropriate candidates for this factor analysis include sensory discriminations and electroencephalographic (EEG) recordings. This finding would support the hypothesis that Pb was a partial determinant of cognitive functioning.

The research studies of Pb effects on cognition of which the Committee is aware generally fail to match any of the above conditions (a), (b), or (c). Rather, the studies mainly report cross-sectional data for which: (1) Pb is correlated with cognitive test scores by a nonzero but modest amount; (2) Pb is correlated with background variables; and (3) background variables are correlated with cognitive performance. In most such studies, efforts are made to separate the influence of Pb on cognitive functions from the influence of confounding variables, using methods of statistical adjustment (e.g., regression analysis).

Statistical adjustment for confounding variables may reduce the residual relation between Pb and cognition to a negligible value. If so, however, it would not necessarily follow that cognitive functioning is not influenced by Pb; the effects of Pb might be masked by one or more of the confounding variables. The research design is generally incapable of providing evidence which permits a clear separation of the magnitude of effect of Pb from effects of the confounding variables.

Statistical adjustment for confounding variables may leave a significant residual correlation between Pb and cognition. If so, however, it would not necessarily follow that cognitive function is influenced by Pb. Perhaps other background variables, not explicitly adjusted for, but correlated both with Pb and cognition, are the effective determinants of cognitive differences. Or, perhaps, less fallible measures of the confounding variables would have further reduced the correlation between Pb and cognition to a non-significant amount. The research design is not sufficiently sensitive to provide guidance concerning which of these alternative conclusions should be embraced.

The controversy over the interpretation of results from a series of recent studies may be attributed to this intrinsic ambiguity regarding the assignment of causal status to the predictor variable of interest (Pb) or to confounding variables (e.g., home and parental measures). Some consider Pb to act as a surrogate for the confounding variables. Others consider the confounding variables to act as a surrogate for Pb. There is no scientific basis for accepting or rejecting either set of interpretations.

In view of these considerations, the Committee concludes that, no matter how carefully designed and executed, cross-sectional studies of relationships between Pb and cognitive functioning are not able to yield definitive conclusions regarding the influence of low-level Pb exposures on human cognitive functioning, when measures of both are correlated with background variables also known to influence cognitive development and performance. At best, such cross-sectional studies may yield evidence suggestive of effects of low-level Pb exposures which would need to be confirmed by studies using more definitive research designs.

The Committee has been charged with evaluating the research reports of Needleman and his associates and of Ernhart and her associates. All of these reports concern essentially cross-sectional studies (although some of the Ernhart data are amenable to longitudinal analysis, a subject to which we return later). Each study is thus subject to the severe reservations expressed above: i.e., from cross-sectional studies with confounding variables, it is not possible to draw definitive conclusions about the role of low-level Pb exposures as a determinant of cognition. Nevertheless, we do present more detailed critiques of these studies in sections below, recognizing the importance of attempting to resolve apparent inconsistencies in the results and conclusions presented by these sets of investigators and recognizing, also, the value of accumulating even small or suggestive indications of possible relationships, or lack thereof, between Pb and cognitive deficits.

We have judged that, to assess effects of Pb on cognition, (i) randomized clinical trials cannot be conducted and (ii) cross-sectional observational studies cannot adequately disentangle effects of Pb from effects of confounding variables. We now comment on strengths and weaknesses of certain other research designs, longitudinal studies and time-series analyses, which may be capable of yielding more definitive conclusions than cross-sectional designs.

3. Longitudinal Designs

A longitudinal design is characterized as a study of the same individuals over a period of time. For the topic at hand, primary interest would reside in changes over time in relative levels of cognitive functioning as a consequence of earlier levels of systemic Pb or of changes in systemic Pb. Many (but not all) of the confounding variables are usually quite stable over time, and thus may be assumed to have a similar influence on measures obtained at different times. To that extent, the difficulty created by confounding variables will be at least partially alleviated.

Documentation of the history of systemic Pb exposure should begin early, even prior to an infant's birth. Cognitive performance also should be assessed early, as soon as 18 months after birth. Measurements of both sets of variables should be repeated periodically for several years, and other measures, e.g., dentine lead, might be obtained at appropriate times. It is crucial that as complete a history as possible of Pb exposure (as indexed by changes in internal indices) be obtained and that such indices of exposure be evaluated for relationships to dependent variables indicative of cognitive/behavioral development both proximate and distant in time after the exposure measures are obtained. This is important both to increase information on latency periods for Pb effects to be manifested and in regard to augmenting our knowledge of reversibility/irreversibility of Pb effects.

In the study by Perino and Ernhart (1974) discussed below, the same sample of children was assessed both for Pb exposure and cognitive performance at ages 4-5 years, and again five years later, as reported by Ernhart et al. (1981). The authors did inquire about the relationship between later cognitive measures and earlier Pb levels, but they failed to study the possibly revealing relationship between change in cognitive score and change in Pb levels (see comments on these studies in a later section of this report).

A longitudinal design reduces but does not necessarily totally avoid problems of confounding variables. Variables that remain stable over time for a given individual, while creating difficulty in a cross-sectional design, may not be as much of a problem in a longitudinal design. However, confounding variables that change over time would be troublesome in longitudinal as well as in cross-sectional studies. Techniques for statistical adjustment may (and usually should) be employed for such variables. To the degree that they are prominently related to Pb levels and cognition or to changes in these variables, they are as troublesome in longitudinal as in cross-sectional studies. The hope is that their effects will be far smaller in longitudinal studies.

4. Time-Series Design

Fortuitous events, from the research perspective, may occasionally provide the opportunity for a time-series study of the effects of Pb on cognitive functioning. For example, it might be recognized that an environmental change is imminent in Community A, a change anticipated to have a large effect on typical systemic Pb exposure in that community. Prior to that change, Pb and cognitive measures might be obtained for a random sample of 5-year-olds (or 7- or 9-year-olds) in community A, and also in community B, considered similar to A except for the impending change. Later, after the environmental change and its effects have had an opportunity to be exerted, similar measures are again collected on random samples at the same age, in both communities. Differential changes in cognition as a function of different levels of Pb may strongly suggest that Pb has influenced cognition. A specific example of where such a research approach might be applied is a situation whereby an imminent governmental or industry action is anticipated that would lead to substantial reductions in Pb exposure in a particular geographic area.

B. Additional Remarks

The Committee cannot conclude these general introductory remarks without presenting an additional caveat regarding the interpretation of even an unambiguous finding of a significant negative relationship between low Pb levels and measures of children's IQ or other behavioral variables, based on epidemiological observations. If an investigator is able to discount the influence of confounding variables, and if no flaws are found with the research design employed or the conduct of the study, the temptation may exist to conclude that Pb is responsible for the observed lowered IQ levels or other behavioral deficits. Note, however, that such results are, in many cases, equally consistent with the conclusion that increased Pb exposures and associated body burdens are a consequence of low IQ or other observed behavioral deficits. Furthermore, knowledge external to the research study generally would not be such so as to provide an obvious basis for preferring one of these conclusions over the other.

REVIEW OF STUDIES BY DR. CLAIRE ERNHART AND COLLEAGUES

A. Background Information

The Committee undertook detailed review of two studies published by Dr. Ernhart and colleagues (Perino and Ernhart, 1974; Ernhart et al., 1981) and, also, preliminary review of a third study reported in the 1976 doctoral dissertation of J. Yamins at Hofstra University. (The latter doctoral research was conducted under Dr. Ernhart's direction but is not yet published in the peer-reviewed literature).

In the first study, based on the doctoral research of J. Perino (under Dr. Ernhart's direction), inner city black children of low socioeconomic status were recruited for study based on blood Pb values obtained during screening for possible undue Pb exposure by the New York City Health Department during 1972. Children were randomly selected to participate in the study so as to represent a group of subjects with lead exposures ranging from low (<30 μg/dℓ) to moderately elevated (40-70 μg/dℓ) according to then existing screening standards. Because the study was designed to evaluate neuropsychologic deficits associated with moderate lead exposures in non-overtly lead-poisoned children, children with histories of overt signs or symptoms typical of Pb poisoning were excluded from the study. Eighty black children (41 boys and 39 girls) of preschool age (3 yr to 5 yr, 1 mo) from Queens, New York, were included in the study. The McCarthy Scales of Children's Abilities (McCarthy, 1972) were administered to the children in their homes by a trained school psychologist (J. Perino), to yield a General Cognitive Scale score with norms obtained in the same manner as and roughly comparable to IQ scores. The test also provided scores on several subscales, i.e. Verbal, Perceptual Performance, Quantitative, and Motor Abilities. Parental IQ was measured by means of the Quick Test (Ammons and Ammons, 1962) of gross intellectual level. Questions regarding other covariates were administered to the parent by the school psychologist, following a standardized format and recording answers on a typed questionnaire form. In general, the results of the study were such so as to lead the authors to conclude that neuropsychologic deficits (i.e., decreased cognitive, verbal, and perceptual performance abilities) were significantly associated with Pb exposure in the otherwise asymptomatic children studied. The results have also been interpreted (in <u>Air Quality Criteria for Lead</u>, U.S. EPA, 1977) as demonstrating such deficits to be associated with Pb exposures resulting in blood Pb levels of 40-70 μg/dℓ.

The study by Ernhart et al. (1981) is a followup study, in which 63 children (30 boys, 33 girls) from the original cohort of 80 black children studied by Perino and Ernhart (1974) were reexamined 5 years later. Scores were obtained for these school-age children on the McCarthy Scales of Children's Abilities, school reading tests, teacher ratings of classroom behavior, and several neurobehavioral exploratory measures. Hypothesized relationships between performance on these neuropsychologic tests and childhood Pb exposures were first statistically evaluated by means of omnibus multivariate tests (Hotelling's T^2), comparing test scores of "low lead" versus "moderate lead" children defined in terms of (1) pre-school blood Pb levels (low = 10-30 μg/dℓ ; moderate = 40-70 μg/dℓ) and (2)school-age blood Pb levels (low ≤26 μg/dℓ; moderate = 27-49 μg/dℓ). Significant omnibus test results ($p<0.05$) were obtained only for reading scores related to both preschool and school-age Pb levels. Further multivariate and univariate analyses were conducted for these significant neuropsychologic outcome variables. Univariate tests for differences on McCarthy scores between low and moderate Pb subjects (ignoring control variables) suggested that the moderate Pb group performed more poorly on the General Cognitive Index (GCI) and 3 of the 5 McCarthy Test subscales. However, multivariate (regression) analyses revealed that sex and parental IQ were control variables that were significantly correlated with one or more outcome measures. When these control variables were ignored in analyses including Pb exposure measures, preschool Pb was significantly negatively related to scores on the GCI, 4 of 5 McCarthy subscales, and the reading tests. When sex and parental IQ were taken into account, however, preschool Pb was not related to any neuropsychologic outcome measure and school-age Pb was significantly related only to the McCarthy GCI, verbal subscale, and motor subscale scores (with the variance attributable to school-age Pb generally being less than half that found for the same outcome measures when the control variables were excluded). Dentine Pb levels in shed deciduous teeth of 33 children were not significantly related to any outcome measures, even when control variables were ignored.

The Yamins (1976) dissertation study, in part, attempted to replicate the Perino and Ernhart (1974) findings, using a different study population and psychometric tests. Preschool children (aged 2 yr, 4 mo to 5 yr, 9 mo), including 80 black (38 girls, 42 boys) and 20 white (10 girls, 10 boys) children from low to low-middle socioeconomic status communities in the Nassau County (Long Island) Department of Health Clinics lead-screening program catchment area, were included in the study. Children with overt signs or symptoms of Pb intoxication were excluded. Of the included children, 54 black subjects had blood Pb levels below 37 μg/dℓ, whereas 26 fell in a 38-70 μg/dℓ range; of the 20 white children, 19 had blood Pb levels below 37 μg/dℓ. Cognitive performance and language development of the children were assessed by the following tests administered in a set sequence in the home by a school psychologist blind as to the children's Pb levels: a "nonverbal" IQ test (Peabody Picture Vocabulary Test; PPVT); a general information test (Caldwell Preschool Inventory); a concepts test (Block Sort); a perceptual-motor functioning test (Copy Forms); and a sentence repetition task designed for the study. The Ammons Quick Test was used to measure parental IQ, and data were gathered on several other background variables (parental education and occupation, quality of housing, child's medical history, number of siblings, etc.) by means of standardized questionnaire and rating forms. Raw scores for all dependent variable measures were used in multiple regression analyses, which took into account age as well as the other potentially confounding background variables that were measured. For the black children, several such variables (e.g., parental IQ and education, absence of father from home, etc.) were significantly negatively related to children's Pb levels but positively related to each other and the children's cognitive and language variables. Similar results were obtained for the white children. Stepwise multiple regressions were then performed, excluding predictor variables contributing less than 1% of the variance, entering the included predictor variables into the equation before Pb. With all predictor variables controlled, for the black sample, Pb contributed 2.4% of the total variance for nonverbal intelligence (PPVT, $p<0.05$), 3.1% for general information (Preschool Inventory, $p<0.01$), 2.4% for overall level of acquired syntax (Total Repetition Score, $p<0.05$), and 2.5% for ability to repeat nongrammatical verbal stimuli (Ungrammatical Stimuli Test, $p<0.05$). Lead did not contribute significantly to conceptual level (Block Sort), perceptual-motor functioning (Copy Forms), or ability to repeat grammatical stimuli (Grammatical Repetition) scores.

In order to evaluate critically the above studies, the Committee met with Dr. Ernhart at EPA facilities in Research Triangle Park, NC on March 17-18, 1983. At that time, a summary overview presentation was made by Dr. Ernhart on the objectives, design, data collection and analysis procedures, and results for each of the studies. Certain listings of raw data values (provided in coded form to protect the privacy of subjects) and other pertinent published and unpublished materials were examined by the Committee and considered during discussions with Dr. Ernhart regarding diverse aspects of the studies reviewed. Some additional, follow-up information was requested by the Committee and was provided to them subsequent to the March 17-18 meeting with Dr. Ernhart. See Attachment 1 for a list of materials examined by the Committee in connection with their review of the subject studies. The Committee's comments regarding the most salient points of concern and controversy related to methodological and other features of the above studies by Ernhart and associates are presented below.

B. Comments on Perino and Ernhart (1974) and Ernhart et al. (1981) Studies

1. Indicators of Lead Exposure

In the first two studies under consideration, Perino and Ernhart (1974) and Ernhart et al. (1981), the major indicator of exposure was blood Pb, with additional use of erythrocyte porphyrin (EP) measurements and a sub-group of dentine Pb samples in the follow-up study of Ernhart et al. (1981).

On the basis of current criteria of methodology assessment for Pb analyses as noted in Chapter 9 of the EPA draft document Air Quality Criteria for Lead (U.S. EPA, 1983), it can be said that the blood lead values in the Perino and Ernhart study are reasonably reliable. With the follow-up study of Ernhart et al. (1981), blood lead accuracy becomes potentially problematical, due both to: (1) a combined positive bias of capillary blood sampling and choice of analysis, and (2) a bias of negative direction but of possibly variable size, owing to the generally poorly recognized effect of whole blood hematocrit/hemoglobin on blood Pb measurements using filter paper. The latter factor requires making use of the hematocrit measurements for the subjects' blood (which presumably are available, since EP measurements also require knowing the hematocrit) to correct for differential spread or diffusion of blood (and concentration of Pb therein) on the filter paper matrix.

Measurement of Pb in dentine of shed teeth from the subjects in the Ernhart et al. (1981) follow-up study was carried out in the laboratory which both pioneered the analysis of Pb in teeth and probably has the most experience and proficiency with such analyses. The method of analysis for dentine Pb is reasonably reliable, and it appears that analysis error in replicate sampling is at the step of isolating the dentine zones from a given whole tooth sample.

Measurement of erythrocyte porphyrin involved blood collected on filter paper and subsequent elution and micro-fluorometric analysis. Such a "wet" or laboratory analysis is considered much more reliable than the use of the hematofluorometer when applied to blood samples of modest EP content. Analytical error was noted to be less than 15 percent.

In the study of Perino and Ernhart (1974) lead exposure in the pediatric subjects was indexed by analysis of venous whole blood. Characteristics of the specific procedures employed and pertinent evaluative comments are as follows:

(a) Venous blood samples were collected by trained technicians in a lead screening program and analyzed in the laboratory facilities of the New York City Department of Health (NYCHD) during the summer of 1972. Sampling involved standard precautions to minimize sample contamination and collection during the summer months, when blood lead values in the city are known to be maximal. Samples were analyzed within 48 hours of collection and refrigeration.

(b) As a lead analysis facility, the NYCHD laboratory has processed a large volume of samples for lead content over many years, an important consideration in view of the fact that laboratory proficiency is directly related to the level of Pb analysis activity.

(c) The specific method of blood lead analysis employed was the Hessel extraction variation of atomic absorption spectrometry (Hessel, 1968), a macro method using venous blood which still enjoys popularity up to the present time. The periodic surveys by the Centers for Disease Control (CDC) of participating laboratories in their proficiency programs indicate (see Boone et al., 1979) that the Hessel method is somewhat more accurate than the Delves cup technique and more accurate than the other variations of atomic absorption spectrometry. Precision tends to be less than for the Delves cup procedure, consistent with the reported analytical error of ±5-6 µg Pb/dℓ for the NYCHD laboratory when using the Hessel method (communication of B. Davidow to N.B. Schell, see Ernhart, March 11, 1983: summary of conversation between N.B. Schell and C. Ernhart). Compared to isotope dilution mass spectrometry (IDMS), the Hessel method for the range of blood lead in the Perino and Ernhart study shows a modest positive bias of 2.5 µg Pb/dℓ.

(d) At the time of data collection for the Perino and Ernhart study, it was the practice of the NYCHD laboratory to report blood lead values rounded to the nearest decile of blood lead. Hence, a subject blood Pb value of 40 µg/dℓ in the Perino and Ernhart (1974) report would have resulted from a reading of some value between 35 and 44 µg Pb/dℓ.

(e) Internal and external quality control protocols were in place in the NYCHD laboratory at the time of the subject study. The latter consisted of participation in both the CDC and New York State proficiency testing programs, and the NYCHD laboratory met acceptable proficiency standards.

(f) It appears that there are no major difficulties with methodological aspects of the blood Pb data. The moderate positive bias in the Hessel procedure, if corrected for among the Perino and Ernhart (1974) study data, would result in a constant shift downward in all values. Since there was decile rounding, absolute corrections for this bias would require having the original blood Pb values.

(g) The full use of confidence bounds for the lead measurements, if employed in any overall reanalysis of the data, would require having the original blood lead values as well as taking into account the analytical error noted above.

Lead exposure levels for subjects in the follow-up study of Ernhart et al. (1981) were indexed via measurement of Pb in capillary blood (applied to filter paper) and free erythrocyte protoporphyrin (FEP) determination. A sub-group of the follow-up population furnished shed teeth for dentine lead analysis. Characteristics of the procedures used and evaluative comments are as follows:

(a) Capillary blood samples were collected on filter paper and analyzed in the same NYCHD laboratories (as noted above) that assayed blood Pb by the Delves cup variation of atomic absorption spectrometry.

(b) It is now generally accepted that capillary blood samples, as compared to those obtained by venous puncture, manifest a significant positive bias in lead level even under relatively stringent sample collection conditions. The best data base by which to estimate the magnitude of this positive bias is the NHANES II survey, which indicated that for the Delves cup procedure used in the present study under discussion capillary blood Pb was 6 μg/dℓ higher than for venous samples. An additional positive bias of approximately 3 μg/dℓ exists for the Delves cup analysis compared to the definitive IDMS method.

(c) A generally unrecognized problem with the analysis of lead in blood using filter paper spotting has to do with the close dependence of blood flow (on filter paper) on hemoglobin/hematocrit content. The study of Carter (1978) makes it clear that blood flow is increased on filter paper with decreasing hematocrit, resulting in proportionately lower blood lead values contrasted to venous blood analysis.

(d) In view of the documented problem of using blood on filter paper without correction for hematocrit, blood lead values in the Ernhart et al. (1981) report would have to be appropriately corrected, if this was not already done initially by the NYCHD. It should be noted that Carter (1978) observed under-estimation of lead content of blood on filter paper at hematocrit values that would be considered in a normal range.

(e) Dentine lead levels were determined in the laboratories of Dr. Irving Shapiro, School of Dental Medicine, University of Pennsylvania, a facility which pioneered such analyses and is recognized as having the most proficiency in such measures. In this study, dentine was isolated from a given whole tooth sample.

(f) Samples of isolated dentine were dissolved in perchloric acid, buffer added, and lead measured by an electrochemical technique, anodic stripping voltammetry. At the levels of lead being measured in the dentine samples, this technique provides reasonably reliable results for the solubilized analyte.

(g) From data of Shapiro et al. (1973), duplicate analysis of dentine with low and elevated lead exposure of subjects indicates that the analytical error increases with concentration, suggesting greater variation in replicate sectioning than in the instrumental measurement itself.

(h) Identification of an "outlier" by Ernhart in further unpublished analyses of data from the follow-up study (see following sections) was done on the basis of a dentine lead value being 107.4 ppm Pb in the subject sample. A methodological problem accounting for this high value has been discounted by Shapiro (personal communication of I. M. Shapiro to C. B. Ernhart, see Ernhart, February 3, 1983, letter to D. Weil), who indicated that any contamination of the dentine section by inclusion of circumpulpal dentine, a region manyfold higher in lead, would only account for 3-4 percent of the above dentine lead level.

(i) Erythrocyte porphyrin analysis in the present study was carried out in the NYCHD laboratory which simultaneously analyzed capillary blood for lead. It was noted that analytical error was less than 15 percent, using a microfluorometric analysis of blood samples eluted from the filter paper. This laboratory employs internal and external quality control protocols for EP analyses, the latter including the EP proficiency testing program of CDC.

2. Psychometric Measurements and Procedures

Comments on the psychometric measurements employed in the Perino and Ernhart (1974) study and the results obtained are as follows:

(a) The study employed the McCarthy Scales to assess the intellectual development of young children and the Ammons Quick Test to assess parents' IQ levels. The Committee agreed that the McCarthy Scales were appropriate for assessing intellectual abilities of the children in this sample, whose race, age range, and socioeconomic status (SES) were represented in the standardization sample. The Ammons Quick Test, however, has more questionable validity for two reasons: (1) the content of the test is a very limited sample of adult intelligence, and (2) the test was not standardized on low SES black subjects. The reliability of the McCarthy Scales is satisfactory for the present research. The reliability of the Quick Test, however, is only about .75, a low value for an adult measure. More importantly, correlations with the Stanford-Binet and WISC range from only .10 to .80, with a median in the .40 range. Large discrepancies have been observed between Quick Test Scores and individual IQ test scores (Sattler, 1982). Although the Quick Test is considered useful in large scale research studies, where a simple and quick assessment of average ability is needed, the measure is less adequate in the Perino and Ernhart study where it was used as a control for individual differences. A measure such as the short form of the WAIS or the WAIS vocabulary scale would have provided a better estimate of parental IQ.

(b) The administration of the tests by the first author, a school psychologist, was blind with respect to the children's blood Pb levels. Although the tests were administered in the home, under nonstandard conditions, the Committee concluded that the assessments were generally valid. Because of the training and clinical experience of the investigator, the Committee thought it unlikely that the assessments were seriously compromised.

(c) Birth weight, history of birth risk factors, and maternal education were reported by the mothers in an interview, but were not verified by checking of appropriate records. Maternal occupation was recorded from clinic records employing a 1950 census classification that was inadequate to differentiate among urban blacks. Thus, nearly all families in this sample fell within two occupational classes. It is not known whether a finer SES scale would have resulted in greater relations between SES and other variables in the study than those reported by Perino (1973).

(d) In view of the limitations of the Quick Test and the measurements of some of the control variables, it is especially important that new analyses of the data (proposed below) be employed to maximize the efficiency of the control variables.

(e) Descriptions of quality control procedures by Dr. Ernhart regarding the checking of data entries onto computer cards and/or tapes seemed to indicate that reasonable care was taken to ensure accurate encoding of data for statistical analyses. The Committee had no feasible way to confirm this independently, but Dr. Ernhart, in responding to the Committee's request for reanalyses of the data set, reported as follows: "In the course of conducting the reanalyses to include parent education, we found several errors in the data and the previous analyses. One child's age was incorrect by one year (38 months rather than 50 months). This changed his General Cognitive Index (GCI) from 50 to 86; other scores were correspondingly affected. Another child's GCI was incorrect by one point. The degrees of freedom used and reported in the tests of significance of the lead effect (regression analyses) should have been 1 and 75, not 4 and 75."

Comments on the psychometric assessment procedures used by Ernhart et al. (1981) in the follow-up study are as follows:

(a) Because this study is a follow-up of the Perino and Ernhart sample, many of the issues raised about the original study apply here. Also, use of the McCarthy Scales with age groups not included in the standardization sample is questionable. More than half of the children were beyond the age range of the test, and their test scores were determined by linear extrapolation. The Committee believes that a better procedure would have been the use of residuals after regressing raw McCarthy Scale scores on age at testing. The Committee acknowledges that no subject reached the ceiling on the subtests and also that some value exists in re-administering the same measure in a longitudinal study. However, the Committee concludes that it would have been preferable to have chosen an age-appropriate measure of intelligence, such as the WISC-R, which is based on a suitable standardization sample and also taps cognitive performance of older children. Despite reservations about the use of the McCarthy Scales in the follow-up, the Committee does not believe that the assessment of intellectual development was seriously compromised. Because the children in the study were functioning at levels well below those of most children of their ages, the test was more appropriate for them than for most children of the same age but of average intelligence.

(b) Reading test scores were obtained from many different tests. The Committee concludes, however, that the combination of multiple measures does not necessarily compromise the assessment of reading achievement (Scarr and Yee, 1980). The age correction used was appropriate.

(c) More than one psychologist collected the follow-up McCarthy data, generally within the children's schools. The Committee believes that the administration and scoring of the protocols were adequate. The testers were both blind as to the children's Pb levels and well-trained in psychometric assessment.

(d) Correlations between the earlier and later administrations of the McCarthy Scales, as reported in Ernhart et al. (1980), were in the moderate range (from .24 to .61). The highest correlation is for the General Cognitive Index, which is considered by the Committee to be the most important cognitive measure in the study. In light of developmental changes from preschool to school-age skills, and the use of extrapolated scores, this relationship suggests considerable reliablity for the McCarthy Scales in the follow-up study.

(e) The same comments regarding quality assurance checks for data entries as were made under (e) above for the Perino and Ernhart (1974) study also apply here.

3. Statistical Analyses

The Committee believes that the treatment of confounding factors can be handled in a better way than as described in the 1974 and 1981 Ernhart papers. The analyses performed by Ernhart and colleagues can be questioned in light of currently accepted statistical practices for dealing with interaction and confounding, and they should be reworked.

The proper model for assessing the effect of Pb exposure on IQ is one which contains, in addition to the Pb exposure variable, all factors deemed to be potentially confounding. A potentially confounding factor is one which is correlated (in the population) with the exposure variable (Pb) and for which there is reasonable evidence from previous research experience and knowledge that it is predictive (in its own right) of the outcome variable (IQ). Thus, a procedure which chooses potential confounders via a forward selection procedure, ignoring the Pb variable (as Ernhart did), can be misleading and can actually incorrectly drop important confounders from consideration. A backward elimination strategy is designed to obtain the most accurate estimate of the effects of the exposure variable adjusted for all key confounders, whereas a forward selection approach is designed to predict the outcome variable but does not assure an accurate estimate of the effects of the exposure variable.

As an example, parental education was eliminated by a forward selection approach not involving the Pb exposure variable at all; in fact, it is a potential confounder, and should only be eliminated from the full model if its elimination does not alter the Pb exposure regression coefficient. Similarly, SES should be handled in the same fashion.

In theory, assuming that the data set is free from bias (i.e., is a random sample of the population under study), one can lose precision but not validity by including in the full model a true non-determinant of the dependent variable under study. An example in the data under consideration is the variable sex, which is generally agreed to be a predictor of motor skills but not of the other outcome measures being studied. Its inclusion in the full model (using any outcome variable other than motor skills) should not cause a validity problem (even though it is correlated with Pb exposure); this can be demonstrated (assuming that the data are representative of the population) by dropping the sex variable from the full model and noting that the Pb exposure coefficient does not materially change. With motor skill as the outcome variable, sex should be expected to be manifest as a real confounder in these data and hence cannot be dropped from the model. In summary, then, a reanalysis of the Perino and Ernhart (1974) and Ernhart et al. (1981) data should be carried out, based on a model containing the Pb exposure variable and all available confounders measured.

In general, interaction effects between the Pb exposure variable and the covariates should be assessed before confounding issues are considered. A qualitative assessment can be done initially by stratifying on each of the covariates and determining whether the relationship between Pb exposure and IQ is reasonably constant across strata of the covariate under consideration. Nonconsistency suggests the need for one or more interaction terms in the model under study. Ultimately, the modeling of interaction involves the use of cross-product terms in regression models. A strategy for dealing with interaction and confounding in observational epidemiologic data is described by Kleinbaum et al. (1982). Possible interaction effects were not really examined by Ernhart and colleagues, and the presence of any interaction(s) would complicate data interpretation. However, given the highly variable nature of the data and the limited sample size, it may be difficult to deal adequately with interaction assessment for these data.

Comments on specific aspects of the statistical analyses employed in the 1974 and 1981 Ernhart studies include:

(a) Corrections for unreliability (i.e., measurement error) in the variables under study should be made, especially concerning the Pb exposure indices. Such corrections will adjust the observed correlations, and could have a major impact on the ultimate conclusions drawn from the analyses. They provide a limit on the strength of relationships by showing the relationship that would be expected were the measurements made and recorded without error.

(b) The outlier alluded to earlier in the Ernhart et al. (1981) data set should be dropped from all analyses. The dentine Pb value provides enough evidence that, in this case, the past Pb exposure was sufficiently high to conclude that the subject may have earlier been overtly symptomatic but undiagnosed.

(c) It is important that the dependent variables be adjusted appropriately for age. As a suggestion, one could use, for each child, the deviation from the linear regression line of "raw score" on "age at testing".

(d) Dependent variable scores for a given individual on a series of intelligence tests are obviously correlated. Multivariate analysis of covariance is one option, but must be done very carefully. Certain assumptions (e.g., multivariate normality) must hold at least approximately.

(e) Treating the Pb exposure variable as a continuous variable is preferable to categorizing cases into high and low Pb groups.

(f) The Committee notes that, for the 62 children in the 1981 study (excluding the outlier), significant correlations are seen between the identification number assigned to children and other variables of interest. Identification number correlates -0.43 with 1972 level of blood Pb, 0.38 with parent IQ, 0.28 with parent education, and 0.27 with 1972 GCI score from the McCarthy scale. These results would be expected if identification numbers were assigned to children in the order in which they were assessed (by Perino in 1972) and if the children with higher Pb levels tended to be the earlier ones assessed. Results of the Ernhart studies, particularly the one by Perino and Ernhart (1974), could be affected by this nonrandom ordering of assessments, to the extent that any aspects of the assessment process systematically changed over the course of the studies.

An alternative research question to those addressed in the Perino and Ernhart (1974) and Ernhart et al. (1981) analyses can be asked by considering data from the cases assessed both in 1972 and in 1977: namely, is there a relationship between differences in cognitive scores and differences in blood lead concentrations from 1972 to 1977? Do children whose blood lead levels were higher in 1977 than expected from their 1972 levels display cognitive scores in 1977 that differ from those expected from their 1972 cognitive scores?

This question could be addressed by regression analyses. The 1977 Pb level could be predicted from its regression on the 1972 Pb level and residual values obtained. Similarly, the 1977 IQ value could be predicted from its regression on the 1972 IQ value and residual values obtained. The correlation between IQ residuals and Pb residuals would indicate an association between change in IQ and change in Pb. That correlation also could be adjusted for confounders such as parental IQ, parental education, and SES. We would urge such a reanalysis of the data from Perino and Ernhart and Ernhart et al. (after correction of the erroneous values recently discovered in those data). Interpretation of results from such analyses must depend upon careful inspection of the cross-lagged correlations on which the correlation of residuals depends. In addition, of course, even the residual values may be acting as surrogates for unrecognized confounding variables.

4. Committee Conclusions and Recommendations

The Committee's conclusions and recommendations regarding the Perino and Ernhart (1974) study can be summarized as follows. Blood lead levels, as the main index of exposure, appear to be of acceptable reliability. The psychometric measures for children are also acceptable, but confounding variables may not have been adequately measured. The statistical analyses do not adequately deal with confounding variables. In the view of the Committee, the findings of this study, as reported, neither support nor refute the hypothesis that low to moderate lead exposures are associated with cognitive impairments in apparently asymptomatic children. The Committee recommends that the data from the Perino and Ernhart study be used in conjunction with the 1981 Ernhart et al. follow-up study in longitudinal analyses.

The Committee's conclusions and recommendations concerning the Ernhart et al. (1981) study include the following points. Given limitations in both the control and the outcome measures, it is difficult to assess the possible role of less-than-ideal measures in contributing to the generally null results reported. When the authors conclude that there are no significant effects, or very weak effects at best, then that outcome might also be reasonably attributable to unreliable measures or other procedural problems. One major difficulty with this study is the potential unreliability of the blood Pb level measurements, such that the Committee recommends that the blood Pb values be corrected in the fashion specified earlier. The psychometric data were adequately collected but should be readjusted for age. The cross-sectional analyses suffer from the same problems as those of the previous (Perino and Ernhart, 1974) study. More importantly, the analyses failed to exploit fully the longitudinal aspects of the study data set. In the view of the Committee, then, the findings of this study as reported in the published form also neither support nor refute the hypothesis that the reported blood lead levels are associated with cognitive impairments in children. The Committee strongly recommends that longitudinal analyses of these data (from both Perino and Ernhart, 1974, and Ernhart et al., 1981) be carried out.

C. Comments on Yamins (1976) Dissertation Study

1. Indicators Of Lead Exposure

Comments on specific aspects of the Pb exposure measurement methodology used in the Yamins (1976) dissertation study are as follows:

(a) Blood Pb was sampled by finger puncture, using established techniques for blood lead sampling, and the blood samples collected on filter paper presumably provided by the New York City Health Department. Upon collection, samples were transported to the New York City Health Department for lead measurement.

(b) The use of filter paper for blood collection raises the same question that is of concern in the Ernhart et al. follow-up study, i.e., the blood lead value must be corrected for hematocrit. The problem here occurs, actually, with each of the three indices (highest blood lead, mean blood lead, or most recent blood lead), so that it is not possible to assess the actual blood lead level in each case nor to determine the suitability of selecting one exposure index over another, since a given subject may have had variable hematocrit over the multi-sampling time period.

(c) Measurement of EP was also carried out in the laboratories of the New York City Health Department, quality control details for which were discussed among comments on the studies of Ernhart and coworkers (vide supra).

2. Psychometric Measurements and Procedures

In the Yamins dissertation, a variety of psychometric measurement procedures were employed, including some standardized instruments such as the Caldwell Preschool Inventory and other, more-or-less experimental measures such as the verbal repetition tasks.

Specific comments on the psychometric measurements utilized by Yamins (1976) include:

(a) This study employed the Peabody Picture Vocabulary Test (PPVT) to assess the intellectual development of the children. The PPVT is not a nonverbal measure of intelligence; it provides a narrow assessment of verbal abilities. However, the committee agreed that the PPVT was a reasonable measure of cognitive performance to use in the study because the scores correlated well with the other cognitive and experimental language measures, including the Caldwell Preschool Inventory, which is an appropriate measure for this population.

(b) The Ammons Quick Test was used to assess parental IQ in this study. As noted earlier, this measure has questionable validity. However, the pattern of correlations between the Quick Test scores and other variables does establish some credibility for its use in the current study.

(c) The experimental measures of language skill correlated with child's age, as one would predict, and yet appear to measure skills already tapped by the cognitive measures.

(d) There is no obvious explanation for the correlation of 0.30 between parent Quick Test scores and child's age. However, when age is partialled out, the correlation between child IQ and parental IQ is approximately 0.35, a value close to that found in other studies.

(e) The author (Yamins) administered all of the cognitive and language measures in the children's homes and was blind as to Pb levels at the time. She appears to have been appropriately trained and competent to collect the psychometric test data. Quality assurance checks regarding data collection (e.g., scoring, coding, and keypunching) could not be ascertained but are assumed to be the same as for the above Ernhart studies.

3. Statistical Analyses

Many of the same reservations expressed earlier regarding the analyses used in the 1974 and 1981 Ernhart studies also apply here. Specific concerns include the following:

(a) Stepwise multiple regression was employed to choose the set of covariates to be included in the final regression model with mean lead level. In the Committee's view, this is not the appropriate way to deal with potentially confounding factors. A backward elimination strategy starting with a model containing lead and all potential confounders is recommended since confounding involves association with both the outcome variable (e.g., a measure of learning performance) and the exposure variable (e.g., mean lead level). A forward selection approach, as was apparently employed by Yamins, ignores the relationship between the potential confounders and the exposure variable in choosing an appropriate subset for control, and hence can lead to inappropriate adjustment.

(b) The strategy for analysis described on page 79 of the Yamins (1976) dissertation is not appropriate for valid control of confounding effects (see preceding comment). Although it may produce the same results as a backward elimination approach, one cannot know without trying both approaches. In this case, backward elimination of variables does not markedly alter the outcome of the analysis. One Committee member (LH) found that the association between Pb and IQ remained after controlling for father's absence, parental IQ, parental education, birth order, and birth weight, using a backward elimination approach.

(c) The results displayed in Table 9 (page 82) of the dissertation do not, in the Committee's opinion, represent a strong indictment of lead exposure. Finding a few significant partial correlations of lead exposure with various dependent variables just by chance is not at all unlikely when performing several analyses using mutually correlated dependent variables. Apparently, no multivariate (as opposed to multivariable) analyses were employed to account for such correlations.

(d) The results in Table 10 (page 85) of the dissertation are based on a comparison of a "low lead" group to a "moderate lead" group (after dichotomizing lead exposure), with adjustment only for the covariate "age". Given that other potential confounders were apparently ignored, and given that several comparisons were made involving correlated responses, not much importance can be attached to the few significant findings reported.

(e) Results presented in Tables 11 (page 86), 12 (page 88), 14 (page 91), and 15 (page 93) of the dissertation are also based on controlling only for age. The Committee finds it difficult to place much importance on the findings presented in these tables.

4. Committee Conclusions and Recommendations

The Committee concludes that, despite reservations expressed regarding psychometric measurements employed in the Yamins study, the pattern of results obtained (including the intercorrelations between different measurement outcomes and parental measures) lends credibility to the psychometric results as reasonably reflective of the cognitive abilities of the subjects tested. The blood lead data require correction for hematocrit because of the use of filter paper for blood collection; unfortunately, the hematocrit values are not available to make this correction. In addition, the Committee finds that specific statistical analyses employed (including stepwise regressions and covariate analyses controlling only for age) are not the most appropriate for analyzing the Yamins data set. Rather, multivariate analyses should have been used that included other potential confounders besides age and, also, backward elimination of variables having negligible impact on the variance attributed to Pb. The Committee further finds that the very modest residual effects attributed to Pb based on the reported analyses controlling only for age are not convincing evidence for a negative effect of Pb on the cognitive abilities of the subject children.

REVIEW OF STUDIES BY DR. HERBERT NEEDLEMAN AND COLLEAGUES

A. Background Information

The Committee undertook detailed review of an epidemiological study published in 1979 by Dr. Herbert Needleman and associates (Needleman et al., 1979). In addition, limited review was carried out for several other reports (Burchfiel et al., 1980; Needleman, 1981; Needleman, 1982; Needleman, 1983; Bellinger and Needleman, 1983) published as follow-up analyses of the same data set and/or new data constituting extensions of the 1979 study.

Approximately 3329 children attending first and second grades between 1975 and 1978 in Chelsea and Somerville, Massachusetts, constituted the study population in the Needleman et al. (1979) study. Children submitted shed teeth to their teacher, who verified the presence of a fresh socket. The shed deciduous teeth were cleaned ultrasonically (discarding any containing fillings), followed by dissection of a 1-mm central slice and subsequent analysis of Pb in dentine tissue by anodic stripping voltammetry. Teeth were donated from 70% of the population sampled. Almost all children who donated teeth (2146) were rated by their teachers on an eleven-item classroom behavior scale. The results obtained for the rated children were reported to demonstrate a dose-response relationship between increasing dentine Pb levels and increasing percentages of students receiving negative (poorer) ratings on several of the 11 categories of classroom behavior, as shown in figure 1 below.

Following the teachers' ratings of classroom behavior, subsets of the rated students (reported to represent polar groups of children with the lowest and highest 10 percent of dentine Pb levels) were recruited for further, extensive neuropsychological evaluation by means of psychometric tests. Each subject whose initial tooth slice was in the highest 10th percentile (>24 ppm) or lowest 10th percentile (<6 ppm) was provisionally classified as having high or low lead levels, respectively. Repeat dentine lead samples from the same teeth were analyzed, when possible, and attempts were made to obtain and analyze other shed teeth from each subject provisionally classified in either lead exposure group (with more than one analysis being obtained for all but one subject). Parents of children provisionally classified as having either high or low dentine Pb levels were invited to have their children participate in further neuropsychological evaluations. Criteria were established for requisite agreement between replicate dentine sample analyses before the data for a given subject were included in the study; when requisite agreement was not found, then the subject was designated as "unclassified" and excluded from data analyses. Other children were excluded from the study because: (1) their parents were unable or unwilling to participate; (2) they came from bilingual homes;

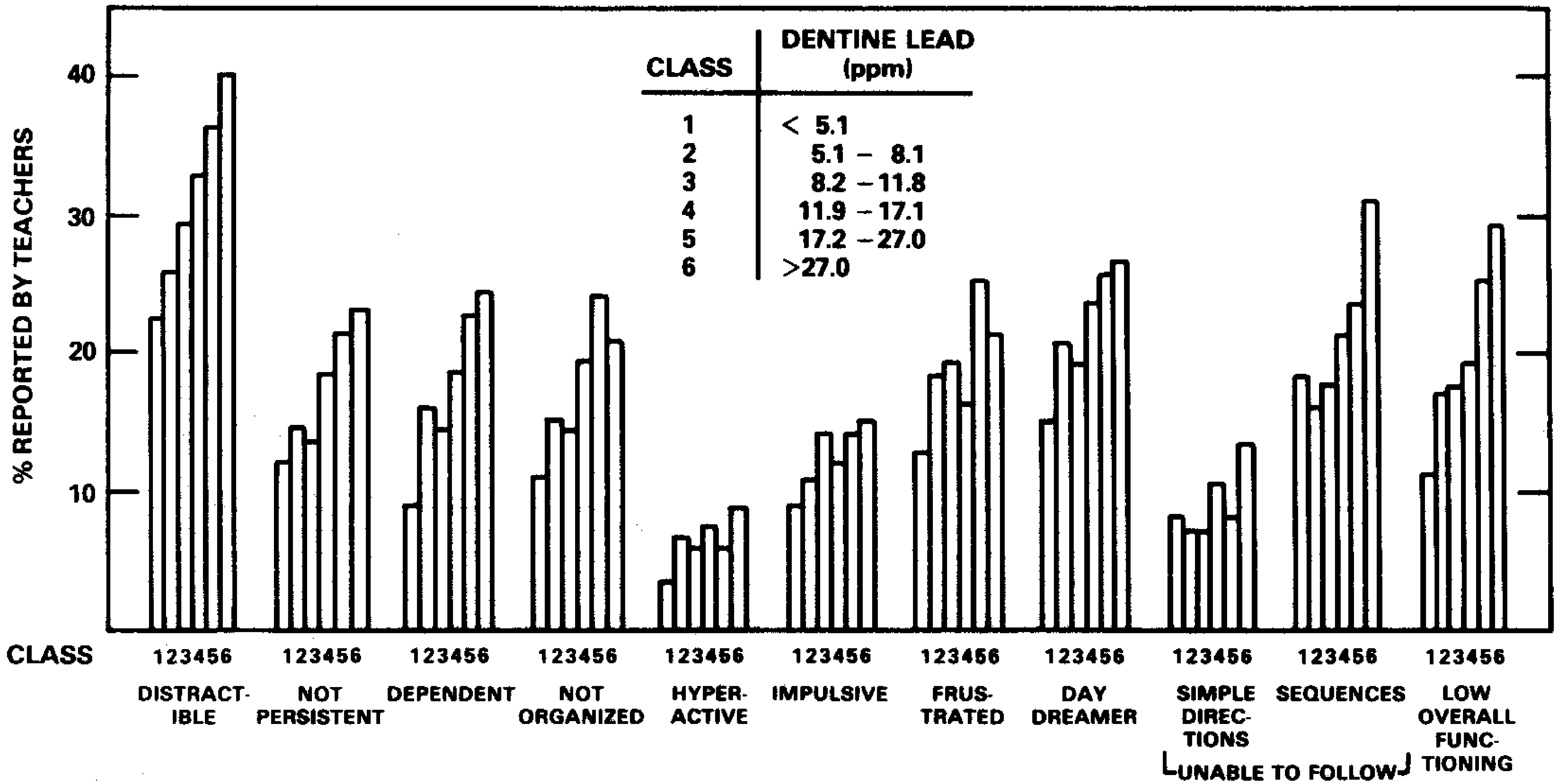

Figure 1. Distribution of negative ratings by teachers on eleven classroom behaviors in relation to dentine lead concentrations.

Source: Needleman et al. (1979).

(3) they had been diagnosed as having been lead poisoned; or (4) their medical history indicated a birth weight of <2500 g, delay in discharge beyond mother's discharge from hospital after birth, or a record of noteworthy head injury (any of which can correlate with slower neurobehavioral development). Table 1 from the Needleman et al. (1979) paper, reproduced below, lists the number of provisionally eligible children, those excluded from neuropsychologic testing, and those undergoing neuropsychologic testing who were retained or excluded from data analyses.

TABLE 1. REASONS FOR EXCLUDING SUBJECTS AND DISTRIBUTION OF FINAL DENTINE LEAD LEVELS IN INCLUDED AND EXCLUDED GROUPS

GROUP		NO.	DENTINE LEAD LEVEL		
			LOW	HIGH	UNCLASSIFIED
Provisionally eligible subjects:		524	258	187	79
Excluded from neuropsychologic testing:		254*	123	101	30
Bilingual home	84				
Not interested	57				
Moved	19				
Other†	94				
Total	254				
Subjects tested		270‡	135	86	49
Excluded from data analysis:		112†	35	28	49
Later tooth discordant	36				
Not discharged from nursery with mother, possible head injury, reported to have plumbism or bilingual home	76				
Total	112				
Cases scored and data analyzed		158	100	58	--

*Teachers' behavioral assessment available on 235.

†Infant at home, two working parents, etc.

‡Teachers' behavioral assessment available on 253.

Source: Needleman et al. (1979)

Mean dentine lead values for the 100 children included in the low-Pb and 58 in the high-Pb exposure groups were not reported by Needleman et al. (1979). However, Bellinger and Needleman (1983), who studied 141 of the 158 subjects of Needleman et al. (1979), reported those subjects to have mean dentine Pb levels of 6.2 ppm and 31.4 ppm, respectively. Mean blood Pb levels reported as having been assayed 4-5 years earlier for approximately 50% of the children in these two groups were 23.8 ± 6.0 μg/dℓ vs. 35.5 ± 10.1 μg/dℓ, respectively;

the highest blood-Pb level recorded was 54 μg/dℓ. The low-Pb and high-Pb group children underwent a comprehensive neuropsychologic evaluation, beginning with the Wechsler Intelligence Scale for Children-Revised (WISC-R), with the examiners blind to the Pb-exposure status of the children. In addition to the WISC-R, the children were administered, in set sequence, tests of: concrete operational intelligence; academic achievement (in mathematics, reading recognition, and reading comprehension); auditory and language processing; visual-motor reflexes; attentional performance; and motor coordination. While each child was being tested, the parents filled out a comprehensive medical and social history, received a 58-item questionnaire on parent attitudes, and took the Peabody Picture Vocabulary Test (PPVT). Also, 39 non-lead variables potentially affecting the children's development were scaled and coded, e.g., estimation of parental socioeconomic status (SES) by a two-factor Hollingshead index.

The scores of the high-Pb and low-Pb children for each of 39 control variables were compared statistically by the Student t-Test, with the two groups differing significantly on such variables as age, father's social class and father's education. Scores from the neuropsychologic evaluations of the high-Pb and low-Pb children were then compared statistically, using an analysis of covariance with dentine-Pb level as the main independent variable and with the following five covariates: father's SES (composed of education and occupation score); mother's age at subject's birth; number of pregnancies; mother's education; and parental IQ score. With the exception of these variables and age, the low-Pb and high-Pb groups were similar in regard to most of the non-Pb control factors.

Results of the neuropsychologic evaluations for the low-Pb and high-Pb groups can be summarized as follows: children in the high-Pb group were reported to have performed significantly less well on the WISC-R (especially on the verbal items), on three measures of auditory and visual processing, on attentional performance as measured by reaction time under varying delay conditions, and on most items of the teachers' behavioral ratings. The high-Pb children appeared to be particularly less competent in areas of verbal performance and auditory processing, having obtained lower scores, for example, on tasks requiring: response to verbal instructions of increasing complexity, immediate repetition of previously uttered sentences of increasing complexity, and discrimination of tone sequences of increasing complexity as either alike or different. Impaired focusing of attention (or distraction) of high-Pb children was also reflected by a significantly higher percentage of high-Pb children rating items being found to be significantly different (i.e., more negative) for high-Pb than low-Pb children at $p < 0.05$. Overall, the sum score (mean) of ratings of classroom behavior were found to be significantly poorer for the high-Pb children based on an analysis of covariance.

Burchfiel et al. (1980), using computer-assisted spectral analysis of recordings from a standard EEG examination on 41 (22 low-Pb and 19 high-Pb) children from the Needleman et al.

(1979) study, reported significant increases in percentages of low frequency delta activity and decreases in percentages of alpha activity in the spontaneous EEG of the high-Pb children. Percentages of alpha and delta frequency EEG activity and results for several psychometric and behavioral testing variables (e.g., WISC-R full-scale IQ and verbal IQ, reaction time under varying delay, etc.) obtained for the same children were then employed as input variables (or "features") in direct and stepwise discriminant analyses. The separation determined by these analyses for combined psychological and EEG variables ($p<0.005$) was strikingly better than the separation of low-Pb from high-Pb children using either psychological ($p <0.041$) or EEG ($p<0.079$) variables alone.

A more recent paper by Needleman (1982) provided a summary overview of findings from the Needleman et al. (1979) study and findings reported by Burchfiel et al. (1980) concerning EEG patterns for a small subset of the children included in the 1979 study. Needleman (1982) also summarized results of an additional analysis of the Needleman et al. (1979) data reported elsewhere by Needleman et al. (1982). More specifically, cumulative frequency distributions of verbal IQ scores for the low-Pb and high-Pb subjects from the 1979 study were reported by Needleman et al. (1982) and reprinted as Figure 2 of the Needleman (1982) paper, as shown below in Figure 2. One key point made by Needleman (1982) was that the average IQ deficit of four points demonstrated by the Needleman et al. (1979) study reflected not just further impairment of cognitive abilities of children with already low IQs but rather a shift downward in the entire distribution of IQ scores across all IQ levels in the high-Pb group, with none of the children in that group having verbal IQs over 125.

The Bellinger and Needleman (1983) paper provided still further follow-up analyses of the Needleman et al. (1979) data set, focusing mainly on comparison of the low-Pb and high-Pb children's observed IQs versus their expected IQs based on their mothers' IQs. Bellinger and Needleman reported that regression analyses showed that the IQs of children with elevated levels of dentine Pb (>20 ppm) fell below those expected based on their mothers' IQs and that the amount by which a child's IQ falls below the expected value increases with increasing dentine Pb levels in a nonlinear fashion (see Figure 3 below, showing plots of IQ residuals by dentine Pb levels as illustrated in Figure 2 of the Bellinger and Needleman paper). In fact, examination of the scatterplot shown in Figure 3 and the discussion of results provided by Bellinger and Needleman (1983) indicate that regressions for the 20-29.9 ppm group did not reveal significant associations between increasing Pb levels in that range and IQ residuals, in contrast to statistically significant ($p<0.05$) correlations found between IQ residuals and dentine Pb in the 30-39.9 ppm range.

In order to evaluate critically the above studies, the Committee met with Dr. Needleman at his University of Pittsburgh (Children's Hospital) office facilities in Pittsburgh, PA, on

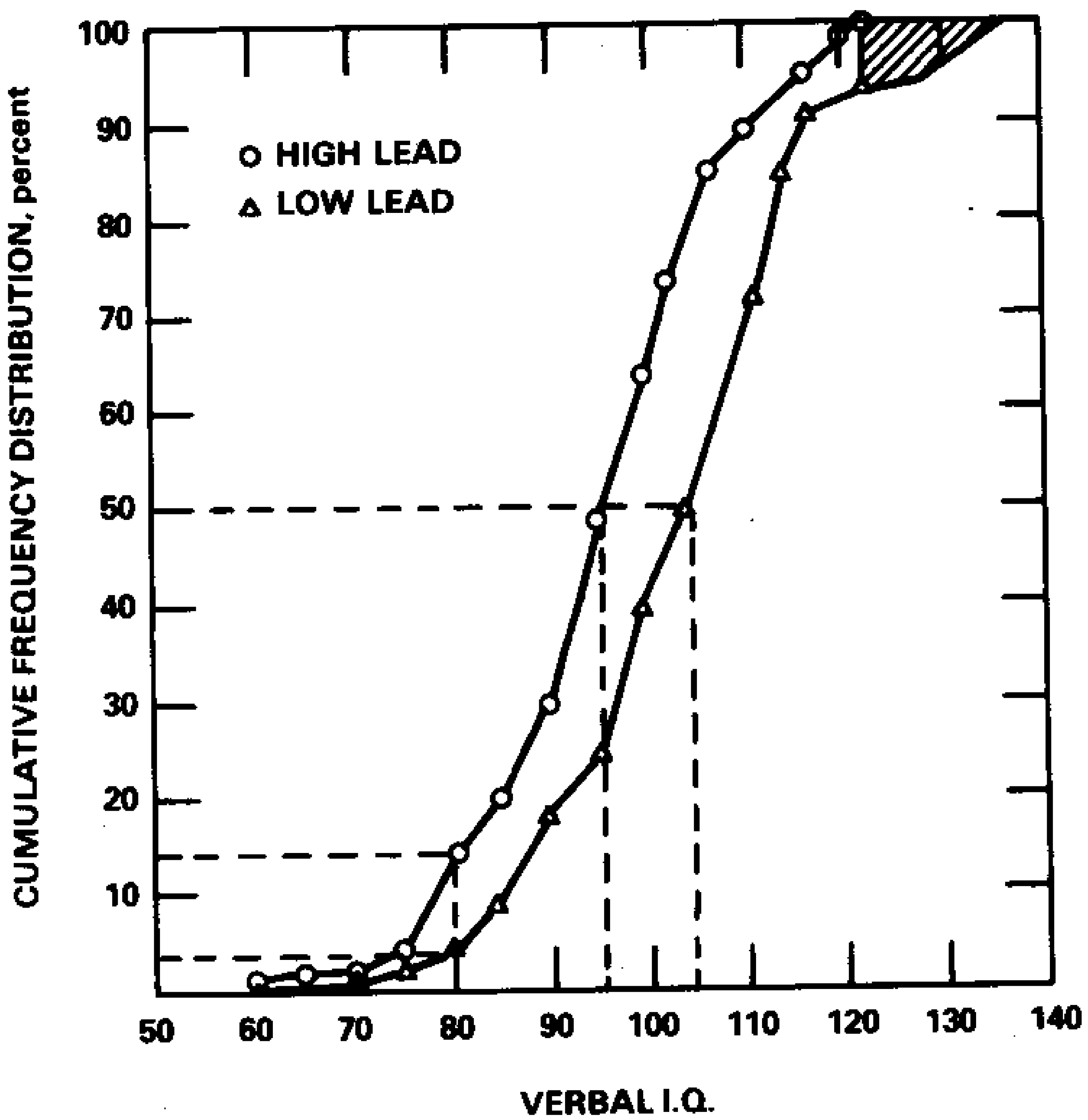

Figure 2. Cumulative frequency distributions of verbal IQ scores in high and low lead subjects.

Source: Needleman (1982) and Needleman et al. (1982).

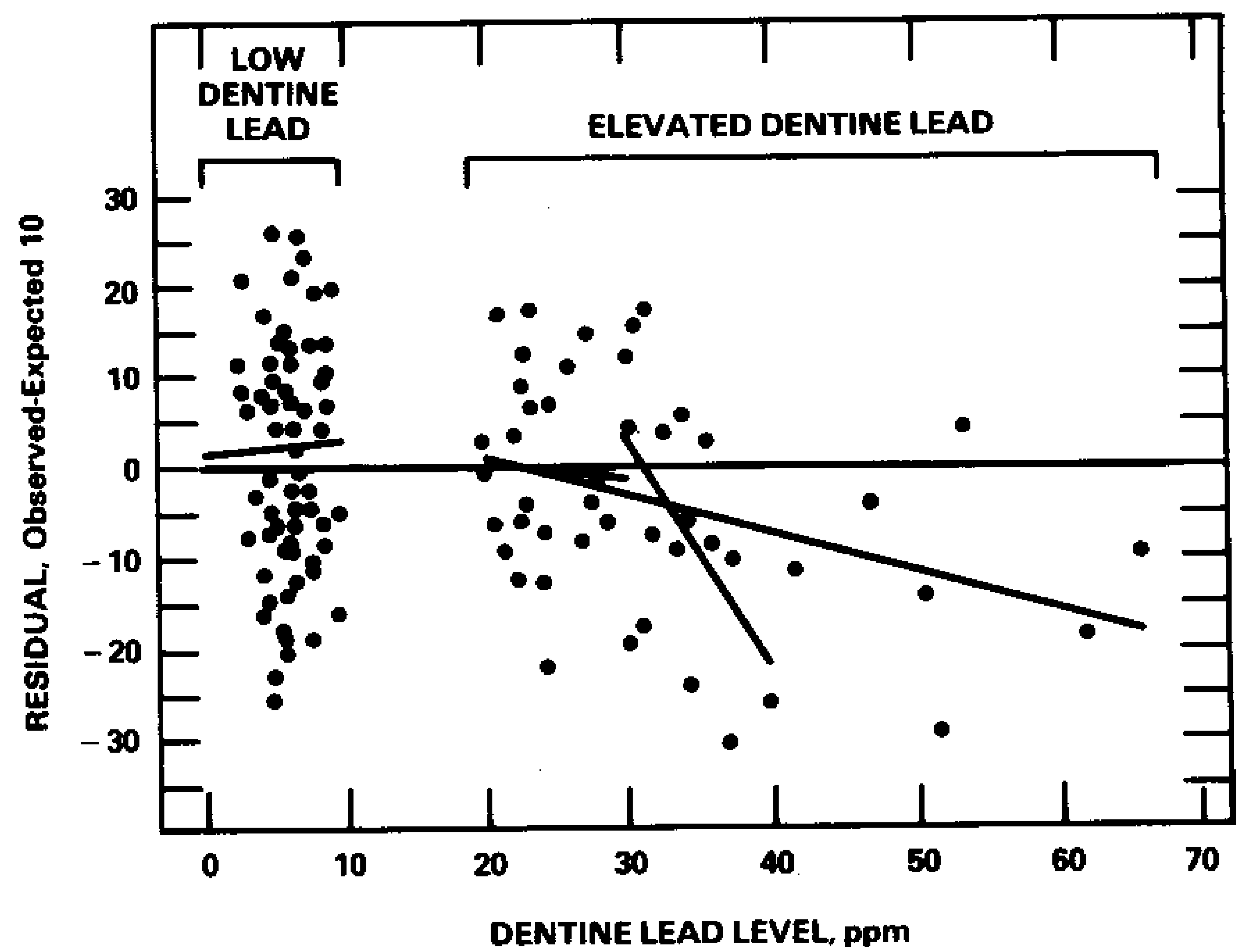

Figure 3. Scatterplot of children's dentine lead level versus residual. Regression lines are shown for four ranges of lead level: low lead, 0.9 to 9.9 ppm; elevated lead, 20.0 to 66.0, 20.0 to 29.9, and 30.0 to 39.9 ppm.

Source: Bellinger and Needleman (1983).

March 30-31, 1983. During that meeting, Dr. Needleman presented an overview focusing mainly on the objectives, design, data collection and statistical analysis procedures, and findings for the original study reported by Needleman et al. (1979). Dr. Needleman also provided additional information regarding follow-up analyses or extensions of the 1979 study either published in other papers referred to above or expected to be published in the near future. This additional information included comments regarding the conduct of a separate study involving the evaluation of teachers' ratings of classroom behavior of children in Lowell, MA (a different population from the one sampled in the 1979 study). Certain listings of raw data (provided in coded form to protect the privacy of subjects), computer printouts summarizing data entries for statistical analyses or results of such analyses, and miscellaneous other pertinent materials were discussed with Dr. Needleman during the March 30-31 meeting. Additional information was requested by the Committee in order to clarify factual points or to help resolve evaluative issues arising from the discussions in March with Dr. Needleman. A portion of the information was provided during the 2-3 months following the March meeting. (See Attachment 1 for a list of materials examined by the Committee in connection with the review of the subject studies.) The Committee's comments regarding the most salient points of concern and controversy related to methodological and other features of the above studies by Needleman and colleagues are presented below.

B. Comments on Needleman et al. (1979) Study

1. Indicators of Exposure

In the principal study (Needleman et al., 1979) as well as in subsequent reports on subsets of subjects from the initial population (e.g., Burchfiel et al., 1980; Bellinger and Needleman, 1983), Pb exposure in the pediatric subjects was assessed by analysis of Pb in the dentine of deciduous teeth. In contrast to blood Pb, which is an exposure marker for relatively recent exposure, whole-tooth or tooth-region analysis for Pb content yields an index of cumulative Pb exposure of the subject up to the time of exfoliation.

In the report of Needleman et al. (1979), blood Pb levels as an additional index of prior exposure were reported as only being available for some (approximately 50%) of the subjects in the highest/lowest deciles, and were discussed only in terms of group means. These measurements were reported to have been obtained as part of a blood screening program in the subject communities 4-5 years prior to collection of tooth samples.

Observations and comments concerning specific aspects of the Pb exposure indices and associated methodological procedures include:

(a) Dentine was isolated from each tooth sample by a procedure described in an earlier report in which the present principal investigator (Dr. Needleman) was also heavily involved (Shapiro et al., 1973). In that procedure, very thin sections of tooth were carefully cut from the central sagittal plane, dentine (coronal plus secondary) was mechanically separated from enamel and circumpulpal dentine, and the dentine samples were dissolved in acid and subsequently analyzed by anodic stripping voltammetry (ASV) for Pb content.

(b) According to Dr. Needleman, the type of tooth selected for analysis was fairly consistent: mainly the incisor, and some bicuspids. Hence, it appears that any variation in Pb content which might arise from random selection of diverse types of dentition (due to variation in Pb content across different types of teeth) would have been minimal.

(c) Unlike blood Pb, there is no external quality control framework by which to evaluate the dentine Pb analyses such as were done in the subject studies. One must therefore consider the specific steps in the analysis against a general body of information. Two steps in the dentine Pb measurement need to be considered. According to Needleman, the homogeneity of dentine in terms of Pb content for a given tooth can vary sufficiently that tooth sectioning was confined to an initial central sagittal sectioning in all cases, the sectioned sample providing two (replicate) samples for analysis. Once dentine was isolated, its subsequent analysis by ASV would be expected to be achieved with good accuracy and precision, given available data for overall performance of ASV assays of Pb in biological matrices and the fact that such Pb levels are relatively high. Since the major determinant of variance in the replicate (single tooth)/duplicate (multiple teeth) analyses was the dentine isolation step, Needleman et al. (1979) attempted to minimize unacceptable variance by use of intra-sample concordance criteria in analyzing relationships between dentine Pb levels and the results of the psychometric test battery.

(d) The impression gained from close reading of the Needleman and other reports, as well as discussions with Needleman, is that use of dentine Pb values entails methodological skill at the step of dentine isolation. From the information available to the Committee as to actual variation in dentine Pb across a given tooth sample, it appears that ±15% represents a reasonable specific estimate of variability for the dentine analysis for subjects from the low-Pb and high-Pb groups included in statistical analyses of neuropsychologic test outcomes (i.e., for subjects with the greatest concordance among their dentine Pb values). Much greater variation existed among replicate or duplicate dentine Pb values for individuals from the low-Pb and high-Pb groups excluded from the statistical analyses. Examination of replicate/duplicate values for measurements for all subjects in the study (including the 2000+ students for which teachers' ratings were obtained) would be necessary to determine an overall estimate of variance for dentine Pb measurements in the study.

(e) Whole-tooth analysis would be simpler technically and has been more often employed than specific tooth-region analysis. However, one can also expect that such a measure would be less sensitive as a biological index of Pb exposure due to the inclusion of enamel, a region that contributes significiantly to tooth mass but has relatively invariant low Pb content regardless of exposure. In the case of whole-tooth Pb, it is known that Pb content is linearly related to age of the subject and that the values of the slopes of Pb content vs. age are better indicators of Pb exposure than just the Pb concentrations alone (Steenhout and Pourtois, 1981). Shapiro and coworkers (1978) have also reported that there is a better correlation of tooth Pb concentration/year with either blood Pb or erythrocyte porphyrin than just Pb concentration unadjusted for age. Expression of dentine Pb as a function of age in the present study ($\mu g\ Pb \cdot g\ dentine^{-1} \cdot yr^{-1}$) would be desirable, especially because the mean age of the high-Pb subjects was greater than that of the low exposure group. While this method of expressing Pb exposure would have minimal effect on the categorization of subjects into high- and low-Pb groups, it might be expected to influence relationships between Pb and other variables, e.g., in Figure 3 (above) from Bellinger and Needleman (1983).

(f) The relative quality of the earlier blood Pb determinations for some low-Pb and high-Pb subjects cannot be ascertained and must be considered more suspect than the main exposure measure used (i.e., dentine Pb). At the time blood Pb levels were measured in the subjects, the quality control for the community Pb-screening programs was minimal, with sampling being done by finger puncture and transfer to capillary tube (communication of V. N. Houk to H. L. Needleman, see Needleman, November 22, 1982: letter to L. D. Grant). These limitations on the relative reliability of such measurements apparently were the reason for Dr. Needleman's discussion of these values only in terms of group means for the low-Pb and high-Pb subjects. Given present knowledge about the impact of sampling protocols on the accuracy of blood Pb measurements, one can reasonably say that finger puncture plus capillary tube versus venous puncture plus low-Pb blood tube would impart a significant positive bias to the blood Pb levels obtained. Hence, the overall means reported for the low-Pb and high-Pb groups in terms of blood Pb would, if anything, likely be higher than their true values for the 50% of the low-Pb and high-Pb children sampled.

(g) Apart from the issue of reliability of the blood Pb measures under consideration is the question of age of the children at the time of blood sampling relative to the known variation of blood Pb with age in children for a given exposure setting: i.e., blood Pb levels in children generally tend to peak at 2-3 years of age and decline in subsequent years. Available information on the ages of the children at the time of psychometric testing, the years when such testing occurred, and the years when blood Pb measurements were made indicate that the ages of some children at the time of blood Pb measurement were probably at or not materially beyond the period of typical peaking in blood Pb, but others may have been sampled at a later age within 1-2 years (during 1973-74) prior to their participation in neuropsychologic testing while in first or second grade (as early as during 1975-76). It would be necessary to know the ages of specific subjects when the blood lead determinations were done and their age at neuropsychologic testing before reliable judgments could be made regarding the representativeness of the reported mean blood Pb values for either the low-Pb or high-Pb subjects.

2. Psychometric Measurements and Procedures

The study employed a comprehensive neuropsychological battery to assess the children's behavioral functioning. The measures included the WISC-R, Piagetian tasks, and selected tests of academic achievement, auditory and language processing, visual and motor performance, reaction time, motor coordination, and teacher ratings of classroom behavior. Mothers' attitudes toward child rearing and parental IQ (indexed by the PPVT) were also assessed. The PPVT is a narrow assessment of mothers' intelligence, but their PPVT scores correlated in expected ways with other variables in the study.

Dr. Needleman administered the PPVT to the mothers, and three other examiners administered the WISC-R and other assessments of the children in a fixed order. The examiners were blind as to children's Pb levels and scored the test immediately after the test session. It is not known how qualified the examiners were to administer individual tests, but Dr. Needleman reported that the examiners were instructed on how to administer and score the test.

In a recent publication (Needleman, 1983, p. 243) additional details of the psychometric procedures were reported. Children with low-Pb exposure were scheduled early in the study, because "I wanted my technicians to get some experience with normal children." In addition, Dr. Needleman told the Committee during the meeting in Pittsburgh that the study began with three technicians, one of whom left during the study and was replaced by a fourth tester. Results of the Needleman et al. (1979) and related publications using the psychometric test scores could be affected by this nonrandom ordering of assessments.

Dr. Needleman reported that quality assurance procedures for ensuring the accuracy of teachers' ratings and neuropsychologic test results used in statistical analyses included: (1) summing of teacher ratings across items and entry of scores for each item and sum scores onto computer cards, followed by verification and transfer onto magnetic tapes; (2) checking of neuropsychologic test scores by a second examiner other than the one doing the original scoring, followed by entry onto cards, verification, and transfer to magnetic tapes; (3) subsequent 5% sampling of computer tape entries to check accuracy against original data listings, with 12 errors in 15,600 columns of entries being found and corrected.

The Committee's inspection of raw data during the March visit with Dr. Needleman revealed some problems of another kind, however. Printouts of parental IQ data for low- and high-Pb subjects included in the statistical analyses (e.g., analyses of covariance) published in the 1979 article revealed errors in calculating parental IQ values for some subjects when their fathers and mothers were both administered the Peabody Picture Vocabulary Test. Instead of an average of mothers' and fathers' scores (midparent IQ), the parents' IQ scores were evidently combined by taking one-half of one parent's score and adding that value to the other parent's score. This erroneous procedure resulted in some parental IQ values that lie well outside probable values. These errors were confirmed later by Dr. Needleman in his letter of October 4, 1983 to Dr. Bernard Goldstein. The impact of this is to introduce error into the results of all of the published analyses of the data set where parental IQ was included as a variable. Correcting these errors would alter the results of the analyses. The precise influence of the errors on the results can be determined only by reanalyzing the data, and the Committee urges that this be done.

During inspection of raw data, the Committee also noted a seemingly higher proportion of large discrepancies between the children's WISC-R Verbal and Performance Scale IQ scores than would be expected in an unselected sample. The discrepancies seemed to be distributed across the high-Pb and low-Pb groups. Neither sufficient time nor facilities were available during the data inspection to carry out an adequate quantitative analysis of the relationship of the verbal to performance IQs, but if the discrepancies are as large and/or numerous as they appeared to be, this may raise questions about the validity of the WISC-R assessment as employed in this study.

3. Statistical Analyses

Comments on specific aspects of the statistical analyses employed in the Needleman et al. (1979) study include the following:

(a) The statistical analyses for teachers' rating scores for classroom behavior were based on classification of the children's Pb exposure levels in terms of first dentine lead values obtained for the first tooth submitted by each of the 2146 subjects included in the analyses (vide supra). Six lead exposure categories were defined as indicated in Figure 1 (i.e., <5.1, 5.1-8.1, 8.2-11.8, 11.9-17.1, 17.2-27.0, and >27.0 ppm dentine Pb), with group boundaries chosen to give symmetrical cell sizes around the median (i.e., 6.8% in Groups 1 and 6, 17.6% in Groups 2 and 5, and 25.6% in Group 4, respectively). However, no statistical analyses that take into account other potentially confounding variables were done on the teachers' rating data shown in Figure 1 and, thus, the dose-response relationships shown in that figure cannot be attributed to Pb exposure alone.

In addition, questions arise regarding the appropriateness or accuracy of classification of subjects in terms of the narrow dentine Pb ranges employed in plotting the dose-response data shown in Figure 1. Given the 15% variability noted for replicate analyses of teeth for those subjects with the most highly concordant dentine Pb values, many subjects who were included in one or another of the six exposure categories based on first dentine Pb analyses could be more appropriately classified as belonging in a different exposure category, according to later replicate/duplicate dentine Pb values. This is particularly likely if the same or analogous concordance criteria used by Dr. Needleman to select low-Pb and high-Pb subjects for inclusion in statistical analyses of later psychometric test scores were used for the teachers' rating analyses. Inspection of raw dentine Pb values for subjects provisionally classified as low-Pb or high-Pb subjects for psychometric testing, but then excluded from final statistical analyses of the psychometric test results because of non-concordance of dentine Pb values, revealed that shifts across the six exposure categories could be substantial if replicate or duplicate dentine Pb values beyond the first dentine Pb value were taken into account.

(b) In regard to the statistical analysis of results from the subsequent psychometric testing phase of the study, comparisions were made only between those children reported to be ranked in the highest 10th percentile for dentine Pb concentrations and those in the lowest 10th percentile. This strategy certainly maximizes the chances for finding significant differences. Reviewing Figure 1 of the 1979 report, the Committee notes that a group with "moderate" exposure might serve to provide evidence for a dose-response relationship (which, if found, would argue more strongly in favor of a causal connection than the polar low-Pb vs. high-Pb group comparision used). An earlier report on the subject (Needleman, 1977) suggests an intention to use low, medium, and high dentine Pb groups, and a very recent report (Needleman, 1983; Table 6; p. 237) does show some results on behavior ratings for a group of 13 subjects with "middle" dentine Pb levels. Psychological test scores have evidently been obtained for a middle group of subjects (Needleman, 1983, p. 242). The Committee recommends that analyses be undertaken to evaluate any available psychometric testing data for "intermediate" lead-exposure subjects.

(c) Many questions about sampling procedures arise from the exclusion of large numbers of potential participants in the psychometric testing phase of this study. From 542 provisionally eligible participants, almost half were excluded from neuropsychological assessment, and 41 percent of those tested were later excluded from data analyses. Although reasons for the exclusions were given (see Table 1 of the 1979 article), the distribution of demographic and psychological outcome measures for those excluded from the low- and high-Pb groups was provided neither in the published article nor by the investigator to the Committee. The Committee could not fully evaluate sources of possible bias due to such exclusions in the selection of the sample reported in this paper and other publications reviewed.

Some of the criteria used to define Pb exposure levels or to exclude subjects from statistical analyses seemed arbitrary, and different results might have been obtained with application of equally good or better alternative criteria for classification of Pb exposure levels and groups. For example, some subjects provisionally classified as low-Pb or high-Pb subjects based on initial dentine Pb values were excluded from final data analyses based on discordances arising from later replicate or duplicate Pb values obtained for the same or different teeth, although certain key "discordant" dentine Pb values were not meaningfully different from the cut-off criteria levels for inclusion as low-Pb or high-Pb subjects. Thus, exclusion of some subjects from the low-Pb group for statistical analyses hinged on a single dentine Pb value (e.g., 10.1

or 10.5 ppm) barely exceeding the 10 ppm criterion ultimately selected and rigidly enforced as defining the low-Pb (or lowest decile) exposure group, although such dentine Pb values were as likely to have true readings below 10 ppm as were certain key values (e.g., 9.5 or 9.8 ppm) for some subjects included in the low-Pb group likely to have true readings above 10 ppm. Also, since inclusion or exclusion of subjects in the statistical analyses was based on dentine Pb values for all teeth submitted by a given subject over the course of the study, some subjects may have been classified as high-Pb or low-Pb children (or excluded from analysis) based on replicate or duplicate dentine Pb values obtained for teeth shed up to 1-3 years after their psychometric testing. The impact of this may not have been symmetrically exerted on the high-Pb and low-Pb groups. That is, it is not likely that "actual" high-Pb exposure children at the time of psychometric testing would have distinctly lower later dentine Pb values; but low-Pb children with initial values <10 ppm could have experienced lead exposures after psychometric testing that substantially increased their later dentine Pb values and resulted in their exclusion from the low-Pb group.

(d) Normalized outcomes for which age-normed scores were not available were constructed by regressing on age before analysis of covariance. Assuming that age effects were accounted for, five covariates (namely, mother's age at subject's birth, mother's educational level, father's socioeconomic status, number of pregnancies, and parental IQ) were considered. Only five covariates were used because that is the limit dictated by a widely used computer software package (SPSS). However, the number of covariates considered should not be arbitrarily dictated by the constraints of a packaged program but should be determined with the goal of properly controlling confounding variables. Father's education (grade) level was not included separately, although (as Dr. Needleman argued) it is part of father's socioeconomic status. It would seem to be better, based on the results shown in Table 5 of the 1979 report, to use father's education directly rather than as part of a diluted "socioeconomic status" variable.

(e) The Committee reviewed computer printouts from numerous SPSS analysis of covariance runs on psychometric testing data indicated by Dr. Needleman as forming the basis for the results and conclusions presented in the 1979 report and noticed many missing data points among the analyses. In fact, the actual number of data points used in certain regression analyses was sometimes as much as 20%

fewer than those for the 158 cases claimed in the 1979 paper to have been analyzed for the low-Pb and high-Pb subjects. For example, the analyses later reported in the Bellinger and Needleman (1983) paper (on the same data set discussed in the 1979 article) are based on 17 fewer cases than the 158 stated to have been included in the final statistical analyses of psychometric test results appearing in the 1979 article, because of missing parental IQ data for the 17 cases. Missing data, not alluded to in the 1979 report, can pose a serious validity problem if the missing observations are not randomly distributed across the important variables. The effects of such missing data are impossible to assess without detailed analysis of the available data set.

(f) Based upon cursory inspection of the numerous statistical analysis computer runs provided by Dr. Needleman (which was all that was possible during the limited time of access to the printouts), the Committee came away with the impression that most runs led to non-significant findings. In a recent publication (Needleman, 1983), the investigator notes that of the 66 outcomes evaluated, 15 were significantly different between the low- and high-lead groups, given the control variables included in the analyses. He notes that 1 in 20 would be expected by chance, <u>if the outcome variables were uncorrelated</u>. Of course, most of the psychological assessments in this study are moderately to highly correlated, so that this probability does not apply. In addition, apparent group differences are affected by the method of handling important covariates.

(g) Of special interest, printouts for several regression analyses in which child's age was entered as a control variable showed reduced and generally non-significant coefficients for Pb levels, but such findings are not presented in the 1979 report or later articles by Needleman and colleagues. This is in contrast to the earlier reporting (Needleman et al., 1978) of statistically significant Pb effects when age was included as a covariate in preliminary statistical analyses performed when the collection of psychometric data for the study was about one-half completed. The standardized psychometric measures with age norms provided do not perfectly correct for age differences in a specific sample. Because there are significant age differences between the high-Pb and low-Pb groups in this study, the regressions of raw test scores on child's own age would have been the more desirable analyses to report. The Committee has reached this conclusion despite the principal investigator's (Dr. Needleman's) argument that it is undesirable to "correct for age twice."

4. Committee Conclusions and Recommendations

Estimation of Pb exposures by dentine Pb measurements is more appropriate than blood Pb as an index of cumulative exposure, and the analytical determination of such dentine Pb levels appears to have generally been done competently in the study. However, it is not possible to estimate the variance of the dentine Pb measurements in replicate/duplicate analyses (beyond the 15% estimate arrived at for replicate analyses for the most concordant samples) without full access to the coded, raw data of all children who participated in the study. The blood Pb measurements, obtained earlier for some of the children, are of unknown reliability. Because the blood data appear to have been obtained at varying ages for the children sampled, the reported blood Pb data probably do not uniformly assess peak exposure levels for them and cannot be accepted as providing quantitive estimates of Pb levels associated with any neurobehavioral deficits demonstrated to exist among the children studied.

Teacher ratings of children's classroom behaviors were collected on more than 2000 children who also contributed shed deciduous teeth for dentine Pb analyses. The failure to revise the lead classification of the children based on discrepancies with later replicate/duplicate dentine Pb values in the analysis of teachers' ratings contrasts sharply with the demand for concordance of dentine Pb readings in the neuropsychological testing phase of the study. Also, the failure to analyze for possible contributions of confounding factors or covariates to the teachers' rating results is disturbing. (The covariance adjustment used for teachers' ratings on the 158 children included in the neuropsychological testing phase of the study is subject to the criticisms noted for other analyses of data for those groups.) The dose-response relationships reported to exist between dentine Pb levels and teachers' rating scores, therefore, cannot be accepted as valid based on the published analyses.

A comprehensive neuropsychological test battery was administered to the children defined as belonging in low-Pb or high-Pb subgroups. Serious questions exist regarding the basis for classification of subjects in these groups or exclusion of others from them. Also, discrepancies between WISC-R verbal and performance scores, if as large or numerous as they seem upon cursory inspection, may raise questions about the test administration or the sample selection. Errors in the calculation of some averaged parental IQ scores, evident in coded materials provided to the Committee, introduced unknown errors into the regression analyses for the psychometric testing results. The use of the PPVT for parental IQ was not ideal, but was still acceptable. Exclusion of large numbers of eligible participants prior to data analysis could have resulted in systematic bias in the results. However, the Committee was unable to evaluate this possibility fully, given the limited information made available by the investigator.

The treatment of covariates in the statistical analysis of the psychometric testing results was unsatisfactory. The failure to report statistical analyses showing generally reduced or non-significant negative correlations between dentine Pb levels and performance on the psychometric tests also lessens the credibility of those few statistically significant effects attributed to Pb in the published version(s) of the Needleman et al. (1979) study.

In summary, at this time, based on the questionable Pb 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. The Committee strongly recommends that the subject data set be reanalyzed to correct for errors in data calculation and entry noted above, that the reanalysis be based on better exposure classification of subjects, and that all potentially confounding variables (including age) be assessed using a backwards elimination approach analogous to that recommended earlier for the reanalysis of Ernhart data.

C. Comments on the Burchfiel et al. (1980) Study

The Committee carried out only a very preliminary review of the Burchfiel et al. (1980) study, focusing mainly on consideration of Pb exposure, neuropsychologic testing, and statistical analysis aspects of the study. Review of the electrophysiological recording aspects of the study would require additional committee members or a separate review committee with recognized expertise in electrophysiology and, in particular, electroencephalography.

In view of the fact that the Pb exposure and psychometric measurement data utilized in the Burchfiel study are subsets of the data underlying the Needleman et al. (1979) article discussed above, most of the preceding comments regarding those aspects apply here as well. Only a few additional remarks are, therefore, felt to be necessary here. Specifically, no definite dentine Pb or blood Pb values were reported for the specific children from the Needleman et al. (1979) study who underwent the EEG evaluations reported by Burchfiel et al. (1980). It is therefore impossible to determine with any confidence the specific Pb exposure levels (including either blood Pb or dentine Pb values) that may have been associated with the reported EEG effects. Nor is it possible to accept with much confidence any reported relationships between the observed brain wave alterations, the psychometric testing scores, and Pb exposure classification as low-Pb or high-Pb, especially in view of the various problems noted above regarding exposure classification, psychometric testing, and statistical treatment of covariates or confounders that preclude acceptance of the findings reported in the 1979 publication.

D. Comments on the Needleman (1982) Report

The 1982 report published by Needleman represents mainly a summary or restatement of findings already reported in the earlier Needleman et al. (1979), Burchfiel et al. (1980), and Needleman et al. (1982) publications. The comments provided above on the first two of the earlier publications, obviously, also apply here.

One additional point worthy of discussion concerns the plot of cumulative frequency distributions of verbal IQ scores for low-Pb and high-Pb subjects shown in Figure 2 of the Needleman (1982) report, as reprinted from the Needleman et al. (1982) article. Given the serious reservations expressed earlier by the Committee regarding the Pb-exposure classification procedures, aspects of the psychometric testing, and statistical treatment of covariates or confounding factors as employed in the analyses reported in the 1979 article, the particular cumulative distribution curves shown in the figure for verbal IQ scores among the low-Pb and high-Pb subjects cannot be accepted at this time as being either qualitatively valid (i.e., as demonstrating lower IQs for high-Pb subjects than for low-Pb subjects) or quantitatively accurate (i.e., in terms of absolute decreases in IQ implied to be associated with Pb exposure). Similarly, the Committee finds certain statements in the discussion (page 731 of the 1982 Needleman paper) of the cumulative distribution curves to be somewhat misleading in noting that none of the included high-Pb subjects had an IQ over 125 (while 5% of the low-Pb subjects did) but failing to mention that at least one subject excluded because of overt plumbism had a full-scale WISC-R IQ over 125.

E. Comments on the Bellinger & Needleman (1983) Study

This paper reports two kinds of reanalyses of the data from the previous (Needleman et al., 1979) report and, again, most of the comments made earlier on aspects of that study also apply here. Certain additional comments are warranted, however. First, child IQ is regressed on mother's IQ separately for the low-Pb and high-Pb groups. The results are that mother-child IQ correlations do not differ for the two Pb exposure groups and the high-Pb group has lower than predicted IQ scores (controlling for maternal IQ).

Second, the residuals of child's IQ regressed on mother's IQ from the first analysis were regressed on dentine Pb levels, arranged by individual values. Four ranges of lead values were used to estimate regression slopes of residual IQ on lead. The sample size for the low-Pb group in this report was N=94; for high-Pb, N=47; and for two subsamples of the high-Pb group, i.e., dentine Pb levels of 20.0-29.9 ppm, N=24, and for 30.0-39.9 ppm, N=17. The latter two groups are far too small to be used to estimate slopes that can be credibly generalized to other samples. The results for the regression of child's IQ residuals on lead in the low-Pb group had, not surprisingly, a slope of zero because of extreme restriction of the range of lead values. The slope for the high-Pb group was -0.36, significantly different from zero. One serious problem in interpreting these results is that only maternal IQ was used as a "covariate" for child IQ. No other background factors, as reported in the earlier paper, were included as adjustments for the residualized IQ scores in this study.

To control, in part, for additional covariates that could affect the relationships between residual IQ and lead level, a stepwise regression was done. Surprisingly, and contrary to all of their other analyses of these data, lead level was allowed to enter the equation second, before two of the three control variables. Table III of the Bellinger and Needleman (1983) article reported results in the form of unstandardized regression coefficients without accompanying standard errors or significance levels. The F ratios reported seemed to be those of the equations, not of the individual coefficients, except of course for the first variable in the first step. Thus, it is not clear that the Pb coefficient is actually reliably different from zero.

Given the above problems and concerns, the reanalyses of the Needleman data set presented in the Bellinger and Needleman (1983) paper cannot be accepted as providing credible or reliable estimates of quantitative relationships between Pb exposure and neuropsychologic deficits in children. Nor can the reported results be taken as either qualitatively supporting or refuting the hypothesis of associations between low-level lead exposure and cognitive deficits in children.

F. Comments on the Needleman (1981) Report

Shed teeth and teacher ratings were collected in 1977-1978 from a new sample of about 1300 first-grade children in Lowell, MA. Children were classified into five groups according to their dentine Pb levels: Group 1, ≦6.4 ppm; Group 2, 6.5 to 8.7 ppm; Group 3, 8.71 to 12 ppm; Group 4, 12.01 to 18.1 ppm; Group 5, ≧18.2 ppm. The association of teacher ratings on 11 behavior scales with Pb levels is displayed separately for males and females in Figures 4 and 5. No effort was made to control for confounding variables in this overall set of results.

Essentially complete follow-up data and Pb levels were obtained on 130 children of the 447 males selected for follow-up. Given the design of this study, the expected analysis would investigate the relationship between teacher ratings and Pb level following adjustment for confounding variables, collected on the follow-up sample. Such an analysis was not reported.

In the Committee's view, these data should be reanalyzed to show clearly the form of the relationship between Pb level and teacher ratings, with appropriate controls for the follow-up subjects.

POSTSCRIPT

In addition to evaluation of the studies of Ernhart and Needleman, the Committee reviewed available reports (some published and others as yet unpublished) of other studies from the United States and Europe. These studies included, for example, those by: Winneke et al. (1982, 1983), Winneke (1983), Yule et al. (1981), Lansdown et al. (1983), Smith (1983), and Harvey et al. (1983). Although an exhaustive, in-depth evaluation of the world literature on low-level Pb exposure was beyond the current charge to the Committee, we note that new studies reported in the spring and summer of 1983, with only a few exceptions, failed to find significant association between low-level Pb exposure and neuropsychologic deficits, once control variables were taken into account.

From its review of the recent research literature covered in this report, the Committee concludes that: (1) in the absence of control for other variables, a negative association between Pb exposure and neuropsychologic functioning has been established; (2) the extent of this negative association is reduced or eliminated when confounding factors are appropriately controlled; and (3) the Committee knows of no studies that, to date, have validly established (after proper control for confounding variables) a relationship between low-level Pb exposure and neuropsychologic deficits in children.

Research addressing possible dose-response relationships between lead and cognitive functioning in children is a worthy effort, and the Committee hopes that future studies can gather data that speak more adequately to this issue. In the view of the Committee, it is unlikely that continued use of cross-sectional epidemiological analyses will produce much credible evidence for or against the hypothesis that low to moderate levels of lead exposure are responsible for neurobehavioral deficits in apparently asymptomatic children. The study design generally does not allow for unambiguous disentangling of possible contributions of such lead exposures to observed cognitive or behavioral deficits versus the contributions of numerous other potentially confounding factors. There is a great need for longitudinal and time-series analyses, which include detailed prospective measurements of Pb exposure indices from early childhood onward and repeated sampling of neurobehavioral endpoints, both during preschool and school-age years.

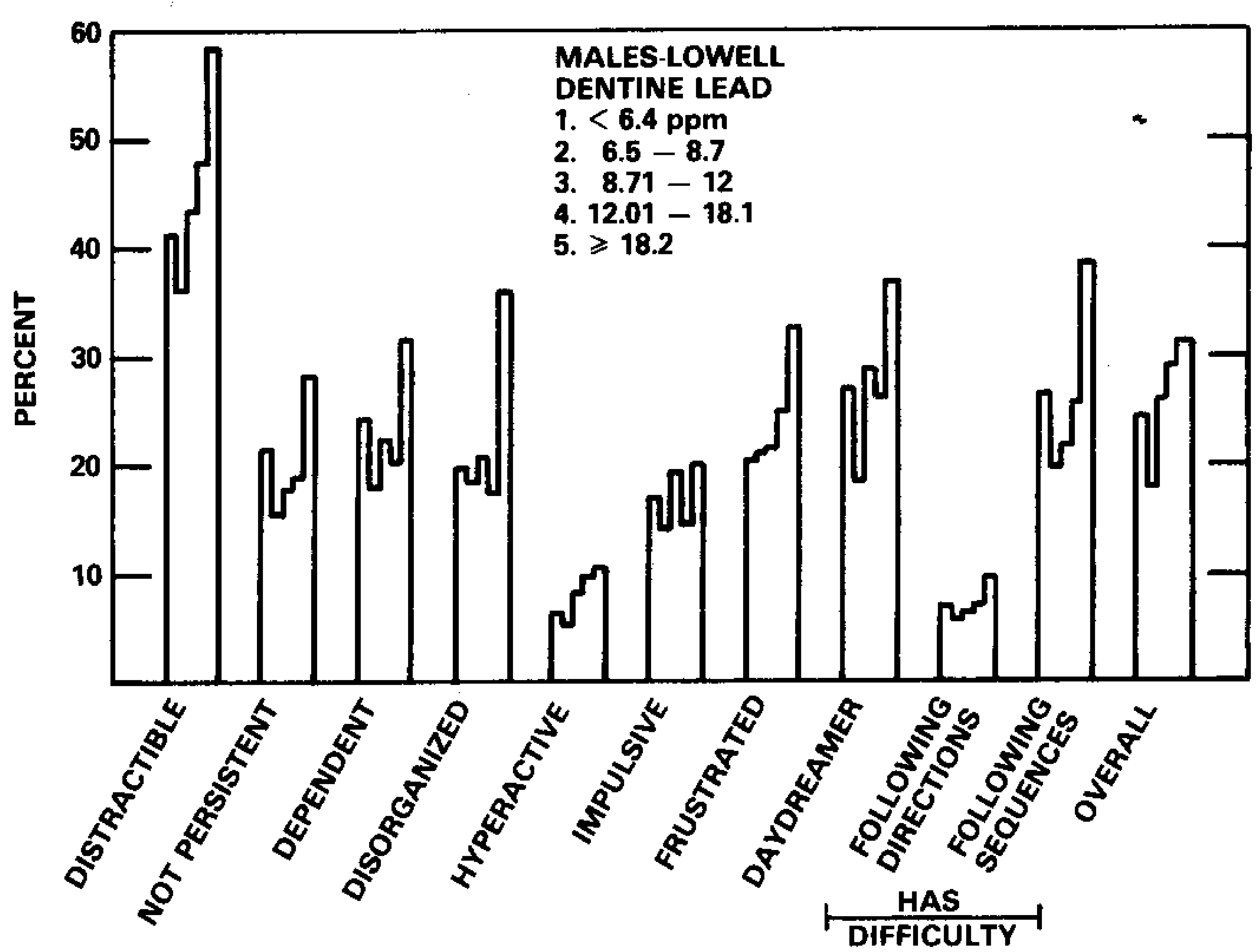

Figure 4. The relationship between negative teachers' ratings and dentine lead level in males. Each sample was classified into 5 groups according to dentine lead level. Each item was then scored. Within each item, Group 1, lowest lead level, is at the left; Group 5, highest lead level, is at the right.

Source: Needleman (1981).

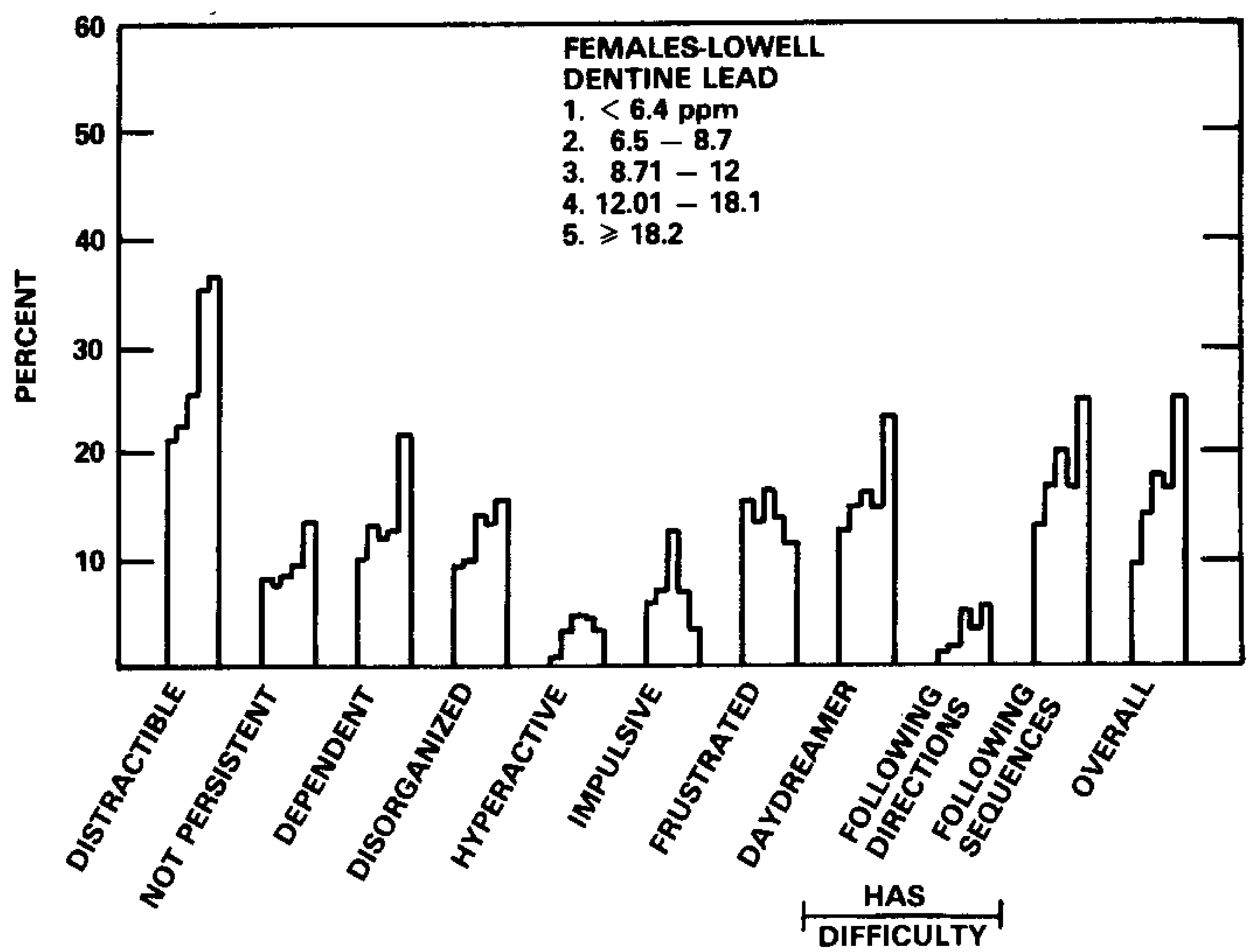

Figure 5. The relationship between negative teachers' ratings and dentine lead level in females. Each sample was classified into 5 groups according to dentine lead level. Each item was then scored. Within each item, Group 1, lowest lead level, is at the left; Group 5, highest lead level, is at the right.

Source: Needleman (1981).

REFERENCES

Ammons, R. B.; Ammons, C. H. (1962) The quick test (QT): provisional manual. Psychol. Rep. 11: 111-161 (monograph suppl. i-viii).

Bellinger, D. C.; Needleman, H. L. (1983) Lead and the relationship between maternal and child intelligence. J. Pediatr. (St. Louis) 102: 523-527.

Boone, J.; Hearn, T.; Lewis, S. (1979) A comparison of interlaboratory results for blood lead with results from a definitive method. Clin. Chem. (Winston Salem, N.C.) 25: 389-393.

Burchfiel, J. L.; Duffy, F. H.; Bartels, P. H.; Needleman, H. L. (1980) The combined discriminating power of quantitative electroencephalography and neuropsychologic measures in evaluating central nervous system effects of lead at low levels. In: Needleman, H. L., ed. Low level lead exposure: the clinical implications of current research. New York, NY: Raven Press; pp. 75-90.

Carter, G. F. (1978) The paper punched disc technique for lead in blood samples with abnormal haemoglobin values. Br. J. Ind. Med. 35: 235-240.

Ernhart, C. B. (February 3, 1983) [Letter to D. Weil]. Account of conversation with I. M. Shapiro. Available for inspection at: U.S. Environmental Protection Agency, Environmental Criteria and Assessment Office, Research Triangle Park, NC.

Ernhart, C. B. (March 11, 1983) Summary of conversation between N. B. Schell and C. Ernhart. Submitted to Committee at March 17-18, 1983 meeting. Available for inspection at: U.S. Environmental Protection Agency, Environmental Criteria and Assessment Office, Research Triangle Park, NC.

Ernhart, C. B.; Landa, B.; Callahan, R. (1980) The McCarthy scales: predictive validity and stability of scores for urban black children. Educ. Psychol. Meas. 40: 1183-1188.

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.

Harvey, P.; Hamlin, M.; Kumar, R. (1983) The Birmingham blood lead study. Presented at: annual conference of the British Psychological Society, symposium on lead and health: some psychological data; April; University of York, United Kingdom. Available for inspection at: U.S. Environmental Protection Agency, Environmental Criteria and Assessment Office, Research Triangle Park, NC.

Hessel, D. W. (1968) A simple and rapid quantitative determination of lead in blood. At. Absorpt. Newsl. 7: 55-56.

Kleinbaum, D. G.; Kupper, L. L.; Morgenstern, H. (1982) Epidemiological research: principles and quantitative methods. Belmont, CA: Lifetime Learning Publishers.

Lansdown, R.; Yule, W.; Urbanowicz, M-A.; Millar, I. B. (1983) Blood lead, intelligence, attainment and behavior in school children: overview of a pilot study. In: Rutter, M.; Russell Jones, R., eds. Lead versus health. New York, NY: John Wiley & Sons, Ltd.; pp.267-296.

McCarthy, D. (1972) McCarthy scales of children's abilities. New York, NY: Psychological Corporation.

Needleman, H. L. (1977) Lead in the child's world, a model for action. In: Hemphill, D. D., ed. Trace substances in environmental health-XI: [proceedings of University of Missouri's 11th annual conference on trace substances in environmental health]; June; Columbia, MO. Columbia, MO: University of Missouri-Columbia; pp. 229-235.

Needleman, H. L. (1981) Studies in children exposed to low levels of lead. Boston, MA Children's Hospital Medical Center; U.S. Environmental Protection Agency report no EPA-600/1-81-066. Available from: NTIS, Springfield, VA; PB82-108432.

Needleman, H. L. (November 22, 1982) [Letter to L. Grant]. Available for inspection at: U.S Environmental Protection Agency, Environmental Criteria and Assessment Office, Research Triangle Park, NC.

Needleman, H. L. (1982) The neurobehavioral consequences of low lead exposure in childhood Neurobehav. Toxicol. Teratol. 4: 729-732.

Needleman, H. L. (1983) Low level lead exposure and neuropsychological performance. In: Rutter, M.; Russell Jones, R., eds. Lead versus health. New York, NY: John Wiley & Sons, Ltd.; pp. 229-248.

Needleman, H. L.; Gunnoe, C.; Leviton, A.; Peresie, H. (1978) Neuropsychological dysfunction in children with "silent" lead exposure. Pediatr. Res. 12:374.

Needleman, H. L.; Gunnoe, C.; Leviton, A.; Reed, R.; Peresie, H.; Maher, C.; Barrett, P. (1979) Deficits in psychological and classroom performance of children with elevated dentine lead levels. N. Engl. J. Med. 300: 689-695.

Needleman, H. L.; Leviton, A.; Bellinger, D. (1982) Lead-associated intellectual deficit [Letter]. N. Engl. J. Med. 306: 367.

Perino, J. (1973) The relation of subclinical lead levels to cognitive and sensorimotor impairment in preschool black children. Hempstead, NY: Hofstra University. Ph.D. Dissertation.

Perino, J.; Ernhart, C. B. (1974) The relation of subclinical lead level to cognitive and sensorimotor impairment in black preschoolers. J. Learn. Dis. 7: 616-620.

Sattler, J. M. (1982) Assessment of children's intelligence and special abilities. 2nd ed. Boston, MA: Allyn and Bacon.

Scarr, S.; Yee, D. (1980) Heritability and educational policy: genetic and environmental effects on IQ, aptitude and achievement. Educ. Psychol. 15:1-22.

Shapiro, I. M.; Dobkin, B.; Tuncay, O. C.; Needleman, H. L. (1973) Lead levels in dentine and circumpulpal dentine of deciduous teeth of normal and lead poisoned children. Clin Chim. Acta 46: 119-123.

Shapiro, I. M.; Burke, A.; Mitchell, G.; Bloch, P. (1978) X-ray fluorescence analysis of lead in teeth of urban children <u>in situ</u>: correlation between the tooth lead level and concentration of blood lead and free erythroporphyrins. Environ. Res. 17: 46-52.

Smith, M.; Delves, T.; Lansdown, R.; Clayton, B.; Graham, P. (1983) The effects of lead exposure on urban children: the Institute of Child Health/Southampton study. London, United Kingdom: Department of the Environment.

Steenhout, A.; Pourtois, M. (1981) Lead accumulation in teeth as a function of age with different exposures. Br. J. Ind. Med. 38: 297-303.

U.S. Environmental Protection Agency, Health Effects Research Laboratory (1977) Air quality criteria for lead. Research Triangle Park, NC: U.S. Environmental Protection Agency, Criteria and Special Studies Office; EPA report no. EPA-600/8-77-017. Available from: NTIS, Springfield, VA; PB-280411.

U.S. Environmental Protection Agency. (August, 1983) Air quality criteria for lead. First External Review Draft. Research Triangle Park, NC: U.S. Environmental Protection Agency, Environmental Criteria and Assessment Office.

Winneke, G. (1983) Neurobehavioral and neuropsychological effects of lead. In: Rutter, M.; Russell Jones, R., eds. Lead versus health. New York, NY: John Wiley & Sons, Ltd.; pp. 249-265.

Winneke, G.; Hrdina, K-G.; Brockhaus, A. (1982) Neuropsychological studies in children with elevated tooth-lead concentrations. Part I: Pilot study. Int. Arch. Occup. Environ. Health 51: 169-183.

Winneke, G.; Kramer, U.; Brockhaus, A.; Ewers, U.; Kujanek, G.; Lechner, H.; Janke, W. (1983) Neuropsychological studies in children with elevated tooth lead concentrations. Part II: Extended study. Int. Arch. Occup. Environ. Health 51: 231-252.

Yamins, J. (1976) The relationship of subclinical lead intoxication to cognitive and language functioning and preschool children. Hempstead, NY: Hofstra University. Dissertation.

Yule, W.; Lansdown, R.; Millar, I. B.; Urbanowicz, M-A. (1981) The relationship between blood lead concentrations, intelligence and attainment in a school population: a pilot study. Dev. Med. Child Neurol. 23: 567-576.

ATTACHMENT 1

Additional materials considered in review of studies by Dr. Claire Ernhart and colleagues

1. Grant, L. D. (March 7, 1983) [Letter to C. Ernhart]. Available for inspection at: U.S. Environmental Protection Agency, Environmental Criteria and Assessment Office, Research Triangle Park, NC.

2. Ernhart, C. B. Coded raw data entries of psychometric test scores and background variable values (sex, parental IQ, etc.) for subjects evaluated in Perino and Ernhart (1974) and Ernhart et al. (1981) studies. Submitted to Committee at March 17-18, 1983 meeting. Available for inspection at: U.S. Environmental Protection Agency, Environmental Criteria and Assessment Office, Research Triangle Park, NC.

3. Ernhart, C. B.; Landa, B. (1980) "Cumulative deficit," a longitudinal analysis of scores on McCarthy scales. Psychol. Rep. 47: 283-286.

4. Ernhart, C. B.; Landa, B.; Schell, N. B. (1981) Lead levels and intelligence [Letter]. Pediatrics 68: 903-904.

5. Spector, S.; Brown, K. E. (1982) Lead study questioned [Letter]. Pediatrics 69: 134-135.

6. Ernhart, C. B.; Landa, B.; Schell, N. B. (1982) Lead study questioned: in reply [Letter]. Pediatrics 69: 135.

7. Ernhart, C. B. (1982) Lead results: no justification [Letter]. Sci. News (Washington, D.C.) 122: 3.

8. Ernhart, C. B. Scatter plots and associated coded data listings of residualized cognitive test performance scores vs. blood lead levels for subjects used in Perino and Ernhart (1979) study. Submitted to Committee at March 17-18, 1983 meeting. Available for inspection at: U.S. Environmental Protection Agency, Environmental Criteria and Assessment Office, Research Triangle Park, NC.

9. Ernhart, C. B. Scatter diagram of parents' IQ vs. child's IQ for low and moderate lead groups in Perino and Ernhart (1979) study (based on Perino dissertation). Submitted to Committee at March 17-18, 1983 meeting. Available for inspection at: U.S. Environmental Protection Agency, Environmental Criteria and Assessment Office, Research Triangle Park, NC.

10. Ernhart, C. B.; Landa, B.; Schell, N. B. (1983) Lead and intelligence - the effect of an outlier. Pediatrics (Submitted for publication).

11. Ernhart, C. B. Summary of conversation between N. B. Schell and C. Ernhart on March 11, 1983. Submitted to Committee at March 17-18, 1983 meeting. Available for inspection at: U.S. Environmental Protection Agency, Environmental and Criteria and Assessment Office, Research Triangle Park, NC.

12. Ernhart, C. B. (1983) Summary report: reanalysis of data of three studies: the effects of lead on children. Available for inspection at: U.S. Environmental Protection Agency, Environmental Criteria and Assessment Office, Research Triangle Park, NC.

13. Ernhart, C. B. (October 7, 1983) Response to [preliminary draft of] Appendix 12-C. Available for inspection at: U.S. Environmental Protection Agency, Environmental Criteria and Assessment Office, Research Triangle Park, NC.

14. Ernhart, C. B. (October 25, 1983) [Letter to L. Grant]. Available for inspection at: U.S. Environmental Protection Agency, Environmental Criteria and Assessment Office, Research Triangle Park, NC.

Additional materials considered in review of studies by Dr. Herbert Needleman and colleagues

1. Grant, L. D. (October 25, 1982) [Letter to H. Needleman]. Available for inspection at: U.S. Environmental Protection Agency, Environmental Criteria and Assessment Office, Research Triangle Park, NC.

2. Needleman, H. L. (November 22, 1982) [Reply to L. D. Grant]. Available for inspection at: U.S. Environmental Protection Agency, Environmental and Criteria and Assessment Office, Research Triangle Park, NC.

3. Needleman, H. L. (November 22, 1982) [List of 250 excluded subjects for Chelsea run]. Available for inspection at: U.S. Environmental Protection Agency, Environmental Criteria and Assessment Office, Research Triangle Park, NC. Attachment to Item 2 above.

4. Needleman, H. L. (November 22, 1982) [Frequency distribution of psychometric scores for high-lead subjects.] Available for inspection at: U.S. Environmental Protection Agency, Environmental Criteria and Assessment Office, Research Triangle Park, NC. Attachment to Item 2 above.

5. Needleman, H. L. (November 22, 1982) [Frequency distribution of psychometric scores for low-lead subjects.] Available for inspection at: U.S. Environmental Protection Agency, Environmental Criteria and Assessment Office, Research Triangle Park, NC. Attachment to Item 2 above.

6. Needleman, H. L. (1977) Lead in the child's world, a model for action. In: Hemphill, D. D., ed. Trace substances in environmental health-XI: [proceedings of University of Missouri's 11th annual conference on trace substances in environmental health]; June; Columbia, MO. Columbia, MO: University of Missouri-Columbia; pp. 229-235.

7. Needleman, H. L.; Leviton, A. (1979) Lead and neurobehavioral deficit in children [Letter]. Lancet 2(8133): 104.

8. Needleman, H. L. (1980) Lead and neuropsychological deficit: finding a threshold. In: Needleman, H. L., ed. Low level lead exposure: the clinical implications of current research. New York, NY: Raven Press; pp. 43-51.

9. Needleman, H. L.; Landrigan, P. J. (1981) The health effects of low level exposure to lead. Ann. Rev. Public Health 2: 277-298.

10. Needleman, H. L.; Bellinger, D.; Leviton, A. (1981) Does lead at low dose affect intelligence in children? Pediatrics 68: 894-896.

11. Needleman, H. L. (1982) The lead debate: a response. Environ. Sci. Technol. 16: 208A.

12. Needleman, H. L. (1983) The prevention of mental retardation and learning disabilities due to lead exposure. In: Jahiel, R. I., ed. The handbook of prevention of mental retardation and learning disabilities. (In preparation)

13. Grant, L. D. (March 14, 1983) [Letter to H. Needleman]. Available for inspection at: U.S. Environmental Protection Agency, Environmental Criteria and Assessment Office, Research Triangle Park, NC.

14. Jones, L. V. (March 23, 1983) [Letter to L. Kupper]. Available for inspection at: U.S. Environmental Protection Agency, Environmental Criteria and Assessment Office, Research Triangle Park, NC.

15. Grant, L. D. (April 8, 1983) [Letter to H. Needleman]. Available for inspection at: U.S. Environmental Protection Agency, Environmental Criteria and Assessment Office, Research Triangle Park, NC.

16. Needleman, H. L. (April 13, 1983) [Reply to L. Grant]. Available for inspection at: U.S. Environmental Protection Agency, Environmental Criteria and Assessment Office, Research Triangle Park, NC.

17. Needleman, H. L. (May 26, 1983) [Lists of all dentine lead levels for included children and for the 30 excluded from neuropsychological testing, and the unadjusted means, SD's and t tests]. Available for inspection at: U.S. Environmental Protection Agency, Environmental Criteria and Assessment Office, Research Triangle Park, NC.

18. Needleman, H. L.; Rabinowitz, M.; Leviton, A. (1983) The risk of minor congenital anomalies in relation to umbilical cord blood lead levels. Pediatr. Res. 17: 300A.

19. Needleman, H. L.; Bellinger, D.; Leviton, A.; Rabinowitz, M.; Nichols, M. (1983) Umbilical cord blood lead levels and neuropsychological performance at 12 months of age. Pediatr. Res. 17: 179A.

20. Needleman, H. L. (June 14, 1983) [Letter to L. Grant]. Final complete dentine lead levels on those subjects excluded because of discordant values. Available for inspection at: U.S. Environmental Protection Agency, Environmental Criteria and Assessment Office, Research Triangle Park, NC.

21. Needelman, H. L. (1983) Listings of coded raw data for dentine lead levels of both "included" and "excluded" subjects who underwent psychometric testing in Needleman et al. (1979) study. Attachment to Item 20 above.

22. Needleman, H. L. (1983) Listings of coded raw data for psychometric test results and background variables (e.g., sex, age, father's education, parental I.Q., etc.) for "included" and "excluded" subjects in Needleman et al. (1979) study. Attachment to Item 20 above.

23. Needleman, H. L. (1983) Computer printouts of frequency distribution of IQ scores and other psychometric test results for high and low lead subjects in Needleman et al. (1979) study. Inspected by Committee at H. L. Needleman's facilities at University of Pittsburgh (Children's Hospital), Pittsburgh, PA.

24. Needleman, H. L. (1983) Computer printouts of results of SPSS version of analysis of covariance for psychometric test scores of high and low lead subjects of Needleman et al. (1979) study, taking into account up to five covariates (including age as a covariate in some runs). Inspected by Committee at H. L. Needleman's facilities at University of Pittsburgh (Children's Hospital), Pittsburgh, PA.

25. Needleman, H. L. (October 4, 1983) [Letter to L. Grant]. Available for inspection at: U.S. Environmental Protection Agency, Environmental Criteria and Assessment Office, Research Triangle Park, NC.

26. Needleman, H. L. (October 4, 1983) [Letter to B. Goldstein]. Available for inspection at: U.S. Environmental Protection Agency, Environmental Criteria and Assessment Office, Research Triangle Park, NC.

27. Grant, L. D. (October 7, 1983) [Reply to H. L. Needleman]. Available for inspection at: U.S. Environmental Protection Agency, Environmental Criteria and Assessment Office, Research Triangle Park, NC.

28. Needleman, H. L. (October 7, 1983) [Letter to L. Grant]. Available for inspection at: U.S. Environmental Protection Agency, Environmental Criteria and Assessment Office, Research Triangle Park, NC.

★U.S. GOVERNMENT PRINTING OFFICE: 1983—659-017/7240