

Principles and Practices of Neurodevelopmental Assessment in Children: Lessons Learned from the Centers for Children's Environmental Health and Disease Prevention Research

Kim N. Dietrich,¹ Brenda Eskenazi,² Susan Schantz,³ Kimberly Yolton,⁴ Virginia A. Rauh,⁵ Caroline B. Johnson,⁶ Abbey Alkon,⁷ Richard L. Canfield,⁸ Isaac N. Pessah,⁹ and Robert F. Berman¹⁰

¹University of Cincinnati College of Medicine, Division of Epidemiology and Biostatistics, Department of Environmental Health, and the Cincinnati Children's Environmental Health Center, Cincinnati, Ohio, USA; ²Center for Children's Environmental Health Research, School of Public Health, University of California at Berkeley, California, USA; ³The Friends Children's Environmental Health Center, University of Illinois, Urbana-Champaign, Illinois, USA; ⁴Children's Environmental Health Center, Cincinnati Children's Hospital Medical Center and University of Cincinnati College of Medicine, Cincinnati, Ohio, USA; ⁵Columbia Center for Children's Environmental Health, Columbia University, New York, New York, USA; ⁶Pediatric Neuropsychology Group, Berkeley, California, USA; ⁷Department of Family Health Care Nursing, School of Nursing, University of California, San Francisco, California, USA; ⁸Division of Nutritional Sciences, College of Human Ecology, Cornell University, Ithaca, New York, USA; ⁹Department of Molecular Biosciences, and ¹⁰Department of Neurological Surgery, University of California at Davis, Davis, California, USA

Principles and practices of pediatric neurotoxicology are reviewed here with the purpose of guiding the design and execution of the planned National Children's Study. The developing human central nervous system is the target organ most vulnerable to environmental chemicals. An investigation of the effects of environmental exposures on child development is a complex endeavor that requires consideration of numerous critical factors pertinent to a study's concept, design, and execution. These include the timing of neurodevelopmental assessment, matters of biologic plausibility, site, child and population factors, data quality assurance and control, the selection of appropriate domains and measures of neurobehavior, and data safety and monitoring. Here we summarize instruments for the assessment of the neonate, infant, and child that are being employed in the Centers for Children's Environmental Health and Disease Prevention Research, sponsored by the National Institute of Environmental Health Sciences and the U.S. Environmental Protection Agency, discuss neural and neurobiologic measures of development, and consider the promises of gene-environment studies. The vulnerability of the human central nervous system to environmental chemicals has been well established, but the contribution these exposures may make to problems such as attention deficit disorder, conduct problems, pervasive developmental disorder, or autism spectrum disorder remain uncertain. Large-scale studies such as the National Children's Study may provide some important clues. The human neurodevelopmental phenotype will be most clearly represented in models that include environmental chemical exposures, the social milieu, and complex human genetic characteristics that we are just beginning to understand. *Key words:* behavior, development, National Children's Study, neurotoxicology, study design. *Environ Health Perspect* 113:1437–1446 (2005). doi:10.1289/ehp.7672 available via <http://dx.doi.org/> [Online 24 June 2005]

The purpose of this article is to present lessons learned in neurodevelopmental assessment of children from the Centers for Children's Environmental Health and Disease Prevention Research (Children's Centers) and from previous research. These comments and recommendations are products of the authors' experiences within their centers and their past works in the area of environmental neurotoxicology and teratology. We present both missteps and achievements in the hope that our collective experience can help guide in the planning and implementation of the National Children's Study.

The Central Nervous System as a Critical Organ

A fundamental lesson learned from studies of the effects of environmental exposures to the fetus and child is that the developing brain is one of the organs in the human body most sensitive to damage. Functional manifestations ranging from frank mental retardation to milder learning disabilities are the most

common class of birth defects (Lipkin 1991), although in most cases the specific etiology is unknown (Kallen 1988). Researchers have speculated that some cases of uncertain etiology may be due to environmental chemical exposures (Rees et al. 1990). The maturation of the central nervous system requires a more complex sequence of processes than any other structure, making this organ uniquely vulnerable to environmental influences (Rodier 1994, 2004).

National and international agencies have focused on the risks of early exposure to several major environmental contaminants such as lead, methylmercury, and polychlorinated biphenyls (PCBs) on neurodevelopment [Agency for Toxic Substances and Disease Registry (ATSDR) 2000; International Programme on Chemical Safety (IPCS) 1995; National Research Council 2000]. Although there is a paucity of data on the risks of lower-level pesticide exposures on human neurodevelopment, there are substantial animal

model data and limited human developmental data that these poisons may affect intrauterine growth and have functional teratogenic properties as well (Guillette et al. 1998; Weiss 1997; Weiss et al. 2004).

Initial Considerations

Timing of neurodevelopmental assessment. The vast and rapid growth of the child's neurobehavioral repertoire from birth through adolescence means that functional expressions of earlier-damaged systems may not be present or accessible at any given moment in time. For example, finding no effects on infant behavior cannot be regarded as conclusive evidence that the toxic agent has had no impact. Negative findings in the preschool period are also inconclusive. A toxicant may damage higher cortical centers that are associated with neurocognitive processes that are not yet functional or only marginally functional in a preschooler. In older children, a wider and more differentiated range of abilities can be examined, scores on psychometric measures are more precise and reliable, and early academic performance and social functioning outside of the home environment can

This article is part of the mini-monograph "Lessons Learned from the National Institute of Environmental Health Sciences/U.S. Environmental Protection Agency Centers for Children's Environmental Health and Disease Prevention Research for the National Children's Study."

Address correspondence to K.N. Dietrich, University of Cincinnati College of Medicine, Department of Environmental Health, Division of Epidemiology and Biostatistics, 3223 Eden Ave., Cincinnati, Ohio 45267-0056. Telephone: (513) 558-0531. Fax: (513) 558-4838. E-mail: Kim.Dietrich@uc.edu

This work was supported in part by the National Center for Environmental Assessment (ES011261, ES009605, ES011263, ES009600, ES011269, ES009584), the U.S. Environmental Protection Agency (R829389, R826709, R829390, R827027, R829388, R827039), and the National Children's Study.

The authors declare they have no competing financial interests.

Received 12 October 2004; accepted 14 March 2005.

be evaluated. This provides the rationale for long-term prospective longitudinal studies of cohorts recruited prenatally or at birth as the optimal design to ascertain neurobehavioral deficits in relation to exposure to environmental chemicals (Krasnegor et al. 1994). The literature on lead provides a model example of early exposure associated with higher-order neuropsychological and behavioral dysfunctions in older children and adolescents (e.g., Bellinger et al. 1992; Burns et al. 1999; Needleman et al. 1990; Tong et al. 1996). Long-term neurodevelopmental consequences have also been observed in a cohort study of children exposed *in utero* to PCBs (Jacobson and Jacobson 1996, 2003). Evaluations at earlier ages can and have revealed changes in cognitive function, but these may be less robust and definitive (Dietrich and Bellinger 1994).

Biologic plausibility. A number of factors determine the transplacental passage of environmental contaminants (Wilson 1977). Toxicants with low molecular weight, lipid affinity, nonpolarity, and low protein binding properties cross the placenta with ease (Slikker and Miller 1994). Unfortunately, most compounds possess one or more of these properties, cross the placenta, and enter the fetal circulation (Beckman and Brent 1999).

The immature blood–brain barrier of the fetus and young infant is more permeable to xenobiotics, and the fetus lacks drug-metabolizing detoxification capacities that are present postnatally (Rozman and Klaassen 1996). The chemical properties of certain toxicants also determine access to brain tissues. Lead imitates calcium ions and therefore crosses the blood–brain barrier with relative ease (Kerper and Hinkle 1997). In the blood stream, methylmercury combines with cysteine, forming a compound that is structurally similar to the essential amino acid methionine. This methylmercury–cysteine compound is actively transported into the endothelial cells in the blood–brain barrier on the methionine carrier and ultimately into the brain on a glutathione carrier (Kerper et al. 1992). The chlorobiphenyls and several classes of pesticides also gain access to brain tissues by virtue of their lipid solubility.

Toxicants linked with growth retardation and maturational delays *in utero* should be considered prime candidates for functional developmental toxicity. Any compound that retards intrauterine somatic growth should be examined as a potential neurobehavioral teratogen (Wilson 1977). Low-level intrauterine exposure to lead, PCBs, and organophosphate pesticides has been associated with lower birth weight, gestational maturity, or reduced head circumference in some prospective studies (e.g., Berkowitz et al. 2004; Dietrich et al. 1987; Eskenazi et al. 2004; Fein et al. 1984; McMichael et al. 1986; Whyatt et al. 2004).

Because of their potential disruption of central nervous system morphoregulation, hormonally active agents should also be considered as candidate neurodevelopmental toxicants. Hydroxylated metabolites of PCBs and related compounds may bind to human transthyretin, the only thyroid-hormone-binding protein synthesized in the brain. By binding to transthyretin, some hydroxylated PCBs may alter brain free thyroxine (T_4) levels and interfere with central nervous system development and function (Cheek et al. 1999; Seegal 2000). Recent evidence also suggests that PCBs up-regulate several thyroid-hormone-responsive genes that are expressed during periods of brain development (Gauger et al. 2004).

Environmental chemicals with excitatory effects on neurons (prolonged depolarization) are potential developmental neurotoxicants, as well. Examples include nicotine, the organophosphate and carbamate pesticides (Slotkin 1999), and some metals, including lead (Bressler et al. 1999).

Population factors. Population factors encompass a host of variables that affect a study's design, plan of execution, and sensitivity to the potentially adverse neurodevelopmental consequences of exposure to environmental toxicants. Several sometimes highly intercorrelated co-factors are involved, such as socioeconomic status, ethnicity, nutrition, access to medical care and educational resources, language spoken in the home, and cultural milieu. These variables will need to be considered when devising recruitment and long-term retention strategies, and especially when selecting the most appropriate methods of neurodevelopmental assessment. The field of neurobehavioral testing has been at the center of this dilemma as psychologists have struggled with the problem of how to design tests that do not measure only culturally specific information that is quite familiar to children from certain groups or social backgrounds but less familiar to those from different backgrounds. Investigators can use tests that have been adapted for non-English-speaking children and their families. In the United States, this typically involves Spanish translations of extant instruments (e.g., intelligence or speech, language, and memory tests), but the inventory of instruments in Spanish is still quite low. Furthermore, different dialects within the U.S. Spanish-speaking population exist, thus adding another complication. For example, the Children's Centers at the University of California at Berkeley, Columbia University, and Mount Sinai School of Medicine had Spanish-speaking populations from Puerto Rico, Mexico, and Dominican Republic with different regional dialects. In these cases, it is important to ask the caregiver what language is spoken in the home. All assessments of non-English-speaking children should

be done by examiners who are bilingual, preferably with the language of concern (e.g., Spanish) as their native tongue. Piloting of previously translated tests in the population of interest is always essential to determine their suitability. Some nonverbal tests have been considered to be culturally neutral, such as the Raven's Progressive Matrices (Raven 1998). However, even nonverbal tests may not be completely culturally neutral if the skills needed to complete the task are outside of the cultural experience of the population being evaluated.

One solution to this problem may lie in the use of so-called culture-neutral electrophysiologic, operant learning, or classical conditioning protocols. These types of tests also have the advantage that parallel animal tests that assess the same or similar functional domains are available. The drawback of using these "culture-fair" procedures is that the functional significance of any finding is sometimes difficult to calculate.

The degree of confounding can also vary greatly from one population to another and significantly affect a study's ability to detect associations between toxicant exposures and various parameters of neurodevelopment. In some populations, the degree of confounding may be so great that, after statistical adjustment, the exposure variable(s) no longer accounts for any further unique (independent) variance in the neurobehavioral data. Lead is a model example where exposure is typically correlated with other suboptimal environmental and sociohereditary factors (e.g., Dietrich et al. 1991). In such cases, any additional attributable developmental risk, beyond that accounted for by confounders, is obscured. This raises the specter of type II error (Needleman and Bellinger 1986).

The problem of confounding is not limited to correlations between a chemical exposure and nonchemical covariables such as socioeconomic status, quality of child rearing, and parental intelligence, among others. In some cases, populations can be exposed to a mixture of compounds that are also intercorrelated. This presents a conceptual and biostatistical challenge when attempting to estimate the independent or combined additive or synergistic effects of multiple chemical exposures (Jacobson 2001). For example, critical reviews of the Faroe Islands study of methylmercury and child development have suggested that the effects attributed to intrauterine exposure to methylmercury might be the consequence of co-exposures to high levels of PCBs [Myers and Davidson 2000; Toxicology Excellence for Risk Assessment (TERA) 1999]. Most other investigations of neurotoxicants in humans have not measured a multitude of exposures but have examined the behavioral toxicity of single toxicants. With the advent of more sensitive biomarkers of environmental exposures,

researchers such as those involved in the Children's Centers and in the National Children's Study will be facing a new challenge of developing both statistical and toxicologic methods for addressing the impact of multiple exposures. The importance of this has been shown in a recent study in animals demonstrating that the effect of the organophosphate chlorpyrifos was exacerbated by the pharmaceutical agent terbutaline commonly used to prevent preterm delivery (Rhodes et al. 2004).

The problem of confounding in environmental neuroepidemiology has led some to speculate about the advantages of studying chemical or drug exposures in lower risk populations for the purpose of reducing confounding and thus strengthening associations between measures of dose and disease (e.g., Bellinger 1995). However, restricting neurobehavioral studies of environmental toxicants to lower risk populations is not without scientific and ethical drawbacks. U.S. citizens at the lowest levels of social strata tend to suffer the greatest toxic burden. Furthermore, the opportunity to assess the interactions of toxicant exposures with other risk factors may be missed. The effects of prenatal and early postnatal exposure to developmental neurotoxicants may be more severe among the disadvantaged, where nutritional deficiencies, lack of adequate prenatal care, and suboptimal psychosocial environmental factors are likely to be more common.

Site factors. Examiners are often compelled to conduct assessments under less than ideal conditions. Inadequate space, lighting, excess ambient noise, and other distractions can invalidate even the most carefully administered protocol. The developmental assessment of an infant or child should be regarded as a controlled experiment, just as in animal models involving behavioral measures.

For neonates and infants, it is especially critical that the clinic environment is conducive to a valid test. For the newborn, lighting, temperature, noise, and any other factors that affect the neonate's state can determine the validity of the test. For an infant or child of any age, a comfortable, quiet, and well-lighted testing environment is a minimal requirement. Furniture, including chairs and table, should be appropriate for the test and developmental stage of the child. Test materials should be out of the child's sight but still easily accessible to the examiner to avoid fumbling and other miscues. Infants and some preschool children will often need the security and support of a caregiver during the examination, so accommodations should be made to have another adult in the room, seated in an area where the subject will not be distracted.

In some studies the geographic dispersion of the population may make home or school place testing the only practical option. Indeed,

testing sometimes must take place in multiple sites to accommodate families and prevent loss to follow-up. For example, studies in large rural communities (e.g., the Berkeley study of children in the Salinas Valley community) find it necessary to use a recreational vehicle in addition to the clinic site to reach families without any means of transportation and to avoid distractions within the home.

Child factors. It would seem obvious that the infant or child should be in an appropriate state or physical condition for assessment. However, this factor is not always given the consideration it deserves. The child should be reasonably well when evaluated with no current infection that is likely to significantly affect performance such as a severe upper respiratory infection or acute otitis media. Any medications the child may be on should also be recorded. It is important to clarify this with the parent before the assessment begins. A child who is too ill to respond appropriately to the demands of the examination should be rescheduled.

Given that many neurodevelopmental testing procedures require normal sensory function, a vision and hearing screen should also precede the examination. In some studies, sensory functions may be core outcome variables. In this case, a more detailed assessment of vision (farsightedness, depth perception, color perception, acuity/contrast sensitivity) and hearing (tympanometry, pure-tone audiometry, central auditory processing) will be indicated. Assessment of audition can be conducted as early as 6 months using methods such as visual reinforcement play audiometry. The auditory brainstem response (ABR) can also be employed to noninvasively measure the electrophysiologic responses of the brainstem auditory pathways (Roizen 1996). The automated ABR is already a routine procedure for screening neonates for hearing problems in a large number of hospitals in the United States, although a program of universal screening has not been implemented [American Academy of Pediatrics (AAP) 1995]. Methods for assessing refractive errors and amblyopia in infants and young children with limited verbal abilities such as the Allen cards and Snellen chart can also be used. However, in the toddler and preschooler, observing the child's response to visually engaging stimuli of various sizes can be just as valuable as a screening tool (Hoon 1996).

These neurosensory data can be used as core outcomes (i.e., if the toxicant in question is expected to affect vision or hearing), as effect modifiers (i.e., a toxicant's effects on tasks involving visual or auditory processing may be more pronounced among youngsters with preexisting deficits in these areas), or as potential confounders. However, extreme care must be taken when using such data as confounding or control variables. If deficits in

neurosensory function are an expression of exposure to the toxicant(s) in question, the resulting overcontrol may lead to a false negative finding or type II error.

Some way of rating the child's behavior and affect during the test session is also desirable. The child's response to examiner and test situation, attitude toward self and test performance, work habits and problem-solving style, motor functioning, visual and auditory acuity, oral communication, and mood are some of the factors that could potentially affect performance and should be noted. These data can be entertained as co-factors in outcome analyses, but this may lead to overcontrol. Indeed, behavioral disturbances that are rated by the examiner as limiting the validity of the examination can also be considered as legitimate behavioral sequelae of exposure.

Another factor that can affect a child's performance is the presence of a parent. A parent provides physical and emotional support for an infant or toddler but can also be a source of disruptive influences on the child's responses. Although the presence of a familiar caregiver is necessary and desirable for assessment of infants < 2 years of age, the parent or other caretaker should be instructed not to interfere with the presentation of stimulus materials, examiner instructions, or the child's behaviors during the test. By the later preschool years, the parent should be encouraged to not be present during the examination. If a child absolutely needs the security of a familiar caregiver, the examiner should seat the adult companion outside of the child's field of vision.

Quality assurance and quality control. Analytical laboratories go to great lengths to assure the validity and reliability of their assessments of toxicants in environmental samples and biologic tissues with quality assurance and quality control protocols. Similar issues are germane to neurodevelopmental assessment. Results of individually administered behavioral assessments depend to a great degree on the interaction between the child and examiner. Therefore, it is imperative that the examiner's contribution be equivalent for all children so that interchild differences in performance can be attributed to child characteristics rather than to some combination of child and examiner characteristics.

Although not always easy to achieve, the following recommendations apply to any longitudinal prospective study of the effects of neurodevelopmental toxicants. Ideally a single examiner should be used at each site or center. The examiner should be well seasoned and familiar with the population of interest. It is not necessary for a doctoral-level psychologist to administer most tests that are likely to be used in these studies (Brandt and van Gorp 1999). Individuals with baccalaureate or master degrees in psychology or related fields with

additional training in the standardized administration and scoring of neuropsychological tests can examine children enrolled in a study, under professional supervision [American Academy of Clinical Neuropsychology (AACN) 1999]. The examiner should be blinded to the group membership or exposure status of the child. Children from high-, medium-, and low-level exposure groups (if known) should be tested in a randomly intermixed order. If more than one examiner is used, their comparability in training and technique should be explicitly checked and regularly monitored. In multicenter studies, examiner training should be standardized across sites and regular meetings, and conference calls should be arranged to discuss issues of administration and scoring as they arise.

Dietrich and Bellinger (1994) cite an example of the difficulties that can arise when the second and last guidelines are neglected. In a study of lead-exposed children, the average IQ scores assessed by two relatively inexperienced examiners differed by more than one standard deviation or 15 points (Gregory et al. 1976).

In long-term prospective studies, it is not always possible to rely on a single experienced examiner to test all children. Attrition of staff and requirements for backup psychometricians in the event of leave usually require that a research center employ the services of two or more examiners. Interexaminer differences can be minimized by having the same developmental neuropsychologist train them and by videotaping practice administration sessions to provide feedback to the trainee and assess the presence of any differences in adherence to standardized administration or style that may result in interexaminer variability and measurement error. Videotaping can also provide the opportunity to assess interexaminer reliability in scoring the protocols (Chandlee et al. 2002). In studies spanning several years, monitoring of interexaminer reliability and proficiency should be practiced at regular intervals. This is particularly critical for multicenter studies employing identical neurodevelopmental assessment protocols. Whether one or several examiners are active at a given site, it is imperative that their reliability and proficiency be monitored over the entire course of the study to prevent any drift away from standardized procedures for administration and scoring.

All studies should carefully document quality assurance and control procedures and report intertester and interscorer reliability coefficients where appropriate. As Bellinger (2002) has noted, "the expectations for reporting the reliability of these [neurobehavioral] measurements should be no different from the expectations for reporting exposure (i.e., biomarker) measurements such as hair-mercury or blood-lead levels."

Nevertheless, although the Children's Centers instituted quality control protocols, particularly for the neurodevelopmental assessments, including direct observations or videotaping, insufficient time and resources made it difficult to fully meet the goals of these quality control programs. In our experience, the effort and cost associated with this are frequently underestimated. The National Children's Study needs to make every effort not to fall short in this critical area.

Sensitivity and specificity of neurodevelopmental measures. The terms "sensitivity" and "specificity" have two different meanings in the context of this article. In the evaluation of diagnostic tests, the sensitivity of a measure is defined as the proportion with the abnormality that the test classifies as abnormal (i.e., the proportion of true positives), whereas the "specificity" is the proportion of normal that the test classifies as normal (i.e., the proportion of true negatives). In the selection of neurodevelopmental measures for environmental studies, it is clearly advantageous to include tests that have the best possible prognostic value. This is particularly critical if neurobehavioral end points are incorporated into risk assessments conducted by regulatory bodies (Bellinger 2002). However, in this context we are also concerned with the selection of instruments that tap into neurodevelopmental domains that have been shown to be sensitive to (affected by) exposure to particular environmental toxicants. Of course, this assumes that some information already exists as to the effects of a given compound on the developing nervous system. Sometimes there are no prior human studies available to help dictate the domains that should be examined, and animal studies may not include domains that are unique to human cognition (e.g., language). In cases where there are multiple potential exposures with different mechanisms or loci of effect or when the exact exposures are not known or hypothesized, it may be necessary to use a broad spectrum of assessment tools. However, to the extent such information is available, selection of tests can be based on hypothesis-based inference. Under these circumstances, the focus shifts from "selecting tests" to the nomination of neurobehavioral domains to tap and subsequently to the employment of the best available measures (Bernstein 1994).

A wide range of neurobehavioral assessments has proven to be sensitive to lower level prenatal and early postnatal exposures to environmental contaminants, for example, lead, methylmercury, PCBs, environmental tobacco smoke, and other chemicals. However, the identification of behavioral phenotypes for specific agents has been an elusive goal. Some attempts have been made to differentiate the effects of alcohol exposure *in utero* from

deficits in neurobehavior associated with PCBs and other toxicants [e.g., Jacobson 1998; National Institute on Drug Abuse (NIDA) 2000]. However, environmental studies have failed to clearly identify a "behavioral signature" for particular compounds. The best evidence for specific effects probably comes from the literature on PCBs, where deficits in cognitive processing involving visual discrimination and memory appear to be a somewhat consistent finding in exposed infants and preschool children (Darvill et al. 2000; Jacobson et al. 1992). However, toxicants such as lead, methylmercury, and even the chlorobiphenyls seem to affect a wide range of neurobehavioral outcomes without illustrating a convincing degree of consistency across domains from one study to another. The sample's socioeconomic status; level, pattern, and timing of exposures; nutritional intake; general health; educational opportunities; and the particular instruments that were employed to examine neurodevelopment probably play an important role in between-study differences (Bellinger 1995; Schantz 1996). This could explain why the wide net provided by global and multiple-domain assessments of cognitive development such as IQ have proven to be so consistently sensitive across studies. Because these tests combine subscales that are representative of a broad number of underlying cognitive functions, they are likely to pick up exposure-related deficits across cohorts that differ in their functional expressions of toxicity. However, despite the sensitivity of these tests, their use in environmental studies has been sharply criticized by some reviewers. Critics of these apical measures note that they cannot readily yield information about the affected neural substrates or the complex executive and regulatory processes involved in learning, problem solving, and behavior (Bernstein 1994; Krasnegor et al. 1994; White et al. 1994). Furthermore, although global measures of intellectual functioning have proven to be sensitive measures of exposure to a number of environmental chemicals, this has not been the case for prenatal exposure to alcohol or drugs (NIDA 2000).

Data safety and monitoring. There is an ethical responsibility that referral protocols be in place to deal with the needs of children who perform poorly in the course of their participation in the study. Criteria need to be established for referral before data collection begins. Because of the sensitivity of certain tests, only performance on some may result in referral. Also, because of the predictive validity of these tests, a more rigorous criterion may be used at older ages (e.g., scores three standard deviations below the mean up to 24 months and 2 standard deviations afterward). Referrals normally take place through the primary care provider with parental permission. All referral

related contacts should be carefully documented and confidentiality closely guarded.

Neurodevelopmental Assessment in Environmental Studies: Initial Challenges

Constructing a battery of tests for an environmental study of developmental neurotoxicants can be a daunting task. First, the investigator is faced with choosing from a large number of potential end points that can be measured in a prospective study of any given agent or mixture of compounds to which the fetus or child may be exposed. Domains of interest in studies of this kind include overall neurologic status, sensorimotor skills, attention, memory, problem solving (executive functions, organization, and planning), visual-spatial and perceptual skills, speech and language abilities, behavioral problems, and adaptive skills as well as more global indices of intellectual attainment and academic achievement. Ideally, as stated above, the choice of neurobehavioral domains and the tests used to index them should be determined by what is known about the impact of a particular environmental agent on the development of specific cognitive, neuromotor, and behavioral features. However, as previously discussed, this is not always easily accomplished because of the lack of evidence for specific behavioral phenotypes. In this regard, animal studies have potential for identifying key functional domains that are affected by exposure and would be relevant to assess in human populations. For example, the Children's Center at the University of Illinois is using studies in animals exposed to PCBs and methylmercury during gestation to identify specific aspects of cognitive, sensory, and motor functions that are affected by combined exposures. These data will be used to guide the assessment of children in a companion cohort study.

Ultimately, the best strategy is to develop a battery that includes both broad-based measures of cognitive and neuromotor status as well as more fine-grained assessments of specific skills. Narrow-band instruments should target domains that, based on the extant human and animal literature, are believed to be affected by exposure to the toxicant(s) in question (e.g., for hypothesis-based battery development from the perspective of childhood lead poisoning, see Dietrich et al. 2004; Rogan et al. 2001).

A second consideration in battery development is the amount of time that can be allocated to the assessment. The energy and attention spans of younger subjects are rapidly exhausted, calling for shorter periods of time for testing, more frequent breaks, and sometimes multiple appointments separated by days or weeks. Infants < 1 year of age can usually tolerate only about 45 min of standardized

testing, whereas a 2-year-old can generally perform adequately over a period of roughly 75 min. For preschool and older children, breaks for a snack or just relaxing with the primary caregiver can fortify the child's endurance. Furthermore, neurodevelopmental evaluations are often part of a larger protocol that includes collection of sociodemographic information from parents, a general medical evaluation of the child, and collection of biologic specimens such as blood, hair, and urine. Retention of children and their families in a long-term study can be compromised if the demands on their time and effort are excessive.

Unique Considerations for Assessment of the Neonate, Infant, and Child

The neonate. The neonatal period extends from birth to roughly 1 month of age. Assessment of the neonate can be used to evaluate gestational maturity, neurologic status, and behavioral style and capacities. Neonatal assessment is a highly specialized area requiring continuous and careful attention to the infant's state as the examination proceeds. Recognized infant states include several sleep states, drowsiness or transitional states, alert states, and crying. Reflexes and muscle tone, signs of stress, and alertness and orientation can all be affected by the newborn's state.

Although neonatal neurodevelopmental assessments are not highly predictive of later functioning, performing a neurologic and neurobehavioral assessment after birth provides an assessment of the immediate effects of prenatal exposure before any influences of the extrauterine environment take place. A repeat exam (i.e., 2 or 4 weeks) allows assessment of neurobehavior after infants have stabilized. Inclusion of both assessment time points may provide insight into those effects of prenatal exposure that are transient and those that persist (Jacobson et al. 1984; Stewart et al. 2000). Neonatal assessment places special demands on quality control and site. Some instruments such as the Neonatal Behavioral Assessment Scale (NBAS) (Brazelton 1984; Brazelton and Nugent 1995) and the Neonatal Intensive Care Unit (NICU) Network Neurobehavioral Scale (Lester et al. 2004) require intensive training of examiners to master the techniques of administration and become reliable in scoring newborn responses. Training may demand travel to institutions with certification programs (for further details, see Brown Medical School Infant Development Center 2005). Each investigator will have to determine if these essential preconditions for neonatal neurobehavioral assessment can be met. The assessment of the neurologic and neurobehavioral status of the newborn will not usually be a practical option in studies involving populations that are widely dispersed geographically

or living in communities lacking the required clinical facilities or trained and certified personnel.

The infant and toddler. Infants and toddlers between 6 and 36 months of age may be the most challenging population for obtaining systematic and reliable behavioral data using paradigms that require the child to sustain attention, follow directions, and direct their cognitive effort toward a task that requires adherence to a somewhat rigid protocol. Although the cognitive capacities of older infants and toddlers are much advanced compared with their prelinguistic and prelocomotor days, directing that cognitive capacity at any particular time is often the most challenging part of the assessment. The clinical skills of the person examining the infant and toddler are of paramount importance for obtaining data that are valid and reliable.

Indices of neurodevelopment in infants and toddlers are less stable over time and, at least before 24 months, lack substantial predictive validity for later morbidity. This is partly because of the means by which infants are able to express their cognitive abilities (i.e., primarily through sensorimotor acts) and the lack of continuity in response modalities from infancy to older childhood and adolescence. However, neurobehavioral test scores in infancy retain strong concurrent validity. Bellinger (2002) suggests that scores on infant neurodevelopmental measures can be understood in a manner analogous to the neonatologist's interpretation of birth weight. Except in very low-birth-weight infants, weight at birth is not predictive of later weight, although birth weight is a very informative index of a newborn's general health status. Also, as with neonatal measures, the time between neurotoxicant exposure *in utero* and postnatally and the assessment of outcome is reduced. Thus, the influence of later intervening and potentially confounding factors is diminished along with increased strength and reduced bias in the estimate of the neurotoxicant's contributions to development (Bellinger 2002). Furthermore, where neurotoxicants such as lead, PCBs, and methylmercury have predicted poorer performance in infants and preschoolers, forward studies have demonstrated that functional deficits in neurobehavior persist into later childhood (Dietrich 2000).

Older preschoolers and school-age children. Because of the greater social maturity, autonomic stability, and endurance of the older preschool and school-age child, substantially more time can be devoted to a single assessment session. Also, owing to the tremendous growth in the range and clarity of a child's response capabilities, assessments of preschoolers and older children generate a more differentiated picture of a subject's developmental strengths and weaknesses. Functional impairments

heretofore unnoticed may become apparent. To investigate the developmental effects associated with exposure to neurotoxicants in the fullness of time, studies should maintain follow-up at least until children attain school age. Indeed, recent studies of the long-term sequelae of early lead exposure would suggest follow-up to adolescence and young adulthood is necessary to reveal the full range of exposure related morbidities. There are a number of reasons for extending follow-up into late childhood and adolescence. Deficits still apparent at later ages are generally thought to be of greater practical significance because the predictive validity or prognostic value of later preschool and school-age performance is considerably greater than that of performance in infancy, thus providing a sounder basis for drawing inferences about the long-term effects of prenatal or early postnatal exposures.

Neurodevelopmental Assessment at the Children's Centers

Neurodevelopmental assessment practices in the longitudinal birth cohort studies at the

Children's Centers were guided by many of the principles outlined above. Most centers conducted neurodevelopmental assessments at various ages after birth (see Table 1). For the first 5 years of funding, most centers conducted assessments at 12 months and 24 months of age.

A diverse group of standardized neurodevelopmental assessment tools were employed. There were no attempts to develop common protocols among the centers conducting neuroepidemiologic studies. Differences in tests and timing of assessment reflect variations in the toxicants under study, their hypothesized effects, and practical considerations. Four centers conducted neonatal assessments; three used the NBAS (Brazelton 1984; Brazelton and Nugent 1995), and one used the NICU Network Neurobehavioral Scale (Lester et al. 2004). The Bayley Scales of Infant Development-II (Bayley 1993), which provides standard scores for mental and psychomotor development, were used by all centers at 12 months of age and by most at 24 months. All centers adjusted the child's chronologic age for prematurity on the Bayley scales until 24 months.

A variety of other assessment tools were used to measure domains including developmental milestones, language, and behavior. A few centers employed more experimental protocols, including assessment of visual recognition memory, the autonomic nervous system, and measures of learning that parallel some of those used in animal studies. The centers at Illinois and the University of Cincinnati also assess the child's hearing.

The choice of neurodevelopmental tests was based first on the age of the child, followed by other considerations including tests employed in previous studies, the domains of behavior thought to be affected by the main toxicant(s) under study, the availability of tests in the language of the study population, time required to administer the examination, and the test's suitability for administration by non-doctoral-level examiners. Some centers, especially those with low-income or bilingual study populations, found it necessary to select assessment tools with a lower end of functioning range, considerably below the ages of the children being assessed.

As discussed above, it can be difficult to determine the language of assessment for

Table 1. Neurodevelopmental assessment in the Children's Centers for the first 5-year funding cycle.

Assessment tool	University of Illinois	University of California, Berkeley	Columbia University	Mount Sinai Medical Center	University of Cincinnati
Neonatal assessment tools					
Brazelton (Brazelton and Nugent 1995)	Birth	Birth		Birth	
Ballard (Ballard et al. 1991)				Birth	
NICU Network Neurobehavioral Scale (Lester et al. 2004)					Birth, 4 weeks
Infant and toddler assessment tools					
Bayley scales (Bayley 1993)	6, 12 months	6, 12, 24 months	12, 24, 36 months	12, 24 ^a months	12, 24 months
PLS (Zimmerman et al. 1992)		6, 12, 24 months			
CBCL (Achenbach and Rescorla 2004)		12, 24 months			
Denver (Frankenburg et al. 1992)			6 months		
Fagan (Fagan et al. 1986)	6 months		6 months		
Infant Behavior Questionnaire (Rothbart 1981)				12 months	
Toddler Behavior Assessment Questionnaire (Goldsmith 1996)				24 ^a months	
Sleep questionnaire (adapted from Morrell 1999)					6, 12, 18, 23 months
Cognitive tasks ^b (Aguilar and Baillargeon 1999)	6, 9, 12 months				
Child assessment tools					
WPPSI-III (Wechsler 2002)		42 ^c months	60 months		
NEPSY (Korkman et al. 1998)		42 ^c months			
McCarthy (McCarthy 1974)		42 ^c months			
CBCL (Achenbach and Rescorla 2004)		42 ^c months			
Other child assessments					
Auditory assessment	Birth, 12 months				
Autonomic Nervous System Assessment (Alkon et al. 2003)		6, 12, 42 ^c months			
HOME Scale (Caldwell and Bradley 1984)		6, 12, 24, 42 ^c months	36 months	12, 24 months	12 months
Maternal intelligence					
PPVT (Dunn and Dunn 1981)/TVIP (Dunn et al. 1986)		6 months		Pregnancy	
WASI (Wechsler 1999)					12 months
TONI-2 (Brown et al. 1990)			24 or 36 months		
WAIS-III (Wechsler 1997)	Enrollment (matrices and block design)	6 months (matrices)			

Abbreviations: CBCL, Child Behavior Checklist; HOME, Home Observation for Measurement of the Environment; NEPSY, A Developmental Neuropsychological Assessment; PLS, Preschool Language Scale; PPVT, Peabody Picture Vocabulary Test; TONI-2, Test of Nonverbal Intelligence-2; TVIP, Test de Vocabulario en Imagenes Peabody; WAIS-III, Wechsler Adult Intelligence Scale-III; WASI, Wechsler Abbreviated Scales of Intelligence; WPPSI-III, Wechsler Preschool and Primary Scales of Intelligence-III.

^aChildren were brought in up to 5 years of age for the "24-month" assessment; if too old for the Bayley, they were assessed with the Batelle Developmental Inventory (Newborg et al. 1984). ^bThis consists of object retrieval tasks that examine working memory, executive control, and cognitive development. ^cCurrently underway as part of second 5-year funding cycle.

children who are raised in the United States in non-English-speaking households. All centers attempted to maximize the child's performance on all tests. One center assessed bilingual children in English. Another center based the child's assessment on what language was usually spoken in the home. In both cases, if the examiner noted language difficulties during the exam, the tasks were readministered in the other language.

The availability of assessment tools increases with the child's age, enhancing the researcher's ability to focus on specific developmental domains that may be affected by the exposure of interest. Because many of the chemicals examined by these centers have not been widely studied in children (e.g., pesticides, polycyclic aromatic hydrocarbons), the choice of the domains likely to be affected may be based solely on animal studies. Finding the human analog to animal behavior is challenging. However, as previously noted, the Illinois center is addressing this problem by conducting concurrent animal studies with exposures similar to those in their human population.

All centers used some quantitative measures of the parent's cognitive functioning and quality of caregiving in the home to control for possible confounding. In centers dealing mainly with multicultural samples, tests of adult intellectual attainment with a minimum of verbal content were used.

Future Directions for Neurodevelopmental Assessments in the National Children's Study

New technologies for assessment of the function and structure of the central nervous system may hold promise for advancing our understanding of the impact of environmental chemicals on neurodevelopment. Some of these are briefly discussed below.

Computer-based experimental measures for children. Numerous attempts have been made to develop computer-administered batteries of tests for children exposed to environmental neurotoxicants. For example, adaptations of the Neurobehavioral Evaluation System (Baker et al. 1985) have been used in studies of children exposed to heavy metals in both the United States and abroad (Dahl et al. 1996). A computer-assisted comprehensive assessment for children enrolled in environmental studies has been developed by Anger and colleagues (Rohlman et al. 2003). The Behavioral Assessment and Research System has been designed for use in studies of children as young as 3 years of age and has been adapted for use with Hispanic populations (Rohlman et al. 2001). The entire battery consists of 11 computer-administered tests assessing sustained and selective attention,

working memory, motivation, new learning, response speed, executive functions, and fine-motor coordination.

The use of computer-assisted tests has several advantages in that examiner effects are reduced and data collection and scoring are automated and objective. However, care must be exercised when applying these methods to populations that have little or no exposure to computers or similar kinds of automated systems. Another potential drawback of computer-assisted tasks is that they can minimize interaction with a friendly, supportive examiner, which can be essential in helping keep the younger child motivated to complete a lengthy battery of tests. If computer-assisted tests are used, it may be best to intersperse them with examiner-administered tasks.

Psychobiologic measures. Studies of young children have included many psychosocial, environmental, and neurodevelopmental factors, but they rarely include individual difference measures of psychobiology. Recent studies show that children's physiologic responses to different stressors are valid measures of individual psychobiology or autonomic reactivity (Alkon et al. 2003; Matthews et al. 1990). These psychobiology measures have been shown to affect how children interact with their family, adults, and peers at home and in school. In addition, children's psychobiology is related to their physical and mental health (Boyce et al. 1998; Kagan et al. 1988; Porges et al. 1994). Autonomic reactivity is being studied as an outcome of organophosphate exposure prenatally and postnatally at the Berkeley Children's Center. Autonomic dysregulation or an imbalance in the parasympathetic and sympathetic branches of the autonomic nervous system may be a sensitive indicator of acute and chronic pesticide exposure (Eskenazi et al. 1999). The Columbia center is using the orthostatic tilt test (Fifer et al. 1999) on a small subsample in the cohort. This is a noninvasive experimental paradigm used to identify infants with individual differences in autonomic regulation and diminished physiologic responses (as measured by heart rate) to blood pressure changes after postural adjustment.

Neurobiologically based markers of development. Several environmental studies of children have used electrophysiologic techniques to assess the effects of neurotoxicants on central nervous system function. Visual as well as auditory evoked potentials have been examined and in many cases have been found to be sensitive to environmental chemical exposures. Another promising area for the future is the use of neuroradiologic techniques such as magnetic resonance imaging (MRI). MRI assessment of brain structure and function is beginning to be used in studies of developmental neurotoxicants. In the Cincinnati center, for example, adults exposed to high levels of lead in early childhood

are being examined using volumetric, functional, and spectroscopic MRI methods. These methods provide data on exposure-related structural changes, brain activation in response to standardized verbal and visual-spatial problems, and brain biochemical processes.

The use of these methods assumes the availability of a clinical facility that can provide the needed equipment and trained personnel. For many centers, particularly those far from large metropolitan areas, this will normally not be an option.

Gene-environment interactions in neurodevelopment. A promising area for future study is the interaction between certain genetic polymorphisms involved in neurotransmitter regulation and metabolism with environmental chemical exposures. For example, the Cincinnati center has examined the joint effects of a dopamine transporter (DAT) polymorphism and maternal prenatal smoking on childhood hyperactivity-impulsivity and inattentiveness (Kahn et al. 2003). Childhood hyperactivity-impulsivity and oppositional behaviors were associated with a DAT polymorphism, but only when the child also had *in utero* exposure to the products of maternal smoking. Cincinnati researchers have also identified a vitamin D receptor gene polymorphism that may be associated with greater absorption of environmental lead and thus an increased risk for neurodevelopmental toxicity (Haynes et al. 2003). The Cincinnati center is also investigating the relationship between early exposure to lead and adult criminality. The interactions of a history of early lead poisoning with four polymorphisms associated with dopamine and serotonin reception and/or transport and genes regulating monoamine oxidase A enzyme activity are being examined (Reif and Lesch 2003). Investigators at the University of California at Davis Children's Center are investigating the interaction between specific GABA_A receptor polymorphisms and autism spectrum disorder (ASD). It is well known that some of the most widely used and environmentally persistent pesticides block GABA_A receptor chloride channels within the mammalian central nervous system, thereby producing hyperexcitability, tremors, and convulsions, depending on the level of exposure (Narahishi et al. 1998). Three genes that encode subunits of ionotropic GABA_A receptors cluster on chromosome 15q11-13. Maternally inherited duplications of this chromosome account for approximately 1-3% of all cases of ASD (Veenstra-VanderWeele et al. 2003). The Children's Centers at Mount Sinai, Berkeley, and the University of Washington are examining the relationship of the paraoxynase 1 (*PONI*) gene in modifying the associations between organophosphate pesticide exposure and neurodevelopment. Recent evidence for a relationship between genetic background and susceptibility to the neurobehavioral toxicity of

neonatal exposure to thimerosal (ethylmercury) in mice (Hornig et al. 2004) further underscores the need to consider gene–environment interactions to identify susceptible populations.

Conclusions and Recommendations

- Long-term studies that follow participants into adolescence and early adulthood are essential to assess the full range of neurodevelopmental consequences of exposure to environmental chemicals. The vast and rapid growth of human neurobehavioral capacities from birth through early adulthood means that the functional effects of earlier damage may not be fully expressed at any given moment in time.
- Compounds targeted for investigation should be those that represent plausible neurodevelopmental hazards to fetal and postnatal central nervous system development. The compounds under investigation should represent a clear hazard to neurodevelopment on the basis of their biochemical properties. Most drugs and environmental chemicals cross the placenta with ease by virtue of their low molecular weight and many gain access to the developing central nervous system. This also requires some knowledge of the exposure levels that are likely to be encountered in the environments occupied by the samples of interest.
- Population factors including ethnicity and language must be considered in the planning of study procedures for recruitment, retention, and selection of neurodevelopmental measures. Population factors can be particularly important if substantial confounding and multicultural issues are involved. These features become crucial when it comes to the choice of appropriate neurodevelopmental measures that will accurately reflect the cognitive and sensorimotor capacities of the population under investigation. Ideally, examiners familiar with the language and culture of the population under study should be employed.
- The proposed work should be compatible with the physical and human resources at hand. Sites faced with populations that are widely dispersed geographically, having limited transportation, or lack of trained personnel will be restricted in the depth of neurodevelopmental assessments that can be realistically implemented.
- Neuropsychologic batteries employed in these studies should include a balance of those with measures both broad and narrow in scope. Batteries too narrow in scope can easily miss deficits associated with environmental chemical exposures. On the other hand, fine-grained measures of more narrowly defined neurobehavioral domains can shed light on exposure-specific effects and brain–behavior relationships. If specific

neurobehavioral domains are targeted for assessment, their selection should be guided by hypothesis-driven inference. An assessment battery that includes both broad- and narrow-band instruments will usually be optimal.

- Procedures for monitoring the quality of data collection and scoring should be maintained throughout the life of the study. Monitoring of interexaminer reliability and proficiency should be practiced on a regular basis. In the event that multiple examiners are required at a single center, their training and examination techniques should be explicitly checked, monitored, and quantitatively recorded. In multicenter studies, training should be standardized across sites. This may require annual meetings and more frequent conference calls to discuss issues related to administrations and scoring of tests.
- Data safety and monitoring procedures must be in place. As health care professionals as well as researchers, we have a duty to treat study participants with respect and concern. Referral of infants and children presenting with suspect neurologic or developmental signs or symptoms should be made through the primary care provider. In every case, consultation with the primary caregiver is mandatory.
- The unique needs of study participants involved in neurodevelopmental assessment from birth through adulthood must be considered. Every age presents special challenges. Assessment of the neonate calls for extensive training of the examiner(s) and a highly controlled environment. Infants and toddlers present special challenges with regard to directing their attention to the demands of cognitive and sensorimotor tasks. Assessment of school-age children and adolescents provides an opportunity to examine a more differentiated picture of participants' developmental strengths and weaknesses. Assessments of the older adolescent and young adult call for an approach that recognizes a participant's legal rights as a research subject and sometimes requires special measures to assure subject protection and confidentiality of any data that may be obtained.
- New approaches and technologies should be exploited in future studies. Computer-based assessment, psychobiologic measures, neural markers of central nervous system activity, and especially the promise of advances in human genetics have the potential to lead to substantial progress in our understanding of the effects of environmental chemicals on fetal and child development.

Final remarks. A battery of tests that can be universally regarded as valid and reliable for evaluating the potential impact of neurotoxins on the developing central nervous system does not exist. Specific approaches and tests have been reviewed and recommended based

on the authors' experiences and ongoing work in the Children's Centers and other studies. The quandary faced by both seasoned investigators and researchers new to the area when planning a study was summarized by one experienced observer: "I cannot recommend any tests in this endeavor; many are appropriate. Overall strategy, a principled theoretical framework, and adequately specified domains are what count, not tests" (Bernstein 1994).

The vulnerability of the developing human central nervous system to environmental chemical compounds has been well established. The contribution of these exposures *in utero* or postnatally to the development of disorders of uncertain etiology such as attentional deficit hyperactivity disorder or pervasive developmental disorder/autism is not known at present. Large-scale human studies such as the National Children's Study may provide some clues. In the final analysis, the human neurodevelopmental phenotype will be more fully revealed in large-scale studies that take account of environmental chemical influences, the social milieu, and complex human genetic characteristics that we are just beginning to understand (Hamer 2002).

REFERENCES

- AACN. 1999. American Academy of Clinical Neuropsychology policy on the use of non-doctoral-level personnel in conducting clinical neuropsychological evaluations [Editorial]. *Clin Psychol* 13:385.
- AAP. 1995. American Academy of Pediatrics Joint Committee on Infant Hearing 1994 position statement. *Pediatrics* 95:152–156.
- Achenbach TM, Rescorla LA. 2004. The Achenbach System of Empirically Based Assessment (ASEBA) for ages 1.5 to 18 years. In: *The Use of Psychological Testing for Treatment Planning and Outcomes Assessment* (Maruish MR, ed). Mahwah, NJ:Lawrence Erlbaum Associates, 179–213.
- Aguiar A, Baillargeon R. 1999. Perseveration and problem solving in infancy. *Adv Child Dev Behav* 17:135–180.
- Alkon A, Goldstein L, Snider N, Essex M, Kupter W, Boyce T, et al. 2003. Developmental and contextual influences on autonomic reactivity in young children. *Dev Psychobiol* 42:64–78.
- ATSDR. 2000. Toxicological Profile for Polychlorinated Biphenyls (PCBs). Atlanta, GA:Agency for Toxic Substances and Disease Registry.
- Baker EL, Letz RE, Fidler AT, Shalat S, Pantamura D, Lyndaom MA. 1985. A computer-based neurobehavioral evaluation system for occupational and environmental epidemiology: methodology and validation studies. *Neurobehav Toxicolol Teratol* 7:369–377.
- Ballard JL, Khoury JC, Wedig K, Wang L, Eilers-Walsman BL, Lipp R. 1991. New Ballard score expanded to include extremely premature infants. *J Pediatr* 119:417–423.
- Bayley N. 1993. Bayley Scales of Infant Development. 2nd ed. San Antonio, TX:Psychological Corporation.
- Beckman DA, Brent RL. 1999. Basic principles of developmental toxicology. In: *Medicine of the Fetus and Mother* (Reece EA, Hobbins JC, eds). 2nd ed. Philadelphia:Lippincott-Raven, 281–288.
- Bellinger DC. 1995. Interpreting the literature on lead and child development: the neglected role of the "experimental system." *Neurotoxicol Teratol* 17:201–212.
- Bellinger DC. 2002. Perspectives on incorporating human neurobehavioral end points in risk assessments. *Risk Anal* 22:487–498.
- Bellinger D, Stiles K, Needleman HL. 1992. Low-level lead exposure, intelligence, and academic achievement: a long-term follow-up study. *Pediatrics* 90:855–861.
- Berkowitz GS, Wetmur JG, Birman-Deych E, Obel J, Lapinski RH,

- Godbold JH, et al. 2004. *In utero* pesticide exposure, maternal paraoxonase activity, and head circumference. *Environ Health Perspect* 112:388–391.
- Bernstein JH. 1994. Assessment of developmental toxicity: neuropsychological batteries. *Environ Health Perspect* 102(suppl 2):141–144.
- Boyce W, Frank E, Jensen P, Kessler R, Nelson C, Steinberg L. 1998. Social context in developmental psychopathology: recommendations for future research from the MacArthur Network on Psychopathology and Development. *Dev Psychopathol* 10:143–164.
- Brandt J, van Gorp W. 1999. American Academy of Clinical Neuropsychology policy on the use of doctoral level personnel in conducting clinical neuropsychological evaluations [Editorial]. *J Clin Exp Neuropsychol* 21:1.
- Brazelton T. 1984. Neonatal Behavioral Assessment Scale. Philadelphia:JB Lippincott Co.
- Brazelton TB, Nugent JK. 1995. Neonatal Behavioral Assessment Scale. 3rd ed. London:McKeith Press.
- Bressler J, Kim KA, Chakraborti T, Goldstein G. 1999. Molecular mechanisms of lead neurotoxicity. *Neurochem Res* 24:595–600.
- Brown L, Sherbenou RJ, Johnsen SK. 1990. Test of Nonverbal Intelligence. 2nd ed. Austin, TX:PRO-ED.
- Brown Medical School Infant Development Center. 2005. Training and Education, NNNS Training Program. Available: <http://www.infantdevelopment.org> [accessed 21 March 2005].
- Burns JM, Baghurst PA, Sawyer MG, McMichael AJ, Tong SL. 1999. Lifetime low-level exposure to environmental lead and children's emotional and behavioral development at ages 11–13 years. The Port Pirie Cohort Study. *Am J Epidemiol* 149:740–749.
- Caldwell BM, Bradley RR. 1984. Home Observation for Measurement of the Environment. Little Rock, AK:University of Arkansas.
- Chandless J, Heathfield LT, Salganik M, Damokosh A, Radcliffe J. 2002. Are we consistent in administering and scoring the Bayley Scales of Infant Development-II? *J Psychoeduc Assess* 20:183–200.
- Cheek AO, Kow K, Chen J, McLachlan JA. 1999. Potential mechanisms of thyroid disruption in humans: interaction of organochlorine compounds with thyroid receptor, transthyretin, and thyroid-binding globulin. *Environ Health Perspect* 107:273–278.
- Dahl R, White RF, Weihe P, Sorensen N, Letz R, Hudnell K, et al. 1996. Feasibility and validity of three computer-assisted neurobehavioral tests in 7-year-old children. *Neurotoxicol Teratol* 18:413–420.
- Darvill T, Lonky E, Reihman J, Stewart P, Pagano J. 2000. Prenatal exposure to PCBs and infant performance on the Fagan Test of Infant Intelligence. *Neurotoxicology* 21:1029–1038.
- Dietrich KN. 2000. Environmental neurotoxicants and psychological development. In: *Pediatric Neuropsychology: Research, Theory and Practice* (Yeates KO, Ris MD, Taylor HG, eds). New York:Guilford Press, 206–234.
- Dietrich KN, Bellinger D. 1994. The assessment of neurobehavioral development in studies of the effects of prenatal exposure to toxicants. In: *Prenatal Exposure to Toxicants: Developmental Consequences* (Needleman HL, Bellinger D, eds). Baltimore, MD:Johns Hopkins University Press, 57–85.
- Dietrich KN, Krafft KM, Bornschein RL, Hammond PB, Berger O, Succop PA, et al. 1987. Low-level fetal lead exposure effect on neurobehavioral development in early infancy. *Pediatrics* 80:721–730.
- Dietrich KN, Succop PA, Berger O, Hammond PB, Bornschein RL. 1991. Lead exposure and the cognitive development of urban preschool children: the Cincinnati Lead Study cohort at age 4 years. *Neurotoxicol Teratol* 13:203–211.
- Dietrich, KN, Ware JH, Salganik M, Radcliffe J, Rogan WJ, Rhoads GG, et al. 2004. Effect of chelation therapy on the neuropsychological and behavioral development of lead-exposed children after school entry. *Pediatrics* 114:19–26.
- Dunn L, Dunn L. 1981. Peabody Picture Vocabulary Test. Rev ed. Circle Pines, MN:American Guidance Service.
- Dunn L, Padilla E, Lugo D, Dunn L. 1986. Test de Vocabulario en Imagenes Peabody. Circle Pines, MN:American Guidance Service.
- Eskenazi B, Bradman A, Castorina R. 1999. Exposures of children to organophosphate pesticides and their potential adverse health effects. *Environ Health Perspect* 107(suppl 3):409–419.
- Eskenazi B, Harley K, Bradman A, Wetzien E, Jewell NP, Barr DB, et al. 2004. Association of *in utero* organophosphate pesticide exposure and fetal growth and length of gestation in an agricultural population. *Environ Health Perspect* 112:1116–1124.
- Fagan JF, Singer LT, Montie JE, Shepherd PA. 1986. Selective screening device for the early detection of normal or delayed cognitive development in infants at risk for later mental retardation. *Pediatrics* 78:1021–1026.
- Fein GG, Jacobson JL, Jacobson SW, Schwartz PM, Dowler JK. 1984. Prenatal exposure to polychlorinated biphenyls: effects on birth size and gestational age. *J Pediatr* 105:315–320.
- Fifer VP, Greene M, Hurtado A, Myers MM. 1999. Cardio-respiratory responses to bidirectional tilts in infants. *Early Hum Dev* 55:265–279.
- Frankenburg WK, Dodds J, Archer P, Shapiro P, Bresnick B. 1992. The Denver II: a major revision and restandardization of the Denver Developmental Screening Test. *Pediatrics* 89:91–97.
- Gauger KJ, Kato Y, Haraguchi K, Lehmler HJ, Robertson LW, Bansal R, et al. 2004. Polychlorinated biphenyls (PCBs) exert thyroid hormone-like effects in the fetal rat brain but do not bind to thyroid hormone receptors. *Environ Health Perspect* 112:516–523.
- Goldsmith HH. 1996. Studying temperament via construction of the Toddler Behavior Assessment Questionnaire. *Child Dev* 67:218–235.
- Gregory R, Lehman R, Mohan P. 1976. Intelligence test results for children with and without undue lead absorption. In: *Shoshone Lead Health Project* (Wegner G, ed). Boise, ID:Idaho Department of Health and Welfare, 120–149.
- Guillette EA, Meza MM, Aquilar MG, Soto AD, Garcia ID. 1998. An anthropological approach to the evaluation of preschool children exposed to pesticides in Mexico. *Environ Health Perspect* 106:347–353.
- Hamer D. 2002. Rethinking behavior genetics. *Science* 298:71–72.
- Haynes EN, Kalkwarf HJ, Hornung R, Wenstrup R, Dietrich K, Lanphear BP. 2003. Vitamin D receptor *FokI* polymorphism and blood lead concentration in children. *Environ Health Perspect* 111:1665–1669.
- Hoon AH. 1996. Visual impairments in children. In: *Developmental Disabilities in Infancy and Childhood* (Capute AJ, Accardo PJ, eds). Baltimore, MD:Brookes Publishing, 461–478.
- Hornig M, Chain D, Lipkin WI. 2004. Neurotoxic effects of postnatal thimerosal are mouse strain dependent. *Mol Psychiatry* 9:833–845.
- IPCS (International Programme for Chemical Safety). 1995. *Environmental Health Criteria 165. Inorganic Lead*. Geneva:World Health Organization.
- Jacobson J, Jacobson S, Schwartz O, Fein G, Dowler J. 1984. Prenatal exposure to an environmental toxin: a test of the multiple effects model. *Dev Psychol* 20:523–532.
- Jacobson JL. 2001. Contending with contradictory data in a risk assessment context: the case of MeHg. *Neurotoxicology* 22:667–675.
- Jacobson JL, Jacobson SW. 1996. Intellectual impairment in children exposed to polychlorinated biphenyls *in utero*. *N Engl J Med* 335:783–789.
- Jacobson JL, Jacobson SW. 2003. Prenatal exposure to polychlorinated biphenyls and attention at school age. *J Pediatr* 143(6):780–788.
- Jacobson JL, Jacobson SW, Padgett RJ, Brumitt GA, Billings RL. 1992. Effects of prenatal PCB exposure on cognitive processing efficiency and sustained attention. *Dev Psychol* 28:297–306.
- Jacobson SW. 1998. Specificity of neurobehavioral outcomes associated with prenatal alcohol exposure. *Alcohol Clin Exp Res* 22:313–320.
- Kagan J, Reznick J, Snidman N. 1988. Biological basis of childhood shyness. *Science* 240:167–171.
- Kahn RS, Khoury J, Nichols WC, Lanphear BP. 2003. Role of dopamine transporter genotype and maternal prenatal smoking in childhood hyperactive-impulsive, inattentive, and oppositional behaviors. *J Pediatr* 143:104–110.
- Kallen B. 1988. Epidemiology of Human Reproduction. Boca Raton, FL:CRC Press.
- Kerper LE, Ballatori N, Clarkson TW. 1992. MeHg transport across the blood-brain by an amino acid carrier. *Am J Physiol* 262:761–765.
- Kerper LE, Hinkle PM. 1997. Lead uptake in brain capillary endothelial cells: activation by calcium store depletion. *Toxicol Appl Pharmacol* 146:127–133.
- Korkman M, Kirk U, Kemp S. 1998. NEPSY—A Developmental Neuropsychological Assessment. San Antonio, TX:Psychological Corporation.
- Krasnegor NA, Otto DA, Bernstein JH, Burke R, Chappell W, Eckerman DA, et al. 1994. Neurobehavioral test strategies for environmental exposures in pediatric populations. *Neurotoxicol Teratol* 16:499–509.
- Lester BM, Tronick EZ, Brazelton TB. 2004. The Neonatal Intensive Care Unit Network Neurobehavioral Scale. *Pediatrics* 113:641–667.
- Lipkin PH. 1991. Epidemiology of developmental disabilities. In: *Developmental Disabilities in Infancy and Childhood* (Capute AJ, Accardo PH, eds). Vol 1. Baltimore, MD:Brookes Publishing, 43–67.
- Matthews K, Woodall K, Stoney C. 1990. Changes in and stability of cardiovascular responses to behavioral stress: results from a four-year study of children. *Child Dev* 61:1134–1144.
- McCarthy D. 1974. McCarthy Scales of Children's Abilities. San Antonio, TX:Psychological Corporation.
- McMichael AJ, Vimpani GV, Robertson EF, Baghurst PA, Clark PD. 1986. The Port Pirie cohort study: maternal blood lead and pregnancy outcome. *J Epidemiol Community Health* 41:18–25.
- Morrell J. 1999. The Infant Sleep Questionnaire: a new tool to assess infant sleep problems for clinical and research purposes. *Child Psychol Psychiatry* 4:20–26.
- Myers GJ, Davidson PW. 2000. Does MeHg have a role in causing developmental disabilities in children? *Environ Health Perspect* 108(suppl 3):413–420.
- Narahishi T, Ginsburg KS, Nagata K, Song JH, Tatebayashi H. 1998. Ion channels as targets for insecticides. *Neurotoxicology* 19:581–590.
- National Research Council. 2000. *Toxicological Effects of MeHg*. Washington, DC:National Academic Press.
- Needleman HL, Bellinger DC. 1986. Type II fallacies in the study of childhood exposure to lead at low dose: a critical and quantitative review. In: *Lead Exposure and Child Development: An International Assessment* (Smith MA, Grant LD, Sors AI, eds). Boston:Kluwer Academic Publishers, 293–304.
- Needleman HL, Schell A, Bellinger D, Leviton A, Allred EN. 1990. The long-term effects of exposure to low doses of lead in childhood. An 11-year follow-up report. *N Engl J Med* 322:83–88.
- Newborg J, Stock JR, Wnek L. 1984. *Batelle Developmental Inventory*. Itasca, IL:Riverside Publishing.
- NIDA. 2000. *Early Childhood Neurobehavioral Assessment for the Differential Diagnosis of Fetal Alcohol Syndrome and Alcohol-Related Neurobehavioral Disorder*. Bethesda, MD:National Institute on Drug Abuse, the Interagency Coordinating Committee on Fetal Alcohol Syndrome and National Institutes of Health.
- Porges S, Doussard-Roosevelt JA, Portales A, Suess P. 1994. Cardiac vagal tone: stability and relation to difficultness in infants and 3-year olds. *Dev Psychobiol* 27:289–300.
- Raven JC. 1998. *Raven's Progressive Matrices*. San Antonio, TX:Psychological Corporation.
- Rees DC, Francis EZ, Kimmel CA. 1990. Scientific and regulatory issues relevant to assessing risk for developmental neurotoxicity: an overview. *Neurotoxicol Teratol* 12:175–181.
- Reif A, Lesch KP. 2003. Toward a molecular architecture of personality. *Behav Brain Res* 139:1–20.
- Rhodes MC, Seidler FJ, Qiao D, Tate CA, Cousins MM, Slotkin TA, et al. 2004. Does pharmacotherapy for preterm labor sensitize the developing brain to environmental neurotoxicants? Cellular and synaptic effects of sequential exposure to terbutaline and chlorpyrifos in neonatal rats. *Toxicol Appl Pharmacol* 195:203–217.
- Rodier PM. 1994. Vulnerable periods and processes during central nervous system development. *Environ Health Perspect* 102(suppl 2):121–124.
- Rodier PM. 2004. Environmental causes of central nervous system maldevelopment. *Pediatrics* 113:1076–1083.
- Rogan WJ, Dietrich, KN, Ware JH, Dockery DW, Salganik M, Radcliffe J, et al. 2001. The effect of chelation therapy with succimer on neuropsychological development in children exposed to lead. *N Engl J Med* 344:1421–1426.
- Rohlman DS, Bailey SR, Anger WK, McCauley L. 2001. Assessment of neurobehavioral function with computerized tests in a population of Hispanic adolescents working in agriculture. *Environ Res* 85:14–24.
- Rohlman DS, Gimenes L, Eckerman DA, Kang SK, Farahat FM, Anger WK. 2003. Development of the Behavioral Assessment and Research System (BARS) to detect and characterize neurotoxicity in humans. *Neurotoxicology* 24:523–531.
- Roizen NJ. 1996. Hearing loss. In: *Developmental Disabilities in Infancy and Childhood* (Capute AJ, Accardo PJ, eds). Baltimore, MD:Brookes Publishing, 479–496.
- Rothbart MK. 1981. Measurement of temperament in infancy. *Child Dev* 52:569–578.
- Rozman KK, Klaassen CD. 1996. Absorption, distribution, and

- excretion of toxicants. In: *Toxicology: The Basic Science of Poisons* (Klaassen CD, Amdur MO, Doull J, eds). 5th ed. New York:McGraw-Hill, 91–109.
- Schantz SL. 1996. Developmental neurotoxicity of PCBs in humans: what do we know and where do we go from here? *Neurotoxicol Teratol* 18:217–228.
- Seegal RF. 2000. The neurotoxicological consequences of developmental exposure to PCBs. *Toxicol Sci* 57:1–3.
- Slikker W, Miller RK. 1994. Placental metabolism and transfer: role in developmental toxicology. In: *Developmental Toxicology* (Kimmel CA, Buelke-Sam J, eds). 2nd ed. New York:Raven Press, 245–283.
- Slotkin TA. 1999. Developmental cholinotoxicants: nicotine and chlorpyrifos. *Environ Health Perspect* 107:71–80.
- Stewart P, Reihman J, Lonky E, Darvill T, Pagano J. 2000. Prenatal PCB exposure and Neonatal Behavioral Assessment Scale (NBAS) performance. *Neurotoxicol Teratol* 22:21–29.
- TERA. 1999. *Comparative Dietary Risks: Balancing Risks and Benefits of Fish Consumption*. Cincinnati, OH:Toxicology Excellence for Risk Assessment.
- Tong S, Baghurst P, McMichael A, Sawyer M, Mudge J. 1996. Lifetime exposure to environmental lead and children's intelligence at 11–13 years: the Port Pirie cohort study. *BMJ* 312:1569–1575.
- Veenstra-VanderWeele J, Cook EH, Lombroso PH. 2003. Genetics of childhood disorders: 46. Autism, part 5: genetics of autism. *J Am Acad Child Psychiatry* 42:116–118.
- Wechsler D. 1997. *Wechsler Adult Intelligence Scale*. 3rd ed. San Antonio, TX:Psychological Corporation.
- Wechsler D. 1999. *Wechsler Abbreviated Scale of Intelligence*. San Antonio, TX:Psychological Corporation.
- Wechsler D. 2002. *Wechsler Preschool and Primary Scales of Intelligence*. 3rd ed. San Antonio, TX:Psychological Corporation.
- Weiss B. 1997. Pesticides as a source of developmental disabilities. *Ment Retard Dev Disabil* 3:244–256.
- Weiss B, Amler S, Amler RW. 2004. Pesticides. *Pediatrics* 113:1030–1036.
- White RF, Gerr F, Cohen RF, Green R, Lezak MD, Lybarger J, Mack J, et al. 1994. Criteria for progressive modification of neurobehavioral batteries. *Neurotoxicol Teratol* 16:511–524.
- Whyatt RM, Rauh V, Barr DB, Camann DE, Andrews HF, Garfinkel R, et al. 2004. Prenatal insecticide exposure and birth weight and length among an urban minority cohort. *Environ Health Perspect* 112:1125–1132.
- Wilson JG. 1977. Current status of teratology—general principles and mechanisms derived from animal studies. In: *Handbook of Teratology* (Wilson JG, Fraser RC, eds). Vol 1. New York:Plenum, 47–74.
- Zimmerman I, Steiner V, Pond R. 1992. *Preschool Language Scale*. 3rd ed. San Antonio, TX:Psychological Corporation.