Future Directions for Neurobehavioral Studies of Environmental Neurotoxicants

David C. Bellinger
Neuropediatrics Unit, Children’s Hospital (Boston), Harvard Medical School, 300 Longwood Avenue, Boston, MA 02115, USA.

Received 28 February 2001; accepted 27 April 2001

Abstract

Three proposals for enriching neurobehavioral toxicology studies are discussed. First, while IQ has proven useful as a primary endpoint, such apical measures are limited; they obscure important individual differences, tend to reflect the product rather than the process of learning, sample a limited range of intelligent behaviors, and are insensitive to critical outcomes such as learning disabilities. In terms of a wide disease burden, behavioral and psychiatric morbidities might be even more important than cognitive vulnerabilities. Such endpoints warrant careful attention. Second, the models of child development can be enriched, increasing our ability both to control for confounding bias and to evaluate potential modification of neurotoxic effects by contextual factors. While the use of the HOME Observation for Measurement of the Environment (HOME scale) and other measures of family-level prenatal factors was an important advance, recent sociological work demonstrates the importance of broader conceptualizations of the ecology of child development (e.g. neighborhood and community characteristics). Third, much effort has been expended in attempts to identify the behavioral signature associated with exposure to a particular neurotoxicant. Given the limited success in identifying behavioral phenotypes even for well-characterized genetic disorders (e.g. Fragile-X, Williams, and/or hearing loss), the prospects seem grim for identifying specific and relatively invariant patterns in the expression of neurotoxicants across diverse dosing regimens and biological and cultural settings. In part this reflects the likely influence of complex, largely unknown, patterns of effect modification on the expressions of neurotoxicants. Efforts to define the nature of these contingencies might be more productive than continued efforts to identify behavioral phenotypes.

Keywords: Neurobehavioral studies; Neurotoxicants; Behavioral signatures; Epidemiology

INTRODUCTION

The methodological quality of investigations into the import of neurotoxicant exposures for child development has advanced dramatically over the past two decades. To appreciate this, one need only compare the sophistication of the cross-sectional studies of lead toxicity conducted in Michigan (Jacobson and Jacobson, 1996) and North Carolina (Elden and Rogen, 1991), or to the prospective studies of methyl mercury toxicity conducted in the Seychelles (Davidson et al., 1998) and Faroe Islands (Grundkje et al., 1997). A critical factor in this evolution was recognition of the need to assemble more inclusive multidisciplinary teams, with representation of the many disciplines whose concepts and methods are needed in order to achieve convincing answers to the difficult questions posed in neurobehavioral toxicology studies. Three methodological advances, in particular, warrant mention. First, the collection of more comprehensive biomarker data has permitted better dose reconstruction, reducing exposure misclassification and its attendant biases. In addition, greater efforts have been made to incorporate toxicokinetic considerations into the process of drawing inferences.
(Masliah, 1992). Second, approaches to statistical analysis have evolved from bivariate tests comparing the performance of “exposed” and “unexposed” groups to multivariate techniques such as generalized estimating equations for modeling longitudinal data and nonparametric regression for identifying nonlinearities in dose-effect relationships. Third, endpoint assessment has evolved from signs of clinical disease to indices of more subtle “subclinical” perturbations of function derived on the basis of current research in cognitive neuropsychology and neuroscience.

Continued improvements in methodologies used in neurobehavioral toxicology studies is both desirable and possible, however. In this paper, three possible avenues are discussed. Two pertain to additional types of data that should be collected. In the realm of endpoints, it is important to continue to expand the breadth of functions measured as outcomes, moving beyond IQ, and indeed, beyond indices of cognition. In the realm of covariate measurement, it is important to enrich the models of child development that serve as the frame of reference within which we work in attempting to identify endpoint variance that is uniquely attributable to a neurotoxicant exposure. The third avenue pertains to the inferences that are drawn from study results, specifically with regard to the existence of “behavioral signatures” specific to neurotoxicant exposures.

ENDPOINT SELECTION

The most frequent choice of primary endpoint continues to be intelligence (IQ) and other similarly global, or apical, indices. The history of research on the effects of low-level lead toxicity demonstrates the utility of such indices in motivating changes in regulatory policy. It is a familiar construct, and most people share the view that more is better. Despite uncertainty as to exactly what it does measure, IQ possesses strong predictive validity for critical outcomes such as an individual’s success in school and at work (Neisser et al., 1996). The absence of an IQ-like apical measure from the Faroe Islands study assessment protocol has complicated the effort to compare the results of that study to those of the Seychelles study. Furthermore, changes in IQ can be expressed in monetary terms (i.e., increased lifetime earnings per additional IQ point), facilitating the task of characterizing the public health significance of factors that alter the IQ distribution within a population (e.g., Schauf, 2000).

As useful as apical indices are, however, they clearly fail to capture much that is critical to the cognitive and psychological health of an individual and of a society. First, they average performance over a broad range of functions, so can obscure important distinctions between individuals. Three children, one with Williams syndrome, one with autism, and one with Down syndrome might achieve the same IQ score yet will likely present very different patterns of cognitive strengths and weaknesses. The child with Williams syndrome might demonstrate a relative strength in language and a relative weakness in spatial cognition (Bellugi et al., 1999), the autistic child the opposite pattern, and the child with Down syndrome a mild decrement in both domains. For this reason, a child’s full-scale IQ score is one of the least informative pieces of data that a clinician gathers in conducting a neuropsychological assessment. Second, an IQ score generally reveals more about the product of a child’s past thinking than about the process of the child’s current, on-line processing, or what is sometimes referred to as “fluid” intelligence. Rather, it conveys information about what a child already knows, or so-called “crys-
tallized” intelligence, which tends to reflect the rich-
ness of opportunity (e.g., family socioeconomic status, quality of the home environment). How a child learns when presented with new material might be less con-
found with opportunity and thus represent a more sensitive index of neurotoxicant exposure. Third, an IQ test tends to sample only some of the behaviors that are usually considered relevant to the concept of “intelli-
gence.” Gardner (1993) proposed the concept of “mul-
tiple intelligences”, arguing that IQ tests assess primarily linguistic and logical-mathematical intelli-
gences, but not bodily-kinesthetic, spatial, or personal intelligences. Similarly, in his triarchic model of intel-
ligence, Sternberg (1985) suggested that IQ tests sam-
ple analytic intelligence, but neither creative nor practical intelligence. Even within the analytic, lin-
guistic and logical-mathematical domains, certain functions are underrepresented among IQ tests items, such as the ability to plan, to organize, and to develop and test hypotheses, what is sometimes referred to as executive functions (Lezak, 1995). Such higher-order integrative and organizational skills might be relatively more vulnerable to neurotoxicant exposure than are the more basic, building block skills (e.g., Fried and Wijckson, 2000). Fourth, some clinically important conditions include a “normal” IQ score as part of the case definition. For example, in most US states, a learning disability is defined as a discrepancy between a child’s ability (i.e., IQ) and his or her achievement in
an academic skill such as reading or mathematics (Reynolds, 1985). If only IQ is measured, this outcome, with its substantial implications for a child's well-being, cannot be assessed. In the Boston prospective lead study, for instance, an exposure-related decrement in academic performance was apparent in the presence of very high IQ scores (Bellinger et al., 1992). Specifically, children's blood lead levels at 2 years of age were inversely related to reading, mathematics, and spelling skills even among children with full-scale IQ greater than 125, or nearly 2 S.D. above the expected population mean (Bellinger, 1995). The possibility that this group of children experienced any adverse effect of lead would probably not have been considered had IQ been the only endpoint assessed.

Important initiatives are underway exploring the application to humans of test methods developed for animals, such as the NCTR Operant Test Battery which includes tests of motivation, color and position discrimination, time estimation, delayed match-to-sample, and learning of response chains (Pande et al., 1999; see also Davidson et al., 2000). There is a critical sense, however, in which we will fail our mission to identify the full spectrum of neurotoxic effects if we continue to focus on IQ, in particular, and cognitive function, in general. Central nervous system toxicity can be expressed in many forms other than cognitive impairments. Although the hypothesis was suggested by Weiss (1985) many years ago, only recently has serious consideration been given to the possibility that exposures to environmental chemicals contribute to behavioral and social pathologies, the study of which has been termed "psychotoxicology" (DuPont, 1989). In view of the prevalence and costs of such disorders, however, it would be foolhardy to overlook this hypothesis. In the recent WHO-sponsored Global Burden of Disease Study (Murray and Lopez, 1996), diseases were compared using a metric called the disability-adjusted life year (DALY), a combination of years of life lost to premature death and years lived with a disability of specified severity and duration. Two findings in particular are noteworthy. First, 5 of the 10 diseases with the greatest DALY totals are mental illnesses: unipolar major depression (#1); alcohol use (#4); bipolar disorder (#6); schizophrenia (#9); obsessive-compulsive disorder (#10). Among women in developed countries, five of the six leading contributors to DALY are mental illnesses: unipolar major depression (#1): schizophrenia (#2); bipolar disorder (#4); obsessive-compulsive disorder (#5); alcohol use (#6). Second, the proportion of disability due to psychiatric and neurological conditions is projected to increase by 50% by the year 2020.

No comparable set of analyses is available for children but the prevalences of the major childhood psychiatric disorders are consistent with a substantial burden of disease. Based on a total birth cohort of 60,000 6- and 9-year-olds in Finland, Arvagast et al. (1999) estimated that 21.8% of children met diagnostic criteria for a psychiatric disturbance. Many cases are likely to have been mild, however, insofar as most estimates of serious emotional disturbance in children are in the range of 5-8% (Costello, 1999). Studies focusing on specific diagnoses suggest a prevalence of 16% for attention deficit disorder (Wolraich et al., 1998), 9.6% for a major affective disorder (Lewisohn et al., 1999), and 2-3% for obsessive-compulsive disorder (Zohar, 1999). Thus, the prevalence of mental disorders worldwide is high, both among adults and children, and the costs of these morbidities are enormous, both to society in terms of lost productivity and to individuals in terms of personal distress. Reducing such morbidities is thus an important public health goal. But do neurotoxic exposures contribute to this problem?

Studies of occupationally-exposed adult cohorts suggest that the answer might be yes (Hartman, 1988). Eireism, a syndrome of emotional lability, irritability, excessive shyness, and memory disturbances, is associated with exposure to inorganic mercury (Sato, 2000). Manganese exposure is associated with a syndrome that involvesmania, insomnia, hallucinations, aggression, incoherent speech, inappropriate affect, and emotional lability (Mergler and Badwin, 1997). Exposure to trimethyl tin is associated with a mood disorder involving alternating bouts of rage and deep depression, sleep disturbance, fatigue, memory loss, and apathy (Rose et al., 1981). Organic forms of lead produce affective disorders (Schottenfeld and Cullen, 1984). In the 1920s, in a tetraethyl lead plant in New Jersey that became known as the "House of Butterflies," workers were frequently seen trying to brush hallucinated insects from their bodies (Needleman, 1973). An encephalopathy associated with exposure to volatile organic compounds involves personality changes and depression (Broadwell et al., 1995), and solvents such as toluene are frequently abused for their psychoactive effects (Balster, 1998).

To date, relatively little work has been done to investigate the potential contributions of environmental neurotoxic exposures to childhood psychiatric morbidity. The hypothesis that autism is associated
with an early prenatal exposure has been raised (London and Etzel, 2000). Some have speculated, on the basis of anecdotal evidence, that ethylmercury, used in the US as a vaccine preservative until recently, is involved in the etiology of autism. At present, it is only lead that has been systematically studied with regard to its role in the origin of childhood behavior disorders. Organic lead poisoning has been associated with a schizophrenia-like psychosis (McCranken, 1987), and increased body burdens of inorganic lead with childhood psychosis, including autism (Cohen et al., 1976; Accardo et al., 1988), although the increased lead exposure seen in such cases is likely to be secondary rather than primary. Several studies have reported associations between increased exposure to inorganic lead and parent or teacher ratings of children’s behavior, particularly indices of attention such as distractibility, impulsivity, and impersistence (Needleman et al., 1979, 1996; Yule et al., 1984; Thomson et al., 1989; Bellinger et al., 1994; Scharlito et al., 1992; Wasserman et al., 1998). Within a subcohort from the Collaborative Perinatal Project, Denno (1993) found that a history of lead poisoning was a significant risk factor for both juvenile and adult crime, as well as for disciplinary problems in school. Needleman and colleagues showed that higher bone lead levels are associated with an increased risk of self-reported antisocial behaviors (Needleman et al., 1996), and more recently, that being an adjudicated delinquent is associated with an increased risk of having an elevated bone lead level (Needleman et al., 2000). In a provocative set of historical ecological analyses, spanning for some endpoints, the last 125 years, Növin (2000) reported strong associations between the amounts of inorganic lead used in paints and gasoline and rates of violent crime and unwed pregnancies.

As a result of the fear exclusive focus on the potential effects of environmental neurotoxins on cognition, and more specifically, on IQ, only preliminary steps have been taken to examine the impact of such exposures on children’s risk of psychiatric disorders. In terms of overall human and financial costs, the latter morbidities might ultimately be more important than the cognitive morbidities. Given the known associations between children’s learning impairments and their psychosocial adjustment (e.g. self-esteem/self-concept, anxiety, depression, and social competence) (Greenbaum, 1999), clarifying the range and severity of the neuropsychiatric expressions of environments’ chemical toxicity is thus a critical research need.

MODELS OF CHILD DEVELOPMENT

Efforts now being undertaken to understand how neurotoxic exposures fit into the complex web of influences on child development are considerably more sophisticated than those of 20 years ago. The construction of veritable models of child outcomes yields many potential benefits. First, the amount of error variance is reduced. Most of the endpoints studied have multiple determinants and thus represent final common pathways for the expression of many factors. Typically, no more than 40-50% of the variance in child cognition can be accounted for by our regression models, and often much less in populations that are relatively homogeneous in sociodemographic characteristics. Building models that are more explanatory would thus, in principle, increase the power of our hypothesis tests involving neurotoxins, permitting the identification of subtler effects than is now possible. Second, richer models would provide greater opportunities to control for potential confounding bias, reducing uncertainty about whether an exposure-related decrement in a test score is merely epiphenomenal, resulting from the correlation between the exposure and some third factor that is causally related to the test score. Third, richer models would also provide greater opportunities to identify factors that modify the association between an exposure and an outcome. To a great extent, our approaches to data analysis reflect a belief that the world is a simple one, consisting of independent main effects (Levitan, 1987) whereas the appropriate models almost certainly involve multiple component causes combining to produce disease. Instances of single sufficient causes seem to be, in fact, the rare exception (Rothman and Greenland, 1998). This perspective leads directly to the conclusion that...
health care, pre-existing disease, psychosocial factors such as stress, nutrition and other lifestyle choices. When the issue of effect modification is addressed at all in environmental epidemiology studies, it is intrinsic factors that tend to be studied.

Endorsing a need for better models is easy. The critical question is, how do we go about building them, especially models that address acquired susceptibility factors? Presently most effort is invested in characterizing proximal influences on child development. These include individual- and family-level factors such as socioeconomic status (SES). There remains considerable room for improvement even at this level, however, SES is typically assessed by means of the Hollingshead Four-Factor Index of Social Status (Hollingshead, 1975). It is perhaps time to agree that time has passed this scale by. Many of the highly technical jobs held by young parents today did not exist in 1975, when this instrument was last revised. Even if the j-ge classification scheme were updated, however, most would probably agree that parental job title and highest level of education are poor proxy measures of the actual forces molding a child’s development. Concepts such as social class, ethnicity, and race cast a very wide measurement net and, “... encapsulate complex information about a person’s life” (Blau, 1995, p. 904). Their status as omnibus variables simultaneously increases their value as means for controlling confounding and reduces their value as means for explaining the “why” of effect causation (Hogue, 1997).

The recent use in neurobehavioral toxicology studies of process-oriented instruments such as the HOME Observation for Measurement of the Environment (HOME scale) represents an important advance in our ability to describe important proximal forces. Nevertheless, many key aspects of the family environment are not captured by the HOME and similar tools and should be incorporated into our models, including the accessibility and involvement of extended family in a child’s life, parental psychopathology, quality of the marital relationship, family coping and negotiation styles, and expressed emotion (Wamboldt and Wamboldt, 2000). Furthermore, the typical “snapshot” taken of family culture as part of data collection ignores the potential reciprocal influences between child status and family process, instead assuming only unidirectionalism, from family to child. Moreover, despite Tolstoy’s clinically astute observation in Anna Karenina that, “Every unhappy family is unhappy in its own way”, a “one size fits all” approach is usually adopted when measuring these factors. For example, the tool used to measure stress is generally a 1 to 5 events inventory that assigns a fixed number of points to a given event (e.g. 20 points for change in residence, 40 for pregnancy, 73 for divorce; Holmes and Rahe, 1967). This fails to take account of the fact that the event is likely to be inibed with different meanings by different families and thus to have different impacts on the children within the families. For that matter, to some extent, the same event will be experienced differently by two children within the same family.

Where the most work remains to be done, however, is in the development of methods to capture influences on children’s development that are more distant than the nuclear family. Although it is acknowledged that societal forces beyond the family influence child development, the assumption is made that effects will be mediated by the family and find expression in family-level variables such as HOME scores or SES. In this view, accurate measurement of family characteristics will result, de facto, in the measurement of the influences of these larger forces. The work of sociologists and anthropologists shows, however, that social structures are more than the sum of their individual parts and that such structures possess emergent properties that cannot be induced from the more proximal or local factors. Whether “it takes a village to raise a child” is arguable, but clearly a child does grow up within a village and is likely to be affected by its properties. McMichael (1990) has characterized investigators pursuing the conventional approach of focusing solely on individual-level factors as, “prisoners of the proximate”, and interest is growing in “eco-epidemiology” or “social-ecologic epidemiology” (Sussner and Sussner, 1996). This view emphasizes the need to identify the social, economic, and political forces that determine the patterns of disease occurrence (e.g. Bailey et al., 1994).

What sorts of analyses are being advocated? Two brief examples are provided. First, in seeking to develop a model to predict crime, Sampson et al. (1997) hypothesized that “... social and organizational characteristics of neighborhoods explain variation in crime rates that are not solely attributable to the aggregated demographic characteristics of individuals” (p. 918). They reasoned that just as individuals vary in their capacity for effective action, neighborhoods vary in their capacity to achieve common goals, a capacity that the investigators called “collective efficacy”. In contrast to formal mechanisms such as police actions, collective efficacy is manifested in the informal means by which residents themselves achieve and maintain order (e.g. monitoring of children’s playgroups, intervening to prevent acts of truancy,
confronting individuals causing disturbances in public areas). The investigators interviewed almost 880 residents of 350 neighborhood clusters in Chicago to identify the factors that foster collective efficacy and to examine patterns of violence and crime within the neighborhoods. The key result was that collective efficacy was negatively associated with variations in neighborhood violence, even when individual-level characteristics, measurement error, and rates of prior violence in the neighborhood were controlled statistically.

Second, Van and Kaplin (1999) used the large Alameda County Study cohort to examine mortality risk as a function of the usual individual risk factors (e.g., age, sex, income, education, smoking, alcohol, body mass index, perceived health status), census data (e.g., percent white-collar employees, percent renters, crowding, per capita income, percent single-family dwellings), and what they termed “neighborhood social environment”. This latter index was operationalized as the number of common commercial stores in the neighborhood, the number of parks, the frequency of motor vehicle crashes, injuries, etc. The key result in this study was that a low neighborhood social environment score was significantly associated with an increased risk of mortality, even after controlling for residents’ individual-level risk factors.

Is the domain of child development, Bronfenbrenner has long advocated the adoption of a broader perspective on the factors that influence child development, once defining academic developmental psychology as, “...the science of the behavior of children with strange adults in strange settings for the shortest periods of time” (Bronfenbrenner, 1977, p. 513). As an alternative, following on the work of Kurt Lewin, he developed a topological model of child development involving a set of nested and interconnected structures. The innermost level he termed the microsystem, which consists of the immediate settings in which a child interacts (home, daycare center, school), the level of which we have factored most of our efforts in developing covariate models in neurotoxicology studies. Beyond the microsystem is the mesosystem. This consists of the relationships among a child’s microsystems, and this has characteristics that might not be discerned even from close observation of the component microsystems. The exosystem is an extension of the mesosystem, encompassing other aspects of social structure that do not directly involve the child but which do impinge on or encompass the immediate settings in which the child functions. They constrain what goes on in those internal lives of children, and define the cultural meanings and values conveyed to the child.

The perspective developed above suggests a need to expand even further the set of collaborations that are considered necessary in order to conduct a state-of-the-art neurobehavioral toxicology study. Psychiatry, sociological, cultural anthropology, and other fields whose practitioners have the tools needed to characterize levels beyond the microsystem should be represented on the study team. The goal is not to transform neurobehavioral toxicology studies into sociological exercises but rather to encourage the construction of models that permit less constrained, statistically more powerful tests of the primary study hypotheses. Just as important is the fact that richer models will provide greater opportunities to appreciate the conditions that foster or suppress the expression of neurobehavioral toxicity, which in turn will permit more precise characterization of susceptible subgroups of the population.

This step in the overall risk assessment process, although viewed as critical, is usually addressed only in very general ways. Indeed, it might become possible to characterize the relative susceptibility of different subgroups on a quantitative rather than qualitative basis, helping us to progress beyond the satisfying practice of deriving exposure standards by the application of vague uncertainty factors which are usually, by defects, tidy round numbers.

**BEHAVIORAL SIGNATURES OR PHENOTYPES**

A question of considerable interest is whether exposure to a particular neurotoxin is associated with a pattern of neurobehavioral deficits that is both unique to that neurotoxin and consistent across individuals, study cohorts, and settings. In a case of the tail wagging the dog, the strong assumption that such behavioral signatures exist has sometimes determined the inferences drawn from study results. For example, judgments as to whether the deficits associated with a neurotoxin are likely to be due to chance, a methodological artifact, or a true toxicological process have depended on the degree to which different studies report the same pattern of neurobehavioral strengths and weaknesses among more highly exposed children. In cohorts exposed to two or more neurotoxins, the pattern of deficits has been used to decide which chemical is responsible for which effects. Some have wondered, for instance, which, if any, of the neurobehavioral deficits identified in the Fierce Islands cohort and to P0 result

In contrast, the toxicology procedure, or review, has tended to favor more detailed analysis of behaviors. The case study procedure, or review, has tended to favor more detailed analysis of behaviors.
due to the use of methyl mercury exposure and which to PCB or other organochlorine exposures that also result from the consumption of marine mammals.

In order to frame the discussion, it is useful to consider in some detail the literatures on pediatric diseases such as chromosomal abnormalities, for which the insult, its timing, and its mechanism of action are known more precisely than is typical for environmental neurotoxicants. In some cases these abnormalities involve a single gene, although most involve deletions or mutations affecting multiple genes. Despite the relatively circumscribed nature of these abnormalities, however, the typical finding is a considerable variability in phenotypic expression. For only three disorders are the signs so specific and reliable across individuals that they have diagnostic utility, i.e., the presence of the defining behaviors is pathognomonic and the appellation of "behavioral phenotype" is appropriate in the rigorous sense: Lesch-Nyan disease, Prader-Willi syndrome, and Rett syndrome (Flinn and Yule, 1994). What distinguishes these disorders from those for which the case for a behavioral phenotype is less tenable is that the critical behaviors are unusual, often stereotyped or ritualistic, and not commonly seen under other circumstances. Children with Lesch-Nyan disease as X-linked disorder involving a deficiency in hypoxanthine phosphoribosyltransferase, engage in severe forms of self-mutilating behavior that are qualitatively different from, and more resistant to intervention than, the self-injurious behaviors associated with other mental retardation syndromes. Prader-Willi syndrome, associated with a deletion of chromosome 15 (at 5q11-q13), is characterized by overeating and abnormal food-seeking behaviors (overeating, leading to extreme obesity. Rett syndrome is an X-linked dominant disorder that appears to involve a highly specific mutation in gene MeCP2 at locus Xq28.

Children with this disorder experience a gradual loss of purposeful hand movements and their replacement by midline hand-wringing stereotypes.

When the behaviors associated with a disorder are merely unusual or extreme versions of behaviors seen in normally developing children, it is harder to make a case for the existence of a behavioral phenotype. For example, children with Down syndrome are often referred to as being friendly and placid in temperament (although the evidence for this is weak; Belmont, 1971). This is not a behavioral signature in the strict sense, however, since not all friendly, placid children have Down syndrome. It is certainly true that certain neurobehavioral findings tend to occur more frequently than expected in association with a given disease or exposure although even this less restrictive conception of a behavioral signature may not hold up very well, especially when viewed from a longitudinal perspective. In interrelationships of metabolism, for instance, the stability and specificity of clinical presentation is only modest. Among children with congenital hypothyroidism, behavioral expression changes over time. A delay in achieving euthyroidism is associated with poor verbal skills in pre-adolescents, but with poor visual skills in adolescence. Low concurrent TV levels are associated mainly with deficits in attention and memory (Rovev, 1999). Among children with phenylketonuria, high concurrent phenylalanine levels are associated with deficits on tests of working memory and inhibitory control, whereas high phenylalanine levels only in first month of life are associated with selective deficits in visual contrast sensitivity (Berenbaum, 1999a,b). Neuropsychological presentation differs for children exposed to excess phenylalanine prenatally because of having a mother with PKU and children exposed to excess phenylalanine postnatally because of being, themselves, phenylketotic (Watson et al., 2000). Children in the former group manifests primarily language deficits and children in the latter group primarily visual-spatial deficits. These findings are not particularly surprising, simply confirming that the timing of a CNS insult determines the manner in which its toxicities are expressed.

Other examples of phenotypic variability in chromosomal disorders are readily available. Fragile-X is caused by a CGG triplet repeat insertion in a region near the FMR-1 gene at Xq27. Given the specificity of the genetic abnormality, the amount of phenotypic heterogeneity in cognitive function is surprising (REF). The FMR-1 gene appears to regulate the expression of other genes so that the link between the genetic lesion and behavior is an indirect one. As Flinn (1996) observed, "The behavior appears to be the outcome of a long and complex pathway, and it is assumed that characterization of a single mutant gene will not take us far," (p. 361). It has also been postulated that environmental factors, such as folic acid level, might affect phenotypic expression of Fragile-X syndrome (Begga et al., 1987). Williams syndrome appears to be caused by a microdeletion on chromosome 7q11.23, a band that contains approximately 20 genes. Although some argue for existence of a characteristic association between language and face processing (strong) and spatial cognition (weak) (Bellugi et al., 1999), other investigators report a tendency for skills to be much more even in level across domains, and emphasize variability in expression within the syndrome.
(Greer et al., 1997; Mervis and Robinson, 2000). Mervis and Robinson (2000) concluded that the wide variability in expressive vocabulary of children with Williams syndrome is likely to be attributable to the same genetic and environmental factors that produce variability in the expressive vocabulary of children in the general population. Moreover, as noted with respect to inborn errors of metabolism, the phenotypic expression of Williams syndrome shifts over time. Numerosity judgments are strong in infancy and weak in adulthood, while language skills show the opposite pattern (Peterson et al., 1999).

Velocardiofacial syndrome (VCFS) is a recently described disorder associated with a microdeletion on chromosome 22q11.2, involving approximately 30 genes. It appears that children formerly diagnosed as cases of DiGeorge syndrome, Shprintzen syndrome, or Conotruncal Face Anomaly syndrome in fact comprise a single population, all with the 22q11.2 microdeletion (Wang et al., 2000). Again, however, the clinical presentation of VCFS is highly variable. A recent large European Collaborative Study revealed that 75% manifest a cardiovascular anomaly (usually tetralogy of fallot or a ventricular septal defect), 60% hypocalcemia, 49% a strabismus abnormality, 36% a renal abnormality, 36% growth impairment, 33% a hearing impairment, 32% a left-sided aortic insufficiency, and 18% a psychiatric diagnosis (usually part of the bipolar spectrum; but also a late onset psychosis) (Ryan et al., 1997). Motzkin et al. (1993) speculated that the size of the 22q11.2 microdeletion accounts for some of the variability in phenotypic severity. In the case of Williams syndrome, however, Mervis and Robinson (2000) reported that the size of the deletion on chromosome 7 was not associated with the severity of language impairment. Finally, the impact of a given genetic abnormality can vary depending on its origin. For instance, the cognitive impairment associated with a 22q11.2 microdeletion is less severe if the deletion is the result of a de novo mutation rather than inherited from a parent (Swillen et al., 1997). A 15q11.2-13 deletion on chromosome 15 will be expressed in Prader-Willi syndrome if the deletion involves the paternally inherited copy of this chromosome, but Angelman syndrome, a distinctly different phenotype, if it involves the maternally inherited copy, as in genomic imprinting (Nicholls, 2000).

Variability in disease expression is the norm rather than the exception, as an individual patient is not expected to express all of the signs and symptoms associated with a particular disease. This principle, while generally true in medicine, is particularly applicable to diseases for which diagnostic laboratory tests are not available and must instead be diagnosed solely on the basis of clinical presentation. In the Diagnostic and Statistical Manual of Mental Disorders (DSM IV) (APA, 1994), for example, six of nine specific behaviors must be present in order for a child to be diagnosed with attention deficit disorder-inattentive subtype. Similarly, six of nine other behaviors must be present in order to diagnose ADD-hyperactive/impulsive subtype. Therefore, for each subtype, 84 unique combinations of behaviors must criteria (i.e., nine features taken six at a time). Criteria for ADD-combined subtype can be met in 842 ways. (Any of the combinations that satisfy criteria for the inattentive subtype can be paired with any of the combinations that satisfy criteria for the hyperactive-impulsive subtype.) Only in a very general sense, then, can ADD be said to have a distinct behavioral phenotype. Examples of the loose application of the concept of behavioral phenotype to neurobehavioral/psychosocial disorders are common. Dills et al. (1996) proposed a behavioral phenotype for neurofibromatosis type I that included the following generalization about IQ, "Average skills are most common. All skill levels occur". In arguing in favor of a behavioral phenotype for Turner syndrome, Temple and Carney (1993) noted that their study resulted, "... support a specific overall intellectual profile in TS, but with considerable inter-individual variation in its expression" (p. 697). It is not clear that the concept of behavioral phenotype, so broadly conceived, has any utility.

Given the apparent phenotypic variation in the expression of relatively discrete chromosomal anomalies, the prospects do not seem encouraging for identifying specific and relatively invariant patterns in the expression of neurotoxic effects across diverse biological and cultural settings. The extreme version of the behavioral signiture hypothesis, that it is possible to diagnose PCB, methyl mercury or lead poisoning based on a child’s neurobehavioral presentation, can be quickly rejected based on the fact that the endpoints typically measured are transpecific and common pathways for the expressions of many influences. A deficit in mean score on a test of visual-spatial skills might result because the children have Tourette syndrome (Schulte et al., 1998), undergo early reparative cardiac surgery (Bellingham et al., 1999), or were of very low birth weight (Pasman et al., 1998). It is conceivable that closer and closer examination of the child’s visual-spatial skills might reveal syndrome or risk factor-specific variations in visual-spatial performance.
but this has not been demonstrated. The key question, then, is whether the inter-individual (and by extension, inter-study) consistency in the neurobehavioral expressions of exposure to PCBs or methyl mercury is great enough to support their use in the two ways cited in the paragraph above. The lead literature provides the best opportunity to address this question simply by virtue of the volume of studies available, but anyone holding out hope for evidence of a behavioral signature for lead will find only a little satisfaction. A case can be made that in the preschool period, children with higher lead burdens tend to show their greatest deficits in nonverbal, particularly visuo-motor/visuo-spatial, abilities (Bellinger et al., 1991; Wasserman et al., 1994, 2000; Dietrich et al., 1991, 1993, 2000; Baghurst et al., 1995). Among the studies that have followed children into school-age, however, the extent of inter-study consistency is the pattern of weaknesses is markedly lower (Bellinger et al., 1992; Baghurst et al., 1992; NRC, 1993; Wasserman et al., 1997; Tong et al., 1996).

In fact, the most consistent finding in the lead literature is an exposure-related generalized decrement on apical tests (e.g. IQ) that average performance over numerous neurobehavioral domains. The variability in specific domains could have many sources, including inter-study differences in the extent of residual confounding and a within-study inflated Type I error rate due to the examination of many endpoints. Another possibility is that not only the severity but also the nature of the deficits associated with lead, or any other neurotoxicant, varies depending on a variety of factors, the "experimental system" of human epidemiological studies (see the earlier discussion of the fallacy of single-cause attribution). This system, which is study-specific, includes, at the least, aspects of the exposure of a cohort (i.e. dose, timing, chronicity), its socioeconomic characteristics (e.g. social class, sex), and the specific assessment strategies used (Bellinger, 1995). It is axiomatic that knowing the timing of exposure to a teratogen or a toxicant is critical in predicting its effects (Rodier, 1995). Even variations in the severity of an insult can produce dramatically different effects. Fetal injuries of different durations during late gestational migration in the rat produce neurobehavioral malformations that differ greatly not only in severity but also in form, ranging from molecular layer ectopias to porencephaly (Rosen and Galaburda, 2000). The role of "system" variables might be especially salient for chemicals such as lead, to which individuals are exposed throughout the lifespan and which act in a very general level mechanistically, affecting diverse aspects of neural development and function. The simple fact that two individuals experienced elevated exposure to lead at some point in their lives provides little information that is useful in predicting the specific forms in which toxicity will be expressed. In part this undoubtedly reflects our poor understanding of the relevant interactions between exposure and these other factors that Peirce referred to as "unknown modifying factors". 

CONCLUSION

Three proposals were discussed in this paper: (1) the need to expand endpoint selection beyond IQ to include other cognitive as well as psychiatric morbidities, (2) the need to enrich our models of child development by considering as potential confounders and effect modifiers broader aspects of the ecological context of neurotoxicant exposure; (3) the need to temper our enthusiasm for the assumption that neurotoxicants leave specific behavioral signatures that are consistent across individuals, study cohorts, and sociocultural settings. Others might identify many other avenues for improvement that should also be implemented. The three avenues identified in this paper should not be viewed as independent of one another. Developing more comprehensive models of child development will become especially important if the scope of endpoint selection is enlarged to include psychiatric, as well as cognitive, morbidities insofar as the sociocultural risk factors are likely to differ considerably for these two types of impairments. Developing methods for including in our models factors general to societal and cultural forces beyond the individual and the family will also become increasingly important if the study population is ethnically and culturally diverse or poorly integrated into the majority culture (e.g. Hispanic children of migrant farm workers exposed to pesticides). Such factors are also relevant to studies of the effects of endocrine disrupting exposures, which might manifest as shifts in the expression of sexually-dimorphic behaviors that are generally considered to be influenced by social and cultural as well as biological factors (Vitez, 1997; Bereznack, 1999a,b). Finally, improvements in our ability to measure the broader ecology of child development will permit us to develop more detailed understanding of how neurotoxicant's effects are expressed in different settings, leading to a more subtle appreciation of the behavioral signature issue. Low-level exposure to a neurotoxicant is likely to have multiple, situation-specific behavioral signatures rather than a single
insurmountable that is expressed under every sce-
ario and that can serve as the arbiter of "truth" in resolv-
ing inconsistencies in findings across studies. The
risks we are attempting to characterize in envir-
onmental neurotoxicology studies are modest, cer-
tainly not comparable in magnitude to those
associated with cigarette smoking, which are suffi-
ciently large to be apparent within each member of
each cohort and longitudinal characteristics. On a popula-
tion basis, even an exposure as prevalent as lead is
likely to explain no more than 5% of the variation in a
health endpoint. In some respects, our default asump-
tion of a neurobehavioral "brain effect" corresponds
caseally to a belief in the validity of the summary
(weighted-average) effect estimate derived by combin-
ing data from different studies in a meta-analysis. In
such an approach, of which Proctor's would have
been proof, it is assumed that all studies estimate an
invariable "true" effect estimate, with the variability in
the estimates from the individual studies attributable
to random processes or to systematic biases associated
with specific study designs. If the unique characteris-
tics of each study setting result in study-specific estimates,
however, the assumption that this variability represents either random or systematic error would
be itself an error, making the pooled estimate of effect
misleading (Belling, 2000). Until we are able to concretely model that are rich enough to permit
exploration of the processes that produce study-spe-
cific effects, we will continue to identify main effects
that at best, are poor summaries of the true dose-effect
relationships.

REFERENCES

Acoado e- Whitman R, Cole J, Role u- Antis and phobia: a
Anderst F, Fawcitt H, Kampmoller K, Toumo-Johannson T, 
Bezemmen J, Halko E, Liina S, RankoE K, Arnes E, 
Kivisilm S, Pisa J, Mattson C, Risman E, Tenhunen T. 
Psychiatric disorders in 8-9-year-old children based on
a diagnostic interview with the parents. Eur Child Adolesc
and statistical manual of mental disorders, 4th ed. Ameri-
Baghurst PA, McMichael AJ, Wigg NR, Vintushi GV, Robertson 
EE, Roberts RJ, Tong EJ. Environmental exposure to lead and
children's intelligence in the age of seven years. New Engl J
Baghurst PA, McMichael AJ, Tong EJ, Wigg NR, Vintushi GV, 
Robertson EE. Exposure to environmental lead and visual-
motor integration at age 7 years: the Port Fife Cohort Study. 
Beasley A, Burgess J, Goodman C. Polluted landscapes: the
epidemiology of environmental lead exposure in Massachusetts
Baker RL. Nutritional basis of inhaled abuse. Drug Alcohol
Bellingcr DC. Interpreting the literature on lead and child
development: the neglected role of the experimental system.
Neurotoxicol Teratol 1995;17:201-12.
Bellingcr DC. Effect modulation in epidemiologic studies of
low-level neurotoxicant exposures and health outcomes.
Bellingcr DC. Effects of lead on IQ, reading and language skills
Bellingcr DC, Stiles KM, Nemitzman HL. Low-level lead
exposure: intellectual and academic achievement: a long-term
Bellingcr D, Leviton A, Alpern E, Ratmanov M. Pre- and
postnatal lead exposure and behavioral problems in school-aged
Bellingcr DC, Wijd J, Kahn KCK, Rapoport LA, Frickey RJ, 
Wrenovsky S, Sonas R, Newburger JW. Developmental and
neuropsychological sequelae of children at 4 years of age after
heart surgery with cardiopulmonary bypass or low-flow
Bellingcr U, Lichtenberg L, Mills D, Galaburda A, Koenberg JR, 
Bridging cognition, the brain and molecular genetics: evolvu-
Bentovim M. Medical behavioral behavioral: metadecision. In: Feltis 
NR, editor. International Review of research in mental retarda-
Brennan SA. Non-psychological following in normal
screening: issues, methods and findings. Znr Psychol Sci
Breslau NA. Effects of early adrenarche on sex-typed activi-
ties and interests in adolescents with congenital adrenal hyperplasia.
Buice D. Social determinations of illness—a sociocultural
Bukinsky BD, Olney E, Lewis A, Maxfield EF, Fragnoli R. 
Biphasic synaptic plasticity of photopic egression. J Am Acad 
Bundell DJ, Dancet DJ, Heshki K, Ott VA, Byers WK, 
Work-site clinical and behavioral assessment of workers
Bunderson W. Toward an experimental ecology of human
Cahen D, Johnson W, Capurano B, Pau and elevated blood lead
Carroll JF. Commentary on prevalence and impact of parent-
Davidson PW, Myers G, Cox C, Astle D, Shandler C, Simon
Reyes J, Iervasi E, Neulson L, Chiu A, Wang Y, Gurrill M, 
Blissett TW. Effects of prenatal and postnatal methylmercury
exposure from fish consumption on neurodevelopmental out-
comes at 6-35 months of age in the Seychelles Child Development 