1

General Issues in Infant Assessment and Development

LYNN TWAROG SINGER

The ability to compile so large a volume devoted solely to the biobehavioral assessment of newborn and young infants is testimony to the continuing and undiminished growth of interest, research, and knowledge in this area of developmental psychology. By necessity, infant psychometrics and the study of individual differences and normative processes in infancy have developed in a mutually dependent fashion, each furthering and fueling greater advances in the other. As noted by Thompson (1990), in a seminal volume edited by Colombo and Fagen, a primary factor inhibiting developmental psychologists from studying individual differences in infancy has been, in addition to the highly canalized nature of some aspects of infant development, the fact that "behavioral differences among infants are difficult to measure reliably and accurately" (p. 45). The impetus for the compilation of this handbook is simple: to describe and critique some of the various infant assessment instruments that have proliferated in the past two decades, and, with a critical lens, to promote further research and refinement of these assessments.

Attempts to assess neonatal neurological and behavioral status have an extensive history (St. Clair, 1978) derived from a combination of medical and psychological/behavioral orientations. The evaluation of reflexes in a neonatal neurological evaluation had several purposes, including diagnosis of an immediate neurological problem, the evaluation of day-to-day changes concomitant with a pathological process, and the determination of long-term prognosis (Parmelee & Michaelis, 1971). Gesell (1925) expanded on this interest in reflex maturation and developmental progression. His develop-

mental schedules were used to study normal infant mental and motor capacities longitudinally. Description of the orderly developmental sequences of reflexes and their relationships to brain injury underlay the introduction of numerous evaluations of the neonate and young infant which followed, such as those of Prechtl and Andre-Thomas (St. Clair, 1978).

By the 1950s and 1960s, behaviorally oriented examinations became prominent, reflecting the belief that maturational status and neuromotor status were interrelated, and reflecting the burgeoning psychological interest in individual differences in behavior. This period saw the development of neonatal behavioral assessments, such as the Graham–Rosenblith Scale (Rosenblith, 1966) and the Cambridge Newborn Behavioral and Neurological Scales (Brazelton & Freedman, 1971). The latter has since become the most popular and well studied neonatal assessment exam in both medical and psychological research under its more well-known name, the Brazelton Neonatal Behavioral Assessment Scale (see Lester & Tronick, Chapter 18, this volume).

In this same period, the methodological advances achieved by Robert Fantz in reliably assessing neonatal capacity to look at and prefer varying stimuli allowed inferences to be drawn about infant visual perception, discrimination, memory, and cognition. Fantz's methodological achievement initiated generations of innovative studies, many of which have culminated in the assessments described in this volume, including the Teller Acuity Cards (see Mayer & Arendt, Chapter 4, this volume); visual recognition memory tests (see Rose & Orlian, Chapter 13, this volume), and visual habituation measures (see Fagen & Ohr, Chapter 12, this volume).

The assessment of change and continuity in neurobehavioral development, one of the central and most challenging tasks of developmental psychology, continues to be rapidly influenced by data derived from the procedures described earlier and elsewhere in this volume, as well as from a growing understanding of basic neurological development. Sameroff and Chandler (1975) introduced the concept of the transactional model of development which stressed the dynamic nature of the mutual influences of child and environment. Two of the bedrocks of an assessment instrument, stability and predictive validity, have often proved to be elusive psychometric properties in neonatal and infant assessments, and they underlie the lengthy ongoing debate among developmentalists about whether there is continuity in cognitive development (Fagan & Singer, 1983; McCall & Carriger, 1993).

Our understanding of infant development and its assessment has been informed by major advances in neurobiology which have delineated the structure and processes of the long trajectory of brain development initiated prenatally. Brain development occurs in a highly complex, orderly sequence, beginning in the earliest days after conception and continuing postnatally, with continued rapid development until 2–3 years of age and extending well into adolescence (Huttenlocher, 1994). Four major stages of development

have been identified (see Nelson, 1999, and Volpe, 2000, for reviews): (1) neural tube induction, which results in the formation of the brain and spinal cord and occurs in a peak period of the third and fourth weeks of gestation; (2) neuronal proliferation, occurring between 2 and 4 months gestation, in which all neurons and glia, the two major classes of brain cells, are derived; (3) neuronal migration, occurring primarily from 3 to 5 months gestation, in which millions of cells leave their places of origin to be permanently settled; and (4) cell differentiation, which includes cell death (apoptosis), synaptogenesis, and myelination, in which a sheathing substance forms to protect nerve cells and to allow faster transmission of information.

Much cortical development is postnatal due to the growth of neurons and their connections. An overabundance of synaptic connections is produced as functionally unspecific to allow for the development of complex neural circuits. Synapse elimination subsequently occurs from the pruning of these same circuits and is thought to be dependent on experience. This process of pruning is thought to underlie the greater plasticity of the brain in infancy and at "sensitive" or "critical" periods (Huttenlocher, 1994; Nelson, 2000).

The predominant causes of neurological or behavioral deficits are known to occur prenatally (Thoresen, 1999), and disruptions in the processes occurring at any of the stages noted previously can have profound developmental consequences. As an example, delayed myelination refers to an immature pattern of myelination commonly caused by prematurity, and characteristic of children with hydrocephalus (Porter & Tennekoon, 2000). Some studies have demonstrated that the quantity of myelin is associated with the child's level of cognitive functioning (Fletcher et al., 1992), but currently these functional correlates can be reliably ascertained only at older ages. Apparently, with treatment by shunting, age-appropriate myelination patterns can be attained in some children with hydrocephalus (Porter & Tennekoon, 2000). The eventual ability to link such treatment changes to functional behaviors in infancy is dependent on the continued psychometric refinement of the biobehavioral assessments described in this volume and others.

PARALLEL NEUROLOGICAL ASSESSMENTS

Our current understanding of developmental neurobiology has also been markedly advanced by technological tools which allow the study of brain development *in vivo*. Among the most salient are X-ray computed tomography (CT), magnetic resonance imaging (MRI), functional magnetic resonance imaging (fMRI), and positron emission tomography (PET). MRI techniques are noninvasive methodologies which provide information about the structure, metabolism, and function of the developing brain. The earliest

MRI has now been done with preterm infants as young as 24 weeks gestation (Battin et al., 1998), whereas, prior to the development of these techniques, only autopsy studies were possible (Inder & Huppi, 2000). Fetal MRI imaging has also been undertaken, although the safety of conducting MRI on both the pregnant mother and her fetus constitutes a major concern (Rivkin, 2000).

MRI provides spatial resolution of brain structure and avoids many limitations posed by CT scan, especially for the study of infants. These limitations include radiation exposure, which limits CT scan to clinically indicated use, and poorer anatomic resolution (Filipek, Kennedy, & Caviness, 1992). Similarly, PET, which measures regional cerebral blood flow directly, requires injecting subjects with radiopharmaceuticals, also limiting its use to adults and seriously ill children. PET studies have demonstrated that the ascending phase of the synaptogenesis curve is paralleled by a concurrent increase in metabolic activity in cortical areas during the postnatal period (Chugani, Phelps, & Mazziotta, 1987). fMRI, like PET, measures brain function through assessment of metabolic changes which occur when regional brain activity increases but is an indirect measure of blood flow. Practical limitations restrict fMRI to older children, because it requires an alert, cooperative child who can overcome the fear of a strange situation and hold his or her head still (Davidson & Slayter, 2000). Studies of sedated infants have been done, however, with passive presentation of visual or auditory stimuli (Bookheimer, 2000).

Additional MRI variations include diffusion sensor magnetic imaging (DTI), which measures diffusion of water in a particular medium. DTI has been used to study the newborn brain by assessing water diffusion in cerebral white matter. MRI spectroscopy allows quantification of metabolic information by employing standard MRI scanners to measure chemical levels in brain regions.

The cerebral Doppler effect is derived from ultrasound waves emitted by a transducer which can measure cerebral blood flow, a factor in several neonatal pathological conditions and central to the fetal behavioral assessments under development (see DiPietro, Chapter 3, this volume). Single photon emission computed tomography (SPECT) is a radionuclear technique of cerebral perfusion which has also been used to measure cerebral blood flow.

Neonatal and infant brain functioning can also be evaluated through assessment of electroencephalic activity reflected in the electroencephalogram (EEG) through electrodes placed on the infant's scalp in a standard pattern. Neurological integrity as reflected in EEG sleep patterns has recently been a focus of research in neurophysiology and behaviors (Scher, 1997; Thoman, Chapter 7, this volume). EEG patterns have proved useful in a number of the assessments featured in this volume as a measure of emotional regula-

tion, language, and communication (see Molfese & Molfese, Chapter 16, and Marshall & Fox, Chapter 17, this volume).

Evoked potentials provide information about infant sensory perceptual processes linked to neurological conditions. They are averaged electrical responses from an EEG evoked by a stimulus. Auditory brainstem responses (ABRs) measure electrical events occurring within the auditory brainstem, while visual evoked potentials are generated by visual cortex responses to visual stimuli, and somatosensory evoked potentials are generated through peripheral brain pathways (Van Baar, 1998).

Refinement of these techniques represents a growing area of research that has already resulted in greater understanding of neurobiological development. Normative brain development using volumetric MRI techniques will be soon studied through a large multisite initiative of the National Institute of Child Health and Human Development. Collaborative efforts by developmental psychologists to develop the assessments of functional behaviors which correlate with the structural and metabolic characteristics of neurobiological processes to be compiled in such efforts will eventually lead to a true understanding of brain–behavior relationships.

ASSESSMENT

Psychological research generally attempts to measure hypothetical constructs rather than concrete physical properties such as height and weight. A psychological test is "an objective and standardized measure of a sample of behavior" (Anastasi & Urbina, 1997). Traditionally, the function of psychological tests has been to measure individual differences. Depending on the number and nature of items in an assessment, the sample of behaviors assessed may or may not be representative of a construct. One way to "assess" an assessment is to determine whether it meets standards required for use in clinical trials (Cronbach, 1984; Singer, 1997), which cannot afford to be compromised by measurement error.

These requirements include that the assessment be standardized, reliable, and valid. Standardization entails the development and publishing of uniform procedures in administrating and scoring the assessment. For comparability of scores, testing conditions must be the same for all subjects. Detailed directions are a major part of test standardization, extending to the stimuli used, which should be exactly described, time limits, how and when testing of limits of performance is allowed, scoring procedures for test items, preliminary explanations, and examiner training requirements. Included under the rubric of standardization is the establishment of norms, because performance on any test is judged based on empirical data. The standardization sample, a large representative sample for whom the procedure is

designed, is used to establish norms to indicate average as well as deviant performance. Norms allow individual raw scores to be interpreted with reference to the distribution of scores obtained by the standardization sample. These raw scores, when converted into a relative measure, can indicate an individual's relative standing in the normative sample, permitting an evaluation of performance in reference to others, and can provide direct comparisons of an infant's performance across different tests. Standard scores, which indicate the infant's distance from the normative mean relative to the standard deviation of the distribution, can be obtained by transformation of the original raw scores. Standard scores are characteristic of the most highly developed infant assessments, including the Bayley Mental and Motor Scales and the Neurobehavioral Assessment of the Preterm Infant.

Reliability refers to the consistency of a measurement. No assessment has perfect reliability, but the range of fluctuation of an infant's test score due to chance factors limits the extent to which an infant's performance can be considered a reflection of the behavioral characteristic measured versus error variance. Several types of reliability are usually considered, most of which rely on correlation coefficients as a measurement.

Test–retest reliability correlates the scores obtained by an infant on two separate administrations of a test. Test–retest reliability has posed special difficulties for infancy assessments. Most childhood intelligence tests, even in the preschool years, yield remarkably stable scores, achieving correlations in the .70s and .80s (Bornstein & Krasnegor, 1989). In contrast, median values of test–retest reliabilities of visual habituation measures are in the .30 to .40 ranges, and for paired comparison performance measures, in the .20s for single novelty preference tasks. Such levels of reliability do not meet the standards for adult psychometric work (Colombo, 1993). Infant assessments are often criticized for their low stability, with correlations declining at younger ages. The Brazelton scales, as well as visual recognition memory and habituation measures, for example, have all raised concerns about their validity due to the low stability of infant performance (Benasich & Bejar, 1992; Fagan & Singer, 1983).

Several factors introduce considerable variability into many infancy assessments resulting in less than acceptable intertest agreement. These factors include biological fluctuations in the infant's physiological status underlying the behavioral construct under study, as well as the attentional variability of the infant. In defense of the poor internal consistency of many infant assessments, it has been noted that some of these assessments, such as cry analysis and the Neonatal Network Neurobehavioral Scales, are designed to be sensitive to the variable physiological states of the infant. Insofar as such measures reflect the rapidly changing neurobiological status of the infant in interaction with the environment, many infant assessments would be expected to have low week-to-week or even day-to-day stability.

Additional reliability data useful to understanding an assessment in-

clude interrater agreements, and intrarater consistency, because examiner characteristics may differentially affect infant performance on some tasks compared to others. Enthusiastic, warm, or gentle examiners may elicit better performance on a neurobehavioral scale with a neonate. The Brazelton (1973) scale, for example, has been demonstrated to be quite susceptible to examiner effects (Richardson, Day, & Taylor, 1989). Even when the infant's actual behavioral output can be judged with high reliability, however, contextual factors may influence infant cooperation. When training examiners for a multisite clinical trial to use the Fagan Test of Infant Intelligence (Fagan & Singer, 1983), for example, many previously unspecified variables needed to be delineated in written form to ensure examiners' uniformity of procedure. Variables such as whether to use a pacifier during the assessment or obtaining infant attention by tapping on the apparatus, which had been incorporated into the procedure, needed specification.

Test validity entails the degree to which an assessment measures what it purports to measure. Construct validity specifies what the test measures and is undergirded by both content and predictive validation (Cronbach & Meehl, 1955).

The construct validity is the extent to which a test measures a theoretical construct or trait. Constructs explain and organize observed response consistencies. Evidence for construct validation is derived from data on the attributes of the construct, including developmental changes, relationship to other measures, and convergent or discriminant relationships with variables with which it should reasonably be related or unrelated (Anastasi & Urbina, 1997).

Predictive validity measures the extent to which an assessment "predicts" performance on an assessment of the same trait over a long period. Because intelligence is a fairly stable construct after the preschool ages and has been so well studied as a psychological construct measured by IQ tests, it may have been a natural progression for psychologists to perceive assessments of infant developmental differences as precursors of childhood IQ. However, decades of follow-up research have indicated that except at the lowest ranges of scores, standardized tests of infant development given prior to 18 months of age have poor predictive validity for later intellectual functioning (Fagan & Singer, 1983; McCall & Mash, 1997) and are close to zero.

This conclusion led to debate about the nature of intelligence, that is, whether or not a unitary construct of intelligence, or *g*, existed and whether it was continuous or discontinuous from infancy. An alternative rationale for the lack of predictive validity of infant developmental assessments was that infant developmental assessments such as the Bayley scales were measuring the wrong behaviors, namely, sensorimotor skills which bear little relationship to higher-order cognitive skills important to school-age learning. Thus, more discrete measures of attention and information processing have

been examined with the hope that they would be more reflective of the attentional, discrimination, memory, and speed of information-processing skills characteristic of later intelligence. Although still evolving, tests of visual information processing in infancy, such as recognition memory, habituation, and attentional duration, have been evaluated in longitudinal studies. Predictive validity of such measures has ranged from .30 to .40 across a variety of studies and populations, suggesting that such measures may be useful in identification of infants at risk for later developmental delays (Bornstein & Sigman, 1986; Colombo, 1997; Fagan & Singer, 1983; McCall & Carriger, 1993). Indeed, in a recent study, Rose and Feldman (1997) found evidence that speed of processing and memory accounted for the relationships that they found from 7-month visual recognition memory and 1-year cross-modal transfer to 11-year IQ. It should be recognized, however, that even with specific processing measures, prediction for an individual infant (i.e., sensitivity and specificity) is still poor, and that other measures, such as parental education, bear greater relationships to child intellectual outcome than any infant assessment.

CURRENT APPLICATIONS

There is a pressing need for the further refinement and development of infant biobehavioral assessments, many of which remain in rudimentary or experimental forms. The implementation of the Education for All Handicapped Children Act Amendments (later retitled the Individuals with Disabilities Education Act), calls for early intervention services for all handicapped infants and their families (Shonkoff & Meisels, 1990; Singer, Minnes, & Arendt, 1999). Eligible for services are infants experiencing developmental delay and those with a diagnosed condition with "high risk" of developmental delay. However, identifying which infants are at risk remains a difficult problem. It is well documented that children younger than school age are highly unlikely to be identified for intervention, as there is a tenfold disparity between prevalence rates of handicapped children from infancy to school age (Meisels & Wasik, 1990). Other explanations for the differences in earlier and later age prevalence rates have been attributed to the absence of the wide range of reliable, valid assessments available for identification of delays at younger ages. Because of the poorer predictive validity of standard global assessments such as the Bayley scales, screening devices employing both psychometric and psychophysiological perspectives have been recommended. But, as yet, although there are a large number of experimental assessments in development, as described earlier, there are none with adequate standardization and predictive properties to be appropriated for large-scale screening efforts (Meisels & Wasik, 1990).

Similarly, Public Law 99-457 paved the way for the need for outcome measures to assess the effects of the early-infancy intervention programs initiated by these amendments. Global outcome measures, such as the mental or motor indices of the Bayley scales, are broad based and represent the conglomerate of a number of diverse infant skills; thus the overall score may mask benefits from intervention on a discrete area of functioning. Moreover, because such outcome measures lack predictive validity in infancy, they provide little information for the researcher about the underlying developmental processes affected by the intervention which resulted in a positive or negative outcome.

Societal concern for and public health focus on the growing numbers of high-risk infants have also contributed to an interest in developing assessments of brain–behavior relationships that can be used as early in life as possible. Currently, about 53,000 very-low-birthweight infants (< 1,500 g) are born annually in the United States, with follow-up studies indicating that as many as 50% will need some type of supportive educational services at school age, and that for some subgroups, such as those with bronchopulmonary dysplasia, 25% will have significant sensory deficits and/or mental or motor retardation (Singer, Yamashita, Lilien, Collin, & Baley, 1997). A landmark randomized clinical intervention trial for low-birthweight infants, the Infant Health and Development Project, found that higher-weight infants in this trial had significant cognitive gains from the intensive intervention at 3 years of age, but very-low-birthweight infants derived no benefit (Brooks-Gunn et al., 1994). It is likely that prematurity leads to changes in brain organizational development which affect learning and developmental outcome. Many of the assessments of sensorimotor capacities and information-processing skills described in this volume have the potential to identify the sensory deficits and behavioral characteristics specifically affected by very-low-birthweight birth and its medical and neurological correlates at earlier ages. Earlier assessment could then lead to environmental interventions specifically directed at remediation of these deficits.

The emergence of behavioral teratology as a distinct area of developmental and neuropsychological interest has also provided impetus for the development of infancy assessments (Fein, Schwartz, Jacobson, & Jacobson, 1983; Voorhees, 1986). This field has flourished concurrent with greater understanding of principles of fetal neurobiological development and the scientific demonstration that many toxic substances could cross the placental barrier to affect fetal development. Teratology was originally conceptualized as the study of gross structural malformations seen after birth and thought to be related to fetal developmental disturbances (e.g., the major limb defects attributable to maternal ingestion of thalidomide as an antinausea drug during pregnancy). This definition is expanded on in behavioral teratology to include a range of possible teratogenic effects on the developing nervous system resulting in behavioral disabilities. Neuropsychological impairments

are now recognized as possible outcomes of toxic fetal exposures even in the absence of gross physical impairments. A wide range of teratogens have now been recognized as causing subtle to significant neurobehavioral disabilities. Among the most prominent are lead (Needleman & Bellinger, 1991), alcohol (Streissguth, Barr, & Sampson, 1990), nicotine and marijuana (Fried, Watkinson, & Gray, 1992), vitamin A (Adams & Lammer, 1995), antiepileptic drugs (Leavitt, Yearby, Robinson, Sells, & Erickson, 1992), and polychlorinated biphenyls (Jacobson, Fein, Jacobson, Schwartz, & Dowler, 1985).

In the early 1990s, in response to widespread public alarm and scientific concern about the epidemic use of crack cocaine by pregnant women in the United States, the National Institute on Drug Abuse initiated funding of a number of longitudinal studies of the development of cocaine-exposed infants (Smeriglio & Wilcox, 1999), many of which incorporated novel or experimental neonatal and infant behavioral assessments as outcomes. These large-scale, methodologically rigorous studies are just beginning to yield findings about the developmental correlates of fetal cocaine and polydrug exposure, but they promise also to provide much information about the psychometric properties of these measures. To date, these studies have examined numerous neonatal and infant behaviors in relation to drug exposure, including habituation (Mayes, Bornstein, Chawarska, & Granger, 1995), neonatal visual preferences (Singer, Arendt, et al., 1999), visual recognition memory (Jacobson, Jacobson, Sokol, Martier, & Chiodo, 1996), sleep EEG (Scher, Richardson, & Day, 2000), neonatal neurobehavioral assessments (Eyler, Behnke, Conlon, Woods, & Wobie, 1998; Richardson, Hamel, Goldschmidt, & Day, 1996; Singer, Arendt, Minnes, Farkas, & Salvator, 2000), and infant movement assessment (Arendt, Singer, Angelopolous, Busdieker, & Mascia, 1998). In addition to the rise in concern about cocaine, increasing numbers of pregnant and postpartum women are receiving pharmacological treatments for depression, a problem which affects about 10% of pregnant women. Data are lacking about the effects of such drugs on fetal development and infant behavioral outcomes (Wisner, Gelenberg, Leonard, Zarin, & Frank, 1999).

An interesting corollary to the concept of behavioral teratogenicity lies in the also relatively recent recognition that nutrient deficiencies prenatally or during infancy can have long-term developmental consequences. Folic acid deficiency, for example, was found to be a causal factor in the occurrence of neural tube disorders such as spina bifida (Wald, 1994). Iron deficiency anemia has been demonstrated to affect infant behavior negatively (Lozoff et al., 1987). Biochemical studies have found that certain long-chain polyunsaturated fatty acids (LCPUFAs) (which are found in human breast milk but not infant formulas) are essential components of infant nutrition. Citing studies which show breast-fed infants to have better intellectual functioning than formula-fed infants, some have proposed that adding

LCPUFAs to infant formulas would be beneficial to infant visual and cognitive functioning (see Dobbing, 1997, for comprehensive review). This proposal has led to a large number of research studies to assess the relationship of LCPUFAs to infant mental, motor, and visual development, including a number of nutritional formula industry-led randomized clinical trials to address efficacy and safety issues of adding these nutrients to infant formulas (Auestad et al., 2001). What is of interest is (1) the notion that in the absence of any known harmful effects to formula-fed infants, these additives might provide some benefit to infants' development, and (2) the use of behavioral assessments as outcomes of these studies, with extensive consultation from developmental psychologists on which functional end points should be examined and which assessments in infancy had adequate psychometric properties to be used (Dobbing, 1997).

Thus, it is clear that recent developments in our understanding of fetal and infant neurobiological development provide impetus for the tremendous research and public health interests in interventions for the developing infant. Accurate measurement of infant biobehavioral status is needed for diagnosis, evaluation of process changes, and long-term prognosis of at-risk infants, a goal which will be reached only through further refinement of specific and global assessments of infant functioning.

ACKNOWLEDGMENTS

This work was supported by Grant No. MCJ 39071 from the Maternal Child and Health Services Bureau (Title V, Social Security Act), Health Resources and Services Administration, and Grant No. RO1-07957 from the National Institute on Drug Abuse.

REFERENCES

Adams, J., & Lammer, E. J. (1995). Human isotretinoin exposure. The teratogenesis of a syndrome of cognitive deficits. *Neurotoxicology and Teratology, 17*, 386–392.

Anastasi, A., & Urbina, S. (1997). *Psychological testing* (7th ed.). Upper Saddle River, NJ: Prentice-Hall.

Arendt, R. E., Singer, L. T., Angelopoulos, J., Busdieker, O., & Mascia, J. (1998). Sensory motor development in cocaine-exposed infants. *Infant Behavior and Development, 21*, 627–640.

Auestad, N., Halter, R. A., Hall, R. T., Blatter, M., Bogle, M., Burks, W., Erikson, J. R., Fitzgerald, K. M., Dobson, V., Innis, S. M., Singer, L. T., Montalto, M. B., Jacobs, J. R., Qui, W., & Bornstein, M. H. (in press). Growth and development in term infants fed long-chain polyunsaturated fatty acids: A double-masked, randomized, parallel, prospective, longitudinal, multivariate study. *Pediatrics*.

Battin, M., Maalouf, E. F., Counsell, S., Herlihy, A., Hall, A., Azzopardi, D., & Edwards, A. D. (1998). Physiological stability of preterm infants during magnetic resonance imaging. *Early Human Development, 52*(2), 101–110.

Benasich, A., & Bejar, I. T. (1992). The Fagan Test of Infant Intelligence: A critical review. *Journal of Applied Developmental Psychology, 13*, 153–171.

Bookheimer, S. (2000). Methodological issues in pediatric neuroimaging. *Mental Retardation and Developmental Disabilities Research Review, 6*, 161–166.

Bornstein, M. H., & Krasnegor, N. A. (Eds.). (1989). *Stability and continuity in mental development: Behavioral and biological perspectives*. Hillsdale, NJ: Erlbaum.

Bornstein, M. H., & Sigman, D. (1986). Continuity in mental development from infancy. *Child Development, 57*, 251–274.

Brazelton, T. (1973). *Neonatal Behavioral Assessment Scale*. Philadelphia: Lippincott.

Brazelton, T., & Freedman, D. (1971). Manual to accompany newborn behavioral and neurological scales. In G. Stoeling & J. VanDerWerffTenBosch (Ed.), *Normal and abnormal development of brain and behaviour*. Baltimore: Williams & Wilkins.

Brooks-Gunn, J., McCarton, C., Casey, P. H., McCormick, M. C., Bauer, C. R., Bernbaum, J. C., Tyson, J., Swanson, M., Bennett, F. C., Scott, D. T., et al. (1994). Early intervention of low-birthweight-premature infants: Results through age 5 years from the Infant Health and Development Program. *Journal of the American Medical Association, 272*, 1257–1262.

Chugani, H. T., Phelps, M. E., & Mazziotta, J. C. (1987). PET study of human brain functional development. *Annals of Neurology, 22*, 487–497.

Colombo, J. (1993). *Infant cognition: Predicting later intellectual functioning*. Newbury Park, CA: Sage.

Colombo, J. (1997). Individual differences in infant cognition. In J. Dobbing (Ed.), *Developing brain and behaviour: The role of lipids in infant formula* (pp. 339–372). San Diego: Academic Press.

Cronbach, L. J. (1984). *The essentials of psychological testing*. New York: Harper & Row.

Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin, 52*, 281–302.

Davidson, R. J., & Slayter, H. A. (2000). Probing emotion in the developing brain: Functional neuroimaging in the assessment of the neural substances of emotion in normal and disordered children and adolescents. *Mental Retardation and Developmental Disabilities Research Reviews, 6*, 166–170.

Dobbing, J. (Ed.). (1997). *Developing brain and behaviour: The role of lipids in infant formula*. San Diego: Academic Press.

Eyler, F. D., Behnke, M., Conlon, M., Woods, N. S., & Wobie, K. (1998). Birth outcome from a prospective, matched study of prenatal crack/cocaine use: II. Interactive and dose effects on neurobehavioral assessment. *Pediatrics, 101*, 237–241.

Fagan, J. F., & Singer, L. T. (1983). Infant recognition memory as a measure of intelligence. In L. Lipsitt (Ed.), *Advances in infancy research* (Vol. II, pp. 31–78). Norwood, NJ: Ablex.

Fein, G. G., Schwartz, P. M., Jacobson, S. W., & Jacobson, J. L. (1983). Environmental toxins and behavioral development: A new role for psychological research. *American Psychologist, 38*(11), 1188–1197.

Filipek, P. A., Kennedy, D. N., & Caviness, V. S. (1992). Neuroimaging in child neuropsychology. In I. Rapin & S. J. Segalowitz (Eds.), *Handbook of neuropsychology: Vol. 6. Child neuropsychology* (pp. 301–329). New York: Elsevier.

Fletcher, J. M., Bohan, T. P., Brandt, M. E., Brookshire, B. L., Beaver, S. R., Francis, D. J., Davidson, K. C., Thompson, N. M., & Miner, M. E. (1992). Cerebral white matter and cognition in hydrocephalic children. *Archives of Neurology 49*, 818–824.

Fried, P. A., Watkinson, B., & Gray, R. (1992). A follow-up study of attentional behavior in 6 year old children exposed prenatally to marijuana, cigarettes, and alcohol. *Neurotoxicology and Teratology, 14,* 299–311.

Gesell, A. (1925). *The mental growth of the preschool child.* New York: Macmillan.

Huttenlocher, P. R. (1994). Synaptogenesis in human cerebral cortex. In G. Davison & K. A. Fischer (Eds.), *Human behavior and the developing brain* (pp. 137–152). New York: Guilford Press.

Inder, T. E., & Huppi, P. S. (2000). *In vivo* studies of brain development by magnetic resonance techniques. *Mental Retardation and Developmental Disabilities Research Reviews, 6,* 59–67.

Jacobson, S. W., Fein, G. G., Jacobson, J. L., Schwartz, P. M., & Dowler, J. K. (1985). The effect of PCB exposure on visual recognition memory. *Child Development, 56,* 853–860.

Jacobson, S. W., Jacobson, J. K., Sokol, R. J., Martier, S. S., & Chiodo, L. M. (1996). New evidence of neurobehavioral effects of *in utero* cocaine exposure. *Journal of Pediatrics, 129*(4), 581–588.

Leavitt, A. M., Yearby, M. S., Robinson, N., Sells, C. J., & Erikson, D. M. (1992). Epilepsy in pregnancy. Developmental outcome of offspring at 12 months. *Neurology, 42,* 141–143.

Lozoff, B., Brittenham, G. M., Wolf, A. W., McClish, D. K., Kuhnert, P. M., Jimenez, E., Jimenez, R., Mora, L. A., Gomez, I., & Krauskoph, D. (1987). Iron deficiency anemia and iron therapy: Effects on infant developmental test performance. *Pediatrics, 79,* 981–985.

Mayes, L. C., Bornstein, M. H., Chawarska, K., & Granger, R. H. (1995). Information processing and developmental assessments in 3 month old infants exposed prenatally to cocaine. *Pediatrics, 95,* 539–545.

McCall, R. B., & Carriger, N. S. (1993). A meta analysis of infant habituation and recognition memory performance as predictors of later IQ. *Child Development, 64,* 57–79.

McCall, R. B., & Mash, C. W. (1997). Long-chain polyunsaturated fatty acids and the measurement and prediction of intelligence (IQ). In J. Dobbing (Ed.), *Developing brain and behaviour: The role of lipids in infant formula* (pp. 295–338). San Diego: Academic Press.

Meisels, S. I., & Wasik, B. (1990). Who should be served? Identifying children in need of early intervention. In S. J. Meisels & J. Shonkoff (Ed.), *Handbook of early intervention* (pp. 605–632). New York: Cambridge University Press.

Needleman, H. L., & Bellinger, D. (1991). The health effects of low level lead. *Annual Review of Public Health, 12,* 111–140.

Nelson, C. A. (1999). Change and continuity in neurobehavioral development: Lessons from the study of neurobiology and neural plasticity. *Infant Behavior and Development, 22,* 415–429.

Nelson, C. A. (2000). Neural plasticity and human development: The role of early experiences in sculpting memory systems. *Developmental Science, 3,* 115–136.

Parmelee, A., & Michaelis, R. (1971). Neurological examination of the newborn. In J. Hellmuth (Ed.), *Exceptional infant* (Vol. 2). New York: Brunner/Mazel.

Porter, B. E., & Tennekoon, G. (2000). Myelin and disorders that affect the formation and maintenance of this sheath. *Mental Retardation and Developmental Disabilities Research Reviews, 6,* 47–58.

Richardson, G. A., Day, N. L., & Taylor, P. M. (1989). The effect of prenatal alcohol, mari-

juana, and tobacco exposure on neonatal behavior. *Infant Behavior and Development, 12*, 199–209.

Richardson, G. A., Hamel, S. C., Goldschmidt, L., & Day, N. L. (1996). The effects of prenatal cocaine use on neonatal neurobehavioral status. *Neurotoxicology and Teratology, 18*, 519–528.

Rivkin, M. (2000). Developmental neuroimaging of children using magnetic resonance techniques. *Mental Retardation and Developmental Disabilities Research Reviews, 6*, 68–80.

Rose, S. A., & Feldman, J. F. (1997). Memory and speed: Their role in the relation of infant information processing to later IQ. *Child Development, 68*(4), 630–641.

Rosenblith, J. (1966). Prognostic value of neonatal assessment. *Child Development, 37*, 623–631.

Sameroff, A., & Chandler, M. (1975). Reproductive risk and the continuum of caretaking casualty. In F. Horowitz, M. Hetherington, S. Scarr-Salpatek, & G. Siegel (Eds.), *Review of child development research* (Vol. 4, pp. 187–244). Chicago: University of Chicago Press.

Scher, M. S. (1997). Neurophysiological assessment of brain function and maturation. I. A measure of brain adaptation in high risk infants. *Pediatric Neurology, 16*, 191–198.

Scher, M. S., Richardson, G., & Day, N. L. (2000). Effects of prenatal cocaine/crack and other drug exposure on electroencephalographic sleep studies at birth and one year. *Pediatrics, 105*(1), 39–48.

Shonkoff, J., & Meisels, S. J. (1990). Who should be served?: Identifying children in need of early intervention. In S. J. Meisels & J. Shonkoff (Eds.), *Handbook of early intervention* (pp. 605–632). New York: Cambridge University Press.

Singer, L. T. (1997). Methodological considerations in longitudinal studies of infant risk. In J. Dobbing (Ed.), *Developing brain and behaviour: The role of lipids in infant formula* (pp. 210–230). San Diego: Academic Press.

Singer, L. T., Arendt, R. E., Fagan, J. F., Minnes, S., Salvator, A., Bolek, T., & Becker, M. (1999). Neonatal visual information processing in cocaine-exposed and non-exposed infants. *Infant Behavior and Development, 22*(1), 1–15.

Singer, L. T., Arendt, R. E., Minnes, S., Farkas, K., & Salvator, A. (2000). Neurobehavioral outcomes of cocaine-exposed infants. *Neurotoxicology and Teratology, 22*, 653–666.

Singer, L. T., Minnes, S., & Arendt, R. E. (1999). Innovations for high risk infants. In D. E. Biegel & A. Blum (Eds.), *Innovations in practice and service delivery across the lifespan* (pp. 57–78). New York: Oxford University Press.

Singer, L. T., Yamashita, T., Lilien, L., Collin, M., & Baley, J. (1997). A longitudinal study of developmental outcome of infants with bronchopulmonary dysplasia and very low birth weight. *Pediatrics, 100*(6), 987–993.

Smeriglio, V., & Wilcox, H. C. (1999). Prenatal drug exposure and child outcome. Past, present, future. *Clinical Perinatology, 26*(1), 1–16.

St. Clair, K. (1978). Neonatal assessment procedures: A historical review. *Child Development, 49*, 280–292.

Streissguth, A., Barr, H. M., & Sampson, P. D. (1990). Moderate prenatal alcohol exposure effects on child IQ and learning problems at 7½ years. *Alcoholism: Clinical and Experimental Research, 14*, 662–669.

Thompson, L. (1990). Genetic contributions to early individual differences. In J. Colombo & J. Fagen (Eds.), *Individual differences in infancy: Reliability, stability, and prevention* (pp. 45–76). Hillsdale, NJ: Erlbaum.

Thoresen, M. (2000). Protecting the perinatal brain. *Seminars in Neonatology, 5*, 1–2.

Van Baar, A. (1998). Evaluation of the human newborn infant. In W. Slikker & L. Chang (Eds.), *Handbook of developmental neurotoxicology* (pp. 439–454). New York: Academic Press.

Volpe, J. (2000). Overview: Normal and abnormal human brain development. *Mental Retardation and Developmental Disabilities Research Reviews, 6,* 1–5.

Vorhees, C. V. (1986). Principles of behavioral teratology in E. P. Riley & C. V. Vorhees (Eds.). *Handbook of behavioral teratology* (pp. 23–48). New York: Plenum.

Wald, N. J. (1994). Folic acid and neural tube defects. In G. Bock & J. Marsh (Eds.), *Ciba Foundation Symposium 8: Neural tube defects* (pp. 70–89). London: Wiley.

Wisner, K. L., Gelenberg, A. J., Leonard, H., Zarin, D., & Frank, E. (1999). Pharmacologic treatment of depression during pregnancy. *Journal of the American Medical Association, 282,* 1264–1269.