The assessment of cognition/intelligence in infancy

By: Terri L. Shelton

Shelton, T.L. (1989). The assessment of cognition/intelligence in infancy. *Infants and Young Children, 1*(3), 10-25.

Abstract:
This article will review the development of infant cognitive assessment and will describe selected tests. Considerations in choosing, administering, and interpreting the results of infant intelligence/cognitive assessment instruments will be outlined. Finally, the usefulness of cognitive assessment will be discussed along with new approaches to assessment.

Article:
The cognitive assessment of infants, like intelligence testing in general, has its roots in the intelligence testing movement of the late 19th and early 20th centuries. Since that time interest in the development and use of infant cognitive assessment instruments has mushroomed. The passage of Public Law 99- 457 in October of 1986, the Education of the Handicapped Act Amendments, established early intervention services for infants from birth through 3 years of age in all states. This legislation also has contributed to an even greater interest in identifying techniques that can be used to document developmental delay, to identify infants at risk, to plan intervention services, and to evaluate their effectiveness.

In light of this growing interest, this article will review the development of infant cognitive assessment and will briefly describe some of the tests in current use. Basic considerations in choosing and administering infant intelligence/cognitive assessment instruments will be discussed along with a review of factors that affect the interpretation of the results of these evaluations. Finally, the usefulness of cognitive assessment will be discussed, including the use of information in an interdisciplinary team, the predictive validity of tests, and new approaches to assessment.

HISTORIC OVERVIEW

Psychometric instruments

Arnold Gesell' was the pioneer of infant assessment. Although most infant cognitive assessment instruments that were later developed included many of the items developed by Gesell, his primary purpose in developing the scales was to identify those infants in need of assistance in the area of welfare and hygiene. The initial scales, published in 1925,[1] consisted of 144 items in four general areas: motor, language, adaptive, and personal/social behavior. The purpose of the scales was broad and clinical. An updated version of the scales was published in 1974 by Knobloch and Pasamanick[2]; it provided more adequate norms for infants from 1 month to 5 years of age. Although Gesell viewed the environment as important, he thought that development was primarily influenced by genetics and the maturational process. While the scales do not yield a specific level of cognitive abilities, but rather developmental quotients in the different areas, the Gesell scales provided the basis for many other assessment instruments, such as the Brunet-Lezine Scale,[3] published in 1951 and largely used in Europe, and the Griffiths Scale of Mental Development, published in 1954.[4]

In contrast to Gesell's broad conceptualization of development, Canal[5] specifically set out to develop a

standardized assessment of infant cognitive abilities that could be a downward extension of the Stanford-Binet. While Cattell used many of the items from the Gesell, she eliminated those thought to be unduly influenced by home training or large muscle control. She made the scoring more objective and standardized the test for infants from birth through 30 months. Like Gesell, the theoretical underpinnings of the test seemed to emphasize the inherent abilities of the infant and to diminish the role of the environment.

Bayley[6,7] also used the Gesell schedules as a starting point in her first California First Year Mental Scale, published in 1933. In its present form the scales include both mental and motor scales and a behavior record for evaluating infants aged 1 month to 30 months. The Bayley scales are perhaps the most widely used of the formal infant intelligence scales, and there is a large body of research documenting the reliability and validity of the scales. Although initially the development of the scales seemed to reflect a more unitary view of intelligence, over time Bayley's concept of intelligence began to change. Intelligence, as measured by these scales, seemed to be viewed as emergent throughout infancy and functionally unique at different ages. In fact, Kohen-Raz's[8] scalogram analysis of the Bayley scales yields five scales that are very similar to the scales typically found on Piagetian-based sensorimotor instruments.

Sensorimotor scales

With the translation of the work of Piaget into English came a new theoretical approach to the study of infant development and to the assessment of cognition/intelligence in infants as well. The previous approaches by Bayley, Cattell, and Gesell assumed that intelligence was more or less a unitary trait and that development was nonhierarchical. In contrast, Piagetian theory viewed development as a series of hierarchical, qualitatively different stages, dependent on the infant's interaction with his or her environment.

Three major assessment scales based on Piagetian theory have been developed: the Casati-Lezine Scale,[9] the Albert Einstein Scales of Sensorimotor Development by Corman and Escalona,[10] and the Infant Psychological Development Scales (IPDS) by Uzgiris and Hunt.[11] The Casati-Lezine test measures object search, the use of intermediaries, object exploration, and the combination of objects. The Einstein scales are designed for infants between 1 month and 2 years of age and assess skills in prehension, object permanence, and functioning in three-dimensional space. The IPDS is appropriate with infants aged 1 month to 2 years and is comprised of six scales: (1) visual pursuit and permanence of objects, (2) means of obtaining desired environmental events, (3a) vocal imitation, (3b) gestural imitation, (4) operational causality, (5) construction of object relations in space, and (6) schemes for relating to objects. In contrast to non-Piagetian-based scales, no developmental quotients or deviation IQ scores are obtained. Rather, infants are characterized with respect to their most advanced performance on each subscale. The IPDS appears to be as reliable and valid as the other psychometric approaches. Although the IPDS has not really been standardized, Uzgiris and Hunt' note that the various levels of cognitive organization are of psychologic significance in their own right and "need not be based on the individual's comparative status in a statistical distribution."[11px]

Concurrent with the development of Piagetian-based instruments in the 1960s and 1970s was the development of assessment instruments in several other areas: (1) neonatal assessment, (2) screening instruments, (3) assessment of skills for special populations, and (4) assessment-based programming instruments. Clearly the impetus for many of these instruments has been and will continue to be the growing interest in early intervention techniques, the need to identify infants at risk, and the survival of infants born prematurely or with disabilities. A brief description of selected and frequently used instruments in each of these categories appears in Appendix A. Appendix B is a checklist for assessing infant cognition. Further accounts of the historic development of infant assessment as well as current trends can be found elsewhere.[12-14]

BASIC CONSIDERATIONS IN INFANT ASSESSMENT

Purpose

The assessment of infant cognition or intelligence can be useful in a variety of ways: to document delay or risk status, to design an intervention program, to provide valuable pre-treatment and posttreatment information for research or treatment regarding the effectiveness of a therapeutic intervention, or to predict future abilities.

Therefore it is extremely important that an examiner first identify the function or purpose of the assessment. The purpose will determine not only the type of test (eg, screening, in-depth assessment) but also the necessary test characteristics (eg, time needed, reliability). Once the examiner's purpose is clear, it also is important to examine the purpose for which the assessment instrument was designed and to try to match these as much as possible. As measurement often deals with intangible variables that are not always directly observable, constructs such as cognition, intelligence, and personality must be measured indirectly. Each test author makes theoretical assumptions about this variable or construct. Information on the theoretical underpinnings of the particular test as well as other test characteristics, such as item selection and reliability and validity estimates, should be available in the test manual's description of the construction of the test.

Skills of the examiner

In choosing an assessment instrument, the examiner should evaluate whether he or she has the necessary training and experience to administer the instrument. Because of the special demands inherent in evaluating infants, examiners should have experience testing very young children and infants in general. For example, it is important to know techniques for eliciting the best performance from the infant, to make the family feel comfortable, and to know when to interrupt or to discontinue testing. In addition to general experience, experience with specific populations is equally important to recognize the limitations of infant assessment, to interpret the test results, and to use the information to assist caregivers or for intervention. For example, specific experience in evaluating infants with cerebral palsy or with sensory impairments and in interpreting test results is critical, particularly in light of the fact that few existing infant assessment techniques have been developed specifically for infants with these disabilities.

Familiarity with the specific assessment instrument is required as well. Because the attention span and interest level of infants are brief, a good working knowledge of the assessment instrument can speed the administration of the items, maximizing the chance that the infant's optimal performance is obtained. Many of the assessment instruments specify the level of expertise needed. For example, the Bayley Scales of Infant Development[7] should not be administered by anyone who is not trained in developmental or intellectual testing. In contrast, the Hawaii Early Learning Profile (HELP)[15] can be administered by an early interventionist who needs no particular training in test construction or developmental testing per se. At a basic level, however, all examiners should know the limitations of the testing instrument, how representative the infant's performance was, and how well they administered the test so that these factors can be considered in the interpretation of results.

The professional discipline of the examiner can vary. There are psychologists, nurses, early interventionists, special educators, among others, who can be competently trained examiners of infant cognitive abilities. However, the ability to interpret the results and to communicate these findings to the infant's parents and to other professionals is as important as the ability to choose and to administer the test. Because of these reasons it is often advisable to choose a developmental or clinical psychologist or other professional with specific knowledge of test construction and limitations of intellectual testing and with the clinical skills necessary to interpret and relate in a clear but compassionate way what may be distressing news to the family. Because the results of a cognitive assessment are important for other developmental areas, the examiner should also have a working knowledge of other develop-mental areas so that the results can be communicated to other professionals, such as a physical therapist or physician, in a meaningful way.

Standardization sample

In addition to matching the purpose of the assessment to the theoretical underpinnings of the test itself, it is important to determine whether the characteristics of the child being evaluated are similar to those of the group for whom the test was developed or standardized. Standardization is a mechanism for attaching meaning to raw scores and for making comparisons between and among individuals. Sometimes it is not possible to find an assessment instrument whose standardization sample is comparable to the characteristics of the infant being evaluated, especially when evaluation of infants with chronic illnesses or multiple disabilities is becoming more common. Nevertheless, when these infants are being evaluated, this factor should enter into the interpretation of the results.

Performance
Another consideration is the type of performance one is interested in obtaining. Cronbach[16] distinguished between measures of maximum or best performance and measures of typical performance. For example, the Brazelton Neonatal Behavioral Assessment Scale[17] is designed to examine the infant's best performance, and specific conditions are outlined that must be met before the infant's performance is considered optimal. Measures of typical performance, such as tests of personality, habits, or interests, are designed to examine an individual's usual, representative, or typical behavior.

Psychometric issues
Psychometric issues, such as reliability and validity, should enter into the choice of an assessment instrument. Reliability refers to the consistency and accuracy of measurement, that is, what proportion of an infant's score can be attributed to chance or error and what proportion is a "true score." There are a number of ways to estimate reliability (eg, test-retest, alternate forms, split half, and coefficient alpha). For example, if an infant is to be reevaluated in the near future, the test-retest reliability of a particular instrument would be an important consideration. A test with good test-retest reliability would give the examiner greater confidence that the results obtained on the second evaluation were a valid representation of the infant's abilities and not reflective of error or chance fluctuations in the test itself. All estimations of reliability are reported in the form of a correlation coefficient (eg, Pearson r) and range from 0 to 1.00. The closer a reliability coefficient is to 1.00 the better. However, because some degree of chance or error enters into every assessment, a reliability of 1.00 is rarely obtained. Reliabilities between .70s and .90s indicate that the test is generally stable and accurate.

While good reliability is an essential characteristic of an assessment instrument, the information it provides is useless if the test lacks validity. Validity is the relationship between the variable or trait being measured and the procedures used to measure it. In other words, does the test measure what it says it does? There are three major types of validity: content, criterion-related, and construct. Content validity is whether the test covers the material it is supposed to cover. Criterion-related validity, such as concurrent or predictive validity, relates to the comparison of test scores with some external variable or "criterion" known or believed to measure the same variable or attribute.

Similar to reliability, the correlation coefficient is the usual index of criterion validity. Unlike the reliability coefficient, the value of the validity coefficient must be squared to determine the percentage of the infant's score that is due to chance fluctuations and the percentage thought to be a true reflection of the infant's abilities. For example, suppose the relationship or correlation between an infant's score on a test of cognitive abilities at 12 months of age and an intellectual assessment at 6 years of age is .70. Squaring .70 results in .49. This number indicates the proportion of variance shared in common by the two tests. Thus 49% of the infant's score is believed to be based on the infant's cognitive abilities; 51% is attributable to error, chance, plus the effects of some other undetermined variables.

How high a validity coefficient should be will vary. Obviously the higher the validity coefficient the more confidence the examiner can place on the results of the test being an accurate reflection of the infant's abilities. Given the choice between two measures where everything else is equal, it would be better to choose the one with the higher criterion validity.

Construct validity examines the meaning of the test. It is determined by examining the relationship of the test to the variables the test is intended to assess as well as the relationship to those that should have no relationship to the domain underlying the instrument. Construct validity cannot be determined on the basis of any one study but is best demonstrated by an accumulation of supportive evidence from different sources over time. For example, the construct validity of an infant assessment instrument would be determined by demonstrating high positive correlations or a strong relationship with other measures of intelligence and a low correlation or minimal relationship with a measure of a presumably unrelated trait or variable, such as temperament.

INTERPRETATION OF TEST RESULTS

Criterion v norm referencing

Assessment instruments differ on whether test scores are left as raw scores or are translated into criterion-referenced or norm-referenced scores. Raw scores alone provide no stable point of comparison between individuals since scores might indicate what is a normal or expected performance on a particular test. To aid in interpreting the infant's score, the raw score can be translated. In criterion-referenced instruments, the raw score is translated into a statement about how that child's performance compares with a specified behavioral criterion established for that specific test. The principal use of criterion referencing has been in the development of mastery learning tests. The tests are designed to measure whether an individual has or has not attained mastery of a specific content area. The content of these criterion-referenced measures is usually designed to test acquisition of a relatively small domain of content and skills. Age-equivalent scores, mental age, ratio intelligence quotients, and some developmental quotients are examples of criterion-referenced scores. The advantage of these conversions is that they are readily understandable and are suitable when one is interested in whether an infant has acquired a specific skill. However, criterion-referenced tests should not be used if one is interested in comparing an infant's performance to the average performance of a group of infants in general. When this type of information is desired, norm-referenced instruments are more appropriate.

Most norm-referenced measures are based on an overview of some broad content domain (eg, intelligence), in contrast to the criterion-referenced tests that examine a specific skill (eg, fine motor). For example, the Early Learning Accomplishment Profile (Early LAP, see Appendix A), a criterion-referenced measure, is designed to provide a record of the child's existing skills in the major developmental areas. The Bayley Scales of Infant Development,[7] a norm-referenced instrument, provides a comparison of an infant's mental and motor abilities with those of other infants. Norm-referenced scores are most commonly reported as percentile ranks, z-scores, t-scores, stanines, and deviation intelligence quotients.

Correction for prematurity

Since an increasing number of the infants whose cognitive abilities are being evaluated have been born prematurely, another consideration in the interpretation of results is the issue of correction for prematurity. There continues to be debate over whether one uses an infant's gestational or chronologic age in the calculation of the scores and how long one uses this conversion. It has typically been the practice to correct for prematurity during the first two years of life.[18-20] However, there is growing evidence that this may not be the most appropriate practice. Caputo et al,[21] Siegel,[22] and others have indicated that correction for the degree of prematurity appears to be appropriate in the first few months or during the first year. However, a slightly more accurate prediction is achieved by using the uncorrected scores. As the Siegel[22] study points out, biologic immaturity alone is not the only consideration. Environmental influences, low birth weight, neurologic insult, among others, influence test scores. Furthermore, the degree of prematurity is a consideration. It might be more appropriate to correct for prematurity longer if the child was born at 28-weeks gestation than if the child was born at 34 weeks. In addition, as the medical conditions under which premature infants are born improve over time, correction for prematurity must be constantly reevaluated.

Whatever the correction method used, it is important to discuss with the child's family the child's current rate of development and what rate would be necessary for the infant to "catch up" in the future. Many parents have reported disappointment when an early prediction of catching up was not accurate. The box is an example of how the infant's rate of development can be explained.

Premature Infant's Rate of Development

Date of evaluation	January		June
Chronologic age (CA)	12 months	+ 6 months	18 months
Gestational age (GA)	10 months	+ 6 months	16 months
Cognitive development	9 months	+ 6 months	15 months
Percentage of delay (CA)	75%		83%
Percentage of delay (GA)	90%		94%

At the January evaluation this infant was only one month behind his or her gestational age and three months behind his or her chronologic age. If the infant gains one month of skills for each month of life until the next evaluation, he or she will still be three months behind his or her chronologic age, but the relative percentage of delay (75% v 83%) would be less. To catch up with his or her chronologic age by the next evaluation, the infant would need to gain nine months of skills in only six months time—an accelerated rate of development. Sharing this type of discussion with the family can be helpful in setting realistic expectations for the infant's development. Continuing to use a child's gestational age for some time (eg, until 3 years of age) without some type of discussion about the infant's rate of development can place the family at risk for bitter disappointment when the child does not make up all the delay. It can, in some cases, lead the family to be disappointed with what may in fact be a good rate of development (eg, month-for-month gain).

Testing situation

When choosing an assessment instrument and when interpreting the results the examiner should consider where the child is being evaluated. Evaluation results obtained during an assessment in familiar surroundings such as the child's home may be very different from those obtained on the same child in unfamiliar or potentially disruptive settings (eg, developmental evaluation clinic). These factors are critical in the interpretation of the results for a child of any age but particularly for an infant.

The timing of the evaluation also is an important factor. If one has any flexibility in scheduling a cognitive evaluation, it would be important to contact the family to determine when the infant normally naps or feeds. Trying to elicit the infant's interest in a structured task during a regular naptime is less likely to yield the best performance.

Since many infants receiving cognitive evaluations may have other health difficulties, possibly as the result of prematurity, the infant's health status should be considered. Was the infant fatigued following several developmental or medical evaluations? Does the infant have an ear infection or some other acute condition that may be affecting his or her performance? While it is not always possible to schedule the evaluation when these factors are not concerns, some mention of them in the interpretation of the results is important. Conferring with other professionals who have ongoing contact with the infant or who may have evaluated the infant earlier in the day or asking the parents if the infant's performance was representative or typical of what they see at home are ways in which the representativeness of the infant's performance can be estimated. Some statement about how valid the test results are thought to be should be included in every developmental evaluation report.

Parent's report

Asking parents about the representativeness of their infant's performance will automatically lead to the question of whether to include, and how to include, the parent's report in the interpretation of results. Historically, professionals' estimation of the validity of a parent report has been that parents overestimate their child's abilities. As a result a parent's report typically has not been included in the analysis of results or in the report.

Furthermore, the differences in perspective are not usually discussed or reconciled, often leading professionals to label the parents as denying the child's true developmental limitations and the parents to discount the examiner's results because they did not account for skills not displayed in the artificial testing situation.

There are many benefits for including a parent's report. First, it is a source of important and accurate information about the child's abilities. While referring to a parent report about parent-child interaction, Maccoby and Martin's' comment is applicable to the assessment of an infant's cognitive abilities as well.

> Using parents as informants has great potential advantages. For assessment of behavior that varies considerably across situations or behavior that is usually not displayed in public, reliable observational data are difficult to obtain and parent inter-views are often the only viable alternative. Parents have an opportunity to observe their children and the patterns of interaction in their families over extended periods

> of time in a broad range of situations. Thus by virtue of their daily participation in the family system, parents have access to a truly unique body of information about the family, and it is reasonable to tap into this information by questioning them.[23(p16)]

There is growing evidence that there is not much difference between the parent and the professional perspective, and where different the parents' perspective may be more accurate. For example, in one study by Honzik et al,[24] use of mothers' reports contributed to more accurate findings. Infants who were suspected, based on birth records, of having neurologic impairment vocalized with greater frequency during their cognitive evaluations at 8 months of age than was true of the normal control group. Honzik et al[24] concluded that the infants in the control group were more inhibited by the strangeness of the test situation and vocalized less in the test situation, thus failing the vocalization items on the Bayley. On the Griffiths and the Gesell scales, the mothers' reports of vocalizing would have been credited, thus leading to a more accurate representation of the infants' skills.

In many cases parents give the child more credit for emerging skills than do professionals and because of their ongoing contact with the infant may be in a better position to note whether the strangeness of the testing situation has hampered the child's test performance. Often the difference between a parent's report and a professional observation may not be in the behavior noted but in the interpretation of the behavior. Parents' definition of "talking" may be quite different from that of a professional. Nevertheless, a skilled interviewer can help the parent to identify what behavior they have interpreted as talking and to discuss the interpretation that the professional would make of this behavior. Another source of disagreement is whether the skill is well-established or whether the skill is emerging, with the infant relying more on the cues of the situation. Again, this is often seen in the interpretation of the infant's understanding of verbal requests. The parent may see the infant make anticipatory moves when the infant is told that he or she will be going outside. The parent reports that the infant understands this statement. However, when administered in the testing situation, without the behavioral environmental cues of the parent standing at the door, holding a coat, and reaching out his or her arms for the infant, the infant does not appear to understand. While the infant may not demonstrate enough understanding to be credited on a developmental evaluation, he or she is demonstrating emerging knowledge, and the correlation between the verbal command and the environmental cues is one way this knowledge of verbal commands can develop. Thus both the parent and professional are accurate, and both pieces of information are important in evaluating the child's development and in planning developmental evaluations.

A parent's report can be included in the developmental report in a number of ways. Some developmental assessments, such as the Brunet-Lezine,[3] do provide for the inclusion of a parent's report. Professionals may also want to score a developmental evaluation in two ways: (1) using only those behaviors that were observable in the testing, and (2) giving credit for a parent's report. In many cases the resulting developmental age or developmental quotient is not that different.

By including a parent's report, the professional has access to valuable information not readily accessible in the testing situation that can serve as a guide in planning interventions and in deciding whether the intervention should take place in the home or in a center-based program. Perhaps more importantly, the inclusion of a parent's report and discussion of different perspectives lays the groundwork for a collaborative relationship between the professional and parent rather than the adversarial one that can often develop when the parent feels the results are widely discrepant from his or her view. While there are some circumstances in which the parent's estimation is extremely overinflated, simply discounting the parent's perspective will not likely end this problem. A sensitive discussion of the behaviors in question will be more likely to lead to a mutual understanding. This agreement is essential if parents are to collaborate on the intervention as well.

USEFULNESS OF INFANT COGNITIVE ASSESSMENT

As mentioned above, the cognitive assessment of infants can serve many purposes. First, it provides an estimate of the way a child currently thinks about the world and processes information. This information, in turn, can be used in planning developmental interventions. Care must be taken that the developmental activity or

intervention (eg, block building) is not the same as the task used to assess the attainment of that skill (eg, builds tower of eight cubes). "Teaching the test," as it is sometimes called, not only leads to inflated test scores and splinter skills but really does not address or remediate the underlying delay or difficulty.

The results of a cognitive assessment can be helpful in other ways as well. Along with results of developmental evaluations in other areas (eg, gross motor, speech-language), it is a critical factor in basing expectations for the development of self-help or adaptive skills. For example, it would be inappropriate to consider beginning toilet training with a 3-year-old whose cognitive, gross motor, and speech-language skills are delayed at a 1-year level. The scores also are helpful in basing expectations for the child's behavior in areas such as attention span or interest in toys.

Professionals from other disciplines may find this information useful as well. In the absence of evaluations in other developmental areas, a skilled examiner can use the results of an infant's performance in a cognitive evaluation as a screen to identify those infants in need of additional testing of motor or speech-language skills. In addition, the results can be used in the planning of developmental interventions in other developmental areas.

The assessment of an infant is itself an intervention. Parental observations of the assessment can help parents and other family members gain an understanding of and insight into the child's needs and strengths. The information can, for example, enable parents to set expectations, to better understand the child's communicative cues, and to incorporate this information into the child's daily routines. The assessment process also provides an opportunity for parents and family members to ask questions and to clarify concerns about the infant's development. The Brazelton Neonatal Assessment Scale is a good example of how an assessment has been used in this manner.

Predictive validity

The results of a cognitive assessment can, in some cases, provide an estimate of the child's overall future intellectual abilities. How well evaluations of cognitive abilities during infancy predict performance on later intellectual tests has been a critical issue in the field since the inception of infant assessment. The method used to assess the predictive validity is interage correlations. As with most assessments, not just infant tests, the longer the time period between evaluations, the lower the correlation or relationship between the test scores. The result of these cumulative findings has been that in some circles infant cognitive assessments have not been regarded as predictive of later intelligence or that the tests were not really measuring cognition or intelligence.

Despite these views there are many theories and much research to support the predictive validity and overall usefulness of infant assessment. Ironically many of the early developers of infant assessment instruments, such as Gesell, indicated that they never intended tests to measure intelligence and therefore they should not be expected to predict later IQ. In addition, many theorists have stated that high correlations between infant assessment and the measurement of later intelligence should not be expected because infant intelligence may not be a unitary concept. Piaget and others have theorized that mental functioning undergoes qualitative changes over time, making the issue of prediction a moot point. Thus the skills being evaluated with infant tests (eg, sensorimotor), while important for later functioning, are not the same skills measured on tests of intellectual functioning at a later age (eg, visual-motor skills, verbal reasoning).

Perhaps the most convincing argument for the predictive validity of infant assessments is the research that documents that for infants with moderate to severe developmental delays the predictive validity is much greater.[13,25-29] These findings are consistent, regardless of whether the "low score is due to chromosomal aberrations (eg, trisomy 21), infection (rubella during pregnancy), injury, perinatal anoxia, or generalized abnormality of unknown etiology."[13(p97)] Honzik[13] also notes that while infant mental tests can be helpful in the diagnosis of neurologic lags or the effects of an impoverished or enriched environment, infant tests have more limited predictive validity when the infant scores are high. According to Honzik, "precocity in infancy may reflect early maturing or the effects of a great deal of stimulation rather than higher potential for later above-average cognitive functioning."[13(p97)]

There are many ways to further improve the diagnostic value of infant assessments. Periodic evaluations as well as the interpretation of the results in light of other developmental tests or related factors (eg, parent's socioeconomic status, degree of prematurity, health) can provide a more complete picture of the infant's current status and of those factors that may have long-term effects on the development of the infant's cognitive abilities.

FUTURE DIRECTIONS

Where is the field of infant cognitive assessment going? One direction seems to be a more thorough evaluation of other factors that inter-act with and relate to the development of cognitive abilities. The previous research on attachment and temperament continues along with more in-depth studies into parent-child relationships and factors that fall under the general heading of infant mental health.[30,31]

A second focus is on improving the ability of infant assessment techniques to predict development later on. McCall[25] and others have suggested several ways in which prediction can be improved, including using existing infant assessment techniques but augmenting them with additional information, such as the health status of the infant or developmental data on other related skills such as speech and language development; the use of frequent periodic assessments to identify abnormal patterns of development that can serve as markers for future developmental delay; or the inclusion of test items in the assessment battery that have been shown to be associated with specific abnormalities or risk conditions (eg, poor social responses, seizures, odd postures or movements).

A third area seems to relate to the age-old question of the continuity/discontinuity of development. There continues to be investigation into techniques that assess abilities in infancy that are assumed to be present in later childhood. The development of measures to assess sensory processing, such as attention, discrimination, and memory, seem to be the most promising at this time. One example of this approach is based on the early research by Fantz and Nevis' and others that demonstrated differences in visual processing among infants without developmental delay, those at risk, and those with identifiable delays. Fagan and his colleagues' have developed a screening device that assesses an infant's ability to discriminate among pictures being shown, to remember a picture previously seen, and to visually fixate to the novel stimulus. Because it is thought that perceptual and memory processes are necessary for successful performance on intellectual evaluations at a later age, documentation of an infant's competence in these areas were hypothesized to be predictive of later functioning. In fact, the available research on this screening instrument has yielded prediction rates clearly better for this approach than for other standardized assessment devices such as the Bayley scales. While the Fagan test does not obviate the need for more thorough developmental evaluations, it does show promise as a reliable and valid screening instrument and perhaps sets the stage for more assessment measures based on this theoretical perspective.

Another possible direction for the field of infant cognitive assessment is evident in the work of Ale[34] and Duffy et al,[35] which uses a combination of behavioral assessment and traditional neurophysiologic measures. Als' APIB,[34] the Assessment of Preterm Infant's Behavior, has been combined with Duffy and colleagues'[35] brain electrical activity mapping (BEAM) technique to examine the diagnostic value of this combined approach. The BEAM uses computerized topographic mapping to develop a visual picture of EEG and evoked potential (EP) data (for additional information see the article by Karniski in the "Technology" section of this issue). In one study by Duffy and Als[36] both measures were used to distinguish between two groups of infants—five neurologist-referred infants and one pediatrician-referred infant judged on the APIB to be the least behaviorally competent (eg, autonomic instability, difficulty in modulating tone, posture, and movement) from five pediatrician-referred infants who were judged to be the most behaviorally competent. These infants were then studied in a neurophysiology laboratory, and the results were used to generate topographic maps. The two groups were successfully distinguished by use of the BEAM. The authors conclude that the data suggest that direct observations of behavioral clustering do appear to have correlates in measures of brain electrical activity and that topographic display can be helpful in identifying differences between these behavioral clusters. Clearly more thorough investigation and development are needed, but the combination of techniques could prove to be a valuable technique in the early identification of those infants at risk for later neurologic and behavioral

difficulties, such as learning disabilities.

The field of infant cognitive assessment has come a long way from Gesell's first scale in 1925. While all the existing instruments can be criticized to some degree, there are a number of well-standardized, reliable, and relatively valid instruments available that are appropriate for use in documenting delay or risk status, for research, for planning intervention, or for predicting future development. Certainly continued refinement and restandardization with existing instruments are needed as well as the development of techniques for use with infants with special needs (eg, cerebral palsy, visual impairment). The recent work of Fantz and Nevis,[32] Fagan et al,[33] Als,[34] Duffy et al,[35] Duffy and Als,[36] and others reflect advancement in this area. As the need for more refined and additional infant assessment techniques grows, particularly with the establishment of more early interventions programs through Public Law 99-457 and the increased survival of infants with special needs become apparent, care must be taken so that the basics of good test construction, the need for solid testing skills and responsible test interpretation of the complexity of infant cognitive development are not disregarded.

REFERENCES

1. Gesell A: *The Mental Growth of the Preschool Child*. New York, Macmillan, 1925.
2. Knobloch H, Pasamanick B: *Gesell and Amatruda's Developmental Diagnosis: The Evaluation and Management of Normal and Abnormal Neuropsychologic Development in Infancy and Early Childhood*, ed 3. New York, Harper & Row, 1974.
3. Brunet O, Lezine PUF*: Le Developpement Psychologique del al Premiere Enfance. Issy-les-Moulineaux*, France, Editions Scientifiques et Psychotechniques, 1951.
4. Griffiths R: *The Abilities of Babies*. New York, McGraw-Hill, 1954.
5. Cattell P: *The Measurement of Intelligence of Infants and Young Children*. New York, Psychological Corporation, 1940.
6. Bayley N: Mental growth during the first three years. *Genetic Psychology Monographs* 1933; 14: 1-92.
7. Bayley, N.: *Bayley Scales of Infant Development*. New York, Psychological Corporation, 1969.
8. Kohen-Raz, R: Scalogram analysis of some develop-mental sequences of infant behavior as measured by the Bayley Infant Scale of Mental Development. *Gen Psychol Monographs* 1967; 76: 3-21.
9. Casati I, Lezine I: *Les Etapes de L'Intelligence Sensori-Motrice*. Paris, Editions du Centre de Psychologie Appliquee, 1968.
10. Corman HH, Escalona SK: Stages of sensorimotor development: A replication study. *Merrill-Palmer Q* 1969; 15: 351-361.
11. Uzgiris IC, Hunt JMcV: *Toward Ordinal Scales of Psychological Development in Infancy*. Champaign, Ill, University of Illinois, 1975.
12. Lewis M, Sullivan MW: Infant intelligence and its assessment, in Wolman BB (ed): *Handbook of Intelligence: Theories, Measurements, and Application*. New York, Wiley, 1985.
13. Honzik MP: Measuring mental abilities in infancy: The values and limitations, in Lewis M (ed): *Origins of Intelligence*. New York, Plenum, 1983, pp 67-106.
14. Brooks-Gunn J, Weinraub M: Origins of infant intelligence testing, in Lewis M (ed): *Origins of Intelligence*. New York, Plenum, 1983, pp 25-66.
15. *Hawaii Early Learning Profile Activity Guide* (HELP). Enrichment Project for Handicapped Infants. Palo Alto, Calif, VORT Corporation, 1979.
16. Cronbach LJ: *Essentials of Psychological Testing*, ed 3. New York, Harper & Row, 1970.
17. Brazelton TB: *Neonatal Behavioral Assessment Scale*. Clinics in Developmental Medicine, Philadelphia, Lippincott, 1973, No. 50.
18. Hunt JV: Predicting intellectual disorders in child-hood for preterm infants with birthweights below 1501 grams, in Friedman SL, Sigman M (eds): *Preterm Birth and Psychological Development*. New York, Academic Press, 1981, pp 84-102.
19. Hunt JV, Rhodes L: Mental development of infants during the first year. *Child Dev* 1977; 49: 204-210.
20. Parmelee A, Schulte F: Developmental testing of preterm and small-for-date infants. *Pediatrics* 1970; 17: 125-146.
21. Caputo DV, Goldstein KM, Taub HB: Neonatal compromise and later psychological development: A

10-year longitudinal study, in Friedman SL, Sigman M (eds): *Preterm Birth and Psychological Development*. New York, Academic Press, 1981, pp 36-73.
22. Siegel LS: Correction for prematurity and its consequences for the assessment of the very low birth weight infant. *Child Dev* 1983; 54: 1176-1188.
23. Maccoby EE, Martin JA: Socialization in the context of the family: Parent-child interaction, in Hetherington EM (ed): *Handbook of Child Psychology*. New York, Wiley, 1983, vol 4.
24. Honzik MP, Hutchings JJ, Burnip SR: Birth record assessment and test performance at eight months. 1965; 109: 416-426.
25. McCall RB: Toward an epigenetic conception of mental development in the first three years of life, in Lewis M (ed): *Origins of Intelligence*. New York, Plenum Press, 1976, pp 97-122.
26. Goodman J, Cameron J: The meaning of IQ constancy in young retarded children. *J Genet Psychol* 1978; 132: 109-111.
27. Vanderveer B, Scheweid E: Infant assessment: Stability of mental functioning in young retarded children. *Am J Ment Defic* 1974; 79: 1-4.
28. Drillien C: A longitudinal study of the growth and development of prematurely born children. Part VII: Mental development 2-5 years. *Arch Dis Child* 1961; 36: 233-240.
29. Ireton H, Thwing E, Gravem H: Infant mental development and neurological status, family socioeconomic status, and intelligence at age four. *Child Dev* 1971; 41: 937-945.
30. Field T: Affective and interactive disturbances in infants, in Osofsky JD (ed): *Handbook of Infant Development*, ed 2. New York, Wiley, 1987, p 972.
31. Emde RN: Infant mental health: Clinical dilemmas, the expansion of meaning, and opportunities, in Osofsky JD (ed): *Handbook of Infant Development*, ed 2. New York, Wiley, 1987, pp 1297-1320.
32. Fantz RL, Nevis S: Pattern preferences and perceptual-cognitive development in early infancy. *Merrill-Palmer Q* 1967; 13: 77-108.
33. Fagan JF, Singer LT, Montie JE, et al: Selective screening device for the early detection of normal or delayed cognitive development in infants at risk for later mental retardation. *Pediatrics* 1986; 78: 1021-1026.
34. Als H: Toward a synactive theory of development: Promise for the assessment and support of infant individuality. *Infant Ment Health J* 1982; 3: 229-243.
35. Duffy FH, Burchfiel JL, Lombroso CT: Brain electrical activity mapping (BEAM): A method for extending the clinical utility of EEG and evoked potential data. *Ann Neurol* 1979; 7: 412-420.
36. Duffy FH, Als H: Neurophysiological assessment of the neonate. An approach combining brain electrical mapping (BEAM) with behavior assessment (APIB), in Brazelton TB, Lester BM (eds): New Approaches to Developmental Screening of Infants. New York, Elsevier Science, 1983, pp 175-196.
37. Frankenburg W, Dodds J: *The Denver Developmental Screening Test*. Denver, University of Colorado Medical School, 1975.
38. Ireton HR, Thwing EJ: *Minnesota Child Development Inventory (MCDI).* Minneapolis, Behavior Science Systems, Inc, 1974.
39. Stillman R (ed): *The Callier-Azusa Scale*. Dallas, Tex, South Central Regional Center for Services to Deaf-Blind Children, 1977.
40. Bluma SM, Shearer MS, Frosham AH, et al: *Portage Guide to Early Education (No. 12).* Portage, Wis, Cooperative Educational Agency, 1976.
41. Glover ME, Preminger J, Sanford A: *Early Learning Accomplishment Profile* (Early LAP). Winston-Salem, Kaplan Press, 1978.

Appendix A
Selected infant assessment instruments
Psychometric instruments
Bayley Scales of Infant Development
The mental scale (162 items) measures sensory-perceptual acuities and discriminations; early acquisition of object constancy and memory, learning and problem-solving ability; vocalizations and the beginning of verbal communications and classifications. The motor scale (81 items) measures the degree of control of the body,

coordination of the large muscles and finer manipulatory skills of the hands and fingers. The tester must be a trained examiner. Test is standardized with established reliability and construct validity. It is appropriate for infants aged 1 month to 30 months.

Sensorimotor instruments

Infant Scales of Psychological Development

Inspired by Piaget's writings on infant intelligence, this instrument assesses an infant's achievement as a sequence of ordered levels of intellectual functioning. A series of six scales include visual pursuit and the permanence of objects, the development of means for obtaining desired environmental events, development of operational causality, the construction of object relations in space, the development of gestural and vocal imitation. The tester should have early intervention experience. Information concerning standardization, reliability, and validity is available in the manual. Appropriate for infants from birth to 2 years of age.

Screening instruments

Denver Developmental Screening Test[37]

Largely based on the Gesell schedules, the test (105 items) is an individually administered screening inventory designed to identify children with developmental delays. It measures four aspects of functioning: adaptive, fine and gross motor, language, and personal-social development. It yields an overall summary label. The tester can be a trained professional or paraprofessional. The test is standardized with established reliability and construct validity. Appropriate for infants and young children from birth through 6 years of age.

Minnesota Child Development Inventory[38]

This 320-item inventory provides a develop-mental profile based on summary scores for each of the following content areas: gross/fine motor, receptive/expressive language, person-social, and situation comprehension. The tester is the parent. The inventory has been standardized, and there is reliability and validity information.

Assessment instruments with special populations[39]

Callier-Azusa Scale

The scale (551 items) is composed of subscales in the following areas: motor development, perceptual and cognitive development, daily living skills, language development, and socialization. Each sub-scale is composed of sequential steps describing developmental milestones. Space is provided for comments by the teacher. The tester is the teacher/ interventionist. The test is an observationally based tool, yielding a developmental profile. The test has been standardized from other tests and reliability is available from the author. There is no information on validity. The scale is appropriate for use with children from birth to 4 years of age who are deaf, blind, or severely and profoundly disabled.

Assessment-based programming instruments Hawaii Early Learning Profile Activity Guide (HELP).

The HELP (585 items) assesses infants in six areas of development: gross/fine motor, expressive language, cognitive, social-emotional, and self-help. The HELP chart allows for a visual representation of the child's functional level in each domain. The accompanying activities guide provides suggested intervention activities accompanying specific developmental ages. Standardization is based on the items drawn from other tests, and there is limited information on reliability and validity. The test is appropriate for children from birth to 3 years of age.

Portage Guide to Early Education[40]

The Portage Guide is comprised of three parts: a checklist, a manual, and cards to be used in teaching behaviors included in the checklist. The checklist (580 items) is to be used as an assessment tool to pinpoint existing developmental strengths as well as areas of need. The tester is the interventionist, and there is limited information on standardization, reliability, and validity. Appropriate for children from birth to 6 years of age.

Early Learning Accomplishment Profile[41]

The Early LAP (412 items) is designed to provide the parent or teacher with a simple criterion-referenced

record of the child's existing skills in all the major developmental domains. It identifies the next appropriate step in development and gives detailed instructions for reaching this objective. The tester is unspecified. Standardization is based on items from other tests, and there is limited information on reliability and validity.

Appendix B
Checklist for assessing infant cognition
Purpose or goals of test

- Screening
- Assessment of cognitive abilities
- Assessment of all developmental areas
- Assessment-based programming instrument

Ages of children for which test is intended
Test administration

- Time to administer
- Who administers (eg, professional, paraprofessional, parent)
- Acceptability
- Cost
- Necessary testing conditions

Test construction

- Type of test or procedure (eg, interview, checklist, inventory, observation)
- Applicability of standardization sample to infant
- Reliability (eg, test-retest, alternate forms)
- Validity (eg, construct, content)

Test scoring and interpretation

- Optimal v typical performance
- Criterion-referenced v norm-referenced