One

OVERVIEW

NEPSY, A Developmental Neuropsychological Assessment (Korkman, Kirk, & Kemp, 1998) is a new neuropsychological instrument composed of 27 subtests designed specifically for children ages 3 to 12. It assesses five domains: Attention/Executive Functions, Language, Sensorimotor, Visuospatial, and Memory and Learning. NEPSY is based on the clinical methods of Luria and on more recent traditions of child neuropsychology. NEPSY offers the advantage of being able to assess a child across functions and modalities. All of the subtests were normed on a large sample of children balanced for age, gender, and parent education level according to the 1995 United States census. Therefore, differences in the child's test performance should reflect true discrepancies, because all subtests are standardized on the same population.

HISTORY AND DEVELOPMENT

Twenty years ago, the scarcity of neuropsychological instruments for children led Marit Korkman, a pediatric neuropsychologist from Finland, to develop *NEPS* (Korkman, 1980), a brief assessment for children 5.0 to 6.11 years old. Various aspects of attention, language, sensorimotor functions, visuospatial functions, and memory and learning were each assessed with two to five tasks similar in content to the tasks in Luria's assessment (Christensen, 1975). These were scaled as 0 to 1 or 0 to 2 in order to preserve Luria's approach. No sum scores were calculated. Although the method proved most useful, the narrow age range was problematic, as was the pass/fail criterion, which was built on the medical model (Korkman, in press).

The NEPS was revised psychometrically by adding more items so that the results could be expressed in total scores. These were converted to z-scores (mean = 0 ± 1) based on age norms. During this revision new subtests were added, derived from tests which had proven useful in pediatric neuropsychology (e.g., Benton, Hamsher, Varney, & Spreen, 1983; Boehm, 1986; Reitan, 1979;

Venger & Holmomskaya, 1978). To complement the test, the shortened versions of the Token Test (De Renzi & Faglioni, 1978), the Motor Free Visual Perception Test (Colarusso & Hammill, 1972) and the Developmental Test of Visual-Motor Integration (Beery, 1983) were used in their original forms and standardized along with NEPSY. Norms were collected for children 4.0 to 7.11 years old. Later norms were extended from ages 3.6 to 9.5. The assessment was called *NEPS-U* in Finnish and *NEPSY* in English (Korkman, 1988a, 1988b, 1988c). The Swedish NEPSY for children aged 4.0 to 7.11 was published in 1990 (Korkman, 1990), and the Danish version for the same age range was published in 1993 (Korkman, 1993).

In the spring of 1987, the decision was made to develop the present American NEPSY, while keeping in mind international needs. It was also planned to expand the age range, from ages 4.0 to 7.11 to ages 2 to 12. We began to collaborate on the present NEPSY, incorporating revisions and new subtests which were based on traditions and views central to contemporary neuropsychological traditions of assessment. To accommodate the development of new material for the expanded age range and to allow for the multicultural populations the test would be designed to serve, an extended period of development and standardization was needed.

During the pilot phase (1987–1989), the original NEPSY subtests were adapted and revised for 3- to 10-year-old children. New items were added, new subtests were developed, and some subtests based on the work of others, such as Fingertip Tapping and Phonemic Fluency (Denckla, 1973; Benton et al., 1983) were included. A pilot version of NEPSY with 41 subtests was administered to 160 children in New York, New Jersey, Connecticut, and Pennsylvania beginning in the fall of 1987. The sample was randomly selected from urban and suburban settings and was stratified for age (3, 5, 7, and 10 years), gender, and educational and socioeconomic backgrounds. Piloting was completed simultaneously in Finland.

By the spring of 1988, when the pilot data were reviewed, some subtests were eliminated, others were modified, and new subtests were developed based on the pilot studies, ongoing literature review, and clinical experiences with NEPSY in Finland and Sweden. Using new items and new subtests, more pilot studies were conducted in the United States and Finland prior to the American tryout. A bias review was undertaken by the Psychological Corporation with very favorable results. Only a few adjustments had to be made on the basis of that information before the tryout phase (1990–1994) began.

The tryout version of NEPSY was composed of 52 subtests. Under the auspices of the Psychological Corporation, it was administered nationwide in 1991 to 1992 to 300 children, 2 to 12 years of age. Due to rapid development in children between 2.0 and 4.0 years of age, participants were grouped at 6-month intervals in those age bands. The sample was also stratified by race/ethnicity, gender, geographical region, and parent education, as a reliable indicator of socioeconomic status. A second bias review was undertaken with a few modifications or deletions of items being made in an effort to keep NEPSY as bias-free as possible.

In the fall of 1992, after a thorough review of the tryout data, subtests with poor reliabilities, including those for 2-year-olds, were eliminated, and floor and ceiling problems were identified and addressed. Psychometric properties were reviewed, and scoring procedures were reviewed and modified. Additional tryouts of the revised and new subtests were completed in the United States and Finland from 1992 to 1994. Standardization of the Finnish edition of the expanded NEPSY (Korkman, Kirk, & Kemp, 1997) began at the end of this period, providing data to guide the American tryout. Subsequently, the items and subtests for standardization of the present NEPSY were selected.

The Psychological Corporation conducted the standardization and validation phase of the present NEPSY from 1994 to 1996. The standardization version of NEPSY consisted of 38 subtests. The standardization sample was 1,000 children (100 at each age level) between the ages of 3.0 and 12.11 years, stratified for race/ethnicity, gender, parent education, and geographical region. Oversampling of minority groups began. Five hundred of the children were part of the bias oversample and validity cases. Validation studies were conducted with clinical populations. When the standardization phase was complete, the final selection of subtests and items was made for each of the five domains. Core and Expanded subtests were selected. Two subtests, Picture Recognition and Memory for Pictures, were eliminated due to poor psychometric properties and high production costs. Orientation and Handedness were reframed as informal supplementary screenings. Components of several subtests were combined, but supplemental scores that preserved the unique information in each component (i.e., immediate; delayed) were included. Base rates for quantified Qualitative Observations were computed.

NEPSY, A Developmental Neuropsychological Assessment was published in January 1998 (Korkman, Kirk, & Kemp, 1998). Just prior to its publication, a corresponding version of NEPSY was published in Finland (Korkman, Kirk, & Kemp, 1997). A revised version was also published in Sweden (Korkman, Kirk, & Kemp, in press). The history of NEPSY publication is shown in Rapid Reference 1.1.

Rapid Reference 1.1

The History of NEPSY

Scandinavia				United States			
Year	Age Range (years)	Country of Publication	Author(s)	Year	Age Range (years)	Phase of Development (U.S.)	Author(s)
1980	5.0–6.11	Finland	Korkman				
1988	3.6–9.5	Finland	Korkman	1987–89	2.0–12.11	Pilot Phase	Korkman, Kirk, & Kemp
1990	4.0–7.11	Sweden	Korkman	1990–94	2.0–12.11	Tryout	Korkman, Kirk, & Kemp
1993	4.0–7.11	Denmark	Korkman	1994–96	3.0–12.11	Standardization	Korkman, Kirk, & Kemp
1997	3.0–12.11	Finland	Korkman, Kirk, & Kemp				
				1998	3.0–12.11	Publication	Korkman, Kirk, & Kemp
2000	3.0–12.11	Sweden	Korkman, Kirk, & Kemp				

THEORETICAL FOUNDATIONS

The theory of A. R. Luria has been a cornerstone of neuropsychology for nearly 40 years (Luria, 1962/1980). Luria conceptualized four interconnected levels of brain/behavior relationships and neurocognitive disorders that the clinician needs to know: the structure of the brain, the functional organization based on structure, syndromes and impairments arising in brain disorders, and clinical methods of assessment (Korkman, 1999).

At the structural and functional levels, Luria's concepts are based on his view of the brain as three functional units or "blocks." Block I is responsible for the basic physiological functions that support life, such as respiration and heartbeat, and for the arousal of attention that is necessary for cognitive functioning. These basic functions of Block I are subserved by the brainstem, diencephalon, and medial regions of each hemisphere. Block II refers to the posterior cortex: the occipital, parietal, and temporal lobes. These areas are purported to subserve the primary intake of information, the processing of that information, and the association of it with other information and experience. It is in this block that visual, auditory, and sensory information is received, processed, and associated across and within these modalities. Block III is purported to regulate the executive functions of planning, strategizing, and monitoring performance needed for efficient problem solving through rich connections to all areas of the brain. Block III regulates the use of information processed in Block II and is affected by, and modulates, the basic attentional/arousal function subserved by Block I (see Rapid Reference 1.2).

Working with adults, Luria also delineated brain regions that are interactively responsible for specific functions (e.g., in adults, motor programming of speech is dependent on left precentral and premotor neural systems, Broca's

Rapid Reference 1.2

Luria's Functional Blocks of the Brain

Block I: Basic physiological functions (brain stem, diencephalon, and mesial regions of each hemisphere)
- respiration, heartbeat, etc.
- arousal of attention

Block II: Primary intake of visual, auditory, and sensory information (posterior cortex: occipital, parietal, and temporal lobes)
- processing of information
- associations of that information and experience across modalities

Block III: Regulation of executive functions of planning, strategizing, and monitoring performance for problem-solving (frontal lobes)
- use of information from Block II
- affected by and modulates attention/arousal function of Block I

area). He viewed the brain as a "functional mosaic," the parts of which interact in different combinations to subserve cognitive processing (Luria, 1963, 1973). One area never functions without input from other areas; thus, integration is a key principle of brain function in the Lurian views.

Cognitive functions, such as attention and executive functions, language, sensory perception, motor function, visuospatial abilities, and learning and memory are complex capacities in the Lurian tradition. They are composed of flexible and interactive subcomponents that are mediated by equally flexible, interactive, neural networks (Luria, 1962/1980). In other words, multiple brain systems contribute to and mediate complex cognitive functions. Multiple brain regions, for instance, interact to mediate attentional processes (Barkley, 1996; Mirsky, 1996). The executive functions subserved by Block III regulate the basic attentional functions of Block I in sustaining optimal levels of arousal and vigilance and in the search for, selection of, and attention to relevant details from a broad array of information (Korkman, Kirk, & Kemp, 1998). For example, the executive function of inhibition makes it possible for a child to resist or inhibit the impulse to respond to salient, but irrelevant, features of a task (Denckla, 1996; Levin et al., 1991; Pennington, Groisser, & Welsh, 1993). Response inhibition allows the child to sustain focused attention throughout the period needed for task performance. (See Rapid Reference 1.3.)

When considering clinical methods and the levels of impairment in neurocognitive functioning, Lurian theory proposes that impairment in one subcomponent of a function will also affect other complex cognitive

Rapid Reference 1.3

Luria's Concept of Interactive Brain Function

- Multiple brain regions interact to mediate complex capacities.
- Complex capacities are composed of flexible, interactive subcomponents.
- Subcomponents are mediated by flexible, interactive neural networks.

Rapid Reference 1.4

Levels of Impairment in Neurocognitive Functioning

Impairment in one subcomponent of a function will also affect other complex cognitive functions to which that subcomponent contributes.

An early-occurring anomaly or event may well affect the chain of development in a basic subcomponent that occurs subsequent to impairment.

functions to which that subcomponent contributes. This is an especially important factor to consider in children, because an early-occurring anomaly or event may well affect the chain of development in a basic subcomponent that occurs subsequent to impairment. (See Rapid Reference 1.4.)

The Lurian approach bases its diagnostic principles on identifying the *primary deficit* underlying impairments of complex functions. For example, a deficiency in auditory decoding may underlie an aphasia. The latter is referred to as a *secondary deficit* (Korkman, 1995, 1999). Both impaired performance and qualitative observations are necessary to detect and distinguish between primary and secondary deficits (Korkman, Kirk, & Kemp, 1998; Luria, 1962/1980). (See Rapid Reference 1.5.)

> **Rapid Reference 1.5**
>
> **Primary and Secondary Deficits**
>
> **Primary Deficit:** a deficit underlying impaired performance in one functional domain (e.g., visuospatial deficit); several primary deficits can be present in different domains
>
> **Secondary Deficit:** a deficit in another functional domain arising from the primary deficit (e.g., visuospatial deficit causing a deficit in visual-motor integration for two- and three-dimensional constructions; a deficit in the comprehension of instructions based on visuospatial words; a deficit in mathematics)

At the level of clinical methods, Luria formulated explicit principles for an indirect, comprehensive review and evaluation of disorders of complex functions, which assesses subcomponents of these functions with carefully focused tests (Christensen, 1984). In accordance with this approach, NEPSY is composed of subtests that assess, as far as is possible, basic subcomponents of a complex capacity within a functional domain, as well as subtests that are designed to assess complex subcomponents that require contributions from several functional domains.

PURPOSES OF NEUROPSYCHOLOGICAL ASSESSMENT OF CHILDREN

For the child with brain damage that is either congenital (e.g., cerebral palsy, hydrocephalus, epilepsy) or acquired (e.g., traumatic brain injury, bacterial meningitis, tumor), neuropsychological assessment is valuable is assessing the effects of damage on brain function. It also evaluates the degree to which damage affects the capacity to process information in a functional domain, and, as a result, the development of competency in other domains (Christensen, 1984; Fischer & Rose, 1994; Levine, 1987; Luria, 1962/1980). Long-term follow-up for these children is as essential as the initial evaluation, because cognitive functioning may change with age (Casey, Rourke, & Picard, 1991; Morris, Blashfield, & Satz, 1996; Olson, Sampson, Barr, Streissguth, & Bookstein, 1992). The clinician needs to follow recovery of function in order to identify improved functioning, as well as persistent deficits, and to adapt interventions to changing needs (Korkman, Kirk, & Kemp, 1998).

Patterns of deficiencies in children with receptive and/or expressive language disorders and developmental disorders such as Autistic Disorders, Nonverbal Learning Disabilities, and Williams Syndrome, to name a few, can also be detected with neuropsychological assessment, thus assisting in diagnosis and intervention planning. Further, subtle deficiencies in children with less severe developmental disorders such as dyslexia, Attention-Deficit/Hyperactivity Disorder (ADHD), or graphomotor problems can be detected. Understanding these deficiencies facilitates the development of behavioral, educational, and cognitive interventions.

CAUTION

Inferences about Brain Pathology

Focal damage is more common in adults, whereas diffuse or multifocal damage is more common in children.

Lateralized or localized damage and neuropsychological findings in children are not usually evident in children with developmental disorders or early neurological insult.

Even with documented lateralized brain damage, the test profiles of children with left damage and with right damage do not differ.

Inferences concerning underlying brain pathology should be drawn with extreme caution, only by neuropsychologists who are trained in brain-behavior relationships.

CAUTIONS AGAINST LOCALIZING

The "working brain," as Luria (1973) termed it, is a very important Lurian concept. This concept is built on the notion that a highly interactive network of multiple brain systems contributes to and mediates complex cognitive functions. The development of more sophisticated imaging techniques, beginning with regional cerebral blood flow (rCBF) studies, through positron emission topography (PET) studies, to the more recent functional magnetic resonance imaging (fMRI) techniques, have demonstrated that many different brain regions are, indeed, activated simultaneously during complex cognitive activities.

Many factors must be kept in mind when assessing children and adults: age at time of event, neural plasticity, recovery of function, et cetera. Children and adults differ in that focal damage is more common in adults, whereas diffuse or multifocal damage is more common in children. For example, Multifocal and diffuse brain abnormalities are typical in very low birth-weight infants (Robertson & Finer, 1993). Such multifocal, diffuse damage also occurs in postasphyxial damage (Truwit, Barkovich, Koch, & Ferreiro, 1992), following exposure to teratogens (Miller, 1986; West & Pierce, 1986), in fetal alcohol exposure (Conry, 1990; Done & Rourke, 1995), and even in developmental disorders such as autism (Gillberg, Bjure, Uvebrandt, & Gillberg, 1993). In addition, because children are developing organisms, damage at a particular moment in time can affect both current and future neurocognitive development.

Although there is some evidence of lateralized or localized damage correlating with neuropsychological findings in children (Duane, 1991; Galaburda et al., 1985; Levin et al., 1994), for the most part, such relationships are not usually evident in children with developmental disorders or early neurological insult. Remarkably, even in children with verified lateralized brain damage, the test profiles of those with left damage and with right damage do not differ consistently (Aram & Ekelman, 1988; Korkman & von Wendt, 1995; Vargha-Khadem & Polkey, 1992). For these reasons, inferences concerning underlying brain pathology should be drawn with extreme caution, and only neuropsychologists who are trained in brain/behavior relationships in children should make such inferences, and even then with great caution. Those who are not trained in neuropsychology can still make extensive use of NEPSY by interpreting it at the cognitive processing level in order to develop modifications and interventions for children in the classroom.

PURPOSES OF NEPSY

NEPSY was developed with four interrelated purposes in mind: (a) to create a reliable and valid instrument sensitive to subtle deficiencies across and within the five functional domains that can interfere with learning in preschool and school-age children, (b) to contribute to understanding the effects of congenital or acquired brain damage, (c) to use in long-term follow-up of children with acquired or congenital brain damage or dysfunction, and (d) to study neuropsychological development in preschool and school-age children, as shown in Rapid Reference 1.6.

Rapid Reference 1.6

Four Purposes of NEPSY

- To be sensitive to the subtle deficiencies across and within the five functional domains and help to formulate interventions
- To aid understanding of the effects of congenital or acquired brain damage so interventions can be planned
- To use in long-term follow-up of children with acquired or congenital brain damage or dysfunction
- To study neuropsychological development in preschool and school age children

Rapid Reference 1.7

The NEPSY Model

- NEPSY has five domains: Attention/Executive Functions, Language, Sensorimotor, Visuospatial, and Memory and Learning.
- Each domain is composed of Core and Expanded Subtests.
- The Core Subtests (mean = 10±3) for each domain yield a Core Domain Score (mean = 100±15) for that domain.
- Subtest scores may be further divided into Supplemental Scores to allow further analysis of performance on the component parts of the subtest.
- Qualitative Observations at the subtest level can be quantified to provide further diagnostic information.

THE NEPSY MODEL

The NEPSY subtests are organized into five domains. Each domain includes Core and Expanded subtests from which clinicians can select additional individual subtests to answer specific referral questions. Thus, in a logical, step-down model, the Expanded Subtests are administered when deficits are revealed in the initial Core Assessment. The Core and Expanded Assessments may be administered in one session. Subtest scores provide a window to a specific function within a domain. The Core Subtest scores also contribute to the composite Core Domain Scores. Supplemental Scores allow for identifying different contributions to the performance (e.g., time vs. accuracy) in order to clarify the child's functioning. Qualitative Observations recorded at the subtest level can be compared to the responses and behavior of normally developing age-mates on the same task. Both Supplemental Scores and Qualitative Observations allow for elucidating the "how" of the child's performance. (See Rapid Reference 1.7.)

It is important to be aware that within the NEPSY model, the Core Assessments for ages 3–4 and ages 5–12 of the Core Domain Subtest differ somewhat. Therefore, the same Core Domain may reflect somewhat different functions between the two age groups.

CAUTION

NEPSY Subtests Differ for Age Ranges

Children 3 to 4 years of age take a different selection of subtests within each Core Domain than do the 5- to 12-year-olds.

Domain	Ages 3–4 only	Ages 5–12 only	Ages 3–4 and 5–12
Attention/ Executive	Statue*	Auditory Attention/Response Set (AARS)	Visual Attention (VA)
Language	Body Part Naming	Tower, Speeded Naming	Comprehension of Instructions
Sensorimotor	N/A	Fingertip Tapping	Phonological Processing, VP, Imitating Hand Positions
Visuomotor	Block Construction*	Arrows	Design Copying (DC)
Memory/ Learning	Sentence Repetition*	Memory for Faces, Memory for Names	Narrative Memory

*Use expanded subtests for ages 5–12

STRENGTHS OF THE NEPSY

1. A Large, Fully Representative Standardization Sample

Unlike many neuropsychological instruments developed prior to NEPSY, a large, fully representative sample of American children from 3 to 12 years of age was used to norm NEPSY. The authors used a stratified random sampling plan to ensure that representative proportions of children from each demographic group are included in the standardization sample. The standardization sample comprises 1,000 children in each of 10 age groups, 50 males and 50 females in each group. Children from four major geographic regions specified in the 1995 Census Bureau Report were selected to participate in the standardization sample in accordance with the proportions of children living in each region. The sample was stratified according to the following parent education levels (a) 11th grade or less, (b) high school graduate or equivalent to three years of college, and (c) four or more years of college. In each age group 3 to 12, Caucasians, African Americans, Hispanics, and other racial or ethnic groups were included in the same proportion as they are found in the United States population according to

Rapid Reference 1.8

Standardization Sample

- Unlike most neuropsychological instruments, NEPSY was standardized using
  —A stratified random sampling plan to ensure that representative proportions of American children ages 3 to 12 years from each demographic group were included in the standardization sample.
  —A standardization sample of 1,000 children in each of 10 age groups with 50 boys and 50 girls to a group.
  —Children from four major geographic regions specified in the 1995 Census Bureau Report in accordance with the proportions of children living in each region.
- Stratified parent education levels:
  —11th grade or less.
  —high school graduate or equivalent to three years of college.
  —four or more years of college.
- The correct racial or ethnic proportions were maintained for Caucasians, African Americans, Hispanics, and other racial and ethnic groups in the United States according to the 1995 census, not only within each age group, but also within each gender, geographic region, and parent education level.
- Any children with diagnosed neurological, psychological, developmental or learning disabilities were excluded (Korkman, Kirk, & Kemp, 1998).

the 1995 census. The racial or ethnic proportions were maintained, not only within each age group, but also within each gender, geographic region, and parent education level. Children with diagnosed neurological, psychological, developmental, or learning disabilities were excluded (Korkman, Kirk, & Kemp, 1998). This meticulous standardization process helped to insure that the norms were based on a representative sample of children in the United States. (See Rapid Reference 1.8.)

This is very important when the clinician is assessing a child whose disadvantaged background or racial experience might influence performance. A large representative sample insures that performance is not being compared only to that of white, middle-class children. The clinician might assume that a child is impaired when he or she actually has not had the same experiences as children from an advantaged background who may comprise an unrepresentative norm group. Reynolds (1997) has noted that too many of our neuropsychological data are based on impaired individuals. Because only normally developing children were included in this standardization sample, it allows for the assessment of pre-morbidly high-functioning individuals (e.g., IQs of 130) who suffer general cerebral trauma, but lose only 20 to 25 IQ points or less. Unless they were compared to a normally developing standardization group, they could appear normal and go untreated or even lose services (Reynolds, 1997). A fully representative standardization sample was a departure for pediatric neuropsychological assessments, and one that the authors feel ultimately will strengthen the field.

2. Over-Sampling of Minority Groups and Bias Review

NEPSY received two bias reviews in order to remove any language or stimuli that might be inadvertently offensive to any group, and a few modifications were made on this basis. Further, over-sampling of 300 children was included for minority groups.

3. Qualified Examiners Collected Standardization Data

Approximately 200 well-qualified examiners were selected to participate in the NEPSY standardization across the United States. Regional training workshops were conducted by the authors or by a member of The Psychological Corporation Development Team.

4. All Subtests Normed on the Same Standardization Sample

A major disadvantage of fashioning a neuropsychological assessment from numerous, brief instruments drawn from different sources is that many of them were normed on small, disparate groups of children. In these circumstances it is difficult to tell whether the between-test differences observed in performance are real differences in an individual child's functioning or merely reflect differences in the standardization sample composition. When the two normative groups are not equally representative, it is impossible to know reliably that a function assessed on a test with one norm group is a weakness for a child when it is compared to a function assessed on a second test with a different normative group. When the clinician sees that differences in a child's performance are evident on various subtests of the NEPSY, he or she can feel more confident that these reflect actual intra-individual differences, because the subtests were all normed on the same standardization sample.

5. Developmental Trends Can Be Observed

Although normative data on tests of a single function or circumscribed group of functions may reflect performance differences across age groups, they cannot reflect developmental differences in one complex, cognitive capacity that may affect development in another (Fischer & Rose, 1994). Because all capacities assessed on NEPSY were normed on the same group, the clinician can assess developmental differences, not only within a function, but across functions. Therefore, a great advantage of using NEPSY in research is the ability to observe developmental differences among different functional areas across age groups. In the normally developing child, one can see how some areas develop along the same trajectory and others show different developmental trajectories. The clinician can compare these developmental trajectories evident in normally developing children to differences in children with delays or deficits. It is important to note, however, that NEPSY subtests differ somewhat between the Core Assessment for ages 3 to 4 and the Core Assessment for ages 5–12.

6. Flexibility

Another advantage to NEPSY is its great flexibility (Kaplan, 1998). It permits assessment at a number of different levels in varying permutations, while retaining the unifying factor of a common normative group, no matter which combination of subtests is used. NEPSY permits a Core Assessment, a brief, overall view of five complex cognitive domains: Attention/Executive Functions, Language, Sensorimotor, Visuospatial, Memory and Learning. It also allows for an in-depth Full Assessment with all subtests, Core and Expanded.

Furthermore, it is possible to do an Expanded Assessment that allows a comprehensive look within a particular domain. After the clinician has administered the Core, and he or she sees that a particular function (language, for instance) shows significant subtest discrepancies, the clinician may administer all of the Expanded subtests within that domain in order to delineate further how pervasive the deficits may be. Finally, NEPSY allows for a Selective Assessment across domains after administering the Core. These may be suggested by performance on the Core, the referral question, or information in the child's history or other records. The clinician may wish to select subtests across domains that he or she knows from the research may elucidate symptoms of a particular disorder hypothesized to be the root of the child's problem. Because of the flexibility of NEPSY, a client-centered investigation using subtests selected to

Because of the flexibility of NEPSY, a client-centered investigation using subtests selected to address a specific referral question is possible (Kaplan, 1998).

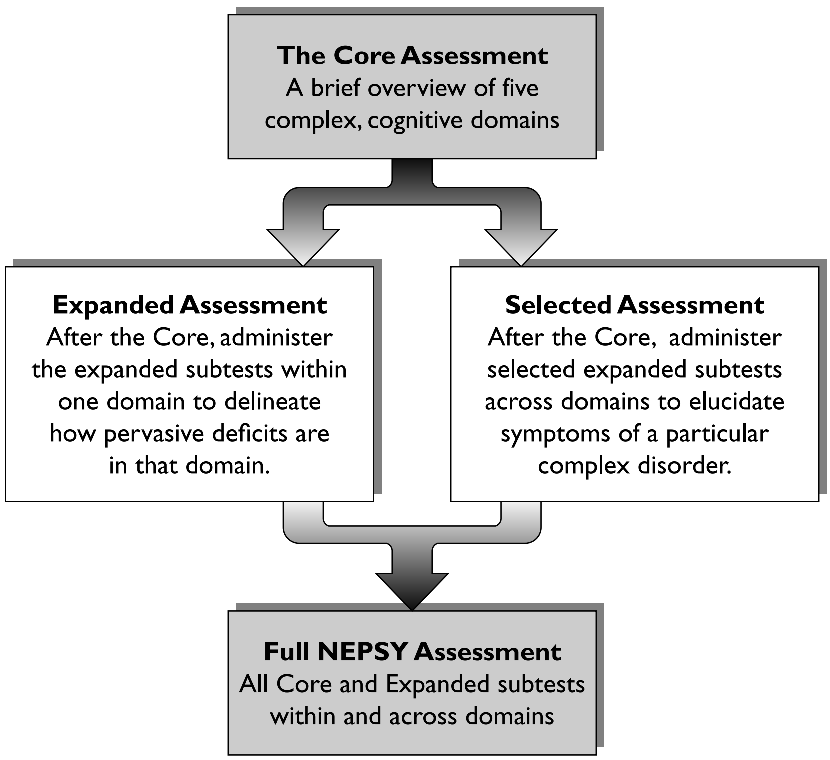

Figure 1.1 The Flexibility of NEPSY

address a specific referral question (Kaplan, 1998) is possible, as shown in Figure 1.1.

7. Standard Scores

The presence of standard scores is unusual on a neuropsychological assessment and raises all sorts of statistical difficulties, because the instrument that is designed to assess children who are impaired must be standardized on normally developing children. For example, a normally developing child at age 8 may easily do a task that the examiner knows from clinical experience an impaired 8-

year-old will not be able to perform. Therefore, the normally developing 8-year-old may ceiling on the task, making it impossible to derive a standard score, because the scores in the standardization sample are not normally distributed across age groups. Nonetheless, because children are developing, and development varies so much within any particular age group, the authors felt that it was important to have a standardized metric to take this into account. A cut score, which has been traditional in neuropsychology, cannot do this.

> Raw scores cannot be compared or assessed directly for a variety of reasons, the most potent being the lack of comparability of the raw score distributions among the tasks of the battery and for any one task across age (Reynolds, 1997).

All subtests in NEPSY that are normally distributed have a scaled score with a mean of 10±3. All Core subtests and many Expanded subtests have scaled scores. Core Domain Scores have a mean of 100±15. Because Subtest Scaled Scores and Core Domain Scores are each expressed in comparable metrics across functions and modalities, a summary of strengths and weaknesses, as well as discrepancies, can be presented in a performance profile (Korkman, Kirk, & Kemp, 1998). The use of these standardized scores helps to solve one of the most difficult current problems of neuropsychological assessment, that of using different brief instruments normed on different groups. Furthermore, because children are functioning within the educational arena, parents and teachers are used to this type of standard score and can understand its function. However, there is some distance to go before solving the scaling problems inherent in trying to develop a standardized pediatric neuropsychological instrument.

Reliability coefficients for NEPSY were calculated for each age separately. Split-half, test-retest, and generalizability procedures were employed, depending on the nature of the subtest. Average reliabilities for ages 3 to 4 for Core Domain Scores are Attention/Executive Function, .70; Language, .90; Sensorimotor, .88; Visuospatial, .88; Memory and Learning, .91. With the exception of two subtests, the subtest reliabilities for ages 3 to 4 range from .74 (Body Part Naming) to .91 (Sentence Repetition). The two subtests at ages 3 to 4 with poor reliabilities are Verbal Fluency (.59) and Statue (.50). Given the developmental variability in attention for very young children, the low reliability for Statue and its effect on the Attention/Executive Core Domain Score is not surprising. During early childhood, the ability to attend to a task is controlled primarily by the reticular activation system (RAS) of the brain. Prefrontal control of the arousal unit of the brain develops over time (Langus & Miller, 1992). At ages 5 to 12, average reliabilities for Core Domain Scores are Attention/Executive Function, .82; Language, .87; Sensorimotor, .79; Visuospatial, .83; Memory and Learning, .87. With the exception of two subtests, the subtest reliabilities for ages 5 to 12 range from .71 (VA; Fingertip Tapping) to .91 (Phonological Processing; Sentence Repetition). The two subtests at ages 5 to 12 with low reliabilities were Design Fluency (.59) and VP (.68), both influenced by graphomotor control that is well-developed in most normally developing children in the standardization sample. See Rapid Reference 1.9 for a summary of reliabilities.

8. Dissociation of Subcomponents of Deficits Possible When Comparing Subtest Performance

It is possible with NEPSY to dissociate components contributing to poor performance in several ways. At the subtest performance level, the clinician can assess attentional problems in the visual and in the auditory modalities. The examiner can assess deficits in phonological awareness by a simple task of recognition of speech sound segments within a word, a task requiring repetition of a nonsense word and a task which requires conceptualizing a sound pattern and manipulating that pattern to form a new word. The naming of body parts can assess a naming deficit in young children. Additionally, rapid naming can be assessed (SN). The clinician can also assess memory and learning of names, which may be related to a naming deficit, not to a memory deficit per se. Fluency can be assessed verbally (Verbal Fluency) and nonverbally (Design Fluency). The Manual Motor Sequences and Oromotor Sequences subtests allow the clinician to observe deficits in motor programming in two different ways (sequences on the Fingertip Tapping subtest also assesses motor programming. See the discussion of SS that follows). Block Design and Design Copying (DC) provide a means to assess constructional apraxia on two-dimensional and three-dimensional construction tasks. Likewise, visuospatial deficits can be dissociated from deficits in graphomotor precision on DC by comparing performance on Arrows and Route Finding, which are nonmotor visuospatial tasks, with performance on VP, a task requiring graphomotor control. The clinician can compare poor performances at the integrative level for visuospatial input and motor output to Imitating Hand Positions, which does not require the use of a pencil, and DC, which does. Fingertip Tapping and Finger Discrimination provide brief comparison of motor and sensory systems. Memory for Faces and Memory for Names provide comparison of two differ-

Rapid Reference 1.9

Reliabilities of Core Domain Scores and Core Subtest Scores

Core Domains and Subtests	Average r^2 Ages 3–4	Average r^2 Ages 5–12
Attention/Executive Function	.70	.82
Tower		.82
Auditory Attention and Response Set		.81
Visual Attention	.76	.71
Statue	.50	
Design Fluency		.59
Language	.90	.87
Body Part Naming	.74	
Phonological Processing	.83	.91
Speeded Naming		.74
Comprehension of Instructions	.89	.73
Repetition of Nonsense Words		.80
Verbal Fluency	.59	.74
Sensorimotor	.88	.79
Fingertip Tapping		.71
Imitating Hand Positions	.89	.82
Visuomotor Precision	.81	.68
Visuospatial	.88	.83
Design Copying	.86	.79
Arrows		.78
Block Construction	.80	.72
Memory and Learning	.91	.87
Memory for Faces		.76
Memory for Names		.89
Narrative Memory	.85	.77
Sentence Repetition	.91	.81
List Learning		.91

Note: Reliability coefficients for NEPSY were calculated for each age separately. Split-half, test-retests, and generalizability procedures were employed, depending on the nature of the subtest.

ent, important aspects of social functioning in children. The clinician can compare auditory short-term memory (Sentence Repetition) to recall of a quantity of details in a structured verbal narrative (Narrative Memory), to verbal learning over trials with short-delay and long-delay recall (List Learning). Thus, NEPSY provides a way to make comparisons between and among subtest performances in order to define more clearly the primary deficit related to the child's performance difficulties.

9. Supplemental Scores (SS)

Another important advantage to NEPSY is the inclusion of SS that provide a further means to tease apart a more global score in order to see what individual component may have adversely affected performance. For instance, the SN Total Score is a global score derived from a speed and accuracy table. However, because of the SS, each of these components can be considered separately. The clinician can see whether the child named accurately but had a problem with rapid access, or whether the child's accuracy in naming was the component affecting performance. Likewise, the clinician can observe whether the child's ability to focus attention was poor on a task of simple, selective auditory attention, a task of complex, shifting auditory attention, or whether he or she performed poorly on both tasks. Further, the clinician can note Omission Errors (OEs) and Commission Errors (CEs) on each auditory and visual attention task and across both simple and complex auditory and visual attention. These are but a few of the comparisons possible with SS on NEPSY.

10. Qualitative Observations (QO)

Closely related to the SS as an adjunctive means of analyzing functioning more diagnostically are the QOs. They provide a way of quantifying the qualitative observations of a child's performance that have traditionally been so important in process-approach assessment (Kaplan, 1988).

> The emphasis in U.S. psychology has been on the psychometric approach, so only recently have developmental neuropsychologists begun to collect data on the qualitative aspects of performance in both normal and abnormal development and attempted to relate these findings to neurologic models of brain development and to cognitive development. (Reynolds, 1997)

On NEPSY the clinician can tally the number of times a child makes a Rule Violation on Tower, for instance, and compare it to base-rate for age. This allows the clinician to take into account the child's development. Was it unusual, for example, for a child of 7 years to make five Rule Violations? Some QOs allow the clinician to examine a behavior across functions, as well. The presence of Recruitment, for instance, can be noted on SN, Verbal Fluency, and Manual Motor Sequences. This allows the clinician to see if the child is recruiting other systems into a task only on a motor task, only on a task that involves the access to or production of language, or on both.

11. Child-Friendly Tasks and Materials

Many pediatric neuropsychological instruments are renormed adult instruments meant to assess acquired brain damage and usually do not have materials specifically designed for children. It has been suggested that instruments meant to assess acquired brain damage in adults may not be suitable for assessing developmental disorders in children (Hynd & Willis, 1987; Kirk, 1983, 1985). The NEPSY subtests and colorful materials have been designed specifically to appeal to children. Because the child is engaged by the task and the materials, he or she will perform better. Although it is never possible to remove boredom and lack of attention completely from the testing situation, they will not be such confounding factors as they might be for a child performing a task meant for an adult. Furthermore, the subtests administration order provides a variety of different tasks in succession to help avoid boredom.

12. Ease of Administration

Materials needed, Start Points, Discontinue Rules, and Reverse Rules are consistently indicated with easy-to-spot icons in the *Manual* and Stimulus Book. The Stimulus Book, which is on an easel for ease of administration, contains the examiner script for the majority of tests. The Manual provides clear drawings when guidance is needed for hand positions, motor sequences, and so forth, and miniature drawings and text guidance are in the Record Form. Most tests can be administered from the Record Form and Stimulus Book. The Record Form is well-designed with plenty of space to record notes and miniature diagrams to record block constructions, sequence of responses on Comprehension of Instructions, and so forth. The record form provides reminders for time on immediate memory trials to ensure the delayed trial is administered at the specified time. All paper pencil items are grouped in a Response Booklet. Those assessing only preschoolers do not need to purchase Record

Rapid Reference 1.10

Strengths of NEPSY

- A large, fully representative standardization sample
- Bias review and over-sampling of minority groups
- 200 qualified examiners collected standardization data
- All subtests normed on the same standardization sample
- Developmental trajectories can be observed
- Flexibility
- Standard scores
- Dissociation of subcomponents of deficits possible when comparing subtest performance
- Supplemental Scores (SS)
- Qualitative Observations (QO)
- Child-friendly tasks and materials
- Ease of Administration

Forms for ages 5 to 12. Rapid Reference 1.10 summarizes the strengths of the NEPSY.

WEAKNESSES OF THE NEPSY

1. Subtests Not Highly Correlated with Core Domain Scores

The NEPSY was based on theory and traditions of neuropsychological assessment rather than on factor analysis. Factor structure inherently suggests that all of the subtests contributing to a particular factor are highly correlated. Therefore, they measure one construct. Neuropsychological functions are not that simple. Complex cognitive functions require contributions from different domains to a greater or lesser degree.

The subtests included in the Core Assessment were chosen based on theory because the combination of those few subtests was most apt to allow brief screening for different acquired and developmental disorders according to current research. For instance, the authors felt it was important to include memory assessments for both names and faces in the brief Memory and Learning Core Domain, because both are significantly important for screening for different types of disorders. Autistic children do not perform well on Memory for Faces, but learn names in a rote manner. Dyslexic children, on the other hand, may perform poorly on Memory for Names due to naming problems, but often have no trouble with Memory for Faces. The two subtests do not correlate, however.

Likewise, Memory for Faces is not correlated with Narrative Memory, but the authors felt the latter measure was an important subtest to include in the brief screening of memory function provided by the Memory and Learning Domain. It allows the examiner to observe the child's ability to recall large quantities of language, including essential details, that are required daily in school. Furthermore, it allows the clinician to determine whether the child is able to recall language details spontaneously or whether he or she must have cues in order to access the details from memory. This has direct implications for instruction. The child who has problems accessing information from memory will perform better on multiple choice tests on which cues are provided by the choices. In addition, Narrative Memory can provide evidence that the structure of the story aids recall for children with attentional problems and certain language disorders. Not including Memory for Faces in the Memory and Learning Domain, and instead including a memory test such as Sentence Repetition, would have caused the subtests within the domain to correlate much more highly. Essentially, however, only one modality would have been measured—verbal memory.

The disadvantage of having moderate and a few low correlations among the two–three subtests within a Core Domain of NEPSY is that if there is significant intradomain variance, the Core Domain Score will not reflect that area of functioning overall in a reliable way. Although that is true of all global scores, such as IQs, it is a principle to which clinicians far too rarely adhere. For example, the verbal subtests on the WISC-III may show significant intrascale discrepancies. The frequencies of those discrepancies may be significant, but the Verbal IQ will be incorrectly reported as reflecting overall verbal ability.

The Core Domain Scores of NEPSY are not meant to reflect a functional area overall. Because the domains were theoretically derived, they are meant to assess *diverse* aspects of a functional area which, based on the research, should be capable of screening for dysfunction on the basis of limited testing. The global Core Domain Scores are reliable (See Rapid Reference Box 5.2). As with most global scores, however, the authors feel that they do not provide the most valuable information for understanding intra-individual strengths and weaknesses. The individual differences are best defined by inspection and reporting of variations in subtest performance where the discrepancies and frequencies of those discrepancies are significant. Beyond the Core Assessment, NEPSY provides Expanded Subtests for further diagnostic testing in areas of concern. NEPSY does not, however, measure all possible constructs. It is very important to interpret findings on NEPSY in the context of results obtained from cognitive and achievement testing, medical, developmental, educational, and psychosocial histories, and the clinician's behavioral observations.

2. Complex Recording and Administration Procedures

Two subtests, AARS and Tower are somewhat difficult to record as the child performs. In both cases, it is just a matter of acquiring the skill. On Tower, the clinician does not actually have to record the child's moves, as long as he or she keeps track of the number of moves. AARS must be recorded, but the secret is to record the color the child touches beside the word on the Record Form that the clinician hears on the tape. The clinician should not try to make any judgement about correctness at that time and should not attempt to score as he or she goes. A little practice will help the clinician relax and score with ease. The hand movements for Manual Motor Sequences also take practice. Sitting so that the drawings in the *Manual* are visible and practicing the movements repetitively with the drawings in the NEPSY Manual (pp. 159–164) will help the clinician master them. The Record Form has written descriptions of the movements to aid administration.

3. Complex Scoring and Different Types of Scores

It is true that the NEPSY is time-consuming to score when the clinician wishes to glean all of the rich diagnostic information available in Expanded Subtests, SS and QOs. There is an intermediate scoring step for three subtests that have an accuracy and a speed component. The speed and accuracy raw scores are used to look up a derived speed/accuracy score. The latter is then used to locate the Total Score in the look-up tables. There is, however, a NEPSY Scoring Assistant (Psychological Corporation, 2000) available that makes scoring very simple. It also provides graphs and tables that can be imported into the child's report. NEPSY includes not only standard scores, but smoothed percentile ranks for selected Expanded Subtests and Cumulative Percentages for SS, and QOs. Percentile ranks and cumulative percentages were calculated for subtests on which the distribution was highly skewed or attenuated due to floor or ceiling problems. This occurs, for the most part, when the clinician is attempting to standardize neuropsychological functions which develop early (e.g., motor skills) or are not a problem in normally developing children (e.g., omission errors on attention tests). Unlike tests that will be given to normally developing children, such as cognitive or achievement tests, neuropsychological tests are designed to evaluate children who are impaired in various ways. It is important to include items for assessment of these impairments even when the clinician is aware that the normally developing children who form the standardization population will be able to perform the task well. This fact causes scaling problems, but to exclude necessary areas of assessment because they cannot generate scaled scores seemed unethical. Therefore, the solution appeared to be percentile rank ranges and cumulative percentages. Actually, 21 of the 27 subtests have scaled scores. All of the Core subtests and six of the Expanded subtests are scaled. The remainder of the Expanded subtests have smoothed percentile ranks. Three of the subtests have Supplemental Scaled Scores and the rest of the SS and all of the QOs as base rates of occurrence (percentage of the standardization sample demonstrating the behavior). Percentile ranks and cumulative percentages use consistent descriptors for performance level (i.e., *At Expected Level,* etc.). On the Domain Analyses page of the Record Form, the graphs have been constructed to allow visual comparison across all types of scores.

4. No Visual Memory Subtest

Memory for Pictures and Picture Recognition, which relied on viewing the final picture on the Memory for Pictures Subtest, were dropped after standardization due to high production costs and their poor psychometric properties. Just before standardization, we changed the picture stimulus from a series of incomplete line drawings that progressed to full drawings of the objects, and substituted computer-generated degraded photographs, progressing to fully visible photographs of the objects. Although they were piloted on a limited basis, it was not sufficient, to produce a reliable subtest. Further, production costs were high.

5. No "Crack-Back" *Manual* with Dividers

We have received feedback from various examiners concerning the paperback *Manual.* It does not make into an easel, as would a "crack-back" manual, and it does not have dividers, as do a number of test manuals. This was an economic decision. A "crack-back" manual adds a significant amount to the cost of a test, and in this day of rapidly rising prices for test kits, the authors wanted very much to keep the cost of NEPSY within reason. The *Manual* does have a special binding, however, which allows the clinician to open it and press it flat without breaking the back. Once the clinician has learned the test administration well, however, he or she will find very little need for the *Manual* during testing, as most instructions are in the Stimulus Book, on the Record Form, or both. The problem of dividers has been addressed with printed side tabs. If holding the manual horizontally, one can see blue printing visible on the edge of pages in certain sections. The thick blue section is the Administration and Scoring Directions. Toward the back of the book are thinner blue sections that indicate the appendices, starting with the marker for Appendix A at the top of p. 269. The markers continue in descending positions for each appendix down to the marker for Appendix D at the bottom of p. 343. The markers start again at the top of p. 361 with Appendix E and again move progressively down to Appendix H on p. 491. Where applicable, there are top tabs of blue to indicate ages as well. Although the NEPSY *Manual* is not the more expensive version, significant effort was made to help clinicians navigate the manual easily. Rapid Reference 1.11 summarizes the weaknesses of NEPSY.

Rapid Reference 1.11

Weaknesses of NEPSY

- Subtests not highly correlated with Core Domain Scores
- Complex Recording or Administration Procedures
- Complex Scoring and Different Types of Scores
- No visual memory subtest
- No "crack-back" manual with dividers

SUMMARY

Although NEPSY has some weaknesses, as do most assessments, it also has significant strengths. It is an attempt to link the best of clinical, process-approach, neuropsychological assessment with psychometrics, without losing too much of value on either front. This is an enormous challenge and one which the authors intend to continue to accept. The next chapter presents a comprehensive discussion of how to administer

Rapid Reference 1.12

The Publisher of NEPSY

NEPSY—A Developmental Neuropsychological Assessment.
Copyright 1998 by

The Psychological Corporation
555 Academic Court
San Antonio, TX 78204

www.PsychCorp.com

NEPSY. It includes a general discussion of good assessment practice, instructions on how to modify NEPSY for special populations, and comprehensive subtest-by-subtest rules for administration.

COMPREHENSIVE REFERENCES

The manual of *NEPSY, A Developmental Neuropsychological Assessment* (Korkman, Kirk, & Kemp, 1998) and the references for this guide provide comprehensive lists of references on NEPSY. The *Manual* also reviews studies performed with NEPSY thus far. The *Manual* further reviews the development of the test and contains descriptions of each subtest, the Core Domains, and standardization, reliability, and validity. Rapid Reference 1.12 gives publication information.

TEST YOURSELF

Fill in the blanks.

1. **NEPSY assesses children in what age range? ____________________**
2. **Upon what theory is NEPSY based? ____________________**
3. **The theory upon which the NEPSY is based proposes that impairment in one subcomponent of a function is likely to affect ____________________ to which that subcomponent contributes.**
4. **When many brief instruments are drawn from different sources and their norm groups are different, it is difficult to tell whether differences in performance merely reflect differences in the ____________________ ____________________.**
5. **Because all capacities assessed on NEPSY have been normed on the same group, ____________________ trends can be assessed, both within and across a function.**
6. **NEPSY permits assessment at a number of different levels. A ___________ is a brief overview of functioning across domains. That might be followed by a ____________________ ____________________ across domains or an ____________________ ____________________ within a particular domain.**
7. **The flexibility of NEPSY permits a ____________________ ________________ investigation (Kaplan, 1998).**
8. **Core Domain Scores on NEPSY have a mean of ____________________, and scaled scores have a mean of ____________________.**
9. **SS allow ____________________ of subcomponents of deficits.**
10. **A brief but thorough assessment of neuropsychological functions can be obtained through a ____________________ ____________________.**

11. Match the Block to the appropriate function.

Block I ___ (a) Subserves the primary intake of information, the processing of that information, and the association of it with other information and experience.

Block II ___ (b) Regulates the executive functions of planning, strategizing, and monitoring performance that are needed for problem solving.

Block III ___ (c) Is responsible for the basic physiological functions that support life, and for the arousal of attention.

12. List the five domains of the 27 NEPSY subtests.

(a) ____

(b) ____

(c) ____

(d) ____

(e) ____

13. NEPSY is an appropriate instrument for assessing localized brain damage. True or False?

14. Secondary deficits are so named because they are not as important as primary deficits. True or False?

15. The NEPSY standardization sample was 500 children drawn from four geographical regions. True or False?

16. The NEPSY standardization sample included children of three parent-education levels. True or False?

17. There were 50 males and 50 females to an age group in the standardization sample. True or False?

18. The 1992 Census Bureau Report was used to determine inclusion of the proportion of racial or ethnic groups in the standardization sample that was found in the U.S. population. True or False?

19. A stratified random sampling plan was used. True or False?

20. The performance of a high-ability child with a mild generalized cerebral trauma is best assessed by comparing performance to other children with mild brain injury. True or False?

21. A bias review insures that a test maintains its restricted cultural approach. True or False?

For questions 22–25, pick the correct letter from the selection below.

22. It is unusual for a pediatric neuropsychological assessment to have ____________________.

23. With two exceptions, the average subtest reliabilities at ages 3 to 4 range from ____________________.

(*continued*)

24. With two exceptions, the average subtest reliabilities at age 5 to 12 range from ___________________.

25. Global scores do not reflect an overall function if there are significant discrepancies between the ___________________.

(a) subtest scores
(b) .74–.91
(c) standard scores
(d) .79–.90
(e) .71–.91

Answers: 1. Ages 3 to 12; 2. Lurian; 3. Any function; 4. In norm groups; 5. Developmental; 6. Core Assessment, Selective Assessment, Expanded Assessment; 7. Client-centered; 8. 100, 10; 9. dissociation; 10. Core Assessment; 11. Block I: c, Block II: a, Block III: b; 12. a. Attention/Executive Functions, b. Language, c. Sensorimotor, d. Visuospatial, e. Memory and Learning; 13. False; 14. False; 15. False; 16. True; 17. True; 18. False; 19. True; 20. False; 21. False; 22. c; 23. b; 24. e; 25. a